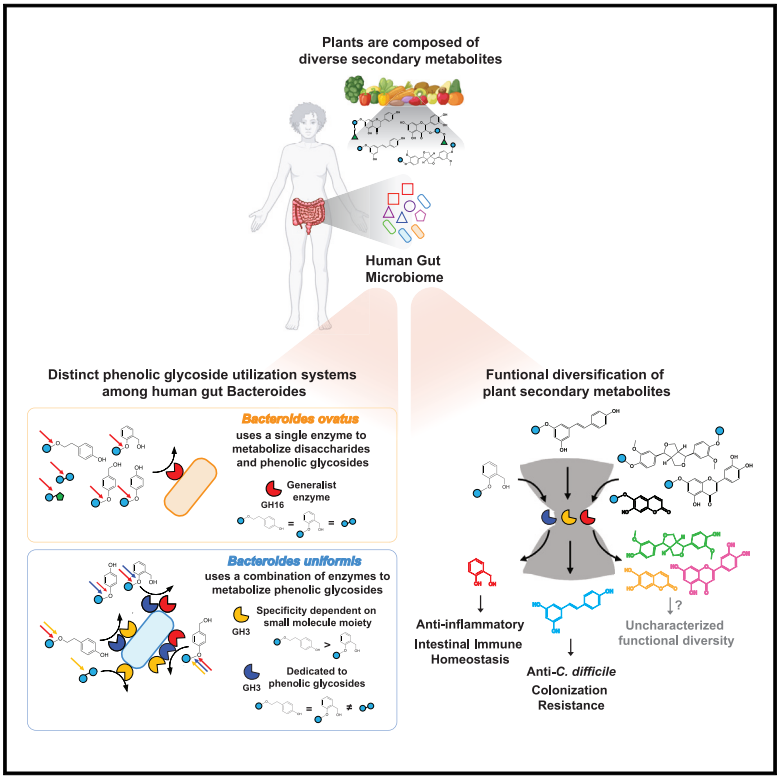
Cell

Functional diversification of dietary plant small molecules by the gut microbiome

Graphical abstract



Highlights

- Dietary phenolic glycoside utilized by specific members of the human gut microbiome
- Dedicated enzymes for plant glycoside metabolism in *Bacteroides uniformis*
- Microbiome-liberated polyphenol aglycones selectively inhibit *Clostridioides difficile*
- *Bacteroides* bioactivation of the aryl glycoside salicin mediates intestinal homeostasis

Authors

Gavin A. Kuziel, Gabriel L. Lozano, Corina Simian, ..., Min Dong, Jing-Ke Weng, Seth Rakoff-Nahoum

Correspondence

seth.rakoff-nahoum@childrens.harvard.edu

In brief

Plants are composed of diverse secondary metabolites. Here, we discover members of the human gut microbiome that have evolved enzymes to metabolize specific dietary phenolic small molecules that activate diverse functions important to the host, such as the regulation of intestinal inflammation and resistance to gut pathogens.

Kuziel et al., 2025, Cell *188*, 1–17
April 3, 2025
https://doi.org/10.1016/j.cell.2025.01.045

CellPress

Please cite this article in press as: Kuziel et al., Functional diversification of dietary plant small molecules by the gut microbiome, Cell (2025), https://doi.org/10.1016/j.cell.2025.01.045

Cell

CellPress

Article

Functional diversification of dietary plant small molecules by the gut microbiome

Gavin A. Kuziel,[1,2,6,7,12] Gabriel L. Lozano,[1,2,6,7,12] Corina Simian,[8,9,10] Long Li,[1,2,6,7] John Manion,[4,5,6] Emmanuel Stephen-Victor,[3,6] Talal Chatila,[3,6] Min Dong,[4,5,6] Jing-Ke Weng,[8,9,10] and Seth Rakoff-Nahoum[1,2,6,7,11,13,*]

[1]Division of Infectious Diseases, Boston Children's Hospital, Boston, MA 02115, USA
[2]Division of Gastroenterology, Boston Children's Hospital, Boston, MA 02115, USA
[3]Division of Immunology, Boston Children's Hospital, Boston, MA 02115, USA
[4]Department of Urology, Boston Children's Hospital, Boston, MA 02115, USA
[5]Department of Surgery, Harvard Medical School, Boston, MA 02115, USA
[6]Department of Pediatrics, Harvard Medical School, Boston, MA 02115, USA
[7]Department of Microbiology, Harvard Medical School, Boston, MA 02115, USA
[8]Whitehead Institute for Biomedical Research, Cambridge, MA 02142, USA
[9]Department of Chemistry and Chemical Biology & Department of Bioengineering, Northeastern University, Boston, MA 02120, USA
[10]Institute for Plant-Human Interface, Northeastern University, Boston, MA 02120, USA
[11]Broad Institute, Cambridge, MA 02139, USA
[12]These authors contributed equally
[13]Lead contact
*Correspondence: seth.rakoff-nahoum@childrens.harvard.edu
https://doi.org/10.1016/j.cell.2025.01.045

SUMMARY

Plants are composed of diverse secondary metabolites (PSMs), which are widely associated with human health. Whether and how the gut microbiome mediates such impacts of PSMs is poorly understood. Here, we show that discrete dietary and medicinal phenolic glycosides, abundant health-associated PSMs, are utilized by distinct members of the human gut microbiome. Within the *Bacteroides*, the predominant gram-negative bacteria of the Western human gut, we reveal a specialized multi-enzyme system dedicated to the processing of distinct glycosides based on structural differences in phenolic moieties. This *Bacteroides* metabolic system liberates chemically distinct aglycones with diverse biological functions, such as colonization resistance against the gut pathogen *Clostridioides difficile* via anti-microbial activation of polydatin to the stilbene resveratrol and intestinal homeostasis via activation of salicin to the immunoregulatory aglycone saligenin. Together, our results demonstrate generation of biological diversity of phenolic aglycone "effector" functions by a distinct gut-microbiome-encoded PSM-processing system.

INTRODUCTION

Diet is a critical determinant of human health.[1] In particular, plant-based diets have strong epidemiological associations with diverse health outcomes ranging from cardiovascular disease,[2,3] inflammatory disease, and autoimmunity[4,5] to the prevention of cancer and neurodegeneration.[6–8] In addition, hundreds of years of the use of medicinal plants across human cultures have provided both empirical support for the impact of plants on human health[9] and inspired pharmaceutical and synthetic chemical approaches for prevention and treatment of disease.[10]

The gut microbiome has emerged as central to mediating the impact of diet, in particular, plant diets, across a broad spectrum of human disease.[11–13] Perhaps, the best appreciated mechanism by which the microbiome mediates the impact of plants on the host is the generation of short-chain fatty acids, the products of gut microbiome fermentation of plant glycans.[14–16] In this process, plant polysaccharides are broken down into monosaccharides by a spectrum of microbiome-encoded metabolic enzymes such as glycoside hydrolases (GHs).[17–19] These monosaccharides enter glycolytic pathways creating energy for the gut microbiome. As a consequence, short chain fatty acids (such as acetate, propionate, and butyrate) are generated as waste products. Such short chain fatty acids serve as the effectors of microbiome metabolism of plant polysaccharides on the host.

In addition to polysaccharides, plants contain a variety of secondary metabolites (PSMs). These PSMs are chemically and functionally diverse and abundant across dietary and medicinal plants.[20] PSMs play critical roles in plant survival, ranging from physiological roles in development and endocrine function[21] to mediating plant-animal and plant-microbe interactions such as protection from herbivorous predators and bacterial and

Please cite this article in press as: Kuziel et al., Functional diversification of dietary plant small molecules by the gut microbiome, Cell (2025), https://doi.org/10.1016/j.cell.2025.01.045

fungal pathogens[22] via toxic,[23] neuroactive,[24] and anti-microbial[25] activities. Importantly, many PSMs are glycosylated both to tightly regulate functions within the plant during storage and to enhance solubility.[26] These PSM-carbohydrate conjugates, known as glycosides, represent the extensive chemical diversity of PSMs in the form of the aglycone moiety of each glycoside, and encompass broad plant biosynthetic chemical diversity across monoaryl and polyphenols, terpenoids, and alkaloids.[20] Epidemiological studies support substantial benefits associated with the consumption of plant glycosides and human health and disease.[27,28] Phenolic glycosides demonstrate both chemical diversity ranging from monophenolic to polyphenolic aglycones, the chemical class of aglycone ("simple" aryls, coumarins, flavonoids [dihydrochalcones], non-flavonoids [stilbenes], etc.), linkage to monosaccharides, disaccharides, or oligosaccharides, and stereochemistry (α vs. β linkage) and site of linkage (α1, β2, etc.),[29,30] are abundant in fruits, vegetables, and nuts[31,32] consumed by humans over evolutionary time and used in ethnobotanical cultural practice.[29,30,33] Dietary interventions in human clinical trials and animal models have demonstrated roles of plant glycosides on health outcomes across inflammatory disease,[34] cancer,[35] cardiovascular disease,[36] and neurodegenerative disorders.[37] However, the specific functions of plant glycosides, in particular, phenolic glycosides, on the host and both whether and how the gut microbiome mediates such impacts of PSMs in human health are poorly understood.

Here, we sought to define the capacity of the human gut microbiome to metabolize plant phenolic glycosides. Utilizing a multi-disciplinary approach spanning microbiology, bacterial genetics, biochemistry, enzymology, gnotobiotics, and mouse models of disease, we reveal the microbe-specific metabolism of distinct phenolic glycosides, the evolution of specialized enzymes among the human gut *Bacteroides* dedicated to processing distinct phenolic glycosides, and the functional diversification of plant glycosides by this system *in vivo* to mediate colonization resistance and intestinal immune homeostasis.

RESULTS

The landscape of dietary and medicinal phenolic glycoside utilization by the human gut microbiome

To define the capacity of the human gut microbiome to metabolize PSM glycosides, we assayed the ability of a taxonomically broad and representative panel of the membership of the human gut microbiome to utilize glycosides as energy sources serving as a scalable proxy for metabolic activity. To begin to capture the functional capacity of the gut microbiome to metabolize the broad chemical range of dietary and medicinal glycosides, we started with a panel of simple aryl, cyanogenic, or coumarin glycosides. Importantly, while each glycoside in this panel is bound to the same sugar (D-glucose) via the same linkage (β1) (Figure 1A), the aglycones of each compound have subtle yet distinct chemical modifications, such as carbon length of the alcohol group (arbutin vs. gastrodin), functionalization by aldehyde vs. alcohol group (salicin vs. helicin), substitution of aromatic methanol modification (*ortho*-salicin vs. *para*-gastrodin), or aglycone linker length (arbutin vs. salidroside) (Figure 1A). Members of the Enterobacteriaceae such as *Escherichia coli*, *Klebsiella pneumoniae*, *Citrobacter portucalensis* (*C. freundii* complex), and *Salmonella enterica*, and other facultative gut anaerobes such as *Enterococcus faecalis* and *faecium*, demonstrated either broad utilization of each glycoside as "generalists" (*K. pneumoniae*, *E. faecalis*, and *E. faecium*; Figure 1B, orange color) or unable to use any (*E. coli*, *C. portucalensis*, and *S. enterica*; Figure 1B, navy color), consistent with previous studies reporting absent or cryptic metabolism of glycosides in these taxa.[38,39]

Gut Clostridia, such as *Lactobacilli*, *Hungatella hathewayi*, *Clostridium symbiosum*, *C. scindens*, *P. hiranononis*, *C. cadaveris*, *C. sporogenes*, and *Coprobacillus cateniformis*, and the gut Actinobacteria (*Bifidobacterium longum*) demonstrated generalist (used all or most) or inability/cryptic utilization across the panel of glycosides (Figure 1B, green color). However, among other gut Clostridia and Erysipelotrichia members representative of taxonomic diversity across Lachnospiraceae, Clostridiaceae, and Erysipelotrichales families, we observed unique patterns of glycoside utilization. *Ruminococcus gnavus* demonstrated comparable growth on specific glycosides (arbutin, gastrodin, salicin, and salidroside) but not amygdalin and demonstrated strain-level variation (*R. gnavus* 2_1_58FAA vs. CC55_001C) in esculin utilization (Figure 1B, pink color text). Other members of gut Clostridia demonstrated unexpected specialization for specific glycosides (Figure 1B, teal color). *Longicatena innocuum* specifically utilized arbutin, esculin, helicin, and salicin but demonstrated absence of growth in salidroside or amygdalin (Figure 1B). *Blautia massiliensis* demonstrated unique specialization (enhanced growth compared with other aryl glycosides) for arbutin, while *Faecalimonas umbilicata* utilized arbutin and salicin, but not other aryl glycosides.

Among members of the human gut Bacteroidales, the predominant gram-negative bacteria of the human gut microbiome,[40,41] we observed unexpected variation in glycoside utilization. *B. fragilis* S36 L11, *B. eggerthii* DSM, *B. vulgatus* ATCC, and *B. dorei* 9_1_42FAA demonstrated inability/cryptic utilization of each glycoside (Figure 1B, navy color). *B. ovatus CL03T12C18* and each member of *Parabacteroides* species, *P. goldsteinii* CL02T12C30, *P. distasonis* 31_2, *P. johnsonii* CL02T12C29, and *P. merdae* CL03T12C32, exhibited broad utilization of each glycoside (except for helicin, for which we identified only *H. hathewayi*, *C. tertium*, *C. innocuum*, and *Enterococci* as able to utilize this glycoside; Figure 1B, orange color). Comparatively, we find *Bacteroides* members exhibiting unique isolate-specific specialization for specific glycosides (Figure 1B, teal color). *B. salyersiae* DSM and *B. nordii* CL02T12C05 utilized only esculin and amygdalin, while *B. finegoldii* DSM exclusively utilized esculin (Figure 1B). In contrast, *Bacteroides uniformis* ATCC demonstrated robust growth in the aryl glycosides (arbutin, salicin, gastrodin, and salidroside) but limited or no growth for those of greater complexity (esculin or amygdalin). *B. caccae* ATCC utilized all simple glycosides (excluding helicin, as above) but not arbutin. Together, our data demonstrate unexpected variation in utilization of glycosides among the Clostridia, Erysipelotrichia, and Bacteroidales of the human gut microbiome whereby certain members demonstrate "all-or-none" utilization of glycosides while others differentially utilize structurally similar PSMs varying by distinct modifications of phenolic aglycones.

Please cite this article in press as: Kuziel et al., Functional diversification of dietary plant small molecules by the gut microbiome, Cell (2025), https://doi.org/10.1016/j.cell.2025.01.045

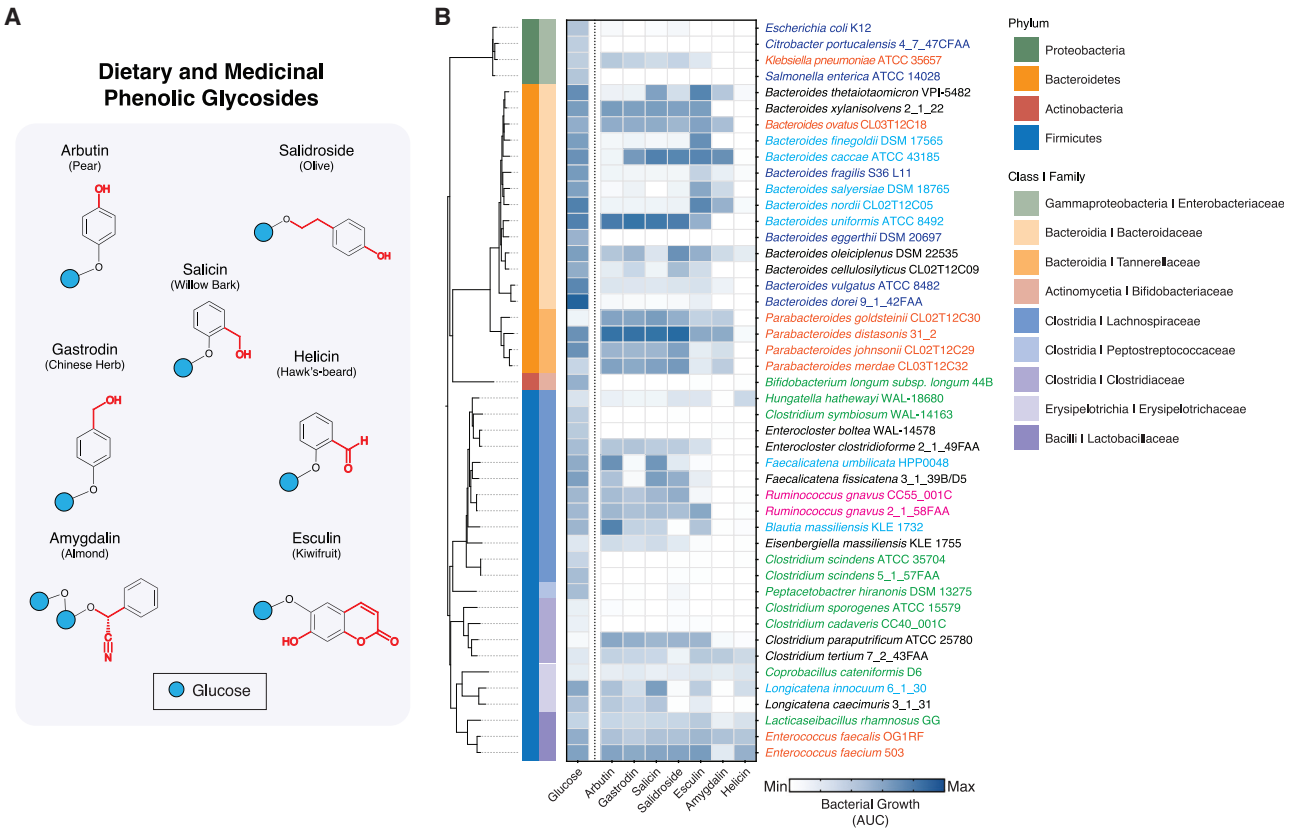


Figure 1. The landscape of dietary phenolic glycoside utilization by the human gut microbiome

(A) Panel of dietary and medicinal phenolic glycosides. Each glycoside is bound to the same sugar (D-glucose) via the same linkage (β1). Each aglycone moiety has subtle yet distinct chemical modifications (highlighted in red) such as carbon length of the alcohol group, functionalization by aldehyde vs. alcohol group, substitution position of the aromatic methanol, or aglycone linker length.

(B) Heatmap of growth of members of the human gut microbiome on glycosides (or glucose) (each at 15 mM) as the sole carbohydrate source. Growth is calculated as the area under each 48-h bacterial growth curve (AUC units). Background growth due to catabolism of media components was subtracted from overall growth of each bacterium in glucose or glycoside to normalize growth across bacteria and substrates. Maximal growth is represented by a deep blue color. Color code: orange are generalist bacteria that utilize most or all glycosides; navy and green are distinct taxa of gut microbes that do not utilize glycosides; teal are specialist bacteria that utilize specific glycosides; pink indicates examples of strain-level variation in glycoside utilization. Throughput screen data in (B) are for one biological replicate, in technical triplicate.

Strain-level variation in phenolic glycoside utilization by the human gut Bacteroidales

Given the abundance of the gut Bacteroidales in the human gut microbiome,[40,41] the links between plant diets and the *Bacteroides* in health,[13,42,43] and our findings demonstrating that the Bacteroidales differentially utilize specific health-relevant plant-derived glycosides, we sought to more deeply investigate phenolic glycoside utilization by the gut Bacteroidales. Defining the functional capacity of 52 *Bacteroides* and *Parabacteroides* strains across 18 species to utilize our panel of simple glycosides, we reveal extensive strain-level and aglycone modification-level variation in glycoside utilization (Figure 2A). In general, we observe similar patterns of glycoside utilization among strains of the same species. For example, each strain of the species *B. ovatus*, *B. xylanisolvens*, and across *Parabacteroides* species were generalists utilizing each of the panel of simple glycosides. Similarly, we find that inability to utilize any of the glycosides was a species-wide trait of *B. vulgatus* and *B. dorei* (Figure 2A). However, among other *Bacteroides* species, we find strain-level variation for distinct glycosides. For example, while each strain of *B. thetaiotaomicron* specialized in esculin and amygdalin utilization, we observed strain-level variation in the dynamics of growth during salicin utilization (*Bt* 1_1_6 vs. *Bt* VPI) (Figure 2B). All strains of *B. fragilis* were universally unable to utilize phenolic glycosides with the exception of *B. fragilis* CL05T00C42, which specifically utilized esculin (Figure 2C). Both *B. uniformis* ATCC and *B. uniformis* D20 specialized in utilization of aryl glycosides, but *B. uniformis* CL03T00C23 did not grow on any phenolic glycoside (Figure 2D). Similarly, *B. caccae* ATCC grew on most glycosides, but *B. caccae* CL03T12C61 was unable to utilize any (Figure 2E). Together our results reveal extensive strain-level and aglycone modification-level variation in glycoside utilization among the human gut *Bacteroides*.

Please cite this article in press as: Kuziel et al., Functional diversification of dietary plant small molecules by the gut microbiome, Cell (2025), https://doi.org/10.1016/j.cell.2025.01.045

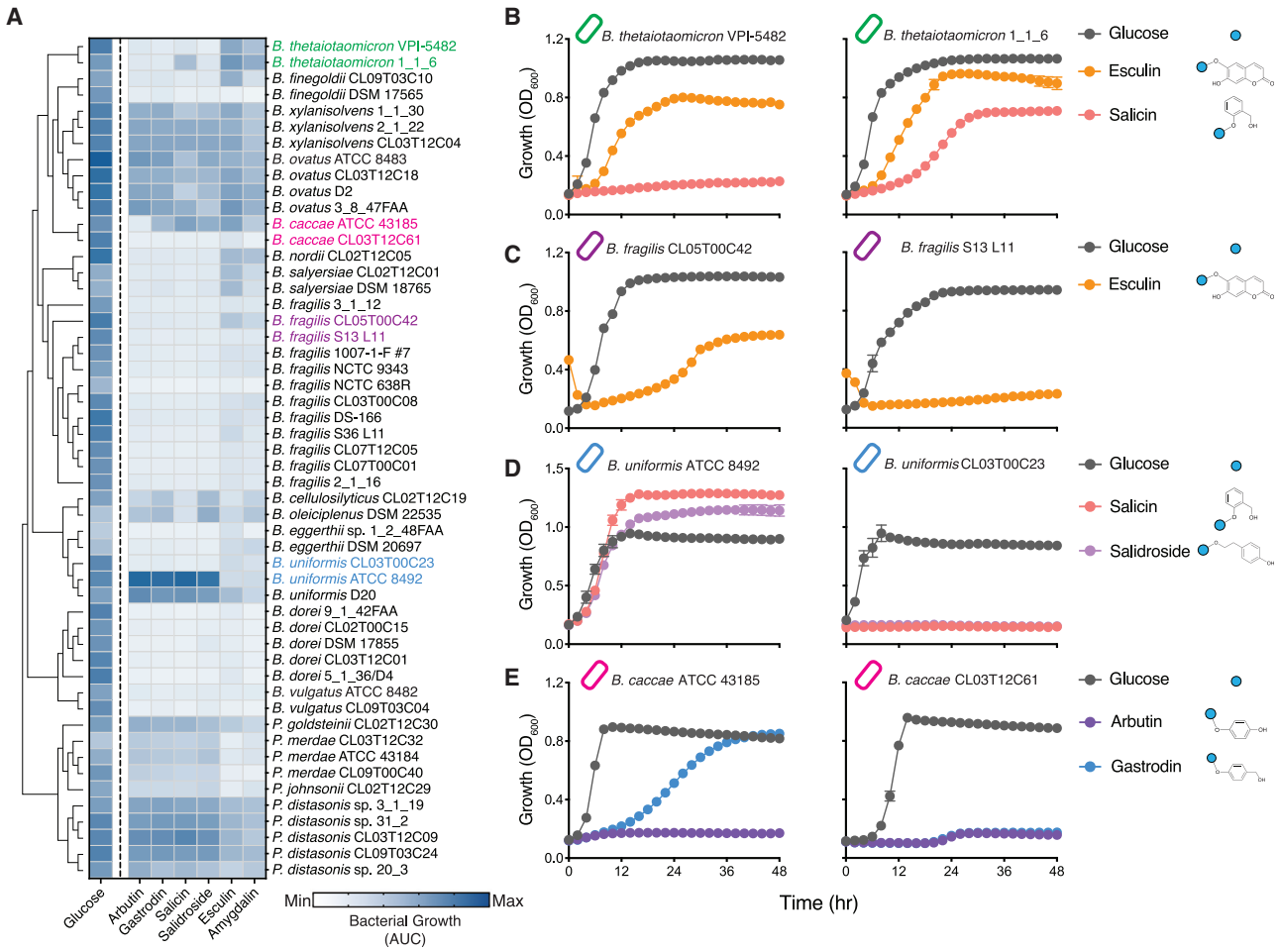


Figure 2. Strain-level variation in dietary phenolic glycoside utilization by the human gut Bacteroidales

(A) Heatmap of Bacteroidales growth as a function of phenolic glycoside (or glucose) (each at 15 mM) utilization. Growth is calculated as the area under each 48-h bacterial growth curve (AUC units). Maximal growth is represented by a deep blue color. Color code: strain-level variation in phenolic glycoside utilization by *B. thetaiotaomicron* (green), *B. fragilis* (purple), *B. uniformis* (blue), and *B. caccae* (pink).

(B) Growth of *Bacteroides thetaiotaomicron* VPI-5482 and 1_1_6 in glucose or phenolic glycosides as the sole carbon source (15 mM) as indicated. Growth is calculated as the area under each 48-h bacterial growth curve.

(C) Growth of *B. fragilis* CL05T00C42 and S13L11 in glucose or phenolic glycosides as the sole carbon source (15 mM) as indicated. Growth is calculated as the area under each 48-h bacterial growth curve.

(D) Growth of *B. uniformis* ATCC 8492 and CL03T12C37 in glucose or phenolic glycosides as the sole carbon source (15 mM) as indicated. Growth is calculated as the area under each 48-h bacterial growth curve.

(E) Growth of *B. caccae* ATCC 43185 and CL03T12C61 in glucose or phenolic glycosides as the sole carbon source (15 mM) as indicated. Growth is calculated as the area under each 48-h bacterial growth curve.

Throughput screen data in (A) are for one biological replicate, in technical triplicate. In (B)–(E), data are representative of two biological replicates, each in technical triplicate. In (B)–(E), error bars represent mean ± SD.

