

## Article

# Flavonoid-Modifying Capabilities of the Human Gut Microbiome—An In Silico Study

Tobias Goris <sup>1,\*</sup>, Rafael R. C. Cuadrat <sup>2</sup> and Annett Braune <sup>1</sup>

<sup>1</sup> Research Group Intestinal Microbiology, Department of Molecular Toxicology, German Institute of Human Nutrition Potsdam-Rehbruecke, Arthur-Scheunert-Allee 114-116, 14558 Nuthetal, Germany; braune@dife.de

<sup>2</sup> Department of Epidemiology, German Institute of Human Nutrition Potsdam-Rehbruecke, Arthur-Scheunert-Allee 114-116, 14558 Nuthetal, Germany; Rafael.Cuadrat@dife.de

\* Correspondence: tobias.goris@dife.de

**Abstract:** Flavonoids are a major group of dietary plant polyphenols and have a positive health impact, but their modification and degradation in the human gut is still widely unknown. Due to the rise of metagenome data of the human gut microbiome and the assembly of hundreds of thousands of bacterial metagenome-assembled genomes (MAGs), large-scale screening for potential flavonoid-modifying enzymes of human gut bacteria is now feasible. With sequences of characterized flavonoid-transforming enzymes as queries, the Unified Human Gastrointestinal Protein catalog was analyzed and genes encoding putative flavonoid-modifying enzymes were quantified. The results revealed that flavonoid-modifying enzymes are often encoded in gut bacteria hitherto not considered to modify flavonoids. The enzymes for the physiologically important daidzein-to-equol conversion, well studied in *Slackia isoflavoniconvertens*, were encoded only to a minor extent in *Slackia* MAGs, but were more abundant in *Adlercreutzia equolifaciens* and an uncharacterized *Eggerthellaceae* species. In addition, enzymes with a sequence identity of about 35% were encoded in highly abundant MAGs of uncultivated *Collinsella* species, which suggests a hitherto uncharacterized daidzein-to-equol potential in these bacteria. Of all potential flavonoid modification steps, *O*-deglycosylation (including derhamnosylation) was by far the most abundant in this analysis. In contrast, enzymes putatively involved in *C*-deglycosylation were detected less often in human gut bacteria and mainly found in *Agathobacter faecis* (formerly *Roseburia faecis*). Homologs to phloretin hydrolase, flavanone/flavanone-cleaving reductase and flavone reductase were of intermediate abundance (several hundred MAGs) and mainly prevalent in *Flavonifractor plautii*. This first comprehensive insight into the black box of flavonoid modification in the human gut highlights many hitherto overlooked and uncultured bacterial genera and species as potential key organisms in flavonoid modification. This could lead to a significant contribution to future biochemical-microbiological investigations on gut bacterial flavonoid transformation. In addition, our results are important for individual nutritional recommendations and for biotechnological applications that rely on novel enzymes catalyzing potentially useful flavonoid modification reactions.

**Citation:** Goris, T.; Cuadrat, R. R. C.; Braune, A. Flavonoid-Modifying Capabilities of the Gut Microbiome—An In Silico Study. *Nutrients* **2021**, *13*, 2688. <https://doi.org/10.3390/nu13082688>

Academic Editor: Elad Tako

Received: 22 June 2021

Accepted: 26 July 2021

Published: 3 August 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

**Keywords:** polyphenols; plant metabolites; phytohormones; isoflavones; personalized nutrition; metagenomes; human gut microbiome

## 1. Introduction

Flavonoids, which are exclusively plant secondary metabolites, are considered to bear many beneficial effects on health [1–5] and are assumed to contribute especially to a lower cardiovascular disease and cancer-related mortality [6,7]. Most flavonoids are taken up as a part of the human diet mainly as glycosides. A portion of dietary flavonoids is absorbed in the small intestine following *O*-deglycosylation by epithelial enzymes. Flavonoid aglycons then undergo phase I/II transformation [8–10]. The major part of dietary

flavonoids (or their phase I/II metabolites) reach the large intestine, where they are subject to transformation by gut bacteria which can enhance or lower their bioavailability and biological activity [11–14]. In particular, members of the genera *Bifidobacterium* and *Lactobacillus* were observed to *O*-deglycosylate, e.g., flavanone and isoflavone glycosides [11]. While *O*-deglycosylation was the most extensively described flavonoid deglycosylation, with many glycosides from flavonols (e.g., rutin in several vegetables), flavanones (often found in citrus fruits) and isoflavones (present mainly in soy products) as substrates, several bacteria also mediate *C*-deglycosylation from compounds such as isovitexin, e.g., from buckwheat. Additional flavonoid transformations involve demethylation, dehydroxylation, reduction, and *C*-ring cleavage [11].

