Table of contents

Supplementary Discussion

High reductase/genome Eggerthellaceae, Erysipelotrichaceae, and Burkholderiaceae clades exhibit dynamic evolutionary trajectories

To gain insight into the evolutionary history of reductase expansion in Eggerthellaceae, Burkholderiaceae, and Erysipelotrichaceae, we performed a phylogenomic analysis of all classified species in each family (including those not associated with the human gut microbiome) and overlaid the reductases per genome data on the resulting tree (**Extended Data Fig. 1A-C, Supplementary Table 3**). 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* (**Extended Data Fig. 1A-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 (**Extended Data Fig. 1A-B)**. These observations are consistent with independent evolutionary trajectories characterized by dramatic expansions in the reductase gene content in ancestral Eggerthellaceae, Burkholderiaceae, and Erysipelotrichaceae species. Furthermore, the variable retention of reductases in descendant lineages suggests that high reductase content genomes confer a context-dependent 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 species (*E. lenta*, *S. wadsworthensis*, and *H. filiformis*) from the Eggerthellaceae, Burkholderiaceae, and Erysipelotrichaceae families. 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 (**Extended Data Fig. 1D**). 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 strain dependent.

Comparison of E. lenta, S. wadsworthensis, and H. filiformis electron acceptor usage patterns

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.[1] 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* (**Fig. 3A-C**, **Extended Data Figs. 2 & 4**). By contrast, *E. lenta* was unique in accommodating carboxyl group modifications, utilizing the cinnamate esters chlorogenate and rosmarinate (**Fig. 3A-C, Extended Data Fig. 2**).

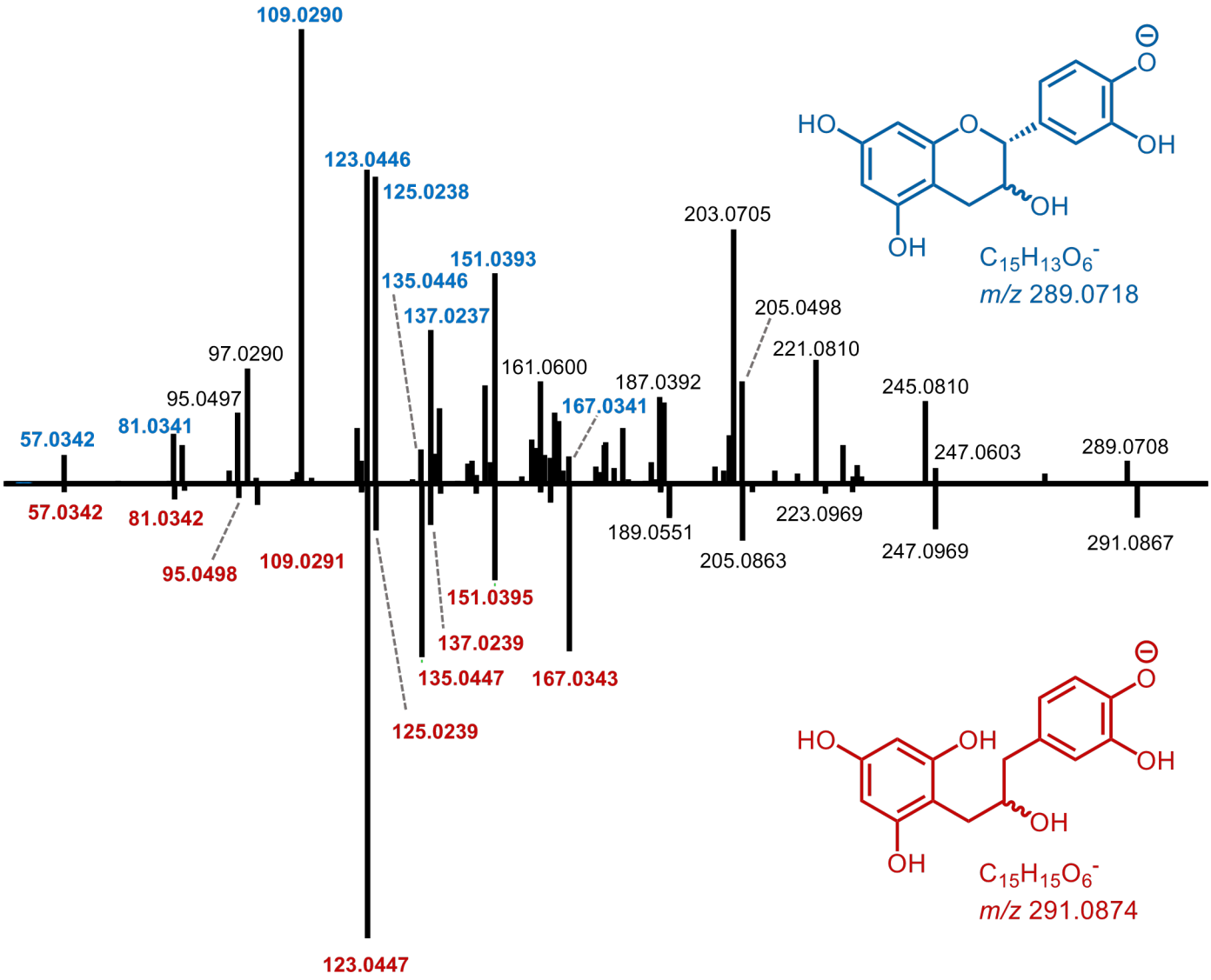
In addition to exhibiting distinct substrate specificities, we observed that *E. lenta* and *H. filiformis* modified cinnamates in different ways. While *H. filiformis* exclusively reduced the enoate group, *E.*