Distinct phenolic glycoside utilization systems among the human gut *Bacteroides*

Across the *Bacteroides*, we reveal unique patterns of aryl glycosides utilization, suggesting differential mechanisms for the metabolism of health-related plant small molecules. We thus next sought to define the mechanistic basis for aryl glycoside utilization among the *Bacteroides*. We began with a forward genetic approach performing transposon insertion mutagenesis coupled with next generational sequencing (Tn-seq) to identify the genetic basis for aryl glycoside utilization in *B. ovatus* ATCC, which we identify as demonstrating broad generalist utilization of these substrates. Identifying Tn insertions resulting in loss of function in arbutin compared with glucose identified 67 genes with $\log_2$-fold-change ≥ 1 and $p < 0.05$. Among these, we find insertions in each gene of a five-gene operon conferring specific fitness disadvantage in arbutin, Bovatus_02231–Bovatus_02227 (Figure 3A). This operon is composed of a transcriptional regulator (Bovatus_02231), two nicotinamide-dependent oxidoreductases (Bovatus_02230 and Bovatus_02229), a protein with structural homology[44] to the class of family-16-like GH (Bovatus_02228),

Please cite this article in press as: Kuziel et al., Functional diversification of dietary plant small molecules by the gut microbiome, Cell (2025), https://doi.org/10.1016/j.cell.2025.01.045

and a sugar-phosphate isomerase (Bovatus_02227) (Figure 3B). We generated isogenic strains with non-polar deletions of each gene of this operon. Deletion of each gene in this operon was defective in utilization of each aryl glycoside (Figure 3C) and restored by complementation (Figure S1A). To date, GHs responsible for hydrolysis of monosaccharide-aglycone linkages within the human gut microbiome have been demonstrated to hydrolyze monosaccharide-monosaccharide linkages for disaccharide or polysaccharide utilization,[45] the latter thought to be the primary evolved substrate of these enzymes.[46] Indeed, we find that Bovatus_02228 and each gene in the Bovatus_02231-02227 operon (except Bovatus_02227, the sugar-phosphate isomerase) is required for utilization of the disaccharides trehalose (α1,1 glucose-glucose homodimer) and palatinose (α1,6 glucose-fructose heterodimer) (Figures 3D, S1B, and S1C), consistent with a recent report demonstrating activity of a homolog of the GH16 family member Bovatus_02228 for both glucosinolates and disaccharides.[47] Together, we conclude that the aryl glycoside generalist of the gut *Bacteroides*, *B. ovatus*, employs a single operonic generalist system (we name the operon glycoside *generalist* hydrolase, *ggh*, composed of *gghR*, Bovatus_02231; *gghA*, Bovatus_02230; *gghB*, Bovatus_02229; *gghC*, Bo_02228; *gghD*, Bo_02227) to utilize both β1-linked aryl glycosides and α1-linked disaccharides.

We next sought to determine the genetic basis for aryl glycoside utilization in the glycoside specialist *B. uniformis*. Compared with *B. ovatus*, which utilizes aryl, coumarin, and cyanogenic glycosides, *B. uniformis* both specializes in aryl glycosides and notably demonstrates greater growth yield in aryl glycosides than in glucose (Figures 2A and 2D). Leveraging the same approach as in *B. ovatus*, we constructed a genome-wide transposon mutagenesis library of *B. uniformis* and performed Tn-seq after selection in glucose vs. arbutin (Figure 3E). Transposon insertions in each gene of an operon composed of four genes (BACUNI_00922–BACUNI_00919), distinct from that in *B. ovatus*, conferred loss of fitness in arbutin, for which BACUNI_00919, encoding a GH3 family member, was the most statistically significant loss-of-fitness Tn insertion (Figure 3E). This operon, which we name glycoside specialist hydrolase *gsh*, is composed of esterases (*gshA*, BACUNI_00922; *gshB*, BACUNI_00921; *gshC*, BACUNI_00920) and a GH3 family member (*gshD*, BACUNI_00919) (Figure 3F). To determine the requirement of each gene in the *B. uniformis gsh* operon for aryl glycoside utilization, we generated non-polar deletions for each gene of the operon. As opposed to the *ggh* generalist utilization locus of *B. ovatus*, deletion of only the GH3 *gshD* in the *B. uniformis gsh* operon resulted in impaired glycoside utilization (Figures S2A and S2B). Importantly, compared with the complete defect in aryl glycoside utilization upon deletion of the GH16 *gghC* in *B. ovatus*, deletion of the GH3 *gshD* in *B. uniformis* resulted in only partially defective growth in aryl glycosides (Figure S2A), suggesting that as opposed to the generalist *B. ovatus*, *B. uniformis* employs multiple GHs to utilize these substrates.

To identify genes acting in concert with *gshD* to utilize aryl glycosides in the specialist *B. uniformis*, we performed a synthetic Tn-seq on a *Bu* Δ*gshD* background under arbutin selection. This screen identified Tn insertions in two genes, BACUNI_01042 (*gshG*, GH3 family member) and BACUNI_0952 (*gghB*, a homolog of the nicotinamide-dependent oxidoreductases identified in *B. ovatus* ATCC directly upstream of a homolog of *Bo* GH16 *gghC*) (Figure 3G; Table S2). We next generated a series of single-, double-, and triple-gene deletion strains lacking combinations of *gshD*, *gshG*, and *gghC*. Deletion of *Bu* GH3 *gshG* or *Bu* GH16 *gghC* alone showed no defects in utilization of aryl glycosides, demonstrating the synergistic requirement of the GH3 *gshD* for aryl glycoside utilization (Figure S2C).

Comparing the utilization of each aryl glycoside of combinations of *gsh* GH3 gene deletions in *B. uniformis* revealed striking findings. Deletion of both *Bu* GH3 *gshD* with *Bu* GH3 *gshG* (*Bu* Δ*gshD*Δ*gshG*) demonstrated impaired growth on arbutin compared with that of *Bu* Δ*gshD* alone (Figure 3H) but notably exhibited residual utilization of this aryl glycoside. In comparison, deletion of *Bu* GH3 *gshD* and the *Bu* GH16 *gghC* (*Bu* Δ*gshD* Δ*gghC*) demonstrated a complete defect in arbutin utilization demonstrating the requirement of a three GH enzymatic system for the utilization of arbutin in *B. uniformis*, with a greater contribution of *Bu* GH16 *gghC* than *Bu* GH3 *gshG* for arbutin utilization (Figure 3H). Similar to arbutin, this three GH enzyme system was required for utilization of gastrodin, which is functionalized with a *para*-hydroxymethylation vs. *para*-hydroxylation as in arbutin. Compared with the three GH enzyme requirement for the utilization of arbutin and gastrodin, *Bu* GH16 *gghC* but not *Bu* GH3 *gshG* was required to utilize salicin. However, for *B. uniformis* utilization of salidroside, we see the opposite, as *Bu* GH3 *gshG* but not *Bu* GH16 *gghC* is required with *Bu* GH3 *gshD* for utilization of this aryl glycoside. Together, our findings demonstrate that while *B. ovatus* uses one GH to utilize aryl glycosides, *B. uniformis* employs unique combinations of GH3 and GH16 to utilize specific PSM glycosides.

Specialized enzymes for phenolic glycoside metabolism in the human gut *Bacteroides*

The requirement for pairs of GH enzymes to utilize distinct aryl glycosides in the glycoside specialist *B. uniformis* suggested that *Bu* GH3 enzymes may be specific and/or dedicated for distinct phenolic glycosides. The requirement of *Bu* GH3 *gshG* for salidroside but not salicin utilization suggested specialization of this enzyme for specific glycosides. To test this hypothesis, we first over-expressed *Bu gshG* in *B. vulgatus*, which is unable to grow on either salidroside or salicin (Figures 2A and S3A). Over-expression of *Bu* GH3 *gshG* resulted in enhanced growth rate and yield on salidroside compared with salicin, with the latter being used as a substrate but to a lesser degree (Figure 4A). This suggested that specificity for hydrolysis of salidroside by *Bu* GH3 *gshG* may be encoded via the unique aglycone of this phenolic glycoside vs. that of salicin. To definitively test the substrate specificity of *Bu* GH3 *gshG*, we purified recombinant *Bu* GH3 GshG (Figure S3B) and measured hydrolytic activity for salicin and salidroside. *Bu* GH3 GshG had 10-fold greater catalytic turnover of salidroside compared with salicin (Figure 4B). These data reveal that the GH3 GshG of *B. uniformis* demonstrates substrate specificity based on discrete structural differences of the aglycone moiety of the glycosides.

For the GH16 family member *gghC* (conserved between *Bo* and *Bu*), we demonstrated roles in utilization by *B. ovatus* for both aryl glycosides and disaccharides (trehalose and palatinose),

Please cite this article in press as: Kuziel et al., Functional diversification of dietary plant small molecules by the gut microbiome, Cell (2025), https://doi.org/10.1016/j.cell.2025.01.045

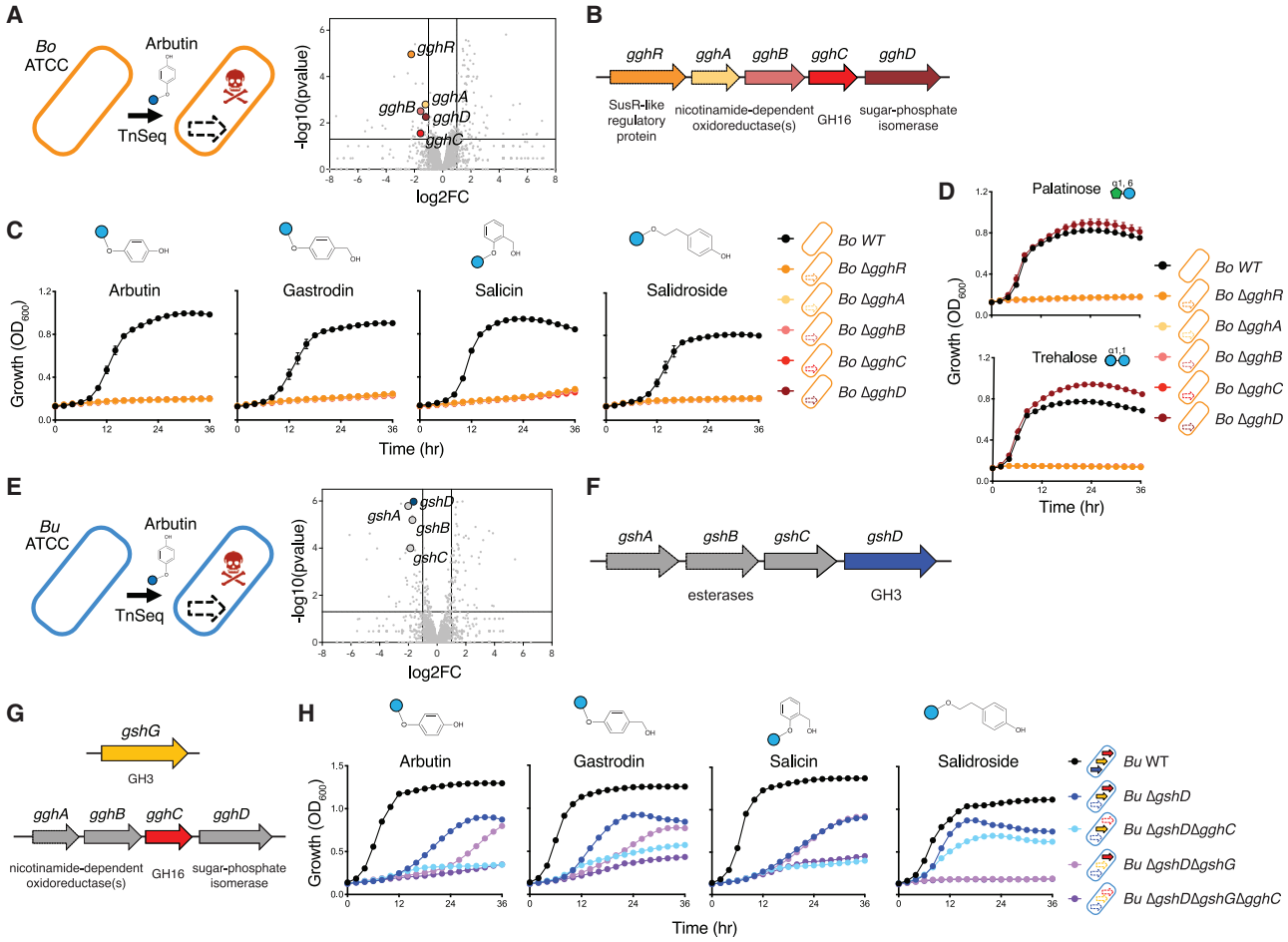


Figure 3. Distinct phenolic glycoside utilization systems among the human gut *Bacteroides*

(A) Tn-seq of *B. ovatus* ATCC grown in arbutin (15 mM) or glucose (15 mM) as the sole carbon source. A five-gene locus is highlighted, and each gene is labeled.

(B) Glycoside utilization locus of *Bacteroides ovatus*.

(C) Growth of *B. ovatus* WT, Δ*gghR*, Δ*gghA*, Δ*gghB*, Δ*gghC*, or Δ*gghD* with each aryl glycoside (15 mM) as the sole carbon source. Note each mutant growth curve overlaps in each condition.

(D) *B. ovatus* WT, Δ*gghR*, Δ*gghA*, Δ*gghB*, Δ*gghC*, or Δ*gghD* grown in palatinose (top) and trehalose (bottom) (15 mM as sole carbon source). Note growth curve of each mutant except *Bo* Δ*gghD* overlaps in both substrates.

(E) Tn-seq of *B. uniformis* ATCC grown in arbutin (15 mM) or glucose (15 mM) as the sole carbon source. A four-gene locus is highlighted, and each gene is labeled.

(F) Glycoside utilization locus identified by Tn-seq in *B. uniformis*.

(G) Secondary glycoside utilization loci identified in *B. uniformis*.

(H) Growth of *B. uniformis* WT, Δ*gshD*, Δ*gshD*Δ*gghC*, Δ*gshD*Δ*gshG*, or Δ*gshD*Δ*gshG*Δ*gghC* in each indicated aryl glycoside (15 mM) as the sole carbon source.

In (A) and (E), fold change and *p* value were calculated by combining two biological replicates. In (C), (D), and (H), data are representative of two to three biological replicates, each in technical triplicate. In (C), (D), and (H), error bars represent mean ± SD.

See also Figures S1, S2, and S7.

demonstrating the broad hydrolytic activity of this GH16 member for PSM glycosides and disaccharides (Figures 3C and 3D). Therefore, we next sought to determine whether *B. uniformis* GH3 (required for aryl glycoside utilization) were also required to utilize disaccharides. We find *Bu* Δ*gshG* partially defective in utilization of gentiobiose, demonstrating the requirement of *Bu* GH3 *gshG* for utilization of this β1,6 glucose-glucose homodimeric disaccharide (Figures 4C and S3C). Purified recombinant *Bu* GH3 GshG protein hydrolyzed both gentiobiose and the β1,4 glucose-glucose homodimeric disaccharide cellobiose, with comparable efficiency to hydrolysis of salidroside (Figure 4D). These data demonstrate that *Bu* GH3 *gshG* enables the utilization of both the aryl glycosides salidroside, arbutin, and gastrodin (and to a lesser extent salicin) and the disaccharide gentiobiose and thus hydrolyzes both aryl glycosides and disaccharides.

We next sought to determine the specificity of *Bu* GH3 GshD, which is required for utilization of each aryl glycoside by *B. uniformis* (Figure 3H). Overexpression of *Bu* GH3 *gshD* in *B. vulgatus* conferred similar capacity to utilize each of these phenolic glycosides (Figure 4E), compared with the salidroside

Please cite this article in press as: Kuziel et al., Functional diversification of dietary plant small molecules by the gut microbiome, Cell (2025), https://doi.org/10.1016/j.cell.2025.01.045

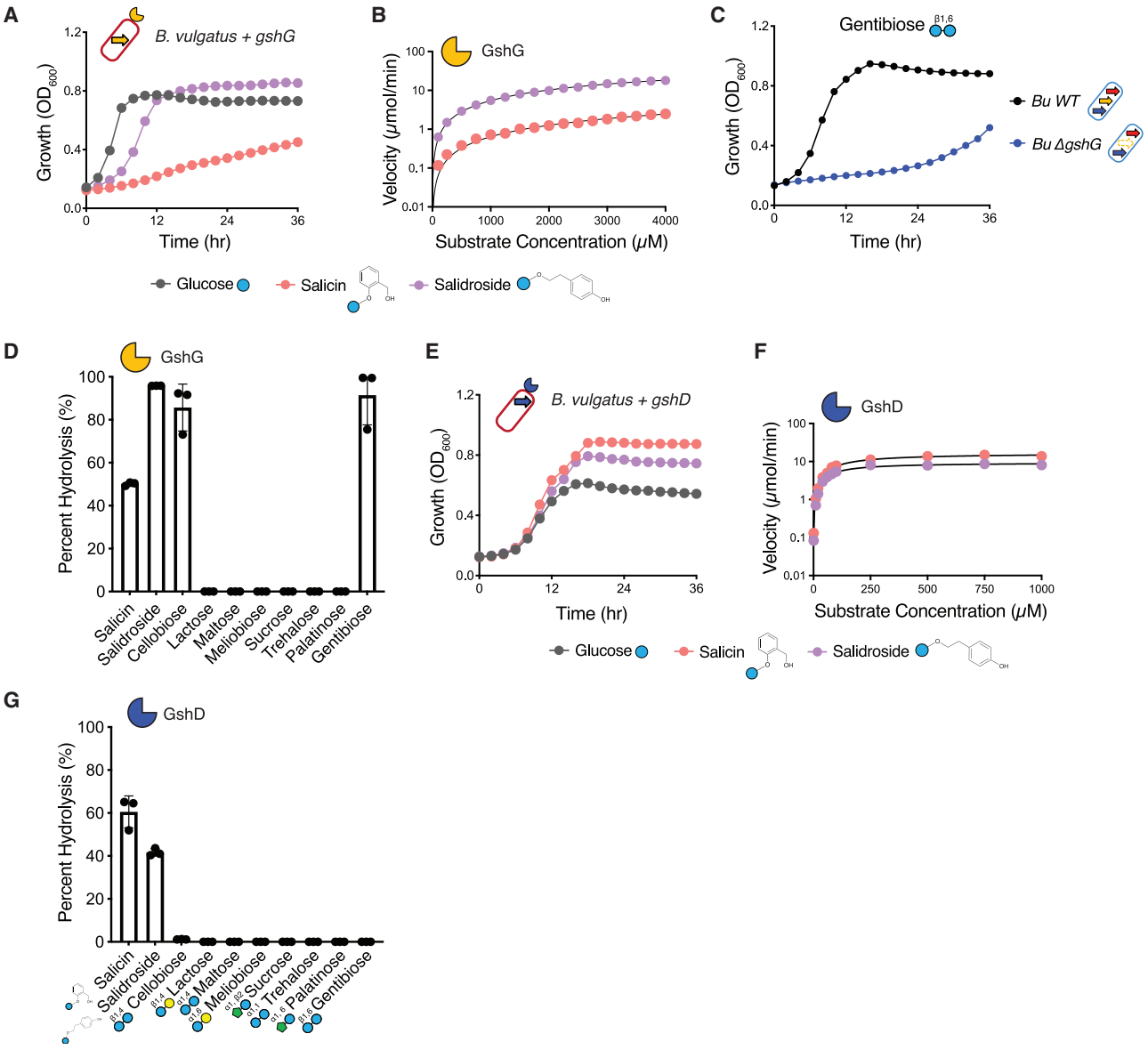


Figure 4. Dedicated enzymes for phenolic glycoside metabolism in the human gut *Bacteroides*

(A) Growth of *B. vulgatus* WT complemented with *gshG* in glucose, salicin, or salidroside as the sole carbon source.

(B) Michaelis-Menten kinetics of gshG hydrolysis of salicin or salidroside.

(C) Growth *B. uniformis* WT and Δ*gshG* grown in gentiobiose in *Bacteroides* minimal media.

(D) Percent hydrolysis of salicin, salidroside or eight disaccharides to their respective aglycones or monosaccharides, respectively, by recombinant GshG.

(E) Growth of *B. vulgatus* WT complemented with *gshD* grown in glucose, salicin, or salidroside as the sole carbon source.

(F) Michaelis-Menten kinetics of GshD hydrolysis of salicin or salidroside.

(G) Percent hydrolysis by recombinant GshD of salicin, salidroside or disaccharides (as indicated) to their respective aglycones or monosaccharides, respectively. All carbon sources at 15 mM.

In (A), (C), and (E), data are representative of two biological replicates, each in technical triplicate. In (B) and (F), data are from one biological replicate, in technical triplicate. In (D) and (G), data are representative of two independent experiments, each in technical triplicate. In all panels, error bars represent mean ± SD. See also Figure S3.

specificity of *Bu* GH3 *gshG* (Figure 4A). As opposed to *Bu* GH3 *gshG*, absence of *Bu* GH3 *gshD* did not affect utilization of disaccharides by *B. uniformis* (Figure S3C). This suggested that *Bu* GH3 *gshD* may be dedicated to the metabolism of PSM glycosides. Indeed, we find that *Bu* GH3 *gshD* is a part of a unique clade of GH3 enzymes among gut bacteria, both distinct from *Bu* GshG (Figure S4D) and present in *B. uniformis* and *P. distasonis* among the Bacteroidales (Figure S3E). To definitively test the substrate specificity of *Bu* GH3 *gshD*, we purified *Bu* GH3 GshD (Figure S3F) and measured hydrolytic activity for

Please cite this article in press as: Kuziel et al., Functional diversification of dietary plant small molecules by the gut microbiome, Cell (2025), https://doi.org/10.1016/j.cell.2025.01.045

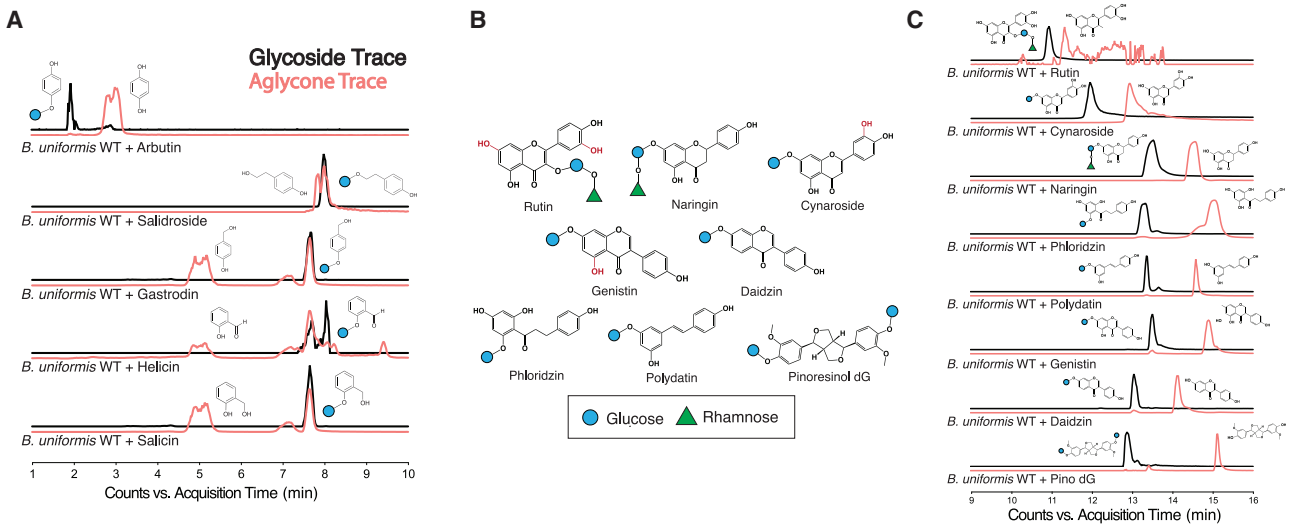


Figure 5. Liberation of diverse dietary phenolic aglycones by the *Bacteroides*

(A) Comparative LC-MS extracted ion chromatograms (EICs) of methanol-water extracts of bacterial cultures supernatants supplemented with each aryl glycoside. The black trace corresponds to the EIC of each glycoside, and the pink trace corresponds to the EIC of each aglycone. The pink trace peak overlapping the black trace peak corresponds to the aglycone detected from fragmentation of the glycoside at the ionization source and is a technical artifact (not biological).

(B) Panel of common dietary plant polyphenolic glycosides. Aglycones are schematically shown glycosylated with glucose (blue circle) and/or rhamnose (green triangle).

(C) Comparative LC-MS EICs of methanol-water extracts of bacterial cultures supplemented with each polyphenolic glycoside after 8 h. The black trace corresponds to the EIC of each glycoside, and the pink trace corresponds to the EIC of each aglycone. The pink trace peak overlapping the black trace peak corresponds to the aglycone detected from fragmentation of the glycoside at the ionization source and is a technical artifact.

See also Figure S7.

salicin and salidroside and a range of disaccharides. *Bu* GH3 GshD had comparable hydrolytic activity for both salicin and salidroside (Figure 4F) but did not hydrolyze any disaccharide (Figure 4G). Together, we demonstrate that *B. uniformis* contains GH3 enzymes with substrate specificity dependent on discrete structural differences of each substrate's aglycone moiety (*Bu* GH3 GshG) and potentially dedicated to the hydrolysis of PSM glycosides (*Bu* GH3 GshD).

Liberation of diverse phenolic aglycones by the *Bacteroides*

Our data have demonstrated the utilization and hydrolysis of aryl glycosides by a specialized metabolic system in the predominant human gut microbiome member *B. uniformis*.[48,49] *B. uniformis* GH enzymes responsible for aryl glycoside metabolism are predicted to localize to the outer membrane.[50] Thus, we next asked if hydrolysis of aryl glycosides by *B. uniformis* resulted in the liberation and extracellular availability of the range of aryl aglycones. To test this, we measured the presence of distinct aglycones by liquid chromatography mass spectrometry (LC-MS) in supernatant after growth of *B. uniformis* in media supplemented with each aryl glycoside. For each aryl glycoside, arbutin, salidroside, gastrodin, helicin and salicin, *B. uniformis* liberated each of the corresponding aglycones, hydroquinone, gastrodigenin, tyrosol, salicylaldehyde, and saligenin, respectively (Figure 5A).

We next asked both whether *B. uniformis* hydrolyzes and liberates aglycones from a broader range of PSMs extending our studies to polyphenolic glycosides, both abundant in diet and associated with wide-ranging impacts across health and disease.[51,52] To address this, we generated a panel of polyphenolic glycosides which, similar to our panel of aryl glycosides, aglycones are bound to glucose via β1-linkages, but for which for each aglycone represents broad chemical diversity spanning flavonoids (flavonols, rutin; flavanones, naringin; flavones, cynaroside; isoflavones, genistein, and daidzein), dihydrochalcones (phloridzin), stilbenes (polydatin), and lignans (pinoresinol diglucoside) (Figure 5B). Using a targeted LC-MS approach, we found that across each polyphenolic glycoside, *B. uniformis* liberated the predicted aglycone deglycosylation product from each polyphenolic glycoside substrate (Figure 5C). Together, these data demonstrate that *B. uniformis* metabolism of aryl glycosides result in the liberation of diverse mono- and polyphenolic aglycones.