Many described gut bacterial flavonoid-converting activities lack biochemical or genetic analyses and, therefore, sequences of the involved enzymes. Of the *O*-deglycosylating enzymes, six beta-galactosidases [15–17] and a set of flavonoid-active rhamnosidases [18] that cleave the terminal rhamnose subunit of flavonoid rhamnosides (e.g., rutin or naringin), resulting in the corresponding flavonoid glucosides [19–24], have been characterized (Table 1). The more complex *C*-deglycosylating enzyme systems were studied in fewer gut bacteria [25–27] (Table 1). The *C*-deglycosylation system described in strain PUE is dependent on at least three enzymes, DgpABC, which are part of a large gene cluster including transporters and other accessory genes whose function in *C*-deglycosylation is unclear [26,27]. The *C*-deglycosylating enzyme system of *Eubacterium cellulosolvens* includes the five enzymes DfgABCDE [25]. Both corresponding enzyme complexes include at least one glycosyltransferase and an oxidoreductase. Phloretin hydrolysis as carried out by phloretin hydrolase [28,29] and the three-step conversion of daidzein to equol [30–32] are relatively well studied, while the gut bacterial enzymes involved in dehydroxylation and *O*-demethylation of flavonoids are not yet described. However, an *O*-demethylase from *Eubacterium limosum* ZL-II demethylates polyphenolic lignans, [33]. As with other characterized *O*-demethylases [34,35], this enzyme system consisted of two methyltransferases (MT), a corrinoid protein (CP) and an activating enzyme (AE), and might also be responsible for the *O*-demethylation of flavonoid carried out by *E. limosum* and *E. ramulus* [36,37].

The enzymes involved in daidzein-to-equol conversion and their encoding genes have been characterized in *Slackia egggerthella* and *Lactococcus* strains [30,32,38,39]. At least three enzymes are involved in the complete conversion of daidzein to equol. All three were proposed to be reductases [30,32], although one of them was recently suggested to be a dismutase [38]. In addition, a dihydrodaidzein (DHD) racemase is employed for stereochemical conversion of DHD in some bacteria [30]. Besides the complete conversion from daidzein to equol estimated to occur in about half of the population [40–42], a large number of individuals carry bacteria catalyzing only the reduction of daidzein to DHD or *O*-desmethylangolensin [42,43].

Phloretin hydrolase (Phy) hydrolytically cleaves the C-C bond adjacent to the aromatic A-ring of phloretin and thereby produces phloroglucinol and 3-(4-hydroxyphenyl)propionic acid. The first Phy was discovered in *E. ramulus* [28,44]. The only other Phy has been found in the intestinal pathogenic *Mycobacterium abscessis*, showing a 30% amino acid sequence identity to that of *E. ramulus* [27].

The flavanone- and flavanol-cleaving reductase (Fcr), cleaving the heterocyclic C-ring of flavanones and flavanonols, was recently characterized in detail [45]. Earlier suggested to be an enoate reductase and to act in a concerted pathway together with a chalcone isomerase (CHI, [46]), the Fcr of *E. ramulus* was described to catalyze the conversion of, e.g., naringenin to phloretin also without CHI. Still, as CHIs [46–49] are responsible for the production of auronol from flavanonols [47], they could play a substantial role in flavonoid transformation in the human gut. Very recently, the characterization of an enereductase (Flr), catalyzing the reduction of the C ring double bond of flavones and flavonols, was published [50].