lenta also reductively dehydroxylated catechol cinnamates – presumably through the action of a previously characterized hydrocinnamate dehydroxylase (**Fig. 3B**, **Extended Data Fig. 7A & 7B**).[2] 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 (**Fig. 1D, Extended Data Fig. 1C**).[2] *E. lenta* however is capable of both dehydroxylating the catechol caffeate to *m*-coumarate and reducing its enoate to hydrocaffeate (**Extended Data Fig. 7A & 7B**). 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) (**Extended Data Fig. 7A & 7B**). Independent experiment with *m*-coumarate and hydrocaffeate support the conclusion that both enoate reduction and catechol dehydroxylation enable respiratory growth (**Extended Data Fig. 7A & 7B**). These results thus reveal a complex and variable use of cinnamates as respiratory electron acceptors by *E. lenta* and *H. filiformis*.

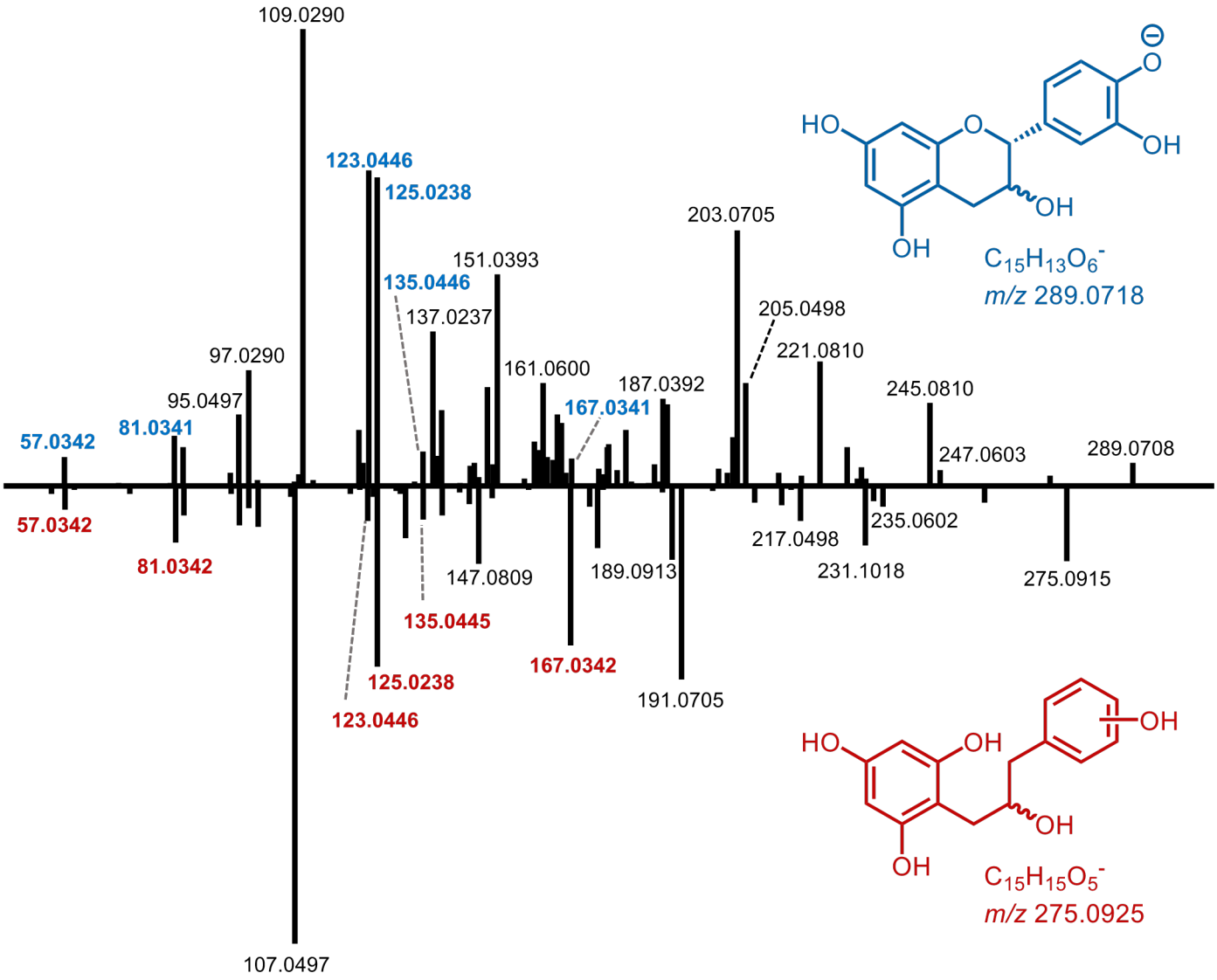
4-carbon-dicarboxylates. Four-carbon (C4)-dicarboxylates (including intermediates in the tricarboxylic acid cycle, bacterial fermentative products, and amino acids) have been shown to be relevant growth substrates within the gut microbiome.[3] We observed that *S. wadsworthensis* selectively utilized the C4-dicarboxylates aspartate, fumarate, malate, and tartrate as respiratory electron acceptors (**Fig. 3A-C, Extended Data Fig. 3**). Each of these C4-dicarboxylates was reduced to succinate, suggesting a metabolic pathway that converges upon fumarate similar to the previously described Enterobacteraceiae mechanism (**Fig. 3C, Extended Data Fig. 3**).[3] 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.[4] Dietary plant phenols and conjugated catecholamines (dopamine, epinephrine, norepinephrine) excreted in the bile are significant sources of catechols within the gut.[5] We observed that *E. lenta* and *S. wadsworthensis* reduced multiple sulfoxides (dimethyl sulfoxide, methionine sulfoxide, and methylcysteine sulfoxide) and that *E. lenta* dehydroxylated catechols (hydrocaffeic acid, catechin, epicatechin) (**Fig. 3B & 3C, Extended Data Figs. 2, 3, & 5**). Consistent with molybdoprotein reductases being responsible for sulfoxide reduction, we observed sulfoxide growth-promoting properties were inhibited by the established molybdoprotein reductase inhibitor tungstate (**Extended Data Fig. 7C & 7D**).