Bioactivation of polyphenolic glycosides into antibiotics by the *Bacteroides*

Having defined the broad capacity of the plant glycoside specialist *B. uniformis* to both metabolize a diverse array of phenolic glycosides and liberate chemically diverse aglycones, we next sought to determine potential functions of such plant small molecules on microbe-microbe and microbe-host functions in the mammalian gut. In plants, deglycosylation of phenolic glycosides by plant exo-glycosidases results in the liberation of aglycones, which play diverse and important functions in defense from microbial pathogens.[25,53,54] Thus, we asked whether *B. uniformis* liberation of diverse aglycones from distinct glycosides activated anti-microbial functions against human gut pathogens. To test this, we determined the

Please cite this article in press as: Kuziel et al., Functional diversification of dietary plant small molecules by the gut microbiome, Cell (2025), https://doi.org/10.1016/j.cell.2025.01.045

fitness effects of aryl or polyphenolic glycosides and each corresponding aglycone liberated by *B. uniformis* (Figure 6A) on a panel of human gut pathogens representing taxa of multidrug resistant (MDR) and opportunistic pathogens. This screen revealed distinct fitness effects on gut pathogens for specific polyphenolic glycosides such as the pro-microbial (growth enhancement) of the *B. uniformis* liberated aglycone of the isoflavone daidzin, daidzein across gut pathogens and anti-microbial effects of *E. faecium* by the liberated aglycone of the flavone cynaroside, luteolin (Figure 6A). Most notably, we reveal the polyphenolic aglycones phloretin and resveratrol liberated from the dihydrochalcone glycoside, phloridzin (abundant in apples[31]), and the stilbene glycoside, polydatin (abundant in grapes and in plants used in traditional medicine notably *Polygonum cuspidatum*; Japanese knotweed[55]), respectively, as potent inhibitors of *Clostridioides difficile* strain M7404 (Figure 6A). Neither the parent polyphenolic glycosides phloridzin or polydatin were anti-microbial against *C. difficile* (Figure 6A).

B. uniformis cultivated with polydatin liberated resveratrol (Figure 5C) and generated potent anti-*C. difficile* activity (Figure S4A), both functions ablated in *B. uniformis* Δ*gshD* (Figures S4B and S4C), demonstrating bioactivation of polydatin and liberation of resveratrol by *B. uniformis* dependent on the glycoside-specific GH3 *Bu* gshD. *C. difficile* infection (CDI) is a major cause of mortality and morbidity, most associated with defects in colonization resistance by the gut microbiome in the setting of use of antibiotics.[56] Across 20 unique *C. difficile* strains representing toxin production and hypervirulence, we find the *B.-uniformis*-liberated polyphenolic aglycone resveratrol, but not its parent glycoside polydatin, as universally anti-microbial against *C. difficile* (Figure 6B). In a mouse model of CDI, the *B. uniformis*-generated aglycones resveratrol and phloretin demonstrated potent anti-microbial effects on *C. difficile in vivo* (Figure 6C) occurring within the context of a restored gut microbiome after antibiotics (Figures S4D and S4E). In addition, we do not find evidence of degradation of resveratrol (Figure S4F) nor metabolism to dihydroresveratrol by the mouse gut microbiome (Figure S4G), which can occur via reduction of resveratrol by specific members of the human gut microbiome.[57,58] Furthermore, we find that dihydroresveratrol maintains anti-microbial activity against *C. difficile* (Figure S4H). Together, our data reveal both a novel function of polyphenolic aglycones against the human gut pathogen *C. difficile* and the role of bioactivation of dietary plant glycosides by members of the gut microbiome in mediating a novel mechanism of colonization resistance to human gut pathogens mediated by microbiome bioactivation of specific phytochemicals.

Functionalization of distinct phenolic glycosides by the *Bacteroides* mediates intestinal immune homeostasis

Plant glycosides, which function in diverse roles in plants in homeostasis to stress,[53,59] have been postulated to play both beneficial and pathological role in human health, notably in inflammation and cancer.[60,61] However, whether and how the gut microbiome mediates such effects of plant glycosides on the host is poorly understood. Given the abundance of the *Bacteroides* in the human gut, in particular *B. uniformis*,[48,49] and our identification of functionalization of plant glycosides by liberation of diverse aglycones by this prominent member of the human gut microbiome, we next sought to identify potential functions of the biotransformation of PSM glycosides by the *B. uniformis* on inflammation. To test this, we assayed the function of aryl and polyphenolic glycosides and the corresponding *B.-uniformis*-liberated aglycones of each glycoside to regulate the production of tumor necrosis factor (TNF) and interleukin-6 (IL-6) by lipopolysaccharide (LPS)-stimulated macrophages. The aryl glycosides salicin and salidroside and the polyphenolic glycosides rutin and naringin did not affect cytokine production by macrophages. In contrast, each of the *B. uniformis*-liberated aglycones: saligenin, tyrosol, quercetin, and naringenin, respectively, demonstrated broad anti-inflammatory function for both TNF and IL-6 (Figures 7A and 7B). The aryl glycoside arbutin was found to selectively negatively regulate TNF, but not IL-6, production, whereas the arbutin aglycone hydroquinone was cytotoxic to macrophages (Figure S5A). The polyphenolic glycoside polydatin enhanced IL-6, but not TNF, production from macrophages, whereby this specific pro-inflammatory function was converted to potent anti-inflammatory function when deglycosylated to the aglycone resveratrol (Figure 7B). While the aryl glycoside gastrodin did not impact macrophage inflammatory cytokine production, the *B.-uniformis*-liberated aglycone of gastrodin, gastrodigenin, demonstrated opposing functions on macrophages, repressing TNF while enhancing IL-6 production (Figure 7A). Both the polyphenolic glycosides naringin and gastrodin had no effect on cytokine production. However, the *B.-uniformis*-liberated aglycone, naringenin repressed IL-6 and enhanced TNF production by macrophages, while we observed the opposite effect (IL-6 enhancement and TNF repression by the *B.-uniformis*-liberated aglycone of gastrodin, gastrodigenin (Figure 7B). We found that the polyphenolic glycoside phloridzin is inert to macrophages while the *B.-uniformis*-liberated aglycone, phloretin, similarly does not affect TNF but suppresses IL-6 production from macrophages. Together, our findings reveal generation of remarkable diversity of inflammatory function of *B.-uniformis*-liberated aglycones across diverse PSM glycosides.

The bioactivation of the anti-inflammatory effect of PSM glycosides such as salicin, salidroside, rutin, and naringin by *B. uniformis* suggested that anti-inflammatory functions of these dietary and medicinal plant glycosides[62] may be mediated by metabolism and liberation of their respective bioactive aglycones by the gut microbiome. For example, the aryl glycoside salicin has long been known to have broad analgesic and anti-inflammatory effects across health and disease and is the basis for development of salicylic acid and acetylsalicylic acid as aspirin.[63] We found that the liberated aglycone saligenin, the product of salicin metabolism by *B. uniformis*, demonstrated potent inhibition of both TNF and IL-6 production by LPS-stimulated macrophages (Figures 7A and S5B). Saligenin liberation was dependent on the *Bu* glycoside metabolic system as conversion of salicin to saligenin *in vitro* was absent in *Bu* deficient in all three GH (GHΔΔΔ; Δ*gshD*Δ*gshG*Δ*gghC*) mutants compared with *Bu* wild type (WT) (Figure S6A). Colonization of germ-free mice with either *Bu* WT or *Bu* GHΔΔΔ followed by oral administration of salicin demonstrated both *Bu*-dependent liberation of saligenin in the gut (as detected in feces) and dependence on the *Bu* glycoside metabolic system (Figure 7C).

Please cite this article in press as: Kuziel et al., Functional diversification of dietary plant small molecules by the gut microbiome, Cell (2025), https://doi.org/10.1016/j.cell.2025.01.045

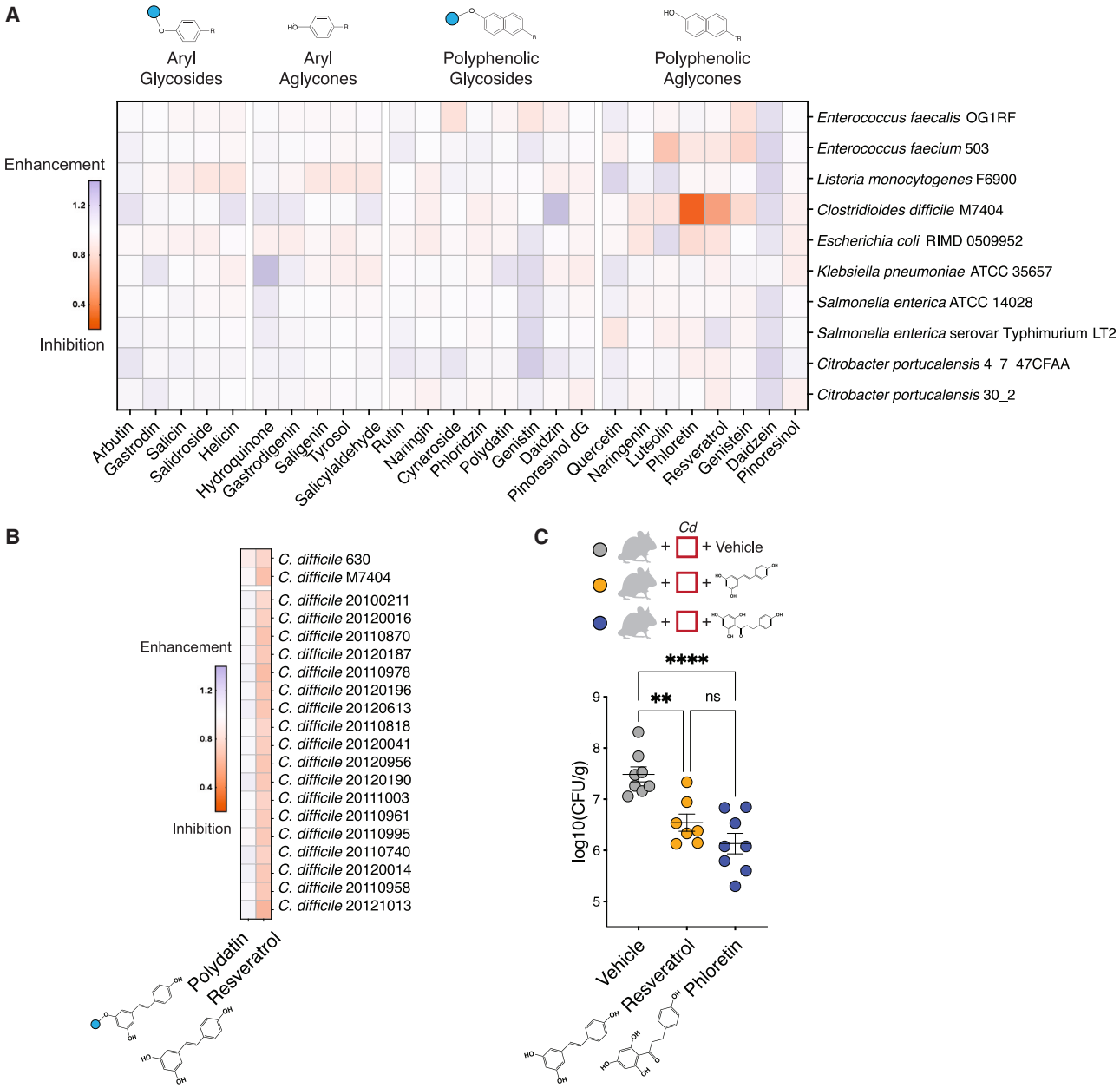


Figure 6. Bioactivation of polyphenolic glycosides into antibiotics by the *Bacteroides*

(A) Heatmap of the impact of polyphenol glycoside and aglycones on growth of a panel of gut bacterial pathogens. Growth is measured as the ratio of growth in media (AUC) supplemented with glucose and 150 μM glycoside or aglycone vs. growth in media (AUC) supplemented with glucose and vehicle. Value > 1 is representative of enhanced growth in compound and value < 1 is representative of inhibition of growth in compound.

(B) Heatmap of growth of 20 *Clostridioides difficile* isolates in the absence or presence of the polyphenolic glycoside polydatin or the *Bacteroides*-liberated aglycone resveratrol. Growth induction or inhibition is measured as the ratio of growth in media (AUC) supplemented with glucose and 150 μM polydatin or resveratrol vs. growth in media (AUC) supplemented with glucose and vehicle. Value > 1 is representative of enhanced growth in compound and value < 1 is representative of inhibition of growth in compound.

(C) Amount of *C. difficile* (colony-forming units) at day 2 in mice infected with *C. difficile* 630 treated with vehicle or 500 μM resveratrol or phloretin in their drinking water.

In (A), dG is diglucoside. Throughput screen data in (A) are for one biological replicate, in technical triplicate. In (B), data are representative of one or two biological replicates, each in technical triplicate. In (C), data are combined from two independent experiments (n = 3–5 mice per experiment). In (C), error bars represent mean ± SEM. Non-significant $p > 0.05$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$; one-way ANOVA in (C) with Tukey's (C) post hoc test.

See also Figures S4 and S7.

Please cite this article in press as: Kuziel et al., Functional diversification of dietary plant small molecules by the gut microbiome, Cell (2025), https://doi.org/10.1016/j.cell.2025.01.045

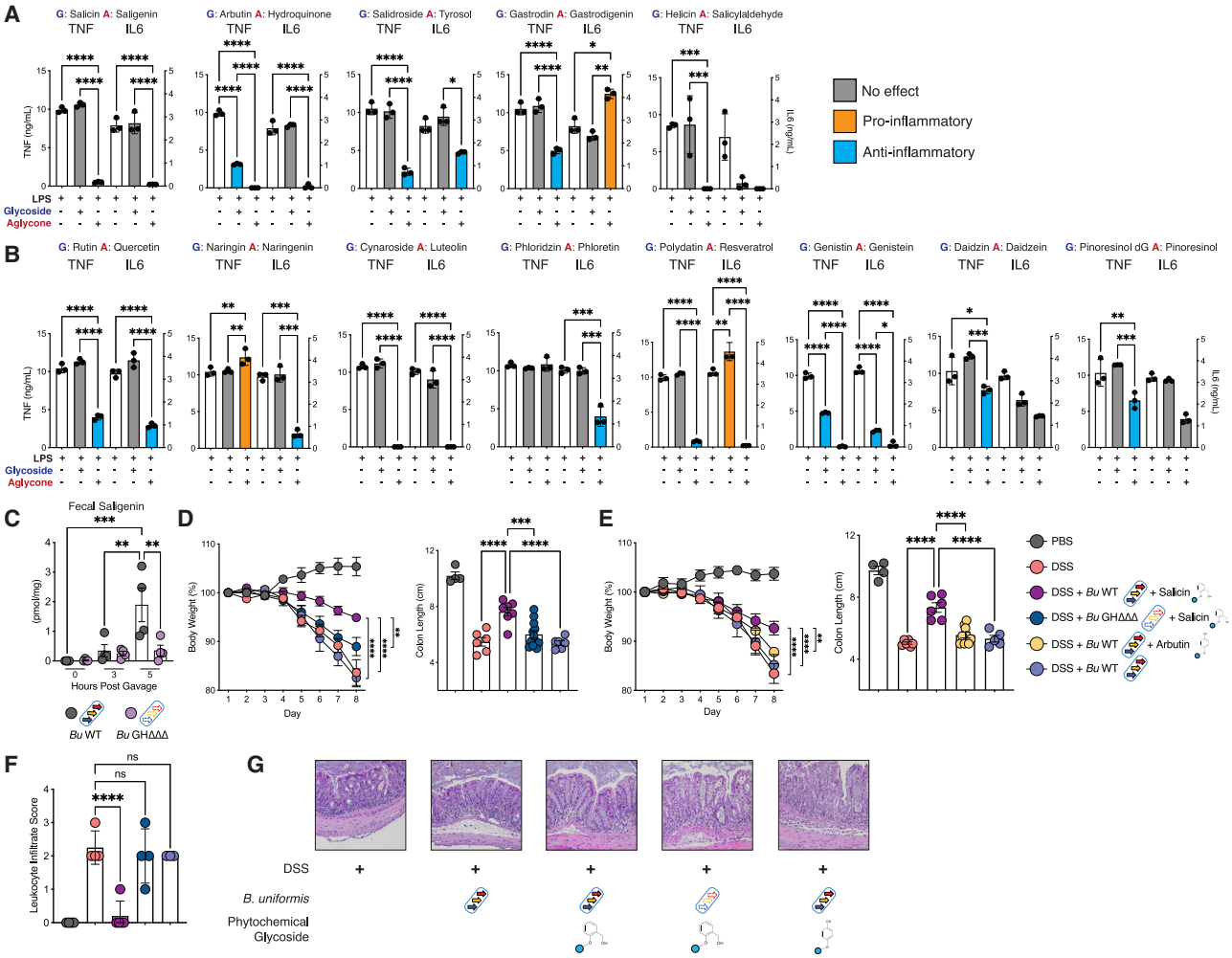


Figure 7. Functionalization of distinct dietary phenolic glycosides by the *Bacteroides* mediates intestinal immune homeostasis

(A) TNF (left) and IL-6 (right) secretion by immortalized murine bone marrow-derived macrophages (iBMDMs) after 24-h LPS (1 μg/mL; *E. coli* O111:B4) stimulation in the absence or presence of specific aryl glycosides (40 h total; 16-h pre-LPS treatment incubation and 24-h LPS stimulation) and their respective aglycones, as indicated. Orange indicates significant induction of cytokine production vs. LPS alone, and blue indicates significant inhibition of cytokine production vs. LPS alone.

(B) TNF (left) and IL-6 (right) cytokine production induction or inhibition by phenolic glycosides and their respective aglycones. Orange indicates significant induction of cytokine production vs. LPS alone, and blue indicates significant inhibition of cytokine production vs. LPS alone.

(C) Fecal saligenin concentration of gnotobiotic mice colonized with *B. uniformis* WT or *B. uniformis* Δ*gshD*Δ*gshG*Δ*gghC* at time points as indicated after intragastric gavage of salicin (100 mg/kg body weight).

(D) Body weight and colon length of PBS- or DSS (2.5%)-treated mice colonized with *B. uniformis* WT or *B. uniformis* Δ*gshD*Δ*gshG*Δ*gghC* with daily i.g. administration of salicin (100 mg/kg body weight) or PBS control.

(E) Body weight and colon length of PBS- or DSS (2.5%)-treated mice colonized with *B. uniformis* WT with daily i.g. administration of salicin (100 mg/kg body weight), arbutin (100 mg/kg body weight) or PBS control.

(F) Histopathologic scoring of leukocyte infiltrate. For the experiments described in Figures 5D and 5E.

(G) Representative H&E images for the experiments described in Figures 5D and 5E.

In (B), dG is diglucoside. In (A) and (B), data are representative of two biological replicates, each in technical triplicate. In (C), data are from one independent experiment ($n = 4$ mice per experiment). In (D)–(F), data are combined from two to three independent experiments ($n = 3$–5 mice per experiment). In (D) and (E), colon lengths were measured at day 8. In (A) and (B), error bars represent mean ± SD. In (C)–(F), error bars represent mean ± SEM. Non-significant $p > 0.05$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, $^{****}p < 0.0001$; one-way ANOVA in (A)–(F) with Tukey's (A–F) and two-way ANOVA in (D) and (E) with Šídák's (D and E).

See also Figures S5, S6, and S7.

We next sought to determine whether *B. uniformis* metabolism of salicin into the anti-inflammatory aglycone saligenin impacted intestinal homeostasis *in vivo*. To address this, we employed a mouse model of intestinal inflammation, dextran sodium sulfate (DSS) in drinking water in specific pathogen-free (SPF) mice colonized with either *Bu* WT or *Bu* GHΔΔΔ. Mice colonized with *Bu* WT and fed salicin were protected from weight loss, colon shortening, and leukocyte infiltration, compared with mice

Please cite this article in press as: Kuziel et al., Functional diversification of dietary plant small molecules by the gut microbiome, Cell (2025), https://doi.org/10.1016/j.cell.2025.01.045

colonized with *Bu* WT without salicin or mice colonized with *Bu* GHΔΔΔ fed salicin (Figures 7D–7F) (in which *Bu* WT and GHΔΔΔ equally colonized mice; Figure S6B). Notably, colons from mice colonized with *Bu* WT fed salicin, compared with *Bu* GHΔΔΔ fed salicin were protected from DSS-induced accumulation of interferon (IFN)-γ-producing $CD4^+$ T cells and T_{bet}^+ T_H1 cells at the lamina propria, suggesting a role of liberation of saligenin from salicin by the *Bu* glycoside metabolic system in regulating pathogenic Th1 mediated intestinal inflammation (Figure S6C). Protection from colitis occurred in the setting of a restored microbiome after antibiotics (allowing *B. uniformis* administration) (Figures S6D and S6E) and without evidence of saligenin degradation by the mouse gut microbiome (Figure S6F). Administration of saligenin in antibiotic-naive mice with intact microbiome diversity and composition (Figures S6D and S6G) protected mice from colitis (Figures S6H–S6J), demonstrating the direct role of saligenin in protection from colitis by *B. uniformis* metabolism of salicin. Similar to salicin, the closely related aryl glycoside arbutin is metabolized by the combination of *Bu* GH3 *gshD* and *Bu* GH16 gghC (Figure 3H). Importantly, however, liberation of the aglycone hydroquinone by *Bu* (Figure 5A) results in cell toxicity (Figure S5A) compared with the anti-inflammatory properties of the *Bu*-liberated saligenin from salicin (Figures 7A and S5B). Compared with mice fed salicin, dietary arbutin did not protect *Bu* WT-colonized mice from DSS-mediated weight loss, colon shortening, or leucocyte inflammation (Figures 7E–7G and S6K). Together, our data demonstrate both the critical role of microbiome bioactivation of salicin to mediate the anti-inflammatory and intestinal immune homeostatic role *in vivo* of this ancient medicinal factor by the gut microbiome and the role of functional diversification of dietary glycosides mediated by the same gut microbial member and glycoside metabolic system on differential host outcomes.

DISCUSSION

In their capacity to generate energy to replicate and survive within the intestine, members of the human gut microbiome employ hundreds of diverse GHs ultimately transforming a multitude of chemically diverse polysaccharides into a limited number of host-active fermentation products (e.g., formate, acetate, propionate, lactate, succinate, and butyrate). As such, the biosynthetic diversity and chemical complexity of plant metabolism to produce glycans is "funneled" whereby plant chemical diversity via a matched diversity of GHs is reduced to a limited set of effectors metabolites serving as the functional output of gut microbiome metabolism of plants on the host. Here, we show that deglycosylation of diverse glycosides by a defined set of GHs encoded within the gut microbiome results in the bioactivation of diverse host-active functions mediated by liberated aglycone products of microbiome enzymatic transformation. Thus, in contrast to such reduction of the biosynthetic diversity of plant glycans, deglycosylation of glycosides by the gut microbiome maintains the biosynthetic diversity of plant small molecules acting in bowtie architecture[64,65] to generate diverse host effector functions (Figure S7A).

GHs within the human gut microbiome have been demonstrated to have substrate specificity dependent on the type of monosaccharide (e.g., glucose, fructose, arabinose, and xylose), stereochemistry (α vs. β linkage) and site of linkage (α1, β2, etc.) between these two sugars.[17,66–69] Here, we reveal two distinct metabolic systems within the human gut *Bacteroides* to hydrolyze glycoside PSMs (Figure S7B). In *B. ovatus*, a single GH16 family member is required to utilize both a range of phenolic glycosides as well as disaccharides. Such substrate specificity is consistent among members of human associated microbiomes, notably that of *B. thetaiotaomicron* (hydrolyzing both glucosinolates and disaccharides[47]), *Enterococcus*,[39,70] *Lactobacillus*,[45] *Streptococcus*,[46] and *E. coli*[38,71,72] in which the GH genes/enzymes are capable of hydrolyzing both glycosides and disaccharides. Such enzyme systems may have evolved primarily for sugar-sugar hydrolysis with broad specificity for glycosides or later expanded functional capacity to include glycoside hydrolysis. In contrast, in *B. uniformis*, one of the most prevalent members of the human gut microbiota,[48,49] we have revealed a glycoside utilization system composed of GH3 family members dedicated to phenolic glycosides and unable to hydrolyze disaccharides (*Bu* gshD) and demonstrating specificity for distinct phenolic glycosides depending on the chemical structure of the aglycone phenol group (for example, the position of hydroxyl or methanol group) (*Bu* gshG). Interestingly, these two GH3 of *B. uniformis* are present within common clades of GH3 shared with members of plant and soil associated bacteria that hydrolyze plant glycosides[73–76] (Figure S3G). Together, this suggests that our demonstration of glycoside-dedicated and -specific GH3 enzymes within the human gut Bacteroidales evolved specifically for plant glycosides. It will be interesting to identify the plant glycoside utilization systems and substrate specificity for members of the gut microbiome across diverse taxa (Lachnospiraceae, Clostridiaceae, and Erysipelotrichales) for which we have identified unique glycoside utilization patterns (Figure 1B) and of herbivores (birds and non-human primates) and humans subsisting on plant diets abundant in PSMs such as in gathering cultures.[77,78]

PSM glycosides have been associated with a range of effects on human health and disease. Here, we demonstrate the role of the microbiome in mediating such impacts via the bioactivation and liberation of PSM aglycones in the gut. For salicin, long associated with anti-inflammatory effects, we show that immunoregulation in the gut is mediated by microbiome deglycosylation. In comparison, our screen of *Bacteroides* metabolism and liberation of aglycones across a panel of dietary phenolic glycosides revealed both the novel pro-inflammatory functions of specific liberated aglycones and in the case of resveratrol, the aglycone product of microbiome deglycosylation of polydatin, as a specific inhibitor of the gut pathogen *C. difficile*, adding to the benefits of resveratrol on human health.[79,80] Our studies both reveal the impact of the microbiome on known connections between plant dietary glycosides and the host (intestinal inflammation) and of new functions and mechanisms of colonization resistance based on the combination of specific dietary substrates and microbiome bioactivation conferring anti-microbial function. By revealing the microbiome functional activation of specific PSM glycosides across a range of known and as-of-yet unknown host outcomes across physiology and disease, our work presents a novel approach based on pharmacologic dosing and

Please cite this article in press as: Kuziel et al., Functional diversification of dietary plant small molecules by the gut microbiome, Cell (2025), https://doi.org/10.1016/j.cell.2025.01.045

delivery of specific glycosides targeting specific members (endogenous, probiotic, or engineered) to convert and deliver specific functions in prevention and treatment of disease. More broadly, given the diminishing quantities of PSMs in human diets,[81,82] elucidation of the conversion of PSMs into immunoregulatory products by the microbiome presents a novel mechanism of diet-microbiome interactions contributing to the increasing incidence of diseases of industrialization such as autoimmunity and atopy.

Limitations of the study

There are certain limitations of our work that will be the focus of future research. First, as our work focused on plant phenolic glycoside utilization and bioactivation by a panel of members of the human gut microbiome *in vitro* and in the context of the murine gut microbiome *in vivo*, future work will seek to define PSM glycoside metabolism across a broader membership of the human gut microbiome and at the community-level toward defining the role of inter-individual variation in the human gut microbiome on PSM glycoside metabolism utilizing *in vivo* (humanized mice) and *ex vivo* (human fecal culture) approaches. Relatedly, given the expansive metabolic capacity of the human gut microbiome, future studies will focus on defining how liberated and functional aglycones may be transformed by members of the gut microbiome resulting in increased or decreased bioactivity. Targeting such members and/or aglycone metabolism can be harnessed to optimize the functional effects of microbiome-liberated aglycones *in vivo*. While our work has revealed functional roles of specific microbiome-liberated aglycones, future studies will elucidate the molecular mechanisms, such as via specific host chemoreceptors or metabolic pathways, mediating the immunoregulatory functions the salicin aglycone saligenin and the mechanism of action of resveratrol on inhibition of *C. difficile*. Finally, our work focused on metabolism and functionalization of PSM phenolic glycosides administered individually, future work will determine gut microbiome bioactivation of PSM within the natural matrix of dietary plants.

RESOURCE AVAILABILITY

Lead contact

Requests for further information and resources may be directed to and will be fulfilled by the lead contact, Seth Rakoff-Nahoum (seth.rakoff-nahoum@childrens.harvard.edu).