**Table 1.** Characterized flavonoid-modifying enzymes of human gut bacteria.

Accession No.	Enzyme	Reaction	Source Bacterium	Reference
KEF29323.1	Beta-glucosidase BpBluA	7-O-deglycosylation of isoflavones	<i>Bifidobacterium pseudocatenum</i> IPLA36007	[15]
KEF27912.1	Beta-glucosidase BpGluB	7-O-deglycosylation of isoflavones	<i>Bifidobacterium pseudocatenum</i> IPLA36007	[15]
KEF28010.1	Beta-glucosidase BpGluD	7-O-deglycosylation of isoflavones	<i>Bifidobacterium pseudocatenum</i> IPLA36007	[15]
KEF28001.1	Beta-glucosidase BpGluE	7-O-deglycosylation of isoflavones	<i>Bifidobacterium pseudocatenum</i> IPLA36007	[15]
AFS33105.1	Beta-glucosidase Bbg572	7-O-deglycosylation of isoflavones, 3-/4-/7-O-deglycosylation of quercetin glycosides	<i>Bifidobacterium animalis</i> ssp. <i>lactis</i> SH5	[16]
AAO76887.1	Beta-glucosidase BtGlu	7-O-deglycosylation of isoflavones	<i>Bacteroides thetaiotaomicron</i> VPI-5482	[17]
AGS77942.1	Alpha-L-Rhamnosidase BdR	Derhamnosylation of 1→2 and 1→6 O-glycosidic bond of rutinosylated flavonoids	<i>Bifidobacterium dentium</i> K13	[19]
AHJ22585.1	Alpha-L-Rhamnosidase BbR	Derhamnosylation of 1→6 O-glycosidic bond of rutin	<i>Bifidobacterium breve</i> 689b	[20]
WP_011107561	Alpha-L-Rhamnosidase BtR	Derhamnosylation of 1→2 O-glycosidic bonds of rutinosylated flavonoids	<i>Bacteroides thetaiotaomicron</i> VPI-5482	[21]
QBM20340.1	Alpha-L-Rhamnosidase HFM-RhaA (MGR1)	Derhamnosylation of 1→2 and 1→6 O-glycosidic bond of rutinosylated flavonoids	Human gut metagenome	[24]
QBM20341.1	Alpha-L-Rhamnosidase HFM-RhaB (MGR2)	Derhamnosylation of 1→2 O-glycosidic bond of rutinosylated flavonoids	Human gut metagenome	[24]
QBM20342.1	Alpha-L-Rhamnosidase HFM-RhaC (MGR3)	Derhamnosylation of 1→2 and 1→6 O-glycosidic bond of rutinosylated flavonoids	Human gut metagenome	[24]
CCC80440.1	Alpha-L-Rhamnosidase LpR1	Derhamnosylation of 1→6 O-glycosidic bond of rutinosylated flavonoids	<i>Lactobacillus plantarum</i> WCFS1 *	[22]
CCC80442.1	Alpha-L-Rhamnosidase LpR2	Derhamnosylation of 1→6 O-glycosidic bond of rutinosylated flavonoids	<i>Lactobacillus plantarum</i> WCFS1 *	[22]
AAV43293.1	Alpha-L-Rhamnosidase LaR	Derhamnosylation of 1→2 and 1→6 O-glycosidic bond of rutinosylated flavonoids	<i>Lactobacillus acidophilus</i> NCFM	[22]
EUA80835.1	Phloretin hydrolase	Phloretin hydrolysis	<i>Mycobacteroides abscessus</i> ssp. <i>bolletii</i> 103	[28]
AAQ12341.1	Phloretin hydrolase	Phloretin hydrolysis	<i>Eubacterium ramulus</i> DSM 16296	[29]
AKC35075.1	DfgCD	7-O-deglycosylation of isoflavones and flavones	<i>Catenibacillus scindens</i> DSM 106146	[25]
AKC35076.1				
EIM58373.1	DfgABCDE	C-deglycosylation of flavones	<i>Eubacterium cellulosolvens</i> ATCC 43171	[25]
EIM58372.1				
EIM58371.1				

EIM58370.1					
EIM58369.1					
BBG22493.1					
BBG22494.1	DgpABC	C-deglycosylation of isoflavones		<i>Dorea</i> sp. strain PUE	[26]
BBG22495.1					
ANU40626.1	Flavone/flavonol reductase (Flr)	C-ring double bond reduction		<i>Flavonifractor plautii</i> DSM 6740	[50]
ADK16070.1	Flavone/flavonol reductase (Flr)	C-ring double bond reduction		<i>Clostridium ljungdahlii</i> DSM 13528	[50]
AGS82961.1	Flavanone/flavanonol-cleaving reductase (Fcr)	Reductive ring cleavage of flavanones and flavanonols		<i>Eubacterium ramulus</i> DSM 16296	[45]
WP_154024723.1	Flavanone/flavanonol-cleaving reductase (Fcr)	Reductive ring cleavage of flavonones and flavanonols		<i>Flavonifractor plautii</i> DSM 6740	[45]
AIS36173.1	Chalcone isomerase (CHI)	Isomerization of chalcones and flavanonols		<i>Eubacterium ramulus</i> DSM 16296	[46,47,49]
EHM54434.1	Chalcone isomerase (CHI)	Isomerization of chalcones		<i>Flavonifractor plautii</i> ATCC 29863 (formerly <i>Clostridium orbiscindens</i> )	[46,47,49]
AFV15450.1	Tetrahydrodaidzein reductase (TDR)	Reduction of tetrahydrodaidzein		<i>Slackia isoflavoniconvertens</i> DSM 22006	[30]
AFV15451.1	Dihydrodaidzein reductase (DDR)	Reduction of dihydrodaidzein		<i>Slackia isoflavoniconvertens</i> DSM 22006	[30]
AFV15453.1	Daidzein reductase (DZR)	Reduction of daidzein		<i>Slackia isoflavoniconvertens</i> DSM 22006	[30]
BAL46928.1	Tetrahydrodaidzein reductase (TDR)	Reduction of tetrahydrodaidzein		<i>Slackia</i> sp. NATTS	[32]
BAL46929.1	Dihydrodaidzein reductase	Reduction of dihydrodaidzein		<i>Slackia</i> sp. NATTS	[32]
BAL46930.1	Daidzein reductase	Reduction of daidzein		<i>Slackia</i> sp. NATTS	[32]
BAM25050.1	Dihydrodaidzein Racemase	Racemization of dihydrodaidzein		<i>Lactococcus garviae</i> 20-92	[31]
BAJ72744.1	Tetrahydrodaidzein reductase	Reduction of tetrahydrodaidzein		<i>Lactococcus garviae</i> 20-92	[51]
BAJ72745.1	Dihydrodaidzein reductase	Reduction of dihydrodaidzein		<i>Lactococcus garviae</i> 20-92	[51]
BAJ22678.1	Daidzein reductase	Reduction of daidzein		<i>Lactococcus garviae</i> 20-92	[39]
WP_013979960.1	Tetrahydrodaidzein reductase	Reduction of tetrahydrodaidzein		<i>Eggerthella</i> sp. YY7918	[38]
WP_013979959.1	Dihydrodaidzein reductase	Reduction of dihydrodaidzein		<i>Eggerthella</i> sp. YY7918	[38]
WP_013979957.1	Daidzein reductase	Reduction of daidzein		<i>Eggerthella</i> sp. YY7918	[38]
ANI69959.1	O demethylase (ODem) activating enzyme (AE)	Activation of CP		<i>Eubacterium limosum</i> ZL-II	[33]
ANI69960.1	ODem Methyltransferase (MT) 1	O-Demethylation of secoisolariciresinol		<i>Eubacterium limosum</i> ZL-II	[33]
ANI69961.1	ODem Corrinoid protein (CP)	Methyl transfer		<i>Eubacterium limosum</i> ZL-II	[33]
ANI69962.1	ODem MT2	Methyl transfer to CP		<i>Eubacterium limosum</i> ZL-II	[33]