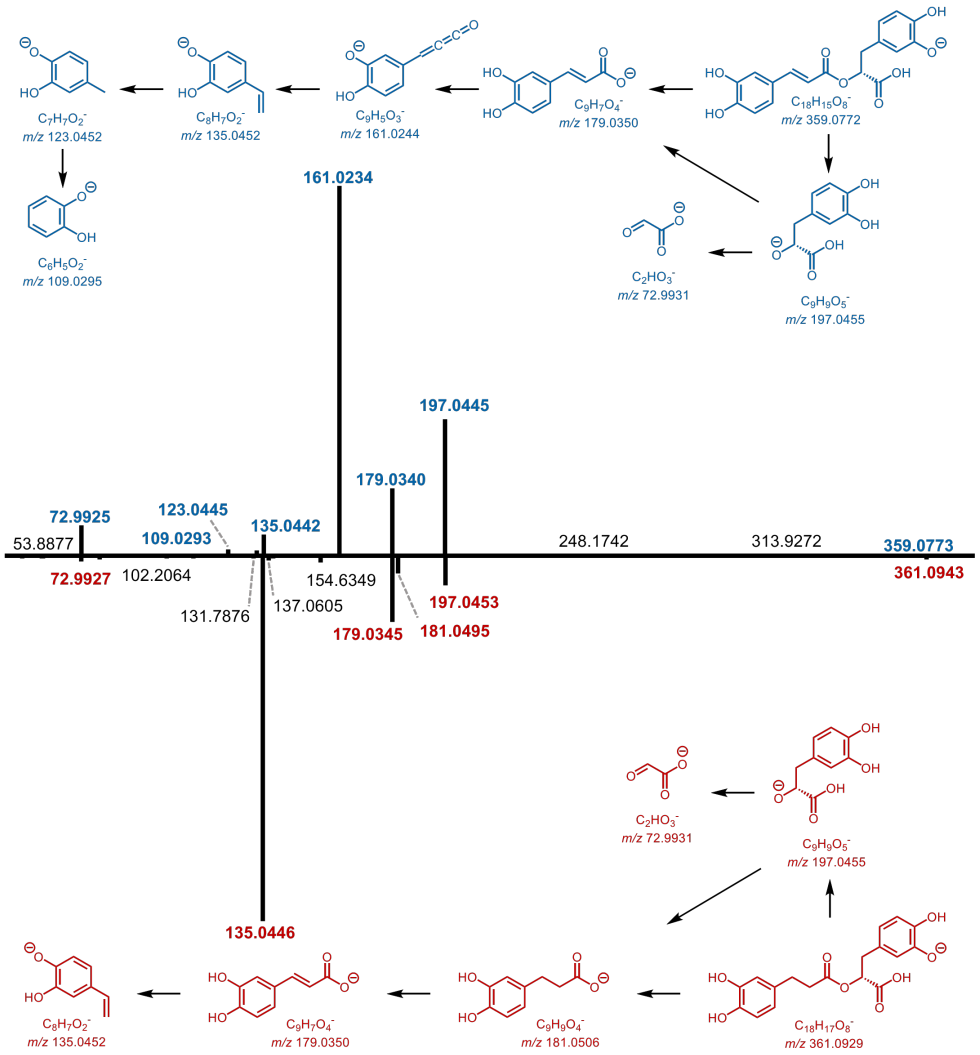
Flavonoids. Flavonoids are a class of plant polyphenolic metabolites with various dietary sources. The flavonoid catechin has previously been shown to undergo a two-step reduction.[2] First, the carbon-oxygen bond in an internal benzyl ether group is reductively cleaved. While the enzyme responsible for this activity has not been identified, an *E. lenta* flavin reductase was reported to catalyze a similar benzyl ether cleavage in a related plant polyphenol.[6] After benzyl ether cleavage, catechin is reductively dehydroxylated by the molybdopterin reductase catechin dehydroxylase.[2] While we failed to detect a respiratory growth phenotype for any of our strains in the presence of catechin or the related flavonoid, epicatechin, we confirmed that *E. lenta* reduced benzyl ether and catechol groups in both compounds (**Extended Data Fig. 5B**). We further found that *S. wadsworthensis* cleaved the benzyl ether bonds in catechin and epicatechin but lacked detectable dehydroxylation activity (**Extended Data Fig. 5B**). Consistent with a respiratory role of flavonoid catechol dehydroxylation (but perhaps not benzyl ether cleavage), we observed that epicatechin promoted *E. lenta* ATP production but had no effect on *S. wadsworthensis* ATP levels (**Extended Data Fig. 5C**).