Materials availability

Plasmids and recombinant strains generated in this study will be distributed on request.

Data and code availability

- This paper does not report original code.
- Replicate data are available at Mendeley data: https://doi.org/10.17632/yx627x7hns.1.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

We thank members of the S.R.-N. lab for helpful comments on the manuscript. We thank C. Vidoudez at the Harvard Center for Mass Spectrometry as well as T. Kunchok and T. Tseyang at the Metabolite Profiling Core Facility of the Whitehead Institute for LC-MS method development and glycoside/aglycone detection. We thank S. Wang and M. Liu for assistance with CDI experiments. We thank E. Reynolds and S. Xu for helpful discussion. This work is supported by an NDSEG Fellowship (to G.A.K.), the Keck Foundation (to J.-K.W.), the Mathers Foundation (to J.-K.W. and S.R.-N.), a Career Award for Medical Scientists from the Burroughs Wellcome Fund (to S.R.-N.), a Pew Biomedical Scholarship (to S.R.-N.), a Pew Latin America Fellowship (to G.L.L.), the NIH-Funded Harvard Digestive Disease Center (P30DK034854 to M.D.), the National Institutes of Health (K08AI130392 and DP2GM136652 to S.R.-N.; R01AI132387 and R01AI139087 to M.D.; R01AI1269151 to T.C. and S.R.-N.), and the Food Allergy Science Initiative (to S.R.-N. and J.-K.W.).

AUTHOR CONTRIBUTIONS

G.A.K. and S.R.-N. conceived the project, and G.A.K., G.L.L., and S.R.-N. designed experiments. G.A.K. generated bacterial mutants and constructs, purified recombinant enzymes, and performed *in vitro* bacterial growth experiments and biochemistry experiments. G.A.K., J.M., and E.S.-V. performed *in vivo* infection and colitis model experiments. G.A.K. and G.L.L. performed *in silico* and phylogenetic analyses. G.A.K. and C.S. designed the enzymology experiments. C.S. performed LC-MS analyses of *in vitro* biochemistry samples. J.M. performed histological scoring of mouse colon sections. L.L. performed mouse fecal cultivation and aglycone metabolism experiments. G.A.K., G.L.L., and S.R.-N. wrote the manuscript with input from all authors.

DECLARATION OF INTERESTS

G.A.K. and S.R.-N. have pending patent applications related to methods or composition for using *Bacteroides* phytochemical metabolism to treat inflammatory bowel disease (IBD) and CDI. G.A.K. has financial interests in Hasana Biosciences. J.-K.W. is a member of the SAB and has equity in DoubleRainbow Biosciences, Galixir, and Inari Agriculture. T.C. is a founder of and has equity in Pareto Bio.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Microbial strains
  - Animal models
- METHOD DETAILS
  - Bacterial culture
  - Gut bacterial glycoside utilization
  - Gut *Bacteroides* glycoside utilization
  - Transposon insertion sequencing
  - Tn-seq library preparation
  - Generation of *Bacteroides* gene deletion mutants
  - Heterologous expression
  - Recombinant protein purification
  - In vitro aglycone liberation
  - Gnotobiotic mouse experiments
  - Glycoside and aglycone detection by LCMS
  - Michaelis-Menten enzyme kinetics
  - Relative enzyme activity
  - Recombinant enzyme activity LCMS
  - Phytochemical microbial killing assays
  - In vitro polydatin bioactivation and *C. difficile* antagonism
  - *C. difficile* infection (CDI) model
  - *C. difficile* colony-forming units (CFUs)
  - In vitro macrophage inflammation screen
  - Mouse model of intestinal inflammation
  - In vivo *Bacteroides* colonization
  - Fecal sample collection and microbiome NGS sequencing

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cell.2025.01.045.

Received: November 1, 2023
Revised: December 9, 2024
Accepted: January 31, 2025
Published: March 7, 2025

REFERENCES

1. Willett, W.C. (1994). Diet and health: what should we eat? Science *264*, 532–537. https://doi.org/10.1126/science.8160011.
2. Wright, N., Wilson, L., Smith, M., Duncan, B., and McHugh, P. (2017). The BROAD study: A randomised controlled trial using a whole food plant-based diet in the community for obesity, ischaemic heart disease or diabetes. Nutr. Diabetes *7*, e256. https://doi.org/10.1038/nutd.2017.3.
3. Hu, F.B. (2003). Plant-based foods and prevention of cardiovascular disease: an overview. Am. J. Clin. Nutr. *78*, 544S–551S. https://doi.org/10.1093/ajcn/78.3.544S.
4. Chicco, F., Magrì, S., Cingolani, A., Paduano, D., Pesenti, M., Zara, F., Tumbarello, F., Urru, E., Melis, A., Casula, L., et al. (2021). Multidimensional impact of Mediterranean diet on IBD patients. Inflamm. Bowel Dis. *27*, 1–9. https://doi.org/10.1093/ibd/izaa097.
5. Sköldstam, L., Hagfors, L., and Johansson, G. (2003). An experimental study of a Mediterranean diet intervention for patients with rheumatoid arthritis. Ann. Rheum. Dis. *62*, 208–214. https://doi.org/10.1136/ard.62.3.208.
6. Graham, S., Schotz, W., and Martino, P. (1972). Alimentary factors in the epidemiology of gastric cancer. Cancer *30*, 927–938. https://doi.org/10.1002/1097-0142(197210)30:4<927::aid-cncr2820300411>3.0.co;2-l.
7. Tantamango-Bartley, Y., Jaceldo-Siegl, K., Fan, J., and Fraser, G. (2013). Vegetarian diets and the incidence of cancer in a low-risk population. Cancer Epidemiol. Biomarkers Prev. *22*, 286–294. https://doi.org/10.1158/1055-9965.EPI-12-1060.
8. Morris, M.C., Tangney, C.C., Wang, Y., Sacks, F.M., Bennett, D.A., and Aggarwal, N.T. (2015). MIND diet associated with reduced incidence of Alzheimer's disease. Alzheimers Dement. *11*, 1007–1014. https://doi.org/10.1016/j.jalz.2014.11.009.
9. Schmidt, B.M., Ribnicky, D.M., Lipsky, P.E., and Raskin, I. (2007). Revisiting the ancient concept of botanical therapeutics. Nat. Chem. Biol. *3*, 360–366. https://doi.org/10.1038/nchembio0707-360.
10. Atanasov, A.G., Waltenberger, B., Pferschy-Wenzig, E.M., Linder, T., Wawrosch, C., Uhrin, P., Temml, V., Wang, L., Schwaiger, S., Heiss, E.H., et al. (2015). Discovery and resupply of pharmacologically active plant-derived natural products: a review. Biotechnol. Adv. *33*, 1582–1614. https://doi.org/10.1016/j.biotechadv.2015.08.001.
11. Asnicar, F., Berry, S.E., Valdes, A.M., Nguyen, L.H., Piccinno, G., Drew, D.A., Leeming, E., Gibson, R., Le Roy, C., Khatib, H.A., et al. (2021). Microbiome connections with host metabolism and habitual diet from 1,098 deeply phenotyped individuals. Nat. Med. *27*, 321–332. https://doi.org/10.1038/s41591-020-01183-8.
12. Lynch, S.V., and Pedersen, O. (2016). The human intestinal microbiome in health and disease. N. Engl. J. Med. *375*, 2369–2379. https://doi.org/10.1056/NEJMra1600266.
13. Kuziel, G.A., and Rakoff-Nahoum, S. (2022). The gut microbiome. Curr. Biol. *32*, R257–R264. https://doi.org/10.1016/j.cub.2022.02.023.
14. Donohoe, D.R., Garge, N., Zhang, X., Sun, W., O'Connell, T.M., Bunger, M.K., and Bultman, S.J. (2011). The microbiome and butyrate regulate energy metabolism and autophagy in the mammalian colon. Cell Metab. *13*, 517–526. https://doi.org/10.1016/j.cmet.2011.02.018.
15. Furusawa, Y., Obata, Y., Fukuda, S., Endo, T.A., Nakato, G., Takahashi, D., Nakanishi, Y., Uetake, C., Kato, K., Kato, T., et al. (2013). Commensal microbe-derived butyrate induces the differentiation of colonic regulatory T cells. Nature *504*, 446–450. https://doi.org/10.1038/nature12721.
16. Koh, A., De Vadder, F., Kovatcheva-Datchary, P., and Bäckhed, F. (2016). From dietary fiber to host physiology: short-chain fatty acids as key bacterial metabolites. Cell *165*, 1332–1345. https://doi.org/10.1016/j.cell.2016.05.041.
17. Martens, E.C., Lowe, E.C., Chiang, H., Pudlo, N.A., Wu, M., McNulty, N.P., Abbott, D.W., Henrissat, B., Gilbert, H.J., Bolam, D.N., et al. (2011). Recognition and degradation of plant cell wall polysaccharides by two human gut symbionts. PLoS Biol. *9*, e1001221. https://doi.org/10.1371/journal.pbio.1001221.
18. Luis, A.S., Briggs, J., Zhang, X., Farnell, B., Ndeh, D., Labourel, A., Baslé, A., Cartmell, A., Terrapon, N., Stott, K., et al. (2018). Dietary pectic glycans are degraded by coordinated enzyme pathways in human colonic Bacteroides. Nat. Microbiol. *3*, 210–219. https://doi.org/10.1038/s41564-017-0079-1.
19. El Kaoutari, A., Armougom, F., Gordon, J.I., Raoult, D., and Henrissat, B. (2013). The abundance and variety of carbohydrate-active enzymes in the human gut microbiota. Nat. Rev. Microbiol. *11*, 497–504. https://doi.org/10.1038/nrmicro3050.
20. Huang, X.-Q., and Dudareva, N. (2023). Plant specialized metabolism. Curr. Biol. *33*, R473–R478. https://doi.org/10.1016/j.cub.2023.01.057.
21. Erb, M., and Kliebenstein, D.J. (2020). Plant secondary metabolites as defenses, regulators, and primary metabolites: the blurred functional trichotomy. Plant Physiol. *184*, 39–52. https://doi.org/10.1104/pp.20.00433.
22. Pieterse, C.M.J., Leon-Reyes, A., Van Der Ent, S., and Van Wees, S.C.M. (2009). Networking by small-molecule hormones in plant immunity. Nat. Chem. Biol. *5*, 308–316. https://doi.org/10.1038/nchembio.164.
23. Wittstock, U., and Gershenzon, J. (2002). Constitutive plant toxins and their role in defense against herbivores and pathogens. Curr. Opin. Plant Biol. *5*, 300–307. https://doi.org/10.1016/s1369-5266(02)00264-9.
24. Nett, R.S., Dho, Y., Low, Y.Y., and Sattely, E.S. (2021). A metabolic regulon reveals early and late acting enzymes in neuroactive Lycopodium alkaloid biosynthesis. Proc. Natl. Acad. Sci. USA *118*, e2102949118. https://doi.org/10.1073/pnas.2102949118.
25. Cowan, M.M. (1999). Plant products as antimicrobial agents. Clin. Microbiol. Rev. *12*, 564–582. https://doi.org/10.1128/CMR.12.4.564.
26. Kren, V., and Martínková, L. (2001). Glycosides in medicine: "The role of glycosidic residue in biological activity.". Curr. Med. Chem. *8*, 1303–1328. https://doi.org/10.2174/0929867013372193.
27. Geleijnse, J.M., Launer, L.J., Hofman, A., Pols, H.A., and Witteman, J.C. (1999). Tea flavonoids may protect against atherosclerosis: the Rotterdam Study. Arch. Intern. Med. *159*, 2170–2174. https://doi.org/10.1001/archinte.159.18.2170.
28. Hertog, M.G., Feskens, E.J., Hollman, P.C., Katan, M.B., and Kromhout, D. (1993). Dietary antioxidant flavonoids and risk of coronary heart disease: the Zutphen Elderly Study. Lancet *342*, 1007–1011. https://doi.org/10.1016/0140-6736(93)92876-u.
29. Kytidou, K., Artola, M., Overkleeft, H.S., and Aerts, J.M.F.G. (2020). Plant glycosides and glycosidases: A treasure-trove for therapeutics. Front. Plant Sci. *11*, 357. https://doi.org/10.3389/fpls.2020.00357.

30. Horowitz, R.M., and Gentili, B. (1969). Taste and structure in phenolic glycosides. J. Agric. Food Chem. *17*, 696–700. https://doi.org/10.1021/jf60164a049.

31. Gosch, C., Halbwirth, H., and Stich, K. (2010). Phloridzin: biosynthesis, distribution and physiological relevance in plants. Phytochemistry *71*, 838–843. https://doi.org/10.1016/j.phytochem.2010.03.003.

32. Karami, A., Fakhri, S., Kooshki, L., and Khan, H. (2022). Polydatin: pharmacological mechanisms, therapeutic targets, biological activities, and health benefits. Molecules *27*, 6474. https://doi.org/10.3390/molecules27196474.

33. Tungmunnithum, D., Thongboonyou, A., Pholboon, A., and Yangsabai, A. (2018). Flavonoids and other phenolic compounds from medicinal plants for pharmaceutical and medical aspects: an overview. Medicines (Basel) *5*, 93. https://doi.org/10.3390/medicines5030093.

34. Okada, K., Shoda, J., Kano, M., Suzuki, S., Ohtake, N., Yamamoto, M., Takahashi, H., Utsunomiya, H., Oda, K., Sato, K., et al. (2007). Inchinkoto, a herbal medicine, and its ingredients dually exert Mrp2/MRP2-mediated choleresis and Nrf2-mediated antioxidative action in rat livers. Am. J. Physiol. Gastrointest. Liver Physiol. *292*, G1450–G1463. https://doi.org/10.1152/ajpgi.00302.2006.

35. Cho, E., Chung, E.Y., Jang, H.Y., Hong, O.Y., Chae, H.S., Jeong, Y.J., Kim, S.Y., Kim, B.S., Yoo, D.J., Kim, J.S., et al. (2017). Anti-cancer effect of cyanidin-3-glucoside from mulberry via caspase-3 cleavage and DNA fragmentation in vitro and in vivo. Anti Cancer Agents Med. Chem. *17*, 1519–1525. https://doi.org/10.2174/1871520617666170327152026.

36. Greenway, F., Liu, Z., Yu, Y., and Gupta, A. (2011). A clinical trial testing the safety and efficacy of a standardized Eucommia ulmoides Oliver bark extract to treat hypertension. Altern. Med. Rev. *16*, 338–347.

37. Shah, F.A., Kury, L.A., Li, T., Zeb, A., Koh, P.O., Liu, F., Zhou, Q., Hussain, I., Khan, A.U., Jiang, Y., et al. (2019). Polydatin attenuates neuronal loss via reducing neuroinflammation and oxidative stress in rat MCAO models. Front. Pharmacol. *10*, 663. https://doi.org/10.3389/fphar.2019.00663.

38. Defez, R., and De Felice, M. (1981). Cryptic operon for beta-glucoside metabolism in Escherichia coli K12: genetic evidence for a regulatory protein. Genetics *97*, 11–25. https://doi.org/10.1093/genetics/97.1.11.

39. Tamura, G., Gold, C., Ferro-Luzzi, A., and Ames, B.N. (1980). Fecalase: a model for activation of dietary glycosides to mutagens by intestinal flora. Proc. Natl. Acad. Sci. USA *77*, 4961–4965. https://doi.org/10.1073/pnas.77.8.4961.

40. Eckburg, P.B., Bik, E.M., Bernstein, C.N., Purdom, E., Dethlefsen, L., Sargent, M., Gill, S.R., Nelson, K.E., and Relman, D.A. (2005). Diversity of the human intestinal microbial flora. Science *308*, 1635–1638. https://doi.org/10.1126/science.1110591.

41. Arumugam, M., Raes, J., Pelletier, E., Le Paslier, D., Yamada, T., Mende, D.R., Fernandes, G.R., Tap, J., Bruls, T., Batto, J.-M., et al. (2011). Enterotypes of the human gut microbiome. Nature *473*, 174–180. https://doi.org/10.1038/nature09944.

42. Sonnenburg, J.L., and Bäckhed, F. (2016). Diet–microbiota interactions as moderators of human metabolism. Nature *535*, 56–64. https://doi.org/10.1038/nature18846.

43. Alexander, M., and Turnbaugh, P.J. (2020). Deconstructing mechanisms of diet-microbiome-immune interactions. Immunity *53*, 264–276. https://doi.org/10.1016/j.immuni.2020.07.015.

44. Lu, S., Wang, J., Chitsaz, F., Derbyshire, M.K., Geer, R.C., Gonzales, N.R., Gwadz, M., Hurwitz, D.I., Marchler, G.H., Song, J.S., et al. (2020). CDD/SPARCLE: the conserved domain database in 2020. Nucleic Acids Res. *48*, D265–D268. https://doi.org/10.1093/nar/gkz991.

45. Theilmann, M.C., Goh, Y.J., Nielsen, K.F., Klaenhammer, T.R., Barrangou, R., and Abou Hachem, M. (2017). Lactobacillus acidophilus metabolizes dietary plant glucosides and externalizes their bioactive phytochemicals. mBio *8*, e01421–e01417. https://doi.org/10.1128/mBio.01421-17.

46. Kiliç, A.O., Tao, L., Zhang, Y., Lei, Y., Khammanivong, A., and Herzberg, M.C. (2004). Involvement of Streptococcus gordonii beta-glucoside metabolism systems in adhesion, biofilm formation, and in vivo gene expression. J. Bacteriol. *186*, 4246–4253. https://doi.org/10.1128/JB.186.13.4246-4253.2004.

47. Liou, C.S., Sirk, S.J., Diaz, C.A.C., Klein, A.P., Fischer, C.R., Higginbottom, S.K., Erez, A., Donia, M.S., Sonnenburg, J.L., and Sattely, E.S. (2020). A metabolic pathway for activation of dietary glucosinolates by a human gut symbiont. Cell *180*, 717–728.e19. https://doi.org/10.1016/j.cell.2020.01.023.

48. Zitomersky, N.L., Coyne, M.J., and Comstock, L.E. (2011). Longitudinal analysis of the prevalence, maintenance, and IgA response to species of the order Bacteroidales in the human gut. Infect. Immun. *79*, 2012–2020. https://doi.org/10.1128/IAI.01348-10.

49. Ishikawa, E., Matsuki, T., Kubota, H., Makino, H., Sakai, T., Oishi, K., Kushiro, A., Fujimoto, J., Watanabe, K., Watanuki, M., et al. (2013). Ethnic diversity of gut microbiota: species characterization of Bacteroides fragilis group and genus Bifidobacterium in healthy Belgian adults, and comparison with data from Japanese subjects. J. Biosci. Bioeng. *116*, 265–270. https://doi.org/10.1016/j.jbiosc.2013.02.010.

50. Teufel, F., Almagro Armenteros, J.J., Johansen, A.R., Gíslason, M.H., Pihl, S.I., Tsirigos, K.D., Winther, O., Brunak, S., von Heijne, G., and Nielsen, H. (2022). SignalP 6.0 predicts all five types of signal peptides using protein language models. Nat. Biotechnol. *40*, 1023–1025. https://doi.org/10.1038/s41587-021-01156-3.

51. Pandey, K.B., and Rizvi, S.I. (2009). Plant polyphenols as dietary antioxidants in human health and disease. Oxid. Med. Cell. Longev. *2*, 270–278. https://doi.org/10.4161/oxim.2.5.9498.

52. Tsao, R. (2010). Chemistry and biochemistry of dietary polyphenols. Nutrients *2*, 1231–1246. https://doi.org/10.3390/nu2121231.

53. Oliver-Bever, B. (1970). Why do plants produce drugs? Which is their Function in the Plants? Q. J. Crude Drug Res. *10*, 1541–1549. https://doi.org/10.3109/13880207009066221.

54. Biere, A., Marak, H.B., and van Damme, J.M.M. (2004). Plant chemical defense against herbivores and pathogens: generalized defense or trade-offs? Oecologia *140*, 430–441. https://doi.org/10.1007/s00442-004-1603-6.

55. Du, Q.H., Peng, C., and Zhang, H. (2013). Polydatin: a review of pharmacology and pharmacokinetics. Pharm. Biol. *51*, 1347–1354. https://doi.org/10.3109/13880209.2013.792849.

56. Leffler, D.A., and Lamont, J.T. (2015). Clostridium difficile infection. N. Engl. J. Med. *372*, 1539–1548. https://doi.org/10.1056/NEJMra1403772.

57. Little, A.S., Younker, I.T., Schechter, M.S., Bernardino, P.N., Méheust, R., Stemczynski, J., Scorza, K., Mullowney, M.W., Sharan, D., Waligurski, E., et al. (2024). Dietary- and host-derived metabolites are used by diverse gut bacteria for anaerobic respiration. Nat. Microbiol. *9*, 55–69. https://doi.org/10.1038/s41564-023-01560-2.

58. Bode, L.M., Bunzel, D., Huch, M., Cho, G.-S., Ruhland, D., Bunzel, M., Bub, A., Franz, C.M.A.P., and Kulling, S.E. (2013). In vivo and in vitro metabolism of trans-resveratrol by human gut microbiota. Am. J. Clin. Nutr. *97*, 295–309. https://doi.org/10.3945/ajcn.112.049379.

59. Kuhn, B.M., Errafi, S., Bucher, R., Dobrev, P., Geisler, M., Bigler, L., Zažímalová, E., and Ringli, C. (2016). 7-rhamnosylated flavonols modulate homeostasis of the plant hormone auxin and affect plant development. J. Biol. Chem. *291*, 5385–5395. https://doi.org/10.1074/jbc.M115.701565.

60. Kepp, O., Menger, L., Vacchelli, E., Adjemian, S., Martins, I., Ma, Y., Sukkurwala, A.Q., Michaud, M., Galluzzi, L., Zitvogel, L., et al. (2012). Anticancer activity of cardiac glycosides: at the frontier between cell-autonomous and immunological effects. Oncoimmunology *1*, 1640–1642. https://doi.org/10.4161/onci.21684.

61. Wang, J., Fang, X., Ge, L., Cao, F., Zhao, L., Wang, Z., and Xiao, W. (2018). Antitumor, antioxidant and anti-inflammatory activities of

kaempferol and its corresponding glycosides and the enzymatic preparation of kaempferol. PLoS One *13*, e0197563. https://doi.org/10.1371/journal.pone.0197563.

62. Fiebich, B.L., and Chrubasik, S. (2004). Effects of an ethanolic salix extract on the release of selected inflammatory mediators in vitro. Phytomedicine *11*, 135–138. https://doi.org/10.1078/0944-7113-00338.
63. Desborough, M.J.R., and Keeling, D.M. (2017). The aspirin story - from willow to wonder drug. Br. J. Haematol. *177*, 674–683. https://doi.org/10.1111/bjh.14520.
64. Friedlander, T., Mayo, A.E., Tlusty, T., and Alon, U. (2015). Evolution of bow-tie architectures in biology. PLoS Comput. Biol. *11*, e1004055. https://doi.org/10.1371/journal.pcbi.1004055.
65. Kitano, H. (2004). Biological robustness. Nat. Rev. Genet. *5*, 826–837. https://doi.org/10.1038/nrg1471.
66. Larsbrink, J., Rogers, T.E., Hemsworth, G.R., McKee, L.S., Tauzin, A.S., Spadiut, O., Klinter, S., Pudlo, N.A., Urs, K., Koropatkin, N.M., et al. (2014). A discrete genetic locus confers xyloglucan metabolism in select human gut Bacteroidetes. Nature *506*, 498–502. https://doi.org/10.1038/nature12907.
67. Martens, E.C., Kelly, A.G., Tauzin, A.S., and Brumer, H. (2014). The devil lies in the details: how variations in polysaccharide fine-structure impact the physiology and evolution of gut microbes. J. Mol. Biol. *426*, 3851–3865. https://doi.org/10.1016/j.jmb.2014.06.022.
68. Koropatkin, N.M., Martens, E.C., Gordon, J.I., and Smith, T.J. (2008). Starch catabolism by a prominent human gut symbiont is directed by the recognition of amylose helices. Structure *16*, 1105–1115. https://doi.org/10.1016/j.str.2008.03.017.
69. Boraston, A.B., Bolam, D.N., Gilbert, H.J., and Davies, G.J. (2004). Carbohydrate-binding modules: fine-tuning polysaccharide recognition. Biochem. J. *382*, 769–781. https://doi.org/10.1042/BJ20040892.
70. Grand, M., Aubourg, M., Pikis, A., Thompson, J., Deutscher, J., Hartke, A., and Sauvageot, N. (2019). Characterization of the gen locus involved in beta-1,6-oligosaccharide utilization by Enterococcus faecalis. Mol. Microbiol. *112*, 1744–1756. https://doi.org/10.1111/mmi.14390.
71. Keyhani, N.O., and Roseman, S. (1997). Wild-type Escherichia coli grows on the chitin disaccharide, N,N′-diacetylchitobiose, by expressing the cel operon. Proc. Natl. Acad. Sci. USA *94*, 14367–14371. https://doi.org/10.1073/pnas.94.26.14367.
72. Schaefler, S., and Malamy, A. (1969). Taxonomic investigations on expressed and cryptic phospho-beta-glucosidases in Enterobacteriaceae. J. Bacteriol. *99*, 422–433. https://doi.org/10.1128/jb.99.2.422-433.1969.
73. Watt, D.K., Ono, H., and Hayashi, K. (1998). Agrobacterium tumefaciens beta-glucosidase is also an effective beta-xylosidase, and has a high transglycosylation activity in the presence of alcohols. Biochim. Biophys. Acta *1385*, 78–88. https://doi.org/10.1016/s0167-4838(98)00046-6.
74. Du, L., Wang, Z., Zhao, Y., Huang, J., Pang, H., Wei, Y., Lin, L., and Huang, R. (2014). A beta-glucosidase from Novosphingobium sp. GX9 with high catalytic efficiency toward isoflavonoid glycoside hydrolysis and (+)-catechin transglycosylation. Appl. Microbiol. Biotechnol. *98*, 7069–7079. https://doi.org/10.1007/s00253-014-5661-3.
75. Faure, D., Desair, J., Keijers, V., Bekri, M.A., Proost, P., Henrissat, B., and Vanderleyden, J. (1999). Growth of Azospirillum irakense KBC1 on the aryl beta-glucoside salicin requires either salA or salB. J. Bacteriol. *181*, 3003–3009. https://doi.org/10.1128/JB.181.10.3003-3009.1999.
76. Quan, L.H., Min, J.W., Yang, D.U., Kim, Y.J., and Yang, D.C. (2012). Enzymatic biotransformation of ginsenoside Rb1 to 20(S)-Rg3 by recombinant beta-glucosidase from Microbacterium esteraromaticum. Appl. Microbiol. Biotechnol. *94*, 377–384. https://doi.org/10.1007/s00253-011-3861-7.
77. Muegge, B.D., Kuczynski, J., Knights, D., Clemente, J.C., González, A., Fontana, L., Henrissat, B., Knight, R., and Gordon, J.I. (2011). Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. Science *332*, 970–974. https://doi.org/10.1126/science.1198719.
78. Smits, S.A., Leach, J., Sonnenburg, E.D., Gonzalez, C.G., Lichtman, J.S., Reid, G., Knight, R., Manjurano, A., Changalucha, J., Elias, J.E., et al. (2017). Seasonal cycling in the gut microbiome of the Hadza hunter-gatherers of Tanzania. Science *357*, 802–806. https://doi.org/10.1126/science.aan4834.
79. Baur, J.A., Pearson, K.J., Price, N.L., Jamieson, H.A., Lerin, C., Kalra, A., Prabhu, V.V., Allard, J.S., Lopez-Lluch, G., Lewis, K., et al. (2006). Resveratrol improves health and survival of mice on a high-calorie diet. Nature *444*, 337–342. https://doi.org/10.1038/nature05354.
80. Lagouge, M., Argmann, C., Gerhart-Hines, Z., Meziane, H., Lerin, C., Daussin, F., Messadeq, N., Milne, J., Lambert, P., Elliott, P., et al. (2006). Resveratrol improves mitochondrial function and protects against metabolic disease by activating SIRT1 and PGC-1alpha. Cell *127*, 1109–1122. https://doi.org/10.1016/j.cell.2006.11.013.
81. Darwin, C.R. (1868). The variation of animals and plants under domestication. Br. Foreign Med. Chir. Rev. *42*, 143–166.
82. Larson, G., Piperno, D.R., Allaby, R.G., Purugganan, M.D., Andersson, L., Arroyo-Kalin, M., Barton, L., Climer Vigueira, C., Denham, T., Dobney, K., et al. (2014). Current perspectives and the future of domestication studies. Proc. Natl. Acad. Sci. USA *111*, 6139–6146. https://doi.org/10.1073/pnas.1323964111.
83. Coyne, M.J., Roelofs, K.G., and Comstock, L.E. (2016). Type VI secretion systems of human gut Bacteroidales segregate into three genetic architectures, two of which are contained on mobile genetic elements. BMC Genomics *17*, 58. https://doi.org/10.1186/s12864-016-2377-z.
84. Simon, R., Priefer, U., and Pühler, A. (1983). A Broad Host Range mobilization system for in vivo genetic engineering: transposon mutagenesis in Gram negative bacteria. Nat. Biotechnol. *1*, 784–791. https://doi.org/10.1038/nbt1183-784.
85. Pal, D., Venkova-Canova, T., Srivastava, P., and Chattoraj, D.K. (2005). Multipartite regulation of rctB, the replication initiator gene of Vibrio cholerae Chromosome II. J. Bacteriol. *187*, 7167–7175. https://doi.org/10.1128/JB.187.21.7167-7175.2005.
86. Smith, C.J., Rogers, M.B., and McKee, M.L. (1992). Heterologous gene expression in Bacteroides fragilis. Plasmid *27*, 141–154. https://doi.org/10.1016/0147-619x(92)90014-2.
87. Park, S.Y., Rao, C., Coyte, K.Z., Kuziel, G.A., Zhang, Y., Huang, W., Franzosa, E.A., Weng, J.K., Huttenhower, C., and Rakoff-Nahoum, S. (2022). Strain-level fitness in the gut microbiome is an emergent property of glycans and a single metabolite. Cell *185*, 513–529.e21. https://doi.org/10.1016/j.cell.2022.01.002.
88. Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C.A., and Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods *6*, 343–345. https://doi.org/10.1038/nmeth.1318.
89. Chen, X., Katchar, K., Goldsmith, J.D., Nanthakumar, N., Cheknis, A., Gerding, D.N., and Kelly, C.P. (2008). A mouse model of Clostridium difficile–associated disease. Gastroenterology *135*, 1984–1992. https://doi.org/10.1053/j.gastro.2008.09.002.
90. Rao, C., Coyte, K.Z., Bainter, W., Geha, R.S., Martin, C.R., and Rakoff-Nahoum, S. (2021). Multi-kingdom ecological drivers of microbiota assembly in preterm infants. Nature *591*, 633–638. https://doi.org/10.1038/s41586-021-03241-8.
91. Katoh, K., Rozewicki, J., and Yamada, K.D. (2019). MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Brief. Bioinform. *20*, 1160–1166. https://doi.org/10.1093/bib/bbx108.
92. Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol. *215*, 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.

93. Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. *30*, 772–780. https://doi.org/10.1093/molbev/mst010.

94. Sela, I., Ashkenazy, H., Katoh, K., and Pupko, T. (2015). GUIDANCE2: accurate detection of unreliable alignment regions accounting for the uncertainty of multiple parameters. Nucleic Acids Res. *43*, W7–W14. https://doi.org/10.1093/nar/gkv318.

95. Darriba, D., Taboada, G.L., Doallo, R., and Posada, D. (2011). ProtTest 3: fast selection of best-fit models of protein evolution. Bioinformatics *27*, 1164–1165. https://doi.org/10.1093/bioinformatics/btr088.

96. Miller, M.A., Pfeiffer, W., and Schwartz, T. (2010). Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In 2010 Gateway Computing Environments Workshop (GCE) https://doi.org/10.1109/gce.2010.5676129.

97. Letunic, I., and Bork, P. (2016). Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res. *44*, W242–W245. https://doi.org/10.1093/nar/gkw290.

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

99. Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. *10*, R25. https://doi.org/10.1186/gb-2009-10-3-r25.

100. Dejesus, M.A., Ambadipudi, C., Baker, R., Sassetti, C., and Ioerger, T.R. (2015). Transit - A software tool for Himar1 TnSeq analysis. PLoS Comp. Biol. *11*, e1004401. https://doi.org/10.1371/journal.pcbi.1004401.

Please cite this article in press as: Kuziel et al., Functional diversification of dietary plant small molecules by the gut microbiome, Cell (2025), https://doi.org/10.1016/j.cell.2025.01.045

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial strains		
B. thetaiotaomicron VPI-5482	ATCC	ATCC 29148
B. thetaiotaomicron 1_1_6	BEI	HM-23
B. finegoldii CL09T03C10	BEI	HM-727
B. finegoldii DSM 17565	DSMZ	DSMZ 17565
B. xylanisolvens 1_1_30	BEI	HM-22
B. xylanisolvens 2_1_22	BEI	HM-18
B. xylanisolvens CL03T12C04	BEI	HM-722
B. ovatus ATCC 8483	ATCC	ATCC 8483
B. ovatus CL03T12C18	BEI	HM-724
B. ovatus D2	BEI	HM-28
B. ovatus 3_8_47FAA	BEI	HM-222
B. caccae ATCC 43185	ATCC	ATCC 43185
B. caccae CL03T12C61	BEI	HM-728
B. nordii CL02T12C05	BEI	HM-721
B. salyersiae CL02T12C01	BEI	HM-725
B. salyersiae DSM 18765	DSMZ	DSMZ 8765
B. fragilis 3_1_12	BEI	HM-20
B. fragilis CL05T00C42	BEI	HM-711
B. fragilis S13 L11	Lab of Laurie Comstock	Coyne et al.[83]
B. fragilis 1007-1-F #7	Lab of Laurie Comstock	Coyne et al.[83]
B. fragilis NCTC 9343	ATCC	ATCC 25285
B. fragilis NCTC 638R	Lab of Laurie Comstock	Coyne et al.[83]
B. fragilis CL03T00C08	BEI	HM-713
B. fragilis DS-166	Lab of Laurie Comstock	Coyne et al.[83]
B. fragilis S36 L11	Lab of Laurie Comstock	Coyne et al.[83]
B. fragilis CL07T12C05	BEI	HM-710
B. fragilis CL07T00C01	BEI	HM-709
B. fragilis 2_1_16	BEI	HM-58
B. cellulosilyticus CL02T12C19	BEI	HM-726
B. oleiciplenus DSM 22535	DSMZ	DSMZ 22535
B. eggerthii sp. 1_2_48FAA	BEI	HM-210
B. eggerthii DSM 20697	DSMZ	DSMZ 20697
B. uniformis CL03T00C23	BEI	HM-715
B. uniformis ATCC 8492	ATCC	ATCC 8492
B. uniformis D20	BEI	HM-189
B. dorei 9_1_42FAA	BEI	HM-27
B. dorei CL02T00C15	BEI	HM-717
B. dorei DSM 17855	DSMZ	DSMZ 17855
B. dorei CL03T12C01	BEI	HM-718
B. dorei 5_1_36/D4	BEI	HM-29
B. vulgatus ATCC 8482	ATCC	ATCC 8482
B. vulgatus CL09T03C04	BEI	HM-720
P. goldsteinii CL02T12C30	BEI	HM-732

(*Continued on next page*)

Please cite this article in press as: Kuziel et al., Functional diversification of dietary plant small molecules by the gut microbiome, Cell (2025), https://doi.org/10.1016/j.cell.2025.01.045

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
P. merdae CL03T12C32	BEI	HM-730
P. merdae ATCC 43184	ATCC	ATCC 43185
P. merdae CL09T00C40	BEI	HM-729
P. johnsonii CL02T12C29	BEI	HM-731
P. distasonis sp. 3_1_19	BEI	HM-19
P. distasonis sp. 31_2	BEI	HM-169
P. distasonis CL03T12C09	BEI	HM-733
P. distasonis CL09T03C24	BEI	HM-734
P. distasonis sp. 20_3	BEI	HM-166
E. coli K12	ATCC	ATCC 29425
S. enterica ATCC 14028	ATCC	ATCC 14028
K. pneumoniae ATCC 35657	ATCC	ATCC 35657
C. portucalensis 4_7_47CFAA	BEI	HM-299
C. tertium 7_2_43FAA	BEI	HM-36
C. sporogenes ATCC 15579	ATCC	ATCC 15579
C. cadaveris CC40_001C	BEI	HM-1039
C. symbiosum WAL-14163	BEI	HM-309
C. scindens ATCC 35704	ATCC	ATCC 35704
C. bolteae WAL-14578	BEI	HM-318
F. umbilicata HPP0048	BEI	HM-794
H. hathewayi WAL-18680	BEI	HM-308
R. gnavus CC55_001C	BEI	HM-1056
P. hiranosis DSM 13275	ATCC	DSM 13275
L. caecimuris 3_1_31	BEI	HM-178
L. rhamnosus GG	ATCC	ATCC 53103
B. massiliensis KLE 1732	BEI	HM-1032
C. scindens 5_1_57FAA	BEI	HM-157
E. massiliensis KLE 1755	BEI	HM-1033
E. clostridioforme 2_1_49FAA	BEI	HM-306
F. fissicatena 3_1_39B/D5	BEI	HM-84
R. gnavus 2_1_58FAA	BEI	HM-154
C. innocuum 6_1_30	BEI	HM-173
C. paraputrificum ATCC 25780	ATCC	ATCC 25780
E. faecalis OG1RF	N/A	N/A
E. faecium 503	N/A	N/A
B. longum subsp. longum 44B	BEI	HM-845
L. monocytogenes F6900	BEI	NR-13227
C. difficile M7404	Lab of D. Lyras	N/A
E. coli RIMD 0509952	BEI	NR-12
S. enterica serovar Typhimurium *LT2*	BEI	NR-174
C. portucalensis 30_2	BEI	HM-34
C. difficile 630 Δerm	Lab of N. Minton	N/A
C. difficile 20100211	BEI	NR-49279
C. difficile 20120016	BEI	NR-49282
C. difficile 20110870	BEI	NR-49288
C. difficile 20120187	BEI	NR-49290
C. difficile 20110978	BEI	NR-49293
C. difficile 20120196	BEI	NR-49296

(*Continued on next page*)

Please cite this article in press as: Kuziel et al., Functional diversification of dietary plant small molecules by the gut microbiome, Cell (2025), https://doi.org/10.1016/j.cell.2025.01.045

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
C. difficile 20120613	BEI	NR-49297
C. difficile 20110818	BEI	NR-49301
C. difficile 20120041	BEI	NR-49303
C. difficile 20120956	BEI	NR-49304
C. difficile 20120190	BEI	NR-49309
C. difficile 20111003	BEI	NR-49315
C. difficile 20110961	BEI	NR-49316
C. difficile 20110995	BEI	NR-49320
C. difficile 20110740	BEI	NR-49323
C. difficile 20120014	BEI	NR-49327
C. difficile 20110958	BEI	NR-49328
C. difficile 20121013	BEI	NR-49329
Bo ATCC *ΔgghR*	This paper	N/A
Bo ATCC *ΔgghA*	This paper	N/A
Bo ATCC *ΔgghB*	This paper	N/A
Bo ATCC *ΔgghC*	This paper	N/A
Bo ATCC *ΔgghD*	This paper	N/A
Bu ATCC *ΔgshA*	This paper	N/A
Bu ATCC *ΔgshB*	This paper	N/A
Bu ATCC *ΔgshC*	This paper	N/A
Bu ATCC *ΔgshD*	This paper	N/A
Bu ATCC *ΔgshE*	This paper	N/A
Bu ATCC *ΔgshF*	This paper	N/A
Bu ATCC *ΔgshG*	This paper	N/A
Bu ATCC *ΔgghC*	This paper	N/A
Bu ATCC *ΔgshDΔgshG*	This paper	N/A
Bu ATCC *ΔgshDΔgghC*	This paper	N/A
Bu ATCC *ΔgshDΔgshGΔgghC*	This paper	N/A
Bo ATCC *ΔgghA*	This paper	N/A
Bo ATCC *ΔgghB*	This paper	N/A
Bo ATCC *ΔgghC*	This paper	N/A
Bo ATCC *ΔgghD*	This paper	N/A
Bo ATCC *ΔgghR* + *pFD340-BoATC_gghR*	This paper	N/A
Bo ATCC *ΔgghA* + *pFD340-BoATC_gghA*	This paper	N/A
Bo ATCC *ΔgghB* + *pFD340-BoATC_gghB*	This paper	N/A
Bo ATCC *ΔgghC* + *pFD340-BoATC_gghC*	This paper	N/A
Bo ATCC *ΔgghD* + *pFD340-BoATC_gghD*	This paper	N/A
Bv ATCC pFD340	This paper	N/A
Bv ATCC pFD340-BuATCC_*gshD*	This paper	N/A
Bv ATCC pFD340-BuATCC_*gshG*	This paper	N/A
E. coli S17-1 lambda pir	Simon et al.[84]	N/A
E. coli DH5α lambda pir	Pal et al.[85]	N/A
E. coli pRK231	Lab of Laurie Comstock	Coyne et al.[83]
E. coli BL21 (DE3)	Novagen	C600003
iBMDM	Lab of J. Kagan	N/A
Chemicals and recombinant proteins		
D-(+)-Glucose	Sigma	Cat# G7021
Arbutin	Sigma	Cat# A4256

(*Continued on next page*)

Please cite this article in press as: Kuziel et al., Functional diversification of dietary plant small molecules by the gut microbiome, Cell (2025), https://doi.org/10.1016/j.cell.2025.01.045

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Hydroquinone	Sigma	Cat# H17902
D-(-)-Salicin	Sigma	Cat# S0625
Saligenin	Sigma	Cat# 166952
Esculin hydrate	Sigma	Cat# E8250
Amygdalin	Sigma	Cat# A6005
Gastrodin	TCI	Cat# G0468
Gastrodigenin	Sigma	Cat# H20806
Helicin	TCI	Cat# H0908
Salicylaldehyde	Sigma	Cat# 84160
Salidroside	ChemImpex	Cat# 33409
Tyrosol	Sigma	Cat# 188255
Rutin	TCI	Cat# R0035
Quercetin	TCI	Cat# P0042
Naringin	TCI	Cat# N0073
Naringenin	TCI	Cat# N0072
Cynaroside	MedChemExpress	Cat# HY-N0540
Luteolin	TCI	Cat# T2682
Phloridzin	TCI	Cat# P0248
Phloretin	TCI	Cat# P1966
Polydatin	TCI	Cat# P1878
Resveratrol	TCI	Cat# R0071
Genistin	Thermo Scientific Chemicals	Cat# AAJ63445MC
Genistein	TCI	Cat# G0272
Daidzin	TCI	Cat# D3920
Daidzein	TCI	Cat# D2668
Pinoresinol diglucoside	Sigma	Cat# SMB00436
Pinoresinol	Sigma	Cat# PHL89525
D-(+)-Cellobiose	Sigma	Cat# C7252
α-Lactose monohydrate	Sigma	Cat# L8783
D-(+)-Maltose monohydrate	Sigma	Cat# M5895
D-(+)-Melibiose	Sigma	Cat# 63630
Palatinose hydrate	Sigma	Cat# P2007
Sucrose	Sigma	Cat# S7903
Gentiobiose	TCI	Cat# G0026
D-(+)-Trehalose dihydrate	Sigma	Cat# T9449
4-Chloro-DL-phenylalanine	ACROS Organics	Cat# A13323
Brain Heart Infusion	BD Bioscience	Cat# 237500
Yeast Extract	Gibco	Cat# 212750
Casitone	Gibco	Cat# 225930
Tyrosol-d^4	TRC Canada	Cat# T947802
C^{13}-Resveratrol	Sigma	Cat# 711004
Platinum SuperFi PCR Master Mix	TFS	Cat#12358010
DreamTaq PCR Master Mix (2x)	TFS	Cat#K1071
IPTG	Sigma	Cat# I6758
Dextran Sodium Sulfate	MP Biomedical	Cat# 216011090
Vancomyin-HCl	Sigma	Cat# V2002
Metronidazole	Sigma	Cat# M1547
Neomycin	Sigma	Cat# N6386
Ampicillin	Sigma	Cat# A8351

(*Continued on next page*)

Please cite this article in press as: Kuziel et al., Functional diversification of dietary plant small molecules by the gut microbiome, Cell (2025), https://doi.org/10.1016/j.cell.2025.01.045

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
LPS (E. coli, O111:B4)	Enzo Life Sciences	ALX-581-012-L001
gshD-His tag	This study	N/A
gshG-His tag	This study	N/A
Critical commercial assays		
Fast SYBR Green Master Mix	Applied Biosystem/ Thermo Fisher	Cat# 4385617
PureLink Genomic DNA mini kit	Thermo Fisher	Cat# K182002
Zymo DNA Clean and Concentrator Kits	Zymo Research	Cat# D4013
Platinum SuperFi PCR Master Mix	Thermo Fisher	Cat# 12358250
HisPur Ni-NTA Resin	Thermo Fisher	Cat# 88221
Halt Protease Inhibitor Cocktail	Thermo Fisher	Cat# PI87786
NuPAGE 4 to 12%, Bis-Tris Mini Protein Gel	Thermo Fisher	Cat# NP0322BOX
Deposited data		
Raw Tn_Seq data	This paper	SRA: PRJNA1228802
Tn-Seq_Bu ATCC/Glucose	This paper	SRA: SAMN47114420
Tn-Seq_Bu ATCC/Arbutin	This paper	SRA: SAMN47114421
Tn-Seq_Bu Δ*gshD*/Glucose	This paper	SRA: SAMN47114422
Tn-Seq_Bu Δ*gshD*/Arbutin	This paper	SRA: SAMN47114423
Tn-Seq_Bo ATCC/Glucose	This paper	SRA: SAMN47114424
Tn-Seq_Bo ATCC/Arbutin	This paper	SRA: SAMN47114425
Oligonucleotides		
Primers used in this study	This study	Table S1
Recombinant DNA		
pRK231	Smith et al.[86]	N/A
pFD340	Smith et al.[86]	N/A
pKNOCK-*bla-erm*Gb	Koropatkin et al.[68]	N/A
pKNOCK-*bla-erm*Gb_*pheS**	Park et al.[87]	N/A
pET-28a	Novagen	Cat# 69864-3
pFD340-*Bu*ATCC_*gshD*	This study	N/A
pFD340-*Bu*ATCC_*gshG*	This study	N/A
pFD340-*Bo*ATCC_*gghR*	This study	N/A
pFD340-*Bo*ATCC_*gghA*	This study	N/A
pFD340-*Bo*ATCC_*gghB*	This study	N/A
pFD340-*Bo*ATCC_*gghC*	This study	N/A
pFD340-*Bo*ATCC_*gghD*	This study	N/A
pKNOCK-*Bu*ATCC_*gshA*	This study	N/A
pKNOCK-*Bu*ATCC_*gshB*	This study	N/A
pKNOCK-*Bu*ATCC_*gshC*	This study	N/A
pKNOCK-*Bu*ATCC_*gshD*	This study	N/A
pKNOCK-*Bu*ATCC_*gshG*	This study	N/A
pKNOCK-*Bu*ATCC_*gghC*	This study	N/A
pKNOCK-*Bo*ATCC_*gghR*	This study	N/A
pKNOCK-*Bo*ATCC_*gghA*	This study	N/A
pKNOCK-*Bo*ATCC_*gghB*	This study	N/A
pKNOCK-*Bo*ATCC_*gghC*	This study	N/A
pKNOCK-*Bo*ATCC_*gghD*	This study	N/A
pET-28a-*Bu*ATCC_*gshD*	This study	N/A
pET-28a-*Bu*ATCC_*gshG*	This study	N/A

(*Continued on next page*)

Please cite this article in press as: Kuziel et al., Functional diversification of dietary plant small molecules by the gut microbiome, Cell (2025), https://doi.org/10.1016/j.cell.2025.01.045

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and Algorithms		
Prism9	GraphPad	https://www.graphpad.com/scientific-software/prism/
Phobius	PMMID_17483518	https://phobius.sbc.su.se/
SignalP 5.0	PMMID_30778233	https://services.healthtech.dtu.dk/service.php?SignalP-5.0/
MEGA	N/A	https://www.megasoftware.net/
Instruments		
Anaerobic Chamber	Coy Laboratory Products	Vinyl Anaerobic Chamber
MicroPlate absorbance reader	BioTek Instruments	Synergy HTX multi-mode reader
MicroPlate Stacker	BioTek Instruments	BioStacker Microplate Stacker
QuantStudio Real-team PCR system	Applied Systems	3
Focused-ultrasonicator	Covaris	M220
MiSeq v3 600 Cycle	Illumina	MS-102-3003

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Microbial strains

Bacterial strains used in this study are listed in the key resources table.

Animal models

Six-week-old female germ-free C57BL/6 mice were used for gnotobiotic mouse experiments. Brigham and Women's Hospital Massachusetts Host-Microbiome Center approved under protocol 2020N000054 by the Brigham and Women's Hospital Institutional Animal Care and Use Committee (IACUC). Mice were housed in Class II Biological Isolator in a temperature-controlled (~21°C) facility on a 12 h light/dark cycle. Mice were fed a standard chow (Laboratory Rodent Diet 5025, LabDiet, St. Louis, MO, USA) unless otherwise indicated.

All in vivo disease models were carried out under protocols approved by the Boston Children's Hospital IACUC (00001465, 00001278). Six-week-old male C57BL/6J *Foxp3*^{YFPCre} mice were used for experimental colitis experiments. Six to eight-week-old male and female C57BL/6J WT mice were used for *C. difficile* infection experiments (bred internally, from mice obtained from Jackson Laboratory). For *C. difficile* infections, for biosafety considerations experiments took place in a biosafety level 2 facility that contains specific pathogens.

METHOD DETAILS

Bacterial culture

All human gut bacterial strains were purchased from the American Type Culture Collection (ATCC), Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ), the Biodefense and Emerging Infections Research Resources Repository (BEI resources), or from collaborators specified above, except for those strains generated in this work. Frozen stock cultures of bacteria were stored at -80°C in cryogenic vials. Individual strains were derived from the same glycerol stocks throughout this study.

Gut bacterial glycoside utilization

All Bacteroidetes, Firmicutes, and Proteobacteria isolates were grown on brain-heart-infusion supplemented with hemin (50 mg/L) and vitamin K1 (0.25 mg/L) (BHIS) agar plates or in YBHIS (brain-heart-infusion powder (37 g/L), yeast (5 g/L), D-(+)-cellobiose (1 g/L), D-(+)-maltose monohydrate (1 g/L), cysteine (0.5 g/L), hemin (50 mg/L), and vitamin K1 (0.25 mg/L). Bifidobacteria isolates were grown on commercial Brucella Blood Agar (Thermo Fisher) or in YBHIS. All isolates were recovered on and grown in pre-reduced media in a Coy anaerobic chamber (Coy Labs). To determine the ability of diverse gut bacteria to utilize glucose (Sigma) or phytochemical glycosides (salicin (Sigma), arbutin (Sigma), salidroside (ChemImpex), gastrodin (TCI), helicin (TCI), esculin (Sigma), amygdalin (Sigma)) as a sole carbon source, each isolate was recovered on a BHIS or BBA agar plate and incubated anaerobically at 37°C for 48 h. A starter culture (1 mL) of YBHIS was inoculated with each isolate and grown for 15 h. This starter culture was then used to inoculate a YBHIS sub-culture (1 mL) at a final dilution of 1:10. This sub-culture was grown for 3 h to an OD600~0.5 before being used to inoculate experimental cultures at a final dilution of 1:50. The experimental media consisted of a modified YCFA media (mYCFA; casitone (2.5 g/L), yeast extract (0.625 g/L), cysteine (0.5 g/L), magnesium sulfate heptahydrate (f.c. 365 μM), calcium chloride dihydrate (f.c. 610 μM), sodium bicarbonate (4 g/L), dipotassium phosphate (0.45 g/L), monopotassium phosphate (0.45 g/L), sodium chloride (0.9 g/L), sodium acetate (2.71 g/L), Hemin (50 mg/L), Vitamin K1 (0.25 mg/L), ferrous sulfate (4 μg/mL), ATCC vitamin

Please cite this article in press as: Kuziel et al., Functional diversification of dietary plant small molecules by the gut microbiome, Cell (2025), https://doi.org/10.1016/j.cell.2025.01.045

mix (1% v/v)) supplemented with glucose (15 mM, Sigma) or glycoside (15 mM). Bacterial growth was monitored by measuring the optical density at 600 nm (OD_{600}) of each culture in a 384 well plate with an Epoch2 spectrophotometer (BioTek Instruments) and MicroPlate stacker (Agilent BioTek) for 24 hours. Growth was calculated as the area under each 48-hour bacterial growth curve (AUC units). Background growth due to catabolism of media components was subtracted from overall growth of each bacterium in glucose of glycoside to normalize growth across bacteria and substrates. Maximal growth is represented by a deep blue color. Growth was considered for when max OD_{600} greater then 0.100 above max OD_{600} in baseline media without carbon source (Figure 1) or greater then 0.100 from starting OD (Figure 2).