\* An identical rhamnosidase was identified in *Lactobacillus plantarum* NCC245 (Avila et al., 2009).

All of the above studies were performed with few flavonoid-modifying bacteria and the corresponding enzymes originate from a limited number of bacterial species. Isolation of these strains from fecal samples could be biased because of selective enrichment of culturable bacteria. Therefore, the research on gut-bacterial flavonoid modification published up to now does not display a complete picture of flavonoid transformation in the human gut: phenotypic studies on human fecal samples rely mainly on (partial) 16S rRNA gene sequencing and do not resolve bacterial species. Thus, flavonoid-converting species might be missed. In addition, quantitative studies of flavonoid-converting bacteria are scarce. A recent study on the quantification of daidzein-to-equol-converting bacteria relied on primers targeting genes of involved reductases [52]. With the advent of metagenomic studies, a closer look into the gene catalog of the human gut microbiome became feasible [53], but still, quantification and especially taxonomic affiliation of the genes of interest are challenging. Current bioinformatic methods offered the opportunity to assemble genomes from metagenomic data [54], which were also recently applied to human gut metagenome studies [55–57], opening up new ways to reveal flavonoid modification in the human gut.

Here, we aim to identify and quantify flavonoid-modifying bacteria in the human gut microbiota. For this, we screened the most recent combined human gut MAGs collection, the Unified Human Gastrointestinal Protein or Genome (UHGP/UHGG) catalog of more than 280,000 assembled genomes, with a BLAST search using protein sequences of characterized flavonoid-modifying enzymes (Table 1) as queries. In addition, the presence, prevalence and abundance of bacterial species described to modify flavonoids was investigated. Based on the obtained data, a possible scenario of flavonoid modification by the identified human gut bacteria is presented.

## 2. Results and Discussion

### 2.1. Distribution of Flavonoid-Modifying Bacterial Species Across Human Gut MAGs

First, we screened the accompanying metadata of the MAGs database for the abundance of bacterial species described to transform flavonoids. For this, a literature-based overview of currently known flavonoid-converting bacterial strains present in the human gut was compiled, resulting in 45 distinct bacterial strains of 43 different species (Table 2). Of these, eleven were not taxonomically validly published, either due to lack of a 16S rRNA gene sequence or because the authors chose an informal taxonomic designation. If a 16S rRNA sequence was available, these unclassified isolates were provisionally assigned to the closest taxonomically described species using a 98% sequence identity cutoff to the respective strain. *O*-deglycosylation was the most abundant flavonoid modification performed by human gut bacteria (31 strains). *C*-deglycosylation, *C*-ring cleavage, reduction, dehydroxylation or *O*-demethylation of flavonoids were far less abundant (between three and seven strains each).

We then performed a text-based filtering and counting of species in the lineage column of the UHGG metadata with the aim of specifying the MAG abundances of described flavonoid-transforming species as percentages of the overall number of MAGs in the human gut. Important to note here is that the taxonomic assignments were based on the genome taxonomy database (GTDB, <https://gtdb.ecogenomic.org/>, accessed on 1 July 2020) in the UHGG metadata. Thus, species without a sequenced genome (*Catenibacillus scindens*, *Lactobacillus leichmanii* and *Slackia* sp. NATTS) or species lacking taxonomic assignment in the GTDB to a given genome (*S. isoflavoniconvertens*) could not be found. Besides the overall abundance of MAGs, we also calculated the prevalence of MAGs across individuals as percentages of the overall number of metagenome samples in the metadata.