Supplementary Figure 1. Butterfly plot of LCMS-MS fragmentation spectra for the catechin/epicatechin starting material at top in blue and putative catechin/epicatechin "derivative 1" product in red at bottom. Characteristic fragments that might support assignment of the reduction product would be indistinguishable isomers of fragments from catechin/epicatechin, so fragment ions that simply match between the spectra are colored and bold to support the products' relatedness to the starting material. These matching fragments and the difference in two Dalton between the precursor ions of the starting material and product support the assignment of m/z 291.0874 as catechin/epicatechin with the B-ring cleaved open, also referred to as "derivative 1". The fragmentation spectra for catechin and its "derivative 1" are the same as the corresponding epicatechin and its "derivative 1", so both stereoisomer pairs are represented in this single figure.

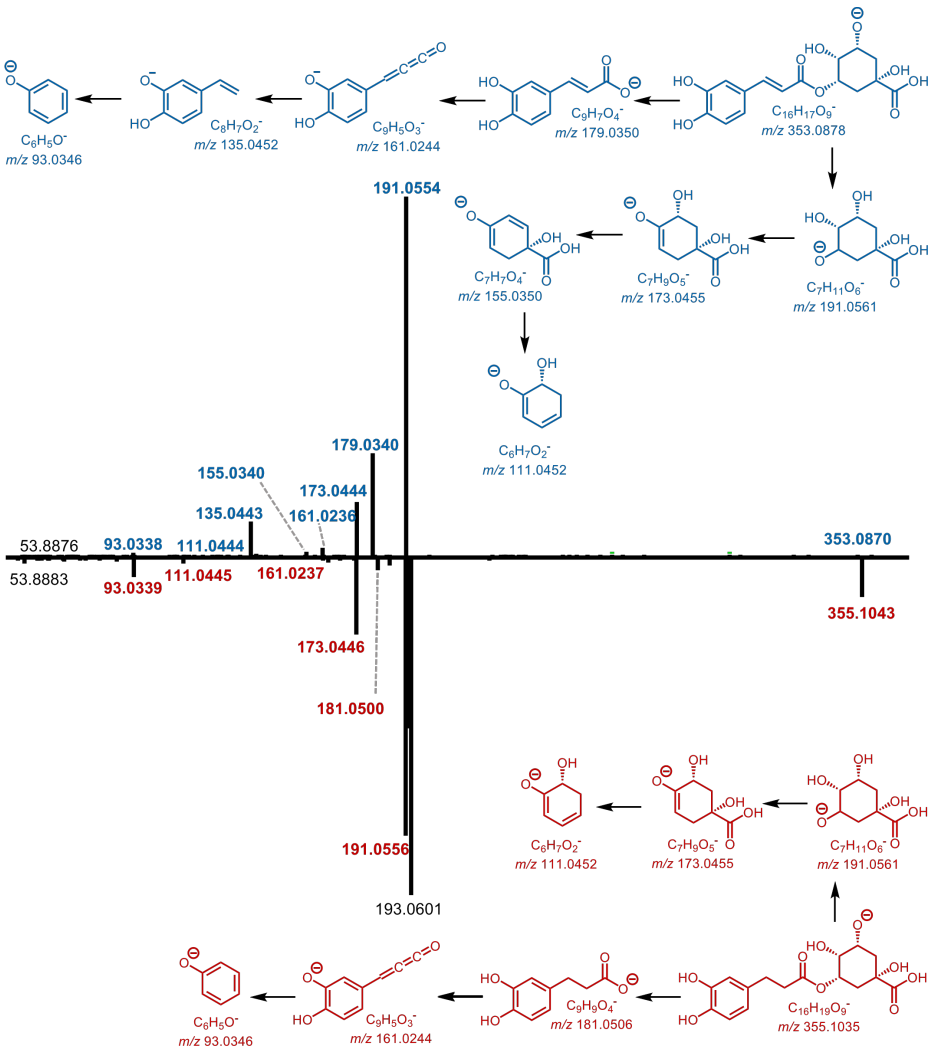


Supplementary Figure 2. Butterfly plot of LCMS-MS fragmentation spectra for the catechin/epicatechin starting material at top in blue and the putative catechin/epicatechin "derivative 2" product in red at bottom. Characteristic fragments that might support assignment of the