Gut *Bacteroides* glycoside utilization

All Bacteroidetes isolates were grown on brain-heart-infusion supplemented with hemin (50 mg/L) and vitamin K1 (0.25 mg/L) (BHIS) agar plates or in basal media (BSG; proteose peptone (20 g/L), yeast (5 g/L), NaCl (5 g/L), glucose (5 g/L), potassium phosphate dibasic (5 g/L), cysteine (0.5 g/L), hemin (50 mg/L), and vitamin K1 (0.25 mg/L). All isolates were recovered on and grown in pre-reduced media in a Coy anaerobic chamber (Coy Labs). To determine the ability of Bacteroidetes strains to utilize glucose (Sigma), phytochemical glycosides (salicin (Sigma), arbutin (Sigma), salidroside (ChemImpex), gastrodin (TCI), helicin (TCI), esculin (Sigma), amygdalin (Sigma)), or disaccharides (D-(+)-cellobiose (Sigma), lactose (Sigma), D-(+)-maltose monohydrate (Sigma), melibiose (Sigma), sucrose (Sigma), trehalose (Sigma), palatinose (TCI), gentiobiose (TCI)) as a sole carbon source, each isolate was recovered on a BHIS agar plate and incubated anaerobically at 37°C for 48 h. A starter culture (1 mL) of BSG was inoculated with each isolate and grown for 15 h. This starter culture was then used to inoculate a BSG sub-culture (1 mL) at a final dilution of 1:10. This sub-culture was grown for 3 h to an OD_{600}∼0.5 before being used to inoculate experimental cultures at a final dilution of 1:50. The experimental media consisted of *Bacteroides* minimal media (BMM; ammonium sulfate (1 g/L), sodium carbonate (1 g/L), potassium phosphate monobasic (0.9 g/L), sodium chloride (0.9 g/L), calcium chloride dihydrate (26.5 mg/L), magnesium chloride hexahydrate (2 mg/L), manganese(II) chloride tetrahydrate (1 mg/L), cobalt(II) chloride hexahydrate (1 mg/L), hemin (50 mg/L), vitamin K1 (0.25 mg/L), ferrous sulfate heptahydrate (4 mg/L), vitamin B_{12} (5 mg/L)) supplemented with glucose (15 mM, Sigma), glycoside (15 mM) or disaccharide (15 mM). Bacterial growth was monitored by measuring the optical density at 600 nm (OD_{600}) of each culture in a 384 well plate with an Epoch2 spectrophotometer (BioTek Instruments) and MicroPlate stacker (Agilent BioTek). Growth was calculated as the area under each 48-hour bacterial growth curve (AUC units). Maximal growth is represented by a deep blue color.

Transposon insertion sequencing

Bacteroides strains (*Bacteroides ovatus* ATCC 8483 and *B. uniformis* ATCC 8492, and *B. uniformis* ATCC 8492 Δ*gshD*) were grown in 20 mL of BSG to the late-lag phase (OD_{600} ∼0.1). The *E. coli* S17-1 lambda pir donor, which harbors the transposon plasmid pWH2-Term6,[87] was grown in 1 mL of LB media supplemented with ampicillin to exponential phase. Both *Bacteroides* and *E. coli* donor cultures were centrifuged (4000 rpm, 5 min, 21C), re-suspended in equal volumes of phosphate buffered saline (PBS), mixed in equal proportions, and spotted onto BHIS agar plates. These plates were incubated aerobically at 37°C for less than 18 hours. The conjugation mix was then selected on BHIS agar plates supplemented with gentamicin (200 μg/mL) and erythromycin (5 μg/mL) anaerobically at 37°C for 48 hours. Over 100,000 transconjugant colonies were scrapped from the plates and pooled in 25% glycerol stocks stored at -80°C. An aliquot of the transposon mutant bank was thawed, recovered overnight in BSG media (starting OD_{600} ∼0.1), back-diluted 10-fold and grown to mid-exponential phase (OD_{600} ∼0.5). This culture was then diluted 50-fold into BMM containing glucose (0.25%, Sigma) or arbutin (0.25%, Sigma). An aliquot of this culture was also saved as an input control. All experimental cultures were grown to late-exponential phase and harvested by spinning down (8000 rpm, 5 min, 21C) and storing at -80C for library preparation and sequencing.

Tn-seq library preparation

To construct TnSeq DNA libraries for sequencing, the collected samples were first subjected to DNA extraction using Invitrogen PureLink Genomic DNA mini kit (Thermo Fisher Scientific, Waltham, MA). An aliquot of 5 ug of DNA was sheared in 100 μL ultrapure water to 300∼500 base pairs using an M220 focused ultrasonicator (Covaris Inc). The sheared DNA was subjected to end repair, A-tailing and ligated with prepared adapters (generated by annealing two oligos: 5'-TACCACGACCA-NH2-3' and 5'-GTGACTGG AGTTCAGACGTGTGCTCTTCCGATCTGGTCGTGGTAT-3') using NEBNext Ultra DNA Library Prep Kit (New England Biolabs). Next, transposon junction fragments containing the insertion sites were enriched by two PCR amplifications using Platinum SuperFi PCR Master Mix (Thermo Fisher Scientific, Waltham, MA). In the first PCR, the fragments were amplified (98°C for 10s, 60°C for 10s, 72°C for 30s with 25 cycles) with universal primers matching the transposon and adaptor sequence: 5'-CCCATTGGG AATAATAACCTTTATACCTG-3' and 5'-GTGACTGGAGTTCAGACGTGTG-3'. In the second PCR, the PCR products were amplified (98°C for 10s, 60°C for 10s, 72°C for 30s with 18 cycles) with Truseq-compatible indexed primers containing specific annealing sequences. The PCR fragments corresponding to 200∼600 bp were selected using AMPure XP beads (Beckman Coulter Inc). These libraries were normalized using NEB- Next Library Quant kit (New England Biolabs) and subjected to HiSeq 2500 High Output sequencing (Illumina Inc, San Diego, CA) with a read depth of 5 million reads per sample.

Generation of *Bacteroides* gene deletion mutants

Clean deletion mutants were created by allelic replacement whereby the following genes were deleted: *Bacteroides uniformis* ATCC 8492 *gshA* (BACUNI_00922), *gshB* (BACUNI_00920), *gshC* (BACUNI_00922), *gshD* (BACUNI_00919), *gshG* (BACUNI_01042), *gghC*

Please cite this article in press as: Kuziel et al., Functional diversification of dietary plant small molecules by the gut microbiome, Cell (2025), https://doi.org/10.1016/j.cell.2025.01.045

(BACUNI_00951) *and B. ovatus* ATCC 8483 *gghR* (Bovatus_02231), *gghA* (Bovatus_02230), *gghB* (Bovatus_02229), *gghC* (Bovatus_02228), *gghD* (Bovatus_02227). We leveraged a modified *pKNOCK-bla-erm*Gb suicide vector, *pKNOCK-bla-erm*Gb_*phe*S*, we previously engineered with a counter-selection system mediated by a variant phenylalanyl-tRNA synthetase (PheSA303G), lethal in the presence of 4-chloro-phenylalanine.[87] To generate deletion constructs, 1kb regions flanking the genomic region-of-interest were amplified by PCR (Platinum SuperFi PCR Master Mix), purified, and cloned into the *pKNOCK-bla-erm*Gb_*phe*S* plasmid in a three-piece ligation reaction.[88]

To mobilize resulting plasmids into *Bacteroides*, a triparental mating strategy was implemented. Overnight cultures of *Bacteroides* grown in BSG were diluted 50-fold into fresh pre-reduced BSG and grown to early exponential phase (OD_{600} ~0.1). The recipient strain was centrifuged at 4,000 rpm for 5 minutes and resuspended in 200 μL of 1x PBS. One-day-old *E. coli* DH5α lambda pir donor strain and *E. coli* DH5α helper strain with pRK231 were scrapped from LB agar plates into 200 μL of 1x PBS. Donor, helper, and recipient strains were combined at a 1:1:1 donor:helper:recipient culture volume ratio at a final volume of 90 μL and spotted on non-selective BHIS agar plate for 18 hours at 37°C under aerobic conditions to allow for conjugation. Mating lawns were scrapped onto BHIS agar plates containing gentamicin (200 μg/mL) and erythromycin (5 μg/mL) to select for transconjugants (merodiploids). Single colonies were isolated by re-streaking colonies onto BHIS agar plates containing erythromycin (5 μg/mL). A single colony was re-streaked onto BMM agar plates supplemented with 4-Chloro-DL-phenylalanine (2 mg/mL, Acros Organics) to select for the loss of the pKNOCK-*bla*-tetQ-pheS* vector. Individual clones were replica-plated onto BHIS agar and BHIS agar supplemented with erythromycin (5 μg/mL) and those clones that did not grow on erythromycin were confirmed for genetic manipulation by PCR and sequencing.

Heterologous expression

Expression of *Bacteroides uniformis* ATCC 8492 genes *gshD* and *gshG* into *B. vulgatus* ATCC 8482, *B. uniformis* ATCC 8492 gene *gshD* into *B. uniformis* ATCC 8492 Δ*gshD*, and *B. ovatus* ATCC 8483 genes *gghR*, *gghA*, *gghB*, *gghC*, and *gghD* into *B. ovatus* ATCC 8483 Δ*gghR*, Δ*gghA*, Δ*gghB*, Δ*gghC*, and Δ*gghD*, respectively, was performed using the extra-chromosomal *Bacteroides* high-copy expression vector pFD340.[86] Target genes plus 50 base pairs upstream of the gene-of-interest containing the native RBS were amplified (Platinum SuperFi PCR Master Mix) using primers listed elsewhere (Table S1), purified, and cloned into pFD340 directly downstream of the vector-borne promoter in a two-piece ligation reaction.[88] A similar triparental conjugation strategy was performed as described above to mobilize the pFD340 constructs into heterologous hosts. Mating lawns were scrapped onto BHIS agar plates supplemented with gentamicin (200 μg/mL) and erythromycin (5 μg/mL) to select for *Bacteroides* containing the constructs. Single colonies were isolated by re-streaking colonies onto BHIS agar plates supplemented with erythromycin (5 μg/mL).

Recombinant protein purification

Bacteroides uniformis ATCC 8492 genes *gshD* and *gshG* were amplified with primers (Table S1) and genomic DNA from *B. uniformis* ATCC 8492 as template. Each gene was cloned into pET-28a (Novagen) by Gibson Assembly.[88] The recombinant plasmids were transformed into *E. coli* BL21 (DE3) (Novagen) strain, grown to OD_{600}~0.5 in LB plus kanamycin at 37°C and induced with IPTG (f.c. 0.1 mM) at 18°C overnight. After harvesting (10,000 g, 10 min, 4°C), cells were suspended in lysis buffer (20 mM K_2HPO_4, 20 mM KH_2PO_4, 500 mM NaCl, 50 mM Imidazole) with 1x Halt Protease Inhibitor Cocktail (Thermo Fisher Scientific) and lysed by sonication. The cell extracts were spun down, filtered through a 0.45 μm filter, mixed with 2 mL of pre-washed Ni-NTA resin (HisPur Ni-NTA Resin, Thermo Fisher Scientific), and incubated on a rotator at 4°C overnight. The resin was loaded onto a column and washed with lysis buffer containing Imidazole (50 mM). The recombinant proteins were eluted using elution buffer (20 mM K_2HPO_4, 20 mM KH_2PO_4, 500 mM NaCl, 500 mM Imidazole) for both gshD and gshG. After concentration in a spin concentrator pre-dialysis, fractions of each protein were dialyzed overnight at 4°C in 1 L of storage buffer (pH 8.0, 25 mM HEPES, 50 mM NaCl) three times. The size and purity of the recombinant proteins were verified by SDS-PAGE gel (NuPAGE 4 to 12%, Bis-Tris Mini Protein Gel, Thermo Fisher Scientific) and the concentration of each protein was measured by absorbance at 280 nm using Nanodrop (Thermo Fisher Scientific).

In vitro aglycone liberation

All *B. uniformis* strains were grown on brain-heart-infusion supplemented with hemin (50 mg/L) and vitamin K1 (0.25 mg/L) (BHIS) agar plates or in basal media (BSG; proteose peptone (20 g/L), yeast (5 g/L), NaCl (5 g/L), glucose (5 g/L), potassium phosphate dibasic (5 g/L), cysteine (0.5 g/L), hemin (50 mg/L), and vitamin K1 (0.25 mg/L). All isolates were recovered on and grown in pre-reduced media in a Coy anaerobic chamber (Coy Labs).

To assess the capacity of *Bacteroides* to liberate aglycones from their parent aryl, coumarin, and cyanogenic glycosides, wildtype *B. uniformis* ATCC 8492 was recovered on a BHIS agar plate and incubated anaerobically at 37°C for 48 h. A starter culture (1 mL) of BSG was inoculated with each isolate and grown for 15 h. This starter culture was then used to inoculate a BSG sub-culture (1 mL) at a final dilution of 1:10. This sub-culture was grown for 3 h to an OD600~0.5 before being used to inoculate experimental cultures at a final dilution of 1:50. The experimental media consisted of *Bacteroides* minimal media (BMM; ammonium sulfate (1 g/L), sodium carbonate (1 g/L), potassium phosphate monobasic (0.9 g/L), sodium chloride (0.9 g/L), calcium chloride dihydrate (26.5 mg/L), magnesium chloride hexahydrate (2 mg/L), manganese(II) chloride tetrahydrate (1 mg/L), cobalt(II) chloride hexahydrate (1 mg/L), hemin (50 mg/L), vitamin K1 (0.25 mg/L), ferrous sulfate heptahydrate (4 mg/L), vitamin B_{12} (5 mg/L)) supplemented with glucose (0.1%,

Please cite this article in press as: Kuziel et al., Functional diversification of dietary plant small molecules by the gut microbiome, Cell (2025), https://doi.org/10.1016/j.cell.2025.01.045

Sigma) and any one G1G (500 μM, Sigma) or glucose (0.1%, Sigma) and DMSO (0.5%, Sigma). At late log, each experimental culture was isolated, diluted (1:50) into 80% methanol-water spiked with tyrosol-D4 (40 nM, TRC Canada), vortexed (2 min), centrifuged (5000 rpm, 5 min), and the supernatant was recovered for down-stream LCMS analyses.

To assess the capacity of *Bacteroides* to liberate aglycones from their parent polyphenolic glycosides, *B. uniformis* ATCC 8492 WT, Δ*gshD*, and Δ*gshD*Δ*gshG*Δ*gghC* were recovered on a BHIS agar plate and incubated anaerobically at 37°C for 48 h. A starter culture (1 mL) of BSG was inoculated with each isolate and grown for 15 h. This starter culture was then used to inoculate a BSG sub-culture (1 mL) at a final dilution of 1:10. This sub-culture was grown for 3 h to an OD600∼0.5 before being used to inoculate experimental cultures at a final dilution of 1:50. The experimental media consisted of *Bacteroides* minimal media (BMM; ammonium sulfate (1 g/L), sodium carbonate (1 g/L), potassium phosphate monobasic (0.9 g/L), sodium chloride (0.9 g/L), calcium chloride dihydrate (26.5 mg/L), magnesium chloride hexahydrate (2 mg/L), manganese(II) chloride tetrahydrate (1 mg/L), cobalt(II) chloride hexahydrate (1 mg/L), hemin (50 mg/L), vitamin K1 (0.25 mg/L), ferrous sulfate heptahydrate (4 mg/L), vitamin B_{12} (5 mg/L)) supplemented with glucose (0.1%, Sigma) and any one G3G (500 μM, Sigma) or glucose (0.1%, Sigma) and DMSO (0.5%, Sigma). At late log, each experimental culture was isolated, diluted (1:50) into 80% methanol-water spiked with C^{13}-Resveratrol (250 nM, Sigma), vortexed (2 min), centrifuged (5000 rpm, 5 min), and the supernatant was recovered for down-stream LCMS analyses.

Gnotobiotic mouse experiments

Six-week-old female GF C57BL/6J mice were colonized with 10^9 CFUs of *Bacteroides uniformis* ATCC 8492 or *Bacteroides uniformis* Δ*gshD*Δ*gshG*Δ*gghC*. At two weeks post-colonization, mice were treated intra-gastrically (i.g.) gavaged with salicin (100 mg/kg) dissolved in sterile 1x PBS. The stool of these mice was collected immediately prior to gavage (T0) and then three (T3) and five (T5) hours post gavage. Stool was immediately stored at -80°C for confirmation of *Bacteroides* colonization and targeted LCMS detection of *Bacteroides*-liberated saligenin.

To extract metabolites from fecal samples, 200 μL of 0.1 mm zirconia/silica beads (BioSpec Products) and 1 mL of organic solvent (methanol:water, 5:1) supplemented with tyrosol-D4 (40 nM, TRC Canada) were added to 5-50 mg of pre-weighed fecal matter. Material was homogenized by vortexing at maximum speed for 30 seconds before mechanical disruption with a bead beater (BioSpec Products) for 5 minutes on high setting at room temperature. Samples were incubated in an ultrasonicator for 5 minutes before incubation at -20°C overnight. Samples were subsequently thawed at room temperature before centrifugation (15,000 rpm, RT) for 5 minutes. 800 μL of supernatant was isolated from each sample for analysis by LCMS.

Glycoside and aglycone detection by LCMS

For aryl glycoside and aglycone-containing samples, samples were dried under nitrogen flow and resuspended in 50 μL of methanol 50% in water. A standard curve was prepared using the same extraction solution and volumes as for the samples. The curve was prepared as a 10 points 1/5 dilution series with 1 mM as the highest concentration. Samples were quantified on a QEplus mass spectrometer coupled to an Ultimate 3000 LC (Thermo fisher). Five microliters were injected on a Luna Omega Polar C18 column (2mm x 150 mm, Phenomenex) maintained at 40°C. The mobile phases were A: water and B: Methanol. The gradient was as follows: 1% B for 4 min, then to 100% B in 6 min. the mobile phases were then maintained at 100% B for 15 min, followed by 5 min re-equilibration at 1% B. The flow rate was 0.15 mL min^{-1}. Ionization for the mass spectrometer was achieved with atmospheric pressure chemical ionization (APCI) with a probe at 350 °C, in switching polarity mode. Quantification was performed using Tracefinder (Thermo fisher), using the ratio of the area under the peak for each compound and of the internal standard, using the accurate mass of the $[M-H]^-$ ions mainly, except for hydroquinone and benzoquinone where the $[M+e]^-$ and the $[M+H]^+$ were used respectively.

For polyphenolic glycoside and aglycone-containing samples, samples were dried under nitrogen flow and resuspended in 50ul of methanol 50% in water. A standard curve was prepared using the same extraction solution as for the samples. The curve was prepared as a 10 points 1/5 dilution series with 100 μM as the highest concentration. Samples were quantified on a QEplus mass spectrometer coupled to an Ultimate 3000 LC (Thermo fisher). Five microliters were injected on a C18Evo column (2mm x 150 mm, Phenomenex) maintained at 30°C. The mobile phases were A: water, 5mM ammonium formate, with pH adjusted to 7 and B: Acetonitrile, 0.1% ammonium hydroxide. The gradient was as follow: 0% B for 5 min, then to 40% B in 9 min, then to 50% B in 6min, and finally to 100% B in 1 min. The mobile phases were then maintained at 100% B for 7 min, followed by 5 min re-equilibration at 0% B. The flow rate was 0.15 mL min^{-1}. Ionization was achieved by heated electrospray ionization (HESI) in negative mode. Quantification was performed using Tracefinder (Thermo fisher), using the ratio of the area under the peak for each compound and of the internal standard, using the accurate mass of the $[M-H]^-$ ions.

Michaelis-Menten enzyme kinetics

Recombinant gshD (10 nM) was combined with salidroside or salicin (substrate concentration range: 0, 1, 10, 20, 40, 60, 80, 100, 250, 500, 750 μM, 1000 μM) in reaction buffer (pH 6.5, 25 mM HEPES, 25 mM NaCl). Reactions were set to incubate at 37°C for 10 minutes before being quenched with sodium carbonate (200 mM, f.c. 100 mM). Each reaction was further diluted in methanol (100%, f.c. 33%) in preparation for down-stream LCMS analyses. Each reaction was performed in triplicate, twice on separate days.

Recombinant gshG (500 nM) was combined with salidroside or salicin (substrate concentration range: 100, 250, 500, 750, 1000, 1250, 1500, 1750, 2000, 2250, 2500, 2750, 3000, 3250, 3500, 3750, 4000 μM) in reaction buffer (pH 6.5, 25 mM HEPES, 25 mM NaCl). Reactions were set to incubate at 37°C for 18 hours before being quenched with sodium carbonate (200 mM, f.c. 100 mM). Each

Please cite this article in press as: Kuziel et al., Functional diversification of dietary plant small molecules by the gut microbiome, Cell (2025), https://doi.org/10.1016/j.cell.2025.01.045

reaction was further diluted in methanol (100%, f.c. 33%) in preparation for down-stream LC-MS analyses. Each reaction was performed in triplicate, twice on separate days.

Relative enzyme activity

Recombinant gshD (10 nM) or gshG (500 nM) were combined with salidroside, tyrosol, salicin, saligenin, cellobiose, lactose, maltose, melibiose, sucrose, trehalose, palatinose, or gentiobiose (substrate concentration of 900 μM) in reaction buffer (pH 6.5, 5 mM HEPES, 5 mM NaCl). Reactions were set to incubate at 37°C for 1 hour (gshD) or 18 hours (gshG) before being quenched with sodium carbonate (200 mM, f.c. 100 mM). Each reaction was further diluted in methanol (100%, f.c. 33%) in preparation for down-stream LCMS analyses. Each reaction was performed in triplicate, twice on separate days.

Recombinant enzyme activity LCMS

Data was collected with a Q-Exactive Orbitrap mass spectrometer (Thermo Fisher Scientific, San Jose, CA) for MS scans in both positive and negative polarity modes and additionally for targeted SIM mode. The samples were run through a SeQuant® ZIC®-pHILIC 150 x 2.1 mm analytical column equipped with a 2.1 x 20 mm guard column (both 5 mm particle size; EMD Millipore). Buffer A was 20 mM ammonium carbonate, 0.1% ammonium hydroxide; Buffer B was acetonitrile. The chromatographic gradient was run at a flow rate of 0.150 mL/min as follows: 0-20 min: linear gradient from 80-20% B; 20-20.5 min: linear gradient form 20-80% B; 20.5-28 min: hold at 80% B.

MS data acquisition was performed in a range of m/z = 70–1000, with the resolution set at 70,000, the AGC target at $1x10^6$, and the maximum injection time (Max IT) at 20 msec. For detection of salidroside, tyrosol, salicin, saligenin, cellobiose, lactose, maltose, melibiose, sucrose, trehalose, palatinose, or gentiobiose, targeted selected ion monitoring (tSIM) scans in negative mode were included. The isolation window was set at 1.0 m/z and tSIM scans were centered at m/z = 299.1136 (salidroside, RT = 3.88 min), m/z = 137.0608 (tyrosol, RT = 2.67 min), m/z = 285.0980 (salicin, RT = 4.08 min), m/z = 123.0452 (saligenin, RT = 2.88 min), m/z = 341.1089 (palatinose, RT =10.04 min; gentiobiose, RT = 11.10 min; sucrose, RT = 9.99 min; melibiose, RT = 11.14 min; maltose, RT = 10.65 min; lactose, RT = 10.85 min; cellobiose, RT = 10.60 min; trehalose, RT = 10.67 min) and m/z = 179.0061 (glucose, RT = 9.63 min; fructose, RT = 8.59 min; galactose, RT = 9.78 min). For all tSIM scans, the resolution was set at 70,000, the AGC target was $1x10^5$, and the max IT was 200 msec. The raw data was processed using Xcalibur (Thermo Fisher Scientific) and the peaks' area and height were calculated by setting the peak algorithm detection as Genesis for all the samples.

Phytochemical microbial killing assays

All gut pathogen isolates were grown on brain-heart-infusion supplemented with hemin (50 mg/L) and vitamin K1 (0.25 mg/L) (BHIS) agar plates or in basal media (BSG; proteose peptone (20 g/L), yeast (5 g/L), NaCl (5 g/L), glucose (5 g/L), potassium phosphate dibasic (5 g/L), cysteine (0.5 g/L), hemin (50 mg/L), and vitamin K1 (0.25 mg/L). Isolates were recovered on and grown in pre-reduced media in a Coy anaerobic chamber (Coy Labs). To determine the fitness effects of salicin (Sigma), saligenin (Sigma), arbutin (Sigma), hydroquinone (Sigma), salidroside (ChemImpex), tyrosol (Sigma), gastrodin (TCI), gastrodigenin (Sigma), helicin (TCI), salicylaldehyde (Sigma), rutin (TCI), quercetin (TCI), naringin (TCI), naringenin (TCI), cynaroside (MedChemExpress), luteolin (TCI), phloridzin hydrate (TCI), phloretin (TCI), polydatin (TCI), resveratrol (TCI), genistin (Thermo Scientific Chemicals), genistein (TCI), daidzin (TCI), daidzein (TCI), pinoresinol diglucoside (Sigma) and pinoresinol (Sigma) on enteric pathogen fitness, each pathogen isolate was recovered on a BHIS agar plate and incubated anaerobically at 37°C for 48 h. A starter culture (1 mL) of BSG was inoculated with *C. difficile* and grown for 15 h. This starter culture was then used to inoculate a BSG sub-culture (1 mL) at a final dilution of 1:10. This sub-culture was grown for 3 h to an OD600∼0.5 before being used to inoculate experimental cultures at a final dilution of 1:50. The experimental media consisted of a modified YCFA media (mYCFA; casitone (2.5 g/L), yeast extract (0.625 g/L), cysteine (0.5 g/L), magnesium sulfate heptahydrate (f.c. 365 μM), calcium chloride dihydrate (f.c. 610 μM), sodium bicarbonate (4 g/L), dipotassium phosphate (0.45 g/L), monopotassium phosphate (0.45 g/L), sodium chloride (0.9 g/L), sodium acetate (2.71 g/L), Hemin (50 mg/L), Vitamin K1 (0.25 mg/L), ferrous sulfate (4 μg/mL), ATCC vitamin mix (1% v/v)) supplemented with glucose (15 mM, Sigma) plus each compound (150 μM) or glucose (15 mM, Sigma) plus DMSO (0.15%, Sigma). Bacterial growth was monitored by measuring the optical density at 600 nm (OD_{600}) of each culture in a 384 well plate with an Epoch2 spectrophotometer (BioTek Instruments) and MicroPlate stacker (Agilent BioTek) for 24 hours.

In vitro polydatin bioactivation and *C. difficile* antagonism

B. uniformis strains were grown on brain-heart-infusion supplemented with hemin (50 mg/L) and vitamin K1 (0.25 mg/L) (BHIS) agar plates or in basal media (BSG; proteose peptone (20 g/L), yeast (5 g/L), NaCl (5 g/L), glucose (5 g/L), potassium phosphate dibasic (5 g/L), cysteine (0.5 g/L), hemin (50 mg/L), and vitamin K1 (0.25 mg/L). Isolates were recovered on and grown in pre-reduced media in a Coy anaerobic chamber (Coy Labs). To assess the effects of resveratrol liberation from polydatin on the fitness of *C. difficile* M7404, wildtype *B. uniformis* ATCC 8492 and *B. uniformis* ATCC 8492 Δ*gshD* were recovered on BHIS agar plates and incubated anaerobically at 37°C for 48 h. A starter culture (1 mL) of BSG was inoculated with each isolate and grown for 15 h. This starter culture was then used to inoculate a BSG sub-culture (1 mL) at a final dilution of 1:10. This sub-culture was grown for 3 h to an OD600∼0.5 before being used to inoculate experimental cultures at a final dilution of 1:50. The experimental media consisted of *Bacteroides* minimal media (BMM; ammonium sulfate (1 g/L), sodium carbonate (1 g/L), potassium phosphate monobasic (0.9 g/L), sodium chloride

Please cite this article in press as: Kuziel et al., Functional diversification of dietary plant small molecules by the gut microbiome, Cell (2025), https://doi.org/10.1016/j.cell.2025.01.045

(0.9 g/L), calcium chloride dihydrate (26.5 mg/L), magnesium chloride hexahydrate (2 mg/L), manganese(II) chloride tetrahydrate (1 mg/L), cobalt(II) chloride hexahydrate (1 mg/L), hemin (50 mg/L), vitamin K1 (0.25 mg/L), ferrous sulfate heptahydrate (4 mg/L), vitamin B_{12} (5 mg/L)) supplemented with glucose (0.1%, Sigma) and polydatin (500 μM, Sigma) or glucose (0.1%, Sigma) and DMSO (0.5%, Sigma). After 8 hours, conditioned media was prepared from experimental cultures by centrifugation (5000 rpm, 5 min), and the supernatant was recovered and filter-sterilized (0.22 μm).