The most prevalent flavonoid-modifying bacteria were *Bacteroides uniformis* (18.5% prevalence across all 21,508 individuals), *Parabacteroides distasonis* (14.5%), *Bifidobacterium infantis* (8.4%), and *Bi. adolescentis* (7.4%; Table 2). All of these species were described to be capable of flavonoid *O*-deglycosylation, mainly flavonols, isoflavones and flavanones [11]. The remaining species were detected in less than 5% of individuals (Table 2). The

overall abundance across all MAGs were similarly distributed (Supplementary Table S1). The numbers for a single *E. coli* strain described to O-deglycosylate flavonoids [58] should be handled with care, since this activity was as-yet not confirmed for other *E. coli* strains and the MAGs database does not allow for strain-level assignment. The lack of *Bi. longum* in the MAGs metadata is remarkable, but can be explained by the fact that *Bi. longum* ssp. *infantis* [59] is classified as *Bi. infantis* in the GTDB. In addition, potential flavonoid-modifying bacterial species from a recently published BRENDA enzyme reaction database search [60] were screened for their abundance in the MAGs metadata. However, these bacteria showed only very low abundance (Supplementary Table S2) and were not included in this study.

**Table 2.** Flavonoid-modifying gut bacterial species and their prevalence (percent of MAGs from all samples involved) as MAGs derived from the taxonomic classification in the metadata of the UHGG [61].

Species	Flavonoid Class	Prevalence
<b>O-Deglycosylation</b>		
<i>Bacteroides ovatus</i>	Flavonols	3.1
<i>Bacteroides uniformis</i>	Flavonols, flavanones	18.5
<i>Bacteroides thetaiotaomicron</i> <sup>1</sup>	Isoflavones	1.9
<i>Bifidobacterium adolescentis</i>	Flavanones, isoflavones	7.4
<i>Bifidobacterium angulatum</i>	Isoflavones	0.3
<i>Bifidobacterium animalis</i>	Anthocyanidins, isoflavones	0.3
<i>Bifidobacterium bifidum</i>	Flavanones, isoflavones	3.3
<i>Bifidobacterium breve</i>	Flavonols, isoflavones, flavanones	1.5
<i>Bifidobacterium catenulatum</i>	Flavonols, isoflavones, flavanones	0.8
<i>Bifidobacterium dentium</i>	Flavonols, flavanones	0.3
<i>Bifidobacterium infantis</i>	Flavonols, isoflavones, flavanones	8.4
<i>Bifidobacterium longum</i>	Isoflavones	0
<i>Bifidobacterium pseudocatenulatum</i>	Flavonols, flavanones	3.5
<i>Bifidobacterium pseudolongum</i>	Isoflavones	<0.1
<i>Blautia producta</i> (MRG-PMF1, 99%)	Flavonols, flavones, flavanones, isoflavones	<0.1
<i>Catenibacillus scindens</i>	Flavones, isoflavones	n.d. (<0.1)
<i>Enterobacter cloacae</i>	Flavanones	0.1
<i>Enterococcus avium</i>	Flavonols	0.1
<i>Enterococcus casseliflavus</i>	Flavonols	<0.1
<i>Enterococcus faecalis</i>	Flavanones	2.9
<i>Eubacterium cellulosolvens</i>	Flavones, isoflavones	0
<i>Eubacterium ramulus</i>	Flavonols, flavones, dihydrochalcones, isoflavones	0.5
<i>Escherichia</i> sp. HGH21 (99% to <i>E. coli</i> (MIDI *))	Isoflavones	(28)
<i>Escherichia</i> sp. 4 ( <i>E. fergusonii</i> )	Flavones	<0.1
<i>Parabacteroides distasonis</i>	Flavonols, flavanones	14.5
<i>Lactobacillus acidophilus</i>	Flavanones	<0.1
<i>Lactobacillus buchmeri</i>	Flavanones	<0.1
<i>Lactobacillus casei</i>	Flavanones, anthocyanidins	0
<i>Lactobacillus leichmanii</i>	Flavanones	n.d.
<i>Lactobacillus plantarum</i>	Flavanones, anthocyanidins	0.2
<i>Lactococcus lactis</i>	Flavonols, isoflavones, flavanones	0.2
<i>Lactococcus paracasei</i> [62]	Isoflavones, flavonols	0.2
<b>C-Deglycosylation</b>		
<i>Eubacterium cellulosolvens</i>	Flavones, isoflavones	0
Strain PUE (1346 nt), <i>Dorea longicatena</i> (98%)	Isoflavones	n.d. (2.9)
<i>Catenibacillus scindens</i>	Flavones, isoflavones	n.d.