reduction product would be indistinguishable isomers of fragments from catechin/epicatechin, so fragment ions that simply match between the spectra are colored and bold to support the products' relatedness to the starting material. These matching fragments and the difference in 14 Dalton between the precursor ions of the starting material and product support the assignment of m/z 291.0874 as catechin/epicatechin with the B-ring cleaved open and dehydroxylation, also referred to as "derivative 2". The fragmentation spectra for catechin and its "derivative 2" are the same as the corresponding epicatechin and its "derivative 2", so both stereoisomer pairs are represented in this single figure.



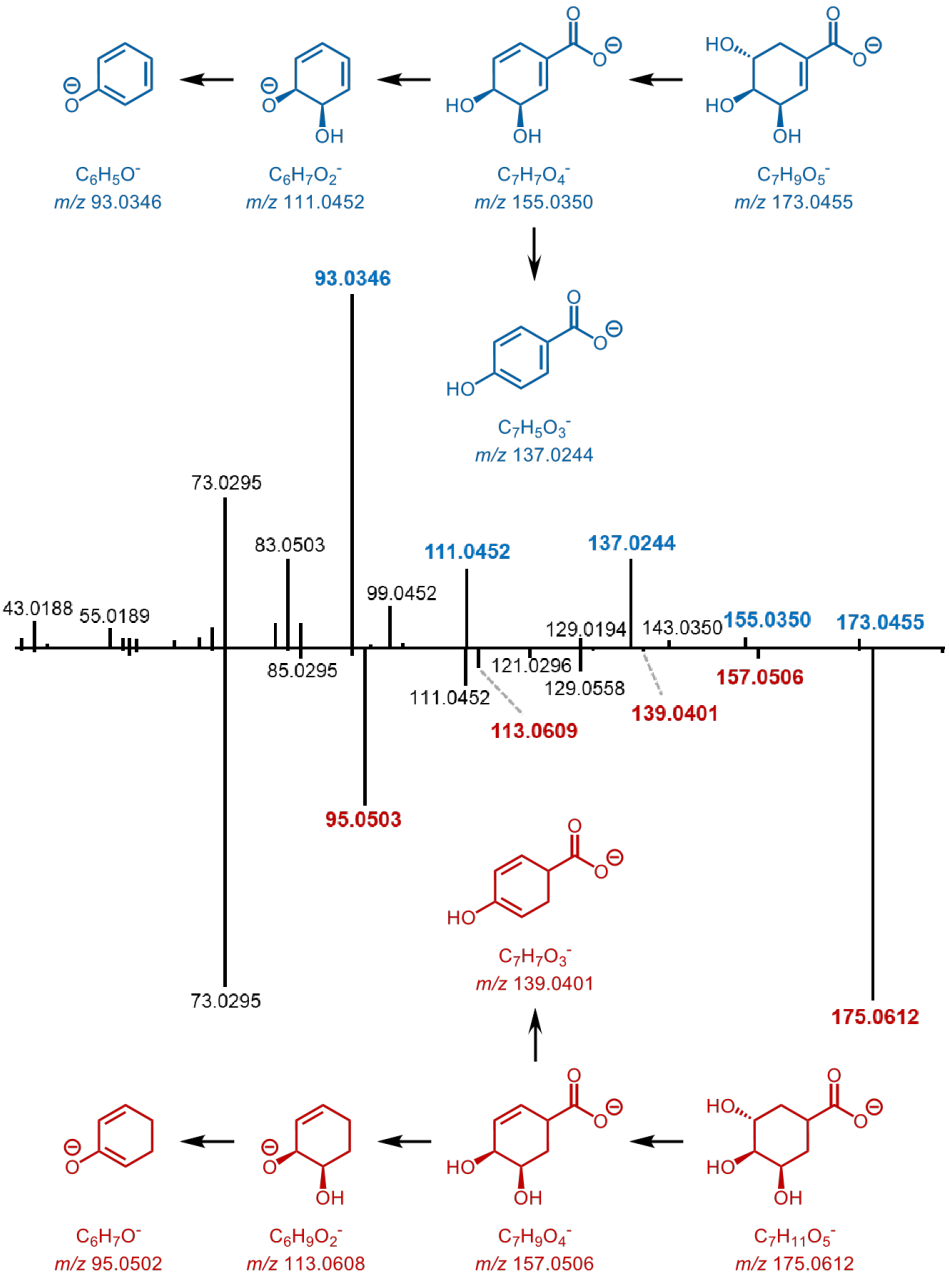
theoretical m/z	observed m/z	error (ppm)	molecular formula
rosmarinate			
359.0772	359.0773	0.28	C18H15O8-
197.0455	197.0445	5.07	C9H9O5-
179.0350	179.0340	5.59	C9H7O4-
161.0244	161.0234	6.21	C9H5O3-
135.0452	135.0442	7.40	C8H7O2-
123.0452	123.0445	5.69	C7H7O2-
109.0295	109.0293	1.83	C6H5O2-
72.9931	72.9925	8.22	C2HO3-
hydro-rosmarinate			
361.0929	361.0943	3.88	C18H17O8-
197.0455	197.0453	1.01	C9H9O5-
181.0506	181.0495	6.08	C9H9O4-
179.0350	179.0345	2.79	C9H7O4-
135.0452	135.0446	4.44	C8H7O2-
72.9931	72.9927	5.48	C2HO3-