To assess the capacity of *B. uniformis* metabolism of polydatin to antagonize the fitness of *C. difficile*, *C. difficile* M7404 was recovered on a pre-reduced BHIS agar plate and incubated anaerobically at 37°C for 48 h. A starter culture (1 mL) of BSG was inoculated with *C. difficile* and grown for 15 h. This starter culture was then used to inoculate a BSG sub-culture (1 mL) at a final dilution of 1:10. This sub-culture was grown for 3 h to an OD600∼0.5 before being used to inoculate experimental cultures at a final dilution of 1:50. The experimental media consisted of a modified YCFA media (mYCFA; casitone (2.5 g/L), yeast extract (0.625 g/L), cysteine (0.5 g/L), magnesium sulfate heptahydrate (f.c. 365 μM), calcium chloride dihydrate (f.c. 610 μM), sodium bicarbonate (4 g/L), dipotassium phosphate (0.45 g/L), monopotassium phosphate (0.45 g/L), sodium chloride (0.9 g/L), sodium acetate (2.71 g/L), Hemin (50 mg/L), Vitamin K1 (0.25 mg/L), ferrous sulfate (4 μg/mL), and ATCC vitamin mix (1% v/v)) supplemented with glucose (15 mM, Sigma) plus ½-diluted *B. uniformis* WT or Δ*gshD* conditioned media from growth with or without polydatin. *C. difficile* growth was monitored by measuring the optical density at 600 nm (OD_{600}) of each culture in a 384 well plate with an Epoch2 spectrophotometer (BioTek Instruments) for 12 hours.

C. difficile infection (CDI) model

We utilized an antibiotic-treatment-induced CDI mouse model developed previously.[89] Briefly, six- to eight-week-old male and female SPF C57BL/6J WT mice were treated for three days with a mixture of antibiotics including vancomycin (0.4 mg/mL, Sigma) colistin (850 U/mL, Sigma), metronidazole (0.215 mg/mL, Sigma), gentamicin (0.035 mg/mL, RPI), and kanamycin (0.045 mg/mL, Sigma) in their drinking water. Mice were then treated with regular water for two days and then injected intraperitoneal (i.p.) with clindamycin (10 mg/kg, Mylan Pharmaceuticals) in normal saline. After one day, mice were gavaged with 10^4 spores of *C. difficile* 630 diluted in PBS. Resveratrol, phloretin, or vehicle (0.1% Tween) were provided in the drinking water at 500 μM the day of infection (day 0) and then each subsequent day for two days (day 1 and day 2). Stool was collected from mice at one day and two days post infection, and immediately taken for assessment of *C. difficile* infection load.

C. difficile colony-forming units (CFUs)

Mouse feces were collected from mice on day one and day two of *C. difficile* infection. Feces were weighed and resuspended to 50 mg/mL in pre-reduced 1x PBS. CFUs were enumerated by plating multiple dilutions on pre-reduced commercial ChromID *C. difficile*-selective plates (Biomerieux) after incubation for 24 hours at 37°C in an anaerobic chamber (Coy Labs).

In vitro macrophage inflammation screen

Immortalized mouse bone marrow-derived macrophages (iBMDM) were plated in a 96-well plate (100,000 cells/well) and individually incubated with all aryl glycosides and their respective aglycones at 5 mM, 2.5 mM, 1.25 mM, and 0.625 mM: salicin (Sigma), saligenin (Sigma), arbutin (Sigma), hydroquinone (Sigma), salidroside (ChemImpex), tyrosol (Sigma), gastrodin (TCI), gastrodigenin (Sigma), helicin (TCI), and salicylaldehyde (Sigma), as well as all polyphenolic glycosides and their respective aglycones at 75 μM: rutin (TCI), quercetin (TCI), naringin (TCI), naringenin (TCI), cynaroside (MedChemExpress), luteolin (TCI), phloridzin hydrate (TCI), phloretin (TCI), polydatin (TCI), resveratrol (TCI), genistin (Thermo Scientific Chemicals), genistein (TCI), daidzin (TCI), daidzein (TCI), pinoresinol diglucoside (Sigma) and pinoresinol (Sigma). Cells were incubated for 16 hours at 37°C and 5% CO_2 and then stimulated with LPS (1 μg/mL, E. coli, Serotype O111:B4, Enzo Life Sciences) or vehicle for 24 hours. Cells were spun down for five minutes at 400 rpm and then the supernatant was recovered and stored at -20°C for future analyses. For in vitro cytokine measurements, TNF-α (Thermo Fisher, 88-7324-88) and IL-6 (Thermo Fisher, 88-7064-88) were measured by ELISA in the supernatant of LPS- or vehicle-stimulated, glycoside- or aglycone-treated iBMDM. Cellular viability was measured at the end-point of each experiment involving salicin, saligenin, arbutin, and hydroquinone via an MTT assay (ATCC, 30-1010K).

Mouse model of intestinal inflammation

We utilized an experimental model of colitis leveraging dextran sodium sulfate (DSS, MP Biomedicals). Briefly, six-week-old male SPF C57BL/6J *Foxp3*^{YFPCre} mice were pre-treated IG via gavage with a mixture of antibiotics including vancomycin (500 mg/L, Sigma), metronidazole (1 g/L, Sigma), ampicillin (1 g/L, Sigma) and neomycin (1 g/L, Sigma) once a day for five days. Mice were then treated with regular water for three days to wash out residual antibiotics. After 24 hours mice were colonized with 10^9 CFUs of wildtype *Bacteroides uniformis* ATCC 8492 or *Bacteroides uniformis* Δ*gshD*Δ*gshG*Δ*gghC* via i.g. administration. After 24 hours, mice were massed and IG administered salicin (100 mg/kg, Sigma) dissolved in sterile 1x PBS, arbutin (100 mg/kg, Sigma) dissolved in sterile 1x PBS, or vehicle and concomitantly treated with 2.5% DSS (MP Biomedicals) in their drinking water. For six subsequent days mice were massed and treated with salicin and DSS, arbutin and DSS, or vehicle and DSS. On the eighth day post-colonization, stool was collected and stored at -80°C. Mice were then sacked, massed, the length of each colon from caecum to rectum was measured, and 1 cm descending colon tissue was isolated into neutral buffered formalin (10%) for histological analyses. After 24 hours, tissue was stored in 70% ethanol for long-term storage.

Please cite this article in press as: Kuziel et al., Functional diversification of dietary plant small molecules by the gut microbiome, Cell (2025), https://doi.org/10.1016/j.cell.2025.01.045

For aglycone protection determination, six-week-old male SPF C57BL/6J *Foxp3*^{YFPCre} mice were pre-treated i.g. with a mixture of antibiotics including vancomycin (500 mg/L, Sigma), metronidazole (1 g/L, Sigma), ampicillin (1 g/L, Sigma) and neomycin (1 g/L, Sigma) once a day for five days. Mice were then treated with regular water for three days to wash out residual antibiotics. After 24 hours, mice were IG administered saligenin (100 mg/kg, Sigma) dissolved in sterile 1x PBS or vehicle and concomitantly treated with 2.5% DSS (MP Biomedicals) in their drinking water. For six subsequent days mice were treated with saligenin and DSS or vehicle and DSS. On the eighth day post-colonization, stool was collected and stored at -80°C. At sacrifice, the length of each colon from caecum to rectum was measured, and 1 cm descending colon tissue was isolated into neutral buffered formalin (10%) for histological analyses. After 24 hours, tissue was transferred to 70% ethanol for long-term storage.

For histological analyses, tissue was delivered to the Beth Israel Deaconess Medical Center (BIDMC) core facility where it was embedded. Briefly, blocks were cut at 6 μm and subjected to Hematoxylin and eosin (H&E) staining. Slides were scored blind, and scores are representative of the entire sample (0 – leukocyte infiltration is normal, 1 – mild localized mild leukocyte infiltrate, 2 – moderate generalized mild leukocyte infiltrate with areas of heavy leukocyte infiltrate, 3 – severe generalized moderate leukocyte infiltrate with areas of high leukocyte infiltrate). Representative images were taken later.

In vivo *Bacteroides* colonization

To confirm *Bacteroides uniformis* colonization in vivo we performed a modification of MK-SpikeSeq which we previously developed.[90] Stool from experimental colitis model mice (Figure 5) treated with (1) vehicle, (2) 2.5% DSS, (3) 2.5% DSS plus wildtype *B. uniformis* ATCC 8492, (4) 2.5% DSS plus *B. uniformis* WT plus salicin, or (5) 2.5% DSS plus *B. uniformis* Δ*gshD*Δ*gshG*Δ*gghC* plus salicin spiked with 0.01 OD of *Salinibacter ruber* DSM 13855 per 10 mg of mouse stool such that the final amount of *S. ruber* per sample would be approximately 1% of the total bacterial 16S abundance. Next, total gDNA was isolated from each sample (ZymoBiomics DNA MiniPrep Kit). Relative quantification of *B. uniformis* using *B. uniformis*-specific primers was performed by real-time PCR (QuantStudio 3) using Fast SYBR Green Master Mix (Applied Biosystems). The abundance of *Bacteroides* colonization was calculated versus PBS-treated mice using *S. ruber* as a standard wherein the resulting -ΔΔCt value was a measure of bacterial colonization in *B. uniformis*-treated versus non-*B. uniformis*-treated mice.

Fecal sample collection and microbiome NGS sequencing

Fecal samples were collected and stored at -80°C for long-term storage. Raw genomic DNA for downstream 16S amplicon next generational sequencing was isolated using the ZymoBIOMICS™ – 96 DNA Kit (Zymo Research). The 16S amplicon library was prepared using dual-index barcodes in a 96-well format and cleaned with the DNA Clean and Concentrator ™ – 5 (Zymo Research) before quantification by qPCR (NEBNext Library Quant Kit). 20 pM of cDNA were loaded on an Illumina MiSeq and sequenced (v3, 600-cycle, 300nt paired-end). To generate the OTU table for downstream analyses of murine gut microbiome composition and diversity, the obtained Illumina raw reads were de-multiplexed, paired end joined, adapter trimmed, quality filtered, dereplicated, and denoised. Sequences were mapped against the publicly available 16S rRNA DNA databases SILVA and UNITE and clustered into OTUs at greater than or equal to 97% nucleotide sequence identity. OTU-based microbial community diversity was estimated by calculating the alpha diversity (Simpson Index).

Cultivation and measurement of aglycones in mouse fecal batch culture

Freshly collected adult SPF B6 mouse fecal pellets transported on ice were taken into a Coy anaerobic chamber (Coy Labs). Fecal pellets (0.5 grams) were suspended in pre-reduced YBHIS (15 mL) with 20% glycerol in a 50 ml sterile conical. The suspension was mixed and filtered with a 70 μm cell strainer (Corning). One milliliter of the filtered suspension was stored in sterile cryogenic vials at −80 °C. A mouse fecal glycerol stock was used to inoculate, with a final dilution of 1:1000, pre-reduced modified YCFA media (mYCFA; casitone (2.5 g/L), yeast extract (0.625 g/L), cysteine (0.5 g/L), magnesium sulfate heptahydrate (f.c. 365 μM), calcium chloride dihydrate (f.c. 610 μM), sodium bicarbonate (4 g/L), dipotassium phosphate (0.45 g/L), monopotassium phosphate (0.45 g/L), sodium chloride (0.9 g/L), sodium acetate (2.71 g/L), Hemin (50 mg/L), Vitamin K1 (0.25 mg/L), ferrous sulfate (4 μg/mL), ATCC vitamin mix (1% v/v)), glucose (0.1 g/L, Sigma), maltose (0.1 g/L, Sigma), cellobiose (0.1 g/L, Sigma), GlcNAc (0.1 g/L, Sigma) and arginine (1 g/L)) supplemented with resveratrol (100 μM, TCI) or saligenin (100 μM, Sigma). Two hundred microliters of the cultures with the corresponding aglycone were grown in a 96-well plate (Corning) for 24 hours at 37 °C in the anaerobic chamber. Cultures at times 0 (start) and 24 (end, stationary phase) hours were stored at -20 °C.

Samples were thawed to room temperature. Each experimental culture was isolated, diluted (1:20) into 80% methanol-water spiked with resveratrol-^{13}C (250 nM, Cambridge Isotopes), vortexed (2 min), centrifuged (5000 rpm, 5 min), and the supernatant was used for down-stream LCMS analyses. A standard curve was prepared using the same extraction solution as for the samples. The curve was prepared as a 10-point 1/5 dilution series with 100 μM as the highest concentration. Dihydroresveratrol standard (Sigma) was used to determine metabolism of resveratrol in mouse fecal cultivation. Samples were quantified on an Agilent 6530 Q-TOF mass spectrometer coupled to an Agilent 1290 Infinity II LC System. One microliter was injected into a ZORBAX Eclipse C18 column (2.1 mm x 50 mm, 1.8-Micron) maintained at 40 °C. The mobile phases were A: water, 5mM ammonium acetate, and B: methanol, 5mM ammonium acetate. The gradient was as follows: 13% B for 1 min, then to 100% B in 4 min, the mobile phases were then maintained at 100% B for 3 min, followed by 1 min re-equilibration at 13% B. The flow rate was 0.3 mL min^{-1}. Ionization was

Please cite this article in press as: Kuziel et al., Functional diversification of dietary plant small molecules by the gut microbiome, Cell (2025), https://doi.org/10.1016/j.cell.2025.01.045

achieved by heated electrospray ionization (HESI) in negative mode. Quantification was performed using Q-TOF Quantitative Analysis (Agilent), using the area under the peak for each compound, using the accurate mass of the $[M-H]^-$ ions.

QUANTIFICATION AND STATISTICAL ANALYSIS

Gut bacterial phylogenetic tree

16S rRNA gene sequences within genomes were located using BLAST, and a multiple gene alignment was carried out using MAFFT version 7. A 16S rRNA tree was reconstructed using the neighbor-joining method implemented in MAFFT version 7, with Jukes-Cantor as the substitution model and a bootstrap value of 1000 replications.[91]

Strains without a complete taxonomic classification were classified at the genus level by comparing the full 16S rRNA genes against curated 16S rRNA databases such as SILVA, and at the species level by comparing the nucleotide identity of several housekeeping genes against several species on the JGI integrated microbial genomes and microbiomes portal.

Bacteroides phylogenetic tree

The phylogenic tree of Bacteroidales strains was generated (Figure 2A), based on the core genes (526 for Bacteroides and 1114 for Parabacteroides) which have 80% identity in amino acid sequence levels, using the concatenated alignment for a neighbor-joining tree in MEGA (https://www.megasoftware.net/).

GH3 Phylogenetic Tree

Proteins belonging to the GH3 family were identified by BLAST alignment tools[92] using the gshD protein sequence on 126 genomes from diverse gut bacteria (Table S3). In addition, 41 previously characterized bacterial GH3 were identified and included in the analysis (http://www.cazy.org/GH3_characterized.html). Protein sequence alignments of GH03 were performed with MAFFT, version 7,[93] and were manually adjusted using as a guide the residue-wise confidence scores generated by GUIDANCE2.[94] ProtTest 3 was used to select the best-fit model of amino acid replacement.[95] Phylogenetic relationships were inferred by maximum likelihood using RAxML-HPC2 on XSEDE conducted on the CIPRES project[96] cluster at the San Diego Supercomputer Center. Phylogenetic trees were visualized using the interactive tree of life (iTOL).[97]

Blast alignment

Amino acid sequence alignment of *Bacteroides uniformis* ATCC 8492 gshD as reference was performed using NCBI BLAST service against 42 *Bacteroides* strains (Figure S3E).

Tn-seq data analysis

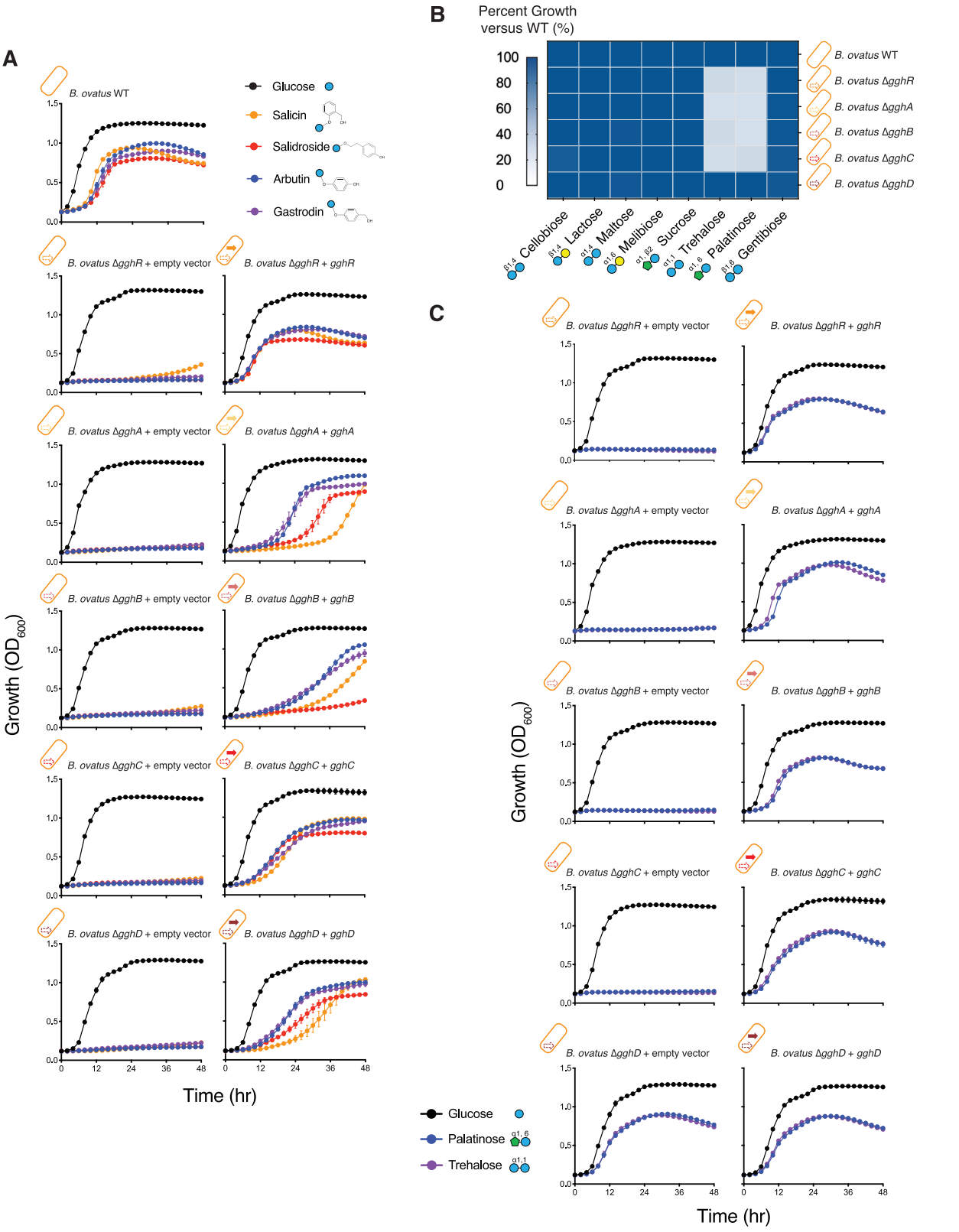
To analyze the fitness of transposon mutants in the library, Illumina raw reads were subjected to the following bioinformatics pipeline: the de-multiplexed reads were first trimmed using cutadapt v1.17[98] to remove transposon and adapter sequences (“-g GACTTA TCATCCAACCTGT -O 17 -e 0.2” and “-a ATACCACGAC -O 5 -e 0.1 -m 15”). These trimmed reads were then mapped to the corresponding *Bacteroides* reference genome using Bowtie v1.2.2[99] with the setting “-n 3 -l 28 -e 120-best”. The derived insertion tallies were subjected to comparative analyses between glucose and arbutin using Transit v2.1.0[100] under the “resampling” mode with the setting “-n TTR -iN 0.05 -iC 0.05”. As each experiment was performed in two biological replicates, to combine fold changes and *p*-values from two replicates, we used the smaller absolute value fold change and the higher *p*-value for each gene, which is more conservative than other methods such as Fisher’s method. As the Transit resampling method had a default *p*-value detection limit of 10e-4, for more significant hits that were marked as *p*-value = 0 we randomly assigned p-values <10e-4 for the sake of volcano plots.

Other statistical analyses

All in vitro bacterial growth experiments (chemical utilization, chemical killing, enzyme activity, in vitro metabolism) encompass three technical triplicates and are representative of two biological duplicates. All in vivo disease model (experimental colitis and *C. difficile* infection) data is derived from at least two different experiments. For pairwise and two independent group comparison Student’s t test was used, while for multiple group comparison one-way and two-way ANOVA were performed. Statistical significance is indicated as follows: non-significant (ns) > 0.05, *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001. Error bars represent the standard deviation (SD) or error (SE) of the mean as indicated per experiment.

CellPress

Supplemental figures



(legend on next page)

Figure S1. The *B. ovatus* glycoside utilization system is required for glycosides and disaccharides, related to Figure 3

(A) Growth curves of *B. ovatus* WT, Δ*gghR*, Δ*gghA*, Δ*gghB*, Δ*gghC*, or Δ*gghD* complemented with each respective gene or empty vector grown in glucose or phenolic glycosides as indicated (each at 15 mM) as sole carbon source.

(B) Heatmap of *B. ovatus* WT, Δ*gghR*, Δ*gghA*, Δ*gghB*, Δ*gghC*, or Δ*gghD* growth as a function of disaccharide utilization in *Bacteroides* minimal media.

(C) Growth curves of *B. ovatus* Δ*gghR*, Δ*gghA*, Δ*gghB*, Δ*gghC*, or Δ*gghD* complemented with each respective gene or empty vector grown in glucose, palatinose, or trehalose (each at 15 mM) as sole carbon source.

In (A)–(C), data are representative of one to two independent experiments, each in technical triplicate. Error bars represent mean ± SD.

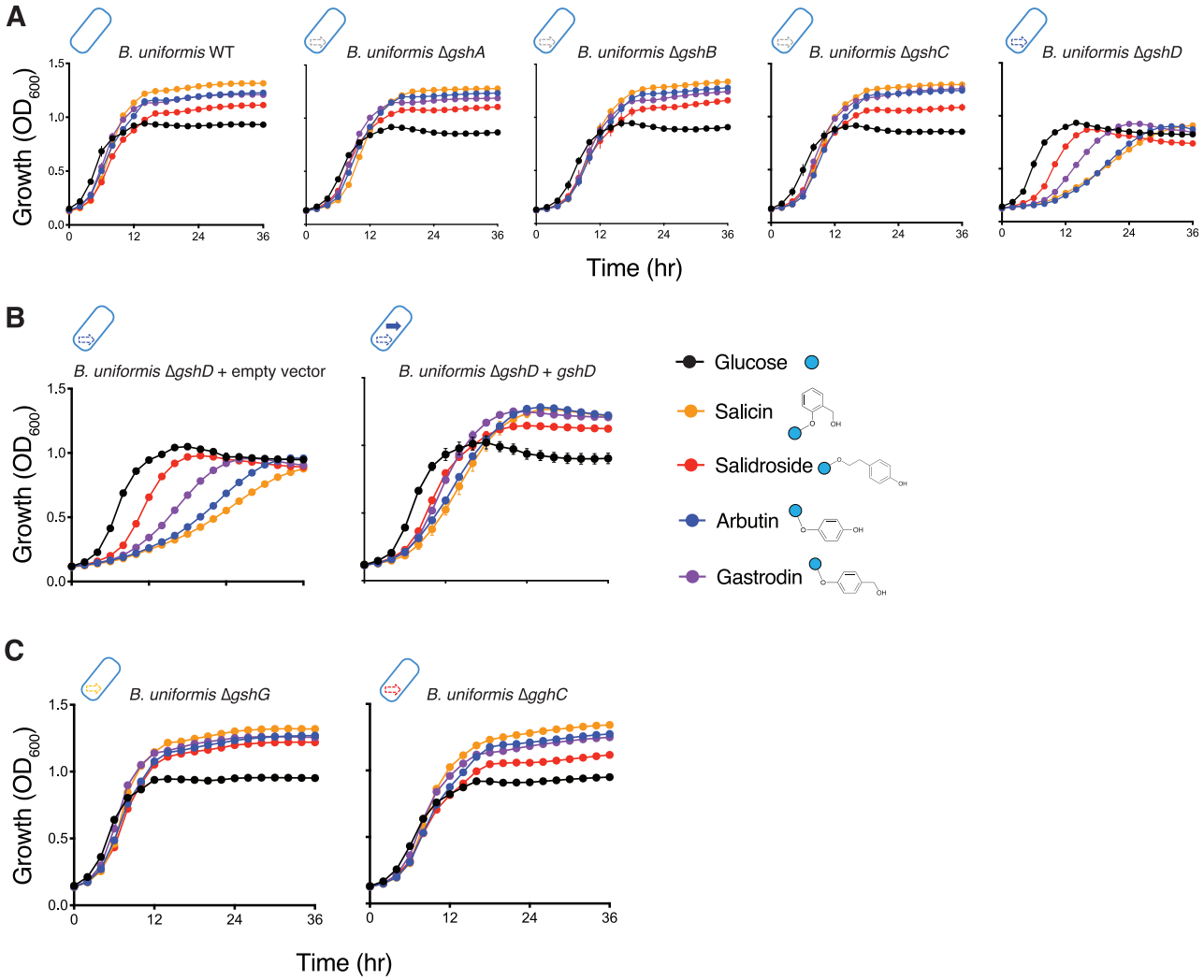


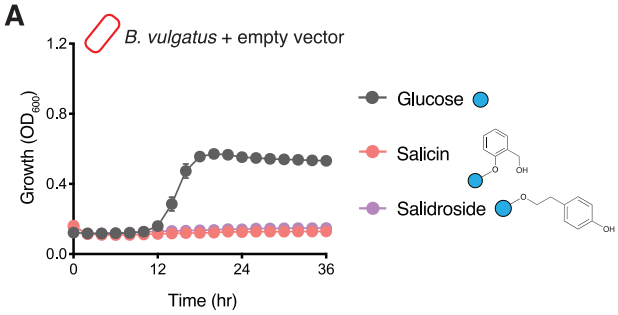
Figure S2. The phenolic glycoside utilization system of *B. uniformis*, related to Figure 3

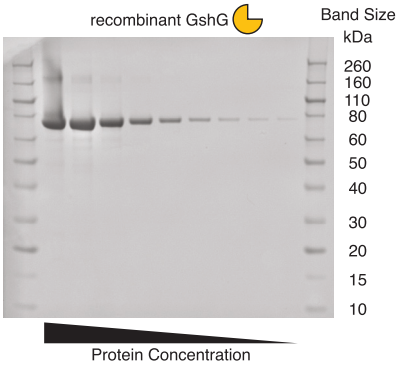
(A) Growth curves of *B. uniformis* WT, Δ*gshA*, Δ*gshB*, Δ*gshC*, or Δ*gshD* grown in glucose or each aryl glycoside.