<i>Enterococcus casseliflavus</i> (sp. 45, 99%)	Flavones	<0.1
<i>Enterococcus faecium</i> (MRG-IFC-2, 99%)	Isoflavones	1.8
<i>Lactococcus lactis</i> (MRG-IFC-1, 99%)	Isoflavones	0.2
<b>C-Ring cleavage</b>		
<i>Flavonifractor plautii</i> (formerly <i>C. orbiscindens</i> )	Flavonols/flavanonols, flavones/flavanones	2.1
<i>Catenibacillus scindens</i>	Flavones	n.d.
<i>Eubacterium ramulus</i>	Flavonols/flavanonols, flavones/flavanones, isoflavones	0.5 <sup>1</sup>
<i>Clostridium butyricum</i>	Flavanones	0.3
<i>Lactobacillus plantarum</i>	Flavan-3-ols	0.2
Strain SY8519 ( <i>Eubacterium_I</i> sp000270305)	Isoflavones	n.d.
<i>Adlercreutzia equolifaciens</i>	Flavan-3-ols	0.8
<i>Eggerthella lenta</i>	Flavanonols (SDG-2), flavan-3-ols	0.9
<i>Enterococcus faecium</i> [63]	Isoflavones	1.8
<b>Reduction</b>		
<i>Bifidobacterium animalis</i>	Isoflavones	0.3
<i>Bifidobacterium longum</i>	Isoflavones	0
<i>Bifidobacterium pseudolongum</i>	Isoflavones	<0.1
<i>Adlercreutzia equolifaciens</i>	Isoflavones	0.8
<i>Lactococcus garvieae</i>	Isoflavones	<0.1
<i>Slackia</i> sp. NATTS	Isoflavones	n.d.
<i>Slackia isoflavoniconvertens</i>	Isoflavones	n.d.
<i>Lactobacillus rhamnosus</i> [64]	Isoflavones	0.4
<i>Enterococcus faecalis</i> [64]	Isoflavones	2.9
<i>Enterococcus faecium</i> [63]	Isoflavones	1.8
<b>Dehydroxylation</b>		
<i>Escherichia</i> sp. 4 ( <i>E. fergusonii</i> )	Flavones	<0.1
<i>Adlercreutzia equolifaciens</i>	Flavan-3-ols	0.8
<i>Eggerthella lenta</i> (SDG-2, 99%)	Flavan-3-ols	0.9
<b>O-Demethylation</b>		
<i>Blautia producta</i> (MRG-PMF1, 99%)	Flavonols, flavones, flavanones, isoflavones	<0.1
<i>Eubacterium limosum</i> (strains DSM 20543T and LMG P23546)	Isoflavones, flavanones	<0.1

Sequence identities are taken from [11] or the reference given in the species column. When the bacterium was classified into the genus, but not into a species, we performed a 16S rRNA gene analysis and list the original species/strain designation and the 16S rRNA sequence identity to the identified strain in brackets behind the species name. \*: as per MIDI technique (Hur Lay et al., 2000). <sup>1</sup>: This number is the combined value from two phylogenetic groups (possibly subspecies) given in the GTDB.

## 2.2. Screening the MAGs Amino Acid Sequence Database for Potential Flavonoid-Modifying Enzymes

In the following, the UHGP database was screened with sequences from characterized flavonoid-modifying enzyme as queries. In total, 1,586,499 hits were found with an e-value of less than 1e-60, 1e-25 or 1e-20, filtered for a coverage of at least 75% and a PID of 30. In particular, O-glycosidases and rhamnosidases gave a large number of hits, while O-demethylases and C-glycosidases resulted in a rather low number of hits, because we filtered for the abundance of several genes from a larger gene cluster in the single genomes (Supplementary Excel File, pivot table).

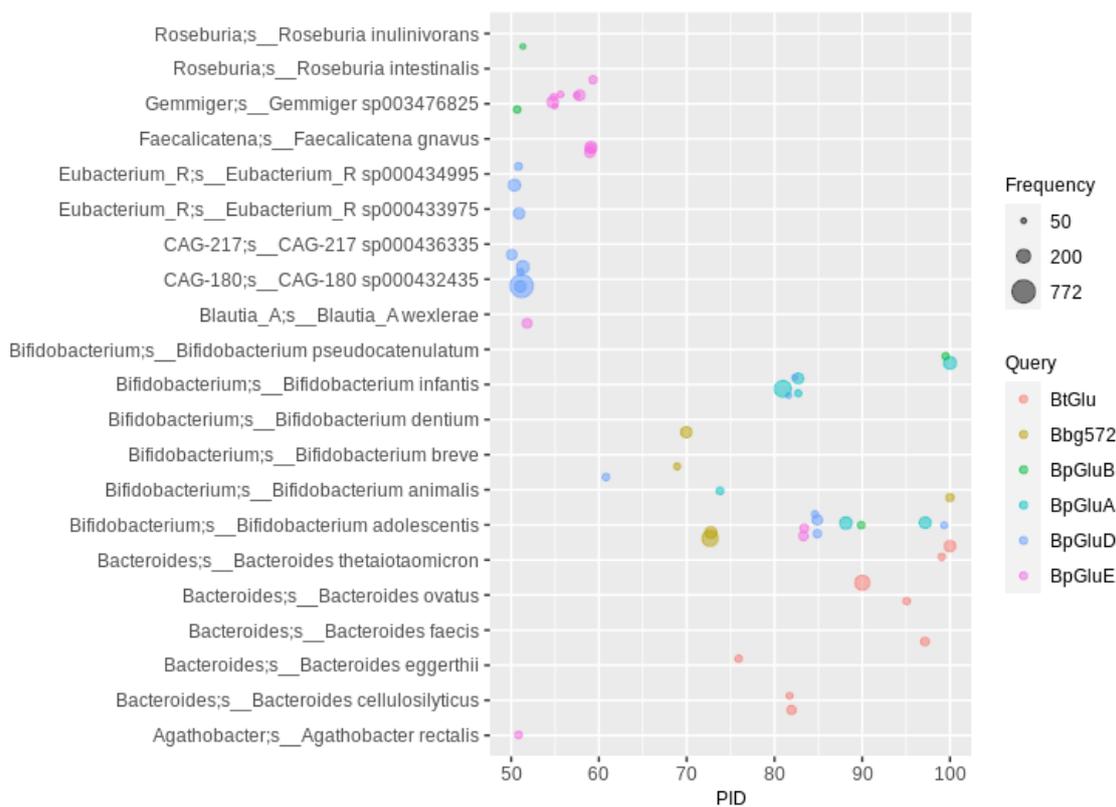
### 2.3. O-Deglycosylation of Flavonoids