Supplementary Figure 3. Butterfly plot of LCMS-MS fragmentation spectra for rosmarinate at top with key putative fragment ions highlighted in blue and hydro-rosmarinate at bottom with key putative fragment ions highlighted in red. The difference in two Dalton between the intact ions and the caffeate and hydrocaffeate fragment ions (*m/z* 179.0340 and *m/z* 181.0495, respectively) in these spectra support the assignment of *m/z* 361.0943 as hydro-rosmarinate as shown.



theoretical m/z	observed m/z	error (ppm)	molecular formula
chlorogenate			
353.0878	353.0870	2.27	C16H17O9-
191.0561	191.0554	3.66	C7H11O6-
179.0350	179.0340	5.59	C9H7O4-
173.0455	173.0444	6.36	C7H9O5-
161.0244	161.0236	4.97	C9H5O3-
155.0350	155.0340	6.45	C7H7O4-
135.0452	135.0443	6.66	C8H7O2-
111.0452	111.0444	7.20	C6H7O2-
93.0346	93.0338	8.60	C6H5O-
hydro-chlorogenate			
355.1035	355.1043	2.25	C16H19O9
191.0561	191.0556	2.62	C7H11O6-
181.0506	181.0500	3.31	C9H9O4-
173.0455	173.0446	5.20	C7H9O5-
161.0244	161.0237	4.35	C9H5O3-
111.0452	111.0445	6.30	C6H7O2-
93.0346	93.0339	7.52	C6H5O-

Supplementary Figure 4. Butterfly plot of LCMSMS fragmentation spectra for chlorogenate at top with key putative fragment ions highlighted in blue and hydro-chlorogenate at bottom with key putative fragment ions highlighted in red. The difference in two Dalton between the intact ions and the caffeate and hydrocaffeate fragment ions (*m/z* 179.0340 and *m/z* 181.0495, respectively) in these spectra support the assignment of *m/z* 355.1035 as hydro-chlorogenate as shown.



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-

Supplementary Figure 5. 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 References

1. Clifford, M. N. Chlorogenic acids and other cinnamates - nature, occurrence, dietary burden, absorption and metabolism. *J Sci Food Agric* **80**, 1033–1043 (2000).
2. Rekdal, V. M. *et al.* A widely distributed metalloenzyme class enables gut microbial metabolism of host-and diet-derived catechols. *Elife* **9**, (2020).
3. 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).
4. Rivera-Chávez, F. & Bäumler, A. J. The Pyromaniac Inside You: Salmonella Metabolism in the Host Gut. *Annu Rev Microbiol* **69**, 31–48 (2015).
5. 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).
6. 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).