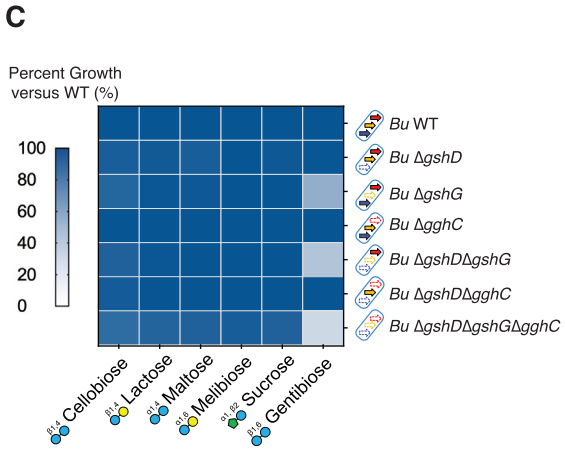
(B) Growth curves of *B. uniformis* Δ*gshD* complemented with *gshD* or empty vector grown in glucose or each aryl glycoside.

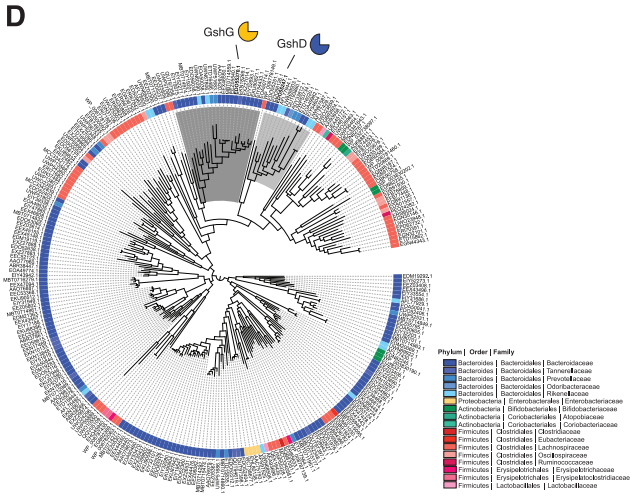
(C) Growth curves of *B. uniformis* Δ*gshG or* Δ*gghC* grown in glucose or each aryl glycoside. For all experiments glycosides or glucose at 15 mM as sole carbon source.

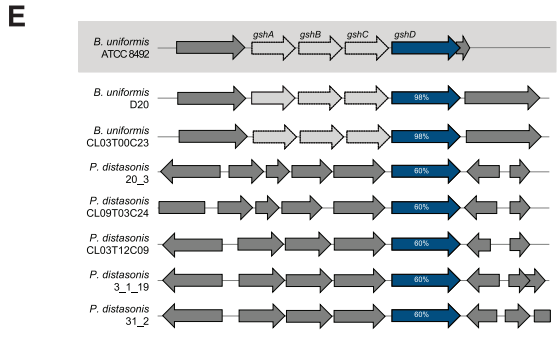
In (A)–(C), data are representative of one to two independent experiments, each in technical triplicate. Error bars represent mean ± SD.

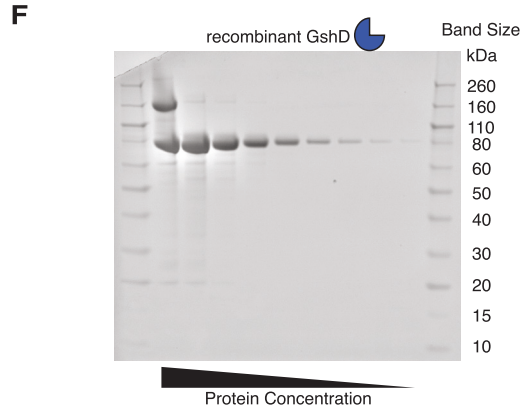


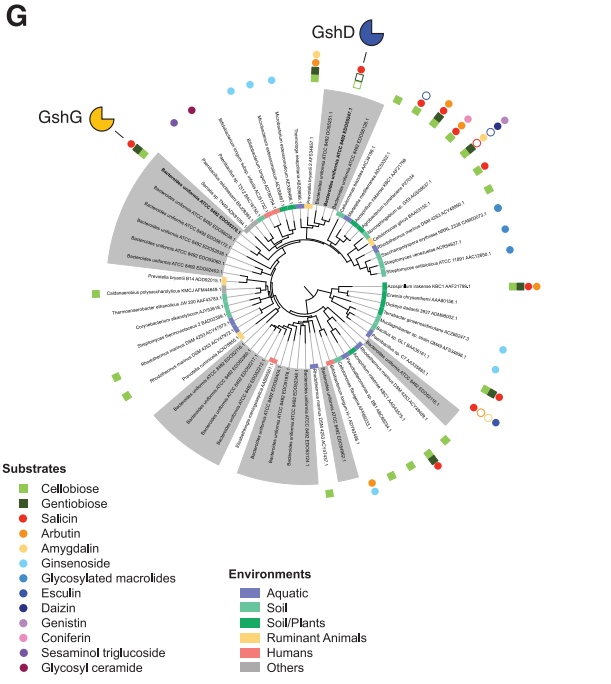












(legend on next page)

Figure S3. Specialization for glycoside metabolism by *B. uniformis*, related to Figure 4

(A) Growth curves of *B. vulgatus* WT complemented with empty vector grown in glucose, salicin, or salidroside.

(B) SDS-PAGE gels of purified, serial-diluted recombinant GshG (treated with a reducing agent).

(C) Heatmap of *B. uniformis* WT, Δ*gshD*, Δ*gshG*, Δ*gghC*, Δ*gshD*Δ*gshG*, Δ*gshD*Δ*gghC*, or Δ*gshD*Δ*gshG*Δ*gghC* growth as a function of disaccharide utilization.

(D) Maximum likelihood phylogenetic tree estimated from the amino acid sequences of 267 GH3 enzymes (including GshD and GshG) recovered from 125 genomes that display the phylogenetically diversity characteristic the human gut.

(E) Loci of *gshD* homologs with greater than 50% amino acid identity across the Bacteroidales isolates assayed in Figure 2A. Percent amino acid homology to *B. uniformis gshD* is noted in each *gshD* homolog. Genes with the same putative functional annotation are displayed in matching colors.

(F) SDS-PAGE gels of purified, serial-diluted recombinant *B. uniformis* GshG (treated with a reducing agent).

(G) Maximum likelihood phylogenetic tree estimated from the amino acid sequences of 60 GH3 enzymes, including the 19 GH3 present in *B. uniformis* ATCC 8492 and 41 GH3 from bacteria of diverse environments that were previously biochemical characterized. Close symbols represented substrates for the corresponding enzyme, open symbol represented substrates tested without activity.

In (A) and (C), data are representative of one to two independent experiments, each in technical triplicate. In (A), error bars represent mean ± SD.

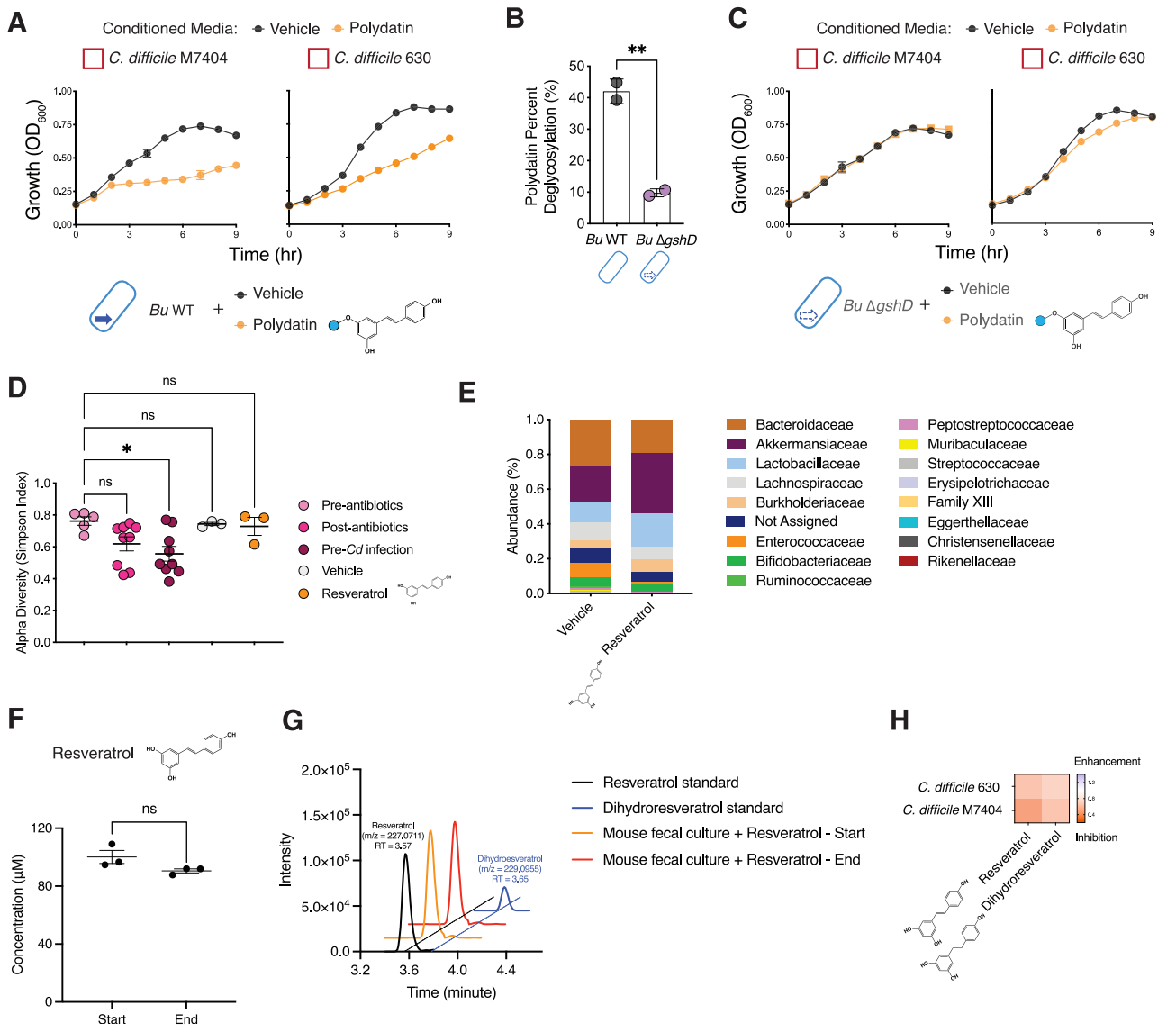


Figure S4. Microbiome-mediated polyphenol aglycone liberation, related to Figure 6

(A) Growth curves of *C. difficile* M7404 or *C. difficile* 630 in 1:1 mix of fresh media supplemented with glucose and mid-log conditioned media. Conditioned media was derived from 8-h cultures of *B. uniformis* WT grown in *Bacteroides* minimal media supplemented with glucose (0.1% v/v) plus polydatin (500 μM) or DMSO (vehicle) and then filter-sterilized.

(B) Percent conversion *in vitro* of polydatin to resveratrol by *B. uniformis* WT or Δ*gshD*.

(C) Growth curves of *C. difficile* M7404 or *C. difficile* 630 in 1:1 mix of fresh media supplemented with glucose and mid-log conditioned media. Conditioned media was derived from 8-h cultures of *B. uniformis* Δ*gshD* grown in *Bacteroides* minimal media supplemented with glucose (0.1% v/v) plus polydatin (500 μM) or DMSO (vehicle) and then filter-sterilized.

(D) Gut microbiome alpha diversity (Simpson index) of mouse stools on the last day of antibiotic administration (post-antibiotics) and after antibiotic washout on the day of *C. difficile* gavage (pre-*Cd* infection) and at time of vehicle and resveratrol administration.

(E) Gut microbiome composition (family level) during vehicle or resveratrol administration in the mouse model of *C. difficile* infection.

(F) Resveratrol concentration at the start and end (stationary phase) points in mice fecal culture with resveratrol.

(G) Combined EICs of resveratrol and dihydroresveratrol for standards and start and end points of mouse fecal cultivation with resveratrol.

(H) Heatmap of growth of *C. difficile* (630 and the hypervirulent M7404) in the presence of resveratrol and dihydroresveratrol. Growth is measured as the ratio of growth in media (AUC) supplemented with Glc and resveratrol or dihydroresveratrol (each 150 μM) over that in Glc and vehicle control (DMSO).

In (A), (C), (G), and (H), data are representative of one to two independent experiments, each in technical triplicate. In (D) and (E), data are combined from one to two independent experiments (n = 3–5 mice per experiment). In (B), data are two biological replicates. In (F), data are three biological replicates. In (A)–(D), (G), and (H), error bars represent mean ± SD. In (F), error bars represent mean ± SEM. Non-significant (ns) $p > 0.05$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$; Student's t test in (B) and (F) and one-way ANOVA in (D) with Dunnett's (D) post hoc test.

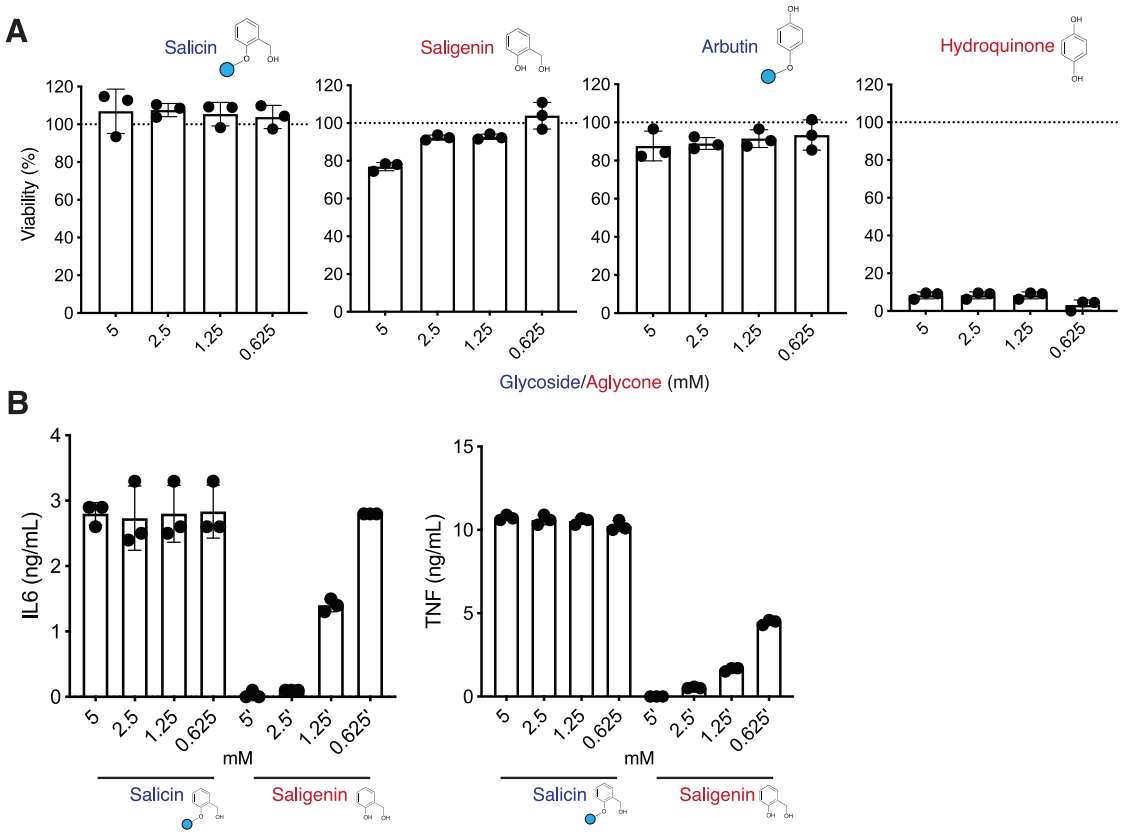


Figure S5. Diverse effects on type 1 inflammation of phenolic glycosides and *Bacteroides*-liberated aglycones, related to Figure 7

(A) Viability of iBMDMs treated with salicin, saligenin, arbutin, or hydroquinone described in Figure 7A.

(B) Dose-response IL-6 (left) and TNF (right) secretion by murine iBMDMs macrophages after 24-h LPS (1 μg/mL; *E. coli* O111:B4) stimulation in the presence of salicin or saligenin (40 h total; 16-h pre-LPS-treatment incubation and 24-h LPS stimulation).

In (A) and (B), data are representative of two independent experiments, each in technical triplicate. In (A) and (B), error bars represent mean ± SD.

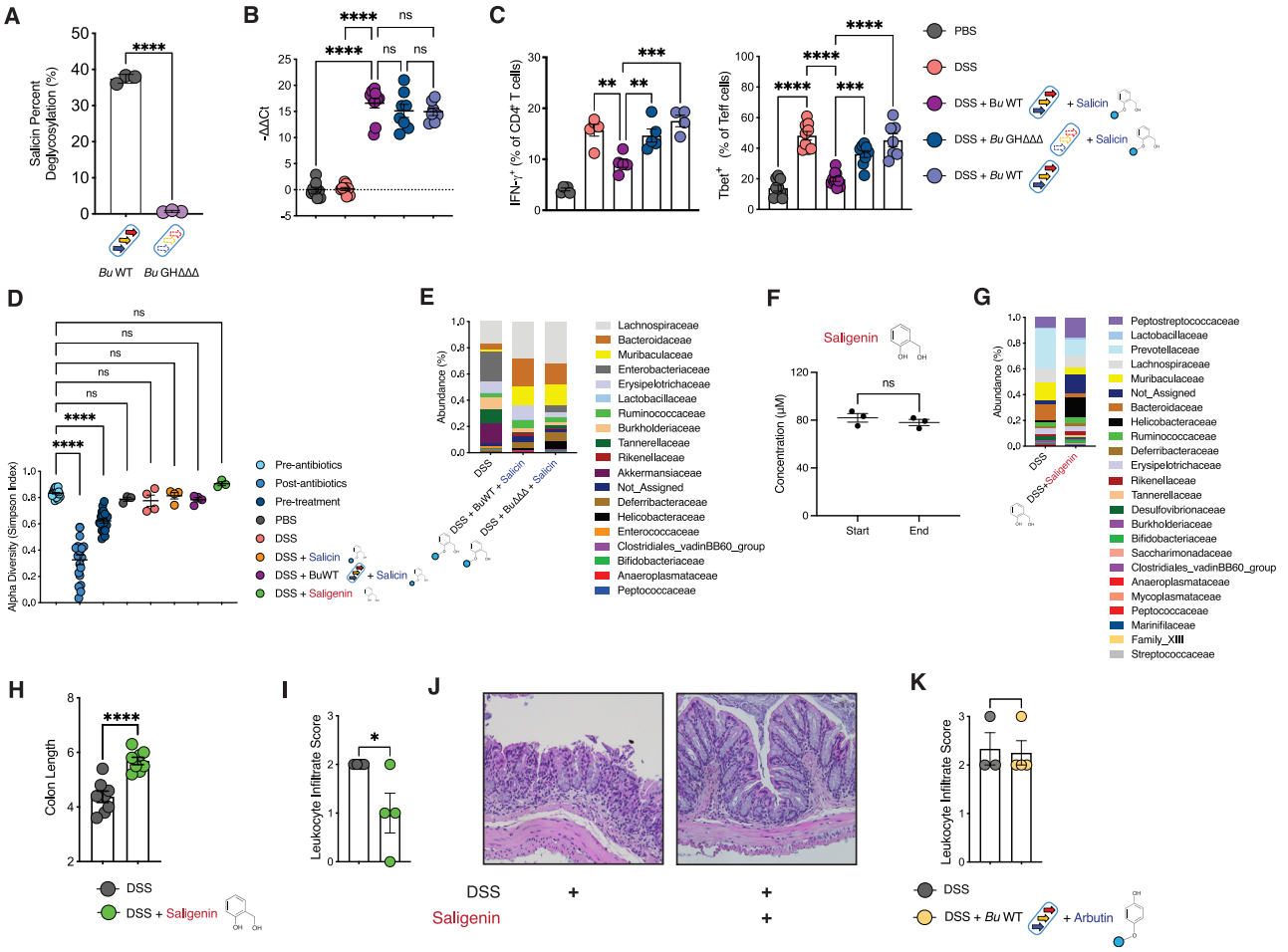


Figure S6. *B. uniformis* bioactivation of salicin to saligenin liberates an anti-inflammatory metabolite, related to Figure 7

(A) Percent conversion *in vitro* of salicin to saligenin by *B. uniformis* WT or *B. uniformis* Δ*gshD*Δ*gshG*Δ*gghC*.

(B) Relative colonization as measured by qPCR of *B. uniformis* at day 8 in mice treated with PBS or DSS (2.5%) and colonized with *B. uniformis* WT or *B. uniformis* Δ*gshD*Δ*gshG*Δ*gghC* with daily i.g. administration of salicin (100 mg/kg body weight) or PBS control. Relative colonization was determined by comparing Ct values against an added internal standard bacterium (*S. ruber*).

(C) Frequency of IFN-γ producing-CD4$^+$ T cells and T_{bet}^+ T_H1 effector cells for the experiments described in Figures 7D and 7E.

(D) Gut microbiome alpha diversity (Simpson index) of stool on the last day of antibiotic administration (post-antibiotics), after antibiotic washout at the time of *B. uniformis* administration before DSS or salicin (pre-treatment) and intervention (PBS, DSS, DSS + salicin, DSS + *B. uniformis* WT + salicin. The last panel (DSS + saligenin) is stool from antibiotic-naive mice treated with DSS and saligenin.

(E) Gut microbiome composition (family level) during DSS intervention in mice colonized with *B. uniformis* WT or Δ*gshD*Δ*gshG*Δ*gghC* treated with salicin compared with DSS alone.

(F) Saligenin concentration of start and end (stationary phase) in mice fecal culture with saligenin.

(G) Gut microbiome composition (family level) during saligenin administration in antibiotic-naive mice in the DSS model of colitis.

(H) Colon length at day 8 of mice treated with DSS (2.5%) plus vehicle or DSS plus saligenin (100 mg/kg body weight) administered i.g. daily.

(I) Histopathologic scoring of colonic leukocyte infiltrate for the experiments described in (H).

(J) Representative H&E images for the experiments described in (H).

(K) Histopathologic scoring of leukocyte infiltrate for the experiments described in Figure 7E.

In (A), data are three biological replicates. In (B)–(E) and (G)–(K), data are combined from one to three independent experiments (n = 3–5 mice per experiment). In (F), data are three biological replicates. In (B), bacterial relative abundance was measured at day 8. In (C), immune cell frequencies were measured at day 8. In (H), colon lengths were measured at day 8. In (A), error bars represent mean ± SD. In (B)–(D), (F), (H), (I), and (K), error bars represent mean ± SEM. Non-significant (ns) $p > 0.05$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$; Student's t test in (A), (F), (H), (I), and (K) and one-way ANOVA in (B)–(D) with Šídák's (B), Tukey's (C), and Dunnett's (D) post hoc tests.

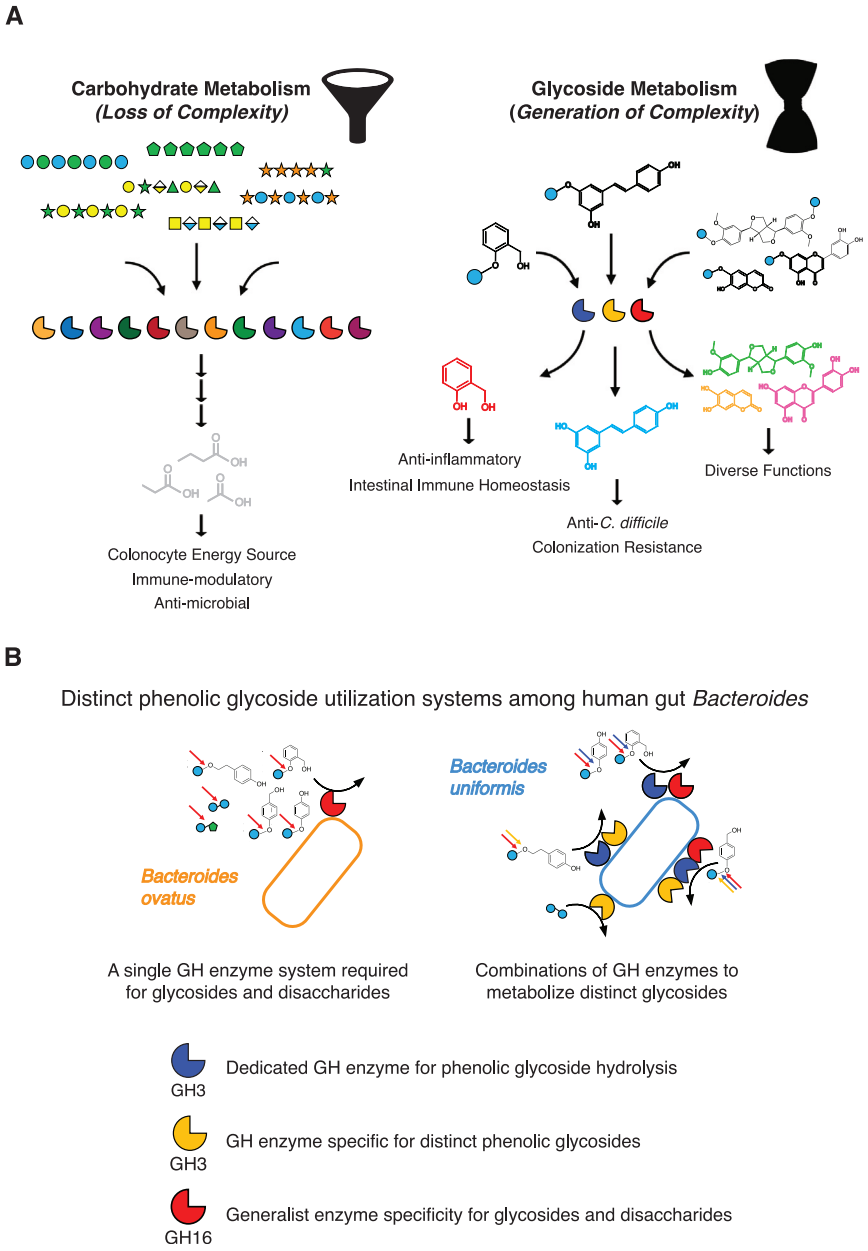


Figure S7. Dietary glycoside metabolism by the gut microbiome as a complexity-generating process, related to Figures 3, 5, 6, and 7

(A) (Left) The human gut microbiome employs a diverse array of glycoside hydrolases to transform the chemical diversity polysaccharides into a limited number of host-active fermentation products ("funneling"). (Right) Deglycosylation of diverse plant small molecule glycosides by a defined set of glycoside hydrolases encoded within the gut microbiome results in the bioactivation of diverse host-active functions generating diverse host effector functions ("bowtie").

(B) Two distinct metabolic systems within the human gut *Bacteroides* to hydrolyze glycoside PSMs. In *B. ovatus* (left), a single GH16 family member is required to utilize both a range of phenolic glycosides as well as disaccharides. In contrast, in *B. uniformis* (right), a glycoside utilization system composed of GH3 family members dedicated to phenolic glycosides and unable to hydrolyze disaccharides (*Bu* gshD) and demonstrating specificity for distinct phenolic glycosides (*Bu* gshG) depending on the chemical structure of each respective aglycone (for example, the substitution of a functionalization).