The five bifidobacterial flavonoid O-deglycosylating enzyme sequences, which showed a maximum amino acid percent sequence identity (PID) of 35 to each other, resulted in a high number of hits in MAGs classified to six different *Bifidobacterium* sp. (Figure 1) (PID threshold of 50, occurrence of at least 50 for each sequence hit). The highest numbers were observed in MAGs classified as *Bi. adolescentis* (overall, 6633 with a PID > 40), *Bi. infantis* and *Bi. pseudocatelunatum* (overall, about 3000 each, see also Supplementary Excel File). The *Ba. thetaiotaomicron* O-glycosidase query resulted mainly in MAGs from six different *Bacteroides* species, with *Ba. ovatus* and *Ba. thetaiotaomicron* (both about 500 occurrences) as the most numerous (Figure 1, Supplementary Excel File). *Bacteroides* members not already characterized as flavonoid O-deglycosylating species include *Ba. cellulosilyticus* and *Ba. eggerthii*.

Other bacterial genera besides *Bacteroides* or *Bifidobacterium* were detected at a PID of 50 to 60, especially to the three glucosidase sequences of *Bi. pseudocatenulatum* (Figure 1). At PID values of 55 to 60 to BpGluE, MAGs from *Faecalicatena gnavus* and the unclassified *Gemmiger* sp003476825 were highly abundant, with 3170 and 8329 overall occurrences, respectively (Supplementary Excel File). No *Gemmiger* bacterium was associated with the deglycosylation of flavonoids up to now, but the high PID to BpGluE and the highest overall abundance of O-glycosidase hits make *Gemmiger* sp003476825 (one of the most prevalent bacteria in the human gut [61]) a potentially highly relevant species for flavonoid deglycosylation. Two other not-taxonomically-described species with a high number of O-glycosidase hits and a PID of more than 50 were CAG-180 sp000432435 (Clostridiales, family *Acutalibacteriaceae* about 2000 occurrences) and CAG-217 sp000436335 of the same family with about 1000 occurrences. All of these show a PID of about 50 to *Bi. pseudocatenulatum* GluD. The high abundance and similarity to BpGluD indicates a potentially prominent role of these *Acutalibacteriaceae* members in deglycosylation of flavonoids. Furthermore, *Blautia wexlerae* and two taxonomically undescribed *Eubacterium* spp. with a high PID to BpGluD were abundant as well. More distantly related glucosidases (PID between 40 and 50 to the queries) include those of *Faecalicatena gnavus*, several *Clostridium* and *Ruminococcus* species and *Gemmiger* sequences more closely related to the *Bifidobacterium breve* enzyme and a second *Blautia* enzyme (Supplementary Figure S1). Since only isoflavones were tested as substrates for the bifidobacterial enzymes, conclusions about the substrate specificity of the homologs detected here cannot be given. Usually, flavonones and flavonols are also deglycosylated by *Bifidobacterium* spp. [11].

To unravel the phylogenetic relationships between the observed enzymes and to see whether the novel sequences can be classified into distinct phylogenetic classes with potential new catalytic properties, a phylogenetic tree was created. The largest part of the *Bl. wexlerae* sequences form a large distinct clade that also includes sequences from *Faecalicatena gnavus*, while the *Gemmiger* glucosidases form another clade, also distinct from all clades harboring characterized flavonoid O-glycosidases (Supplementary Figure S2). The latter forms one phylogenetic clade together with glucosidases of four *Acutalibacteriaceae* species. Only four of the sequenced enzymes (BpGluA, BpGluB, BpGluD) and the *Bacteroides* sequence form larger clades, with the *Bl. wexlerae*, the CAG-180 and the *F. prausnitzii* clades being much larger, suggesting that the O-deglycosylating potential of the human gut microbiota is largely uncharacterized and awaits biochemical investigation.

A second class of O-glycosidases was observed in *Catenibacillus scindens*, where a cluster of two genes, *dfgCD*, was characterized to encode an enzyme that 7-O-deglycosylated isoflavones and flavones [17]. *DfgCD* homologs are found scarcely in human gut MAGs, mainly in *Hungatella hathewayi*, *Faecalicatena* species, and *Eisenbergiella* sp900066775, all with approximately 100 occurrences (Supplementary Figure S3).

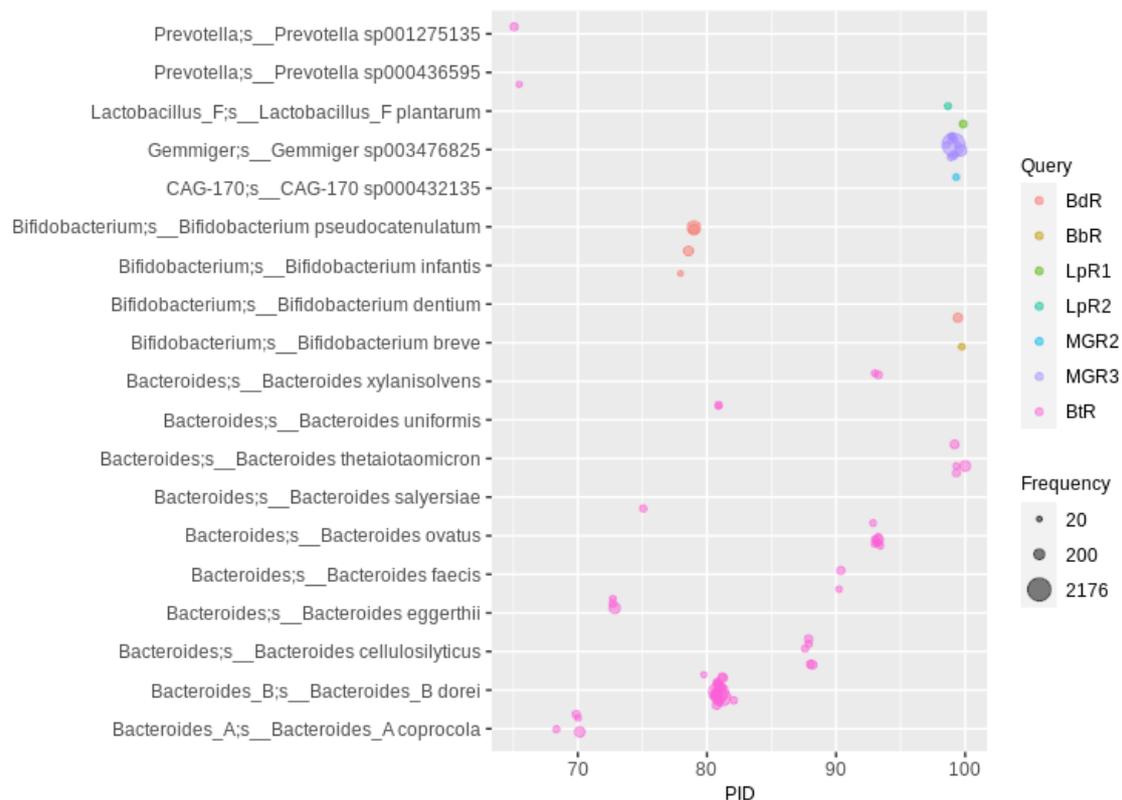


**Figure 1.** Flavonoid O-glycosidase homologs in human gut MAGs. The PID (percent amino acid sequence identity) threshold to the queries (see color code) was set to 50 (For abbreviations and details, see Table 1). Hits were filtered for at least 50 occurrences, so that each bubble represents a number of redundant sequences ranging from 50 to 772.

#### 2.4. Derhamnosylation of Flavonoids

The BLAST search with queries from six rhamnosidases of *Lactobacillus*, *Bifidobacterium* and *Bacteroides* species resulted mainly in hits with MAGs classified as *Bifidobacterium* and *Bacteroides*. The low number of hits to *Lactobacillus* MAGs (only *L. plantarum* was detected with a PID > 65 and Freq > 20) was expected, given the low abundance of *Lactobacillus* species in the gut MAGs metadata (Table 2). One of the most abundant species was *Bacteroides dorei*, not yet described as a flavonoid rhamnoglycoside-hydrolyzing species (Figure 2), whereas the species described to de-rhamnosylate, e.g., rutin, such as *Bi. pseudocatenulatum*, *Bi. dentium*, *Bi. breve*, and *Ba. thetaiotaomicron*, were of comparably lower numbers. The highest number of MAGs in this analysis was observed when using the metagenome-derived rhamnosidase 3 sequence as a query. These MAGs could be assigned to *Gemmiger* sp003476825, which also contains a putative flavonoid O-glucosidase (see above, no sequence similarities to the rhamnosidase described here). With a lower abundance (about 50 appearances), metagenome-derived rhamnosidase 2 orthologs were detected in MAGs of the genus CAG-170 (family *Oscillospiraceae*). Orthologs (PID of about 98) of the metagenome-derived rhamnosidase 1 were detected in only five MAGs from the family *Monoglobaceae*, which could not yet be assigned to a genus and differ from the only described member of this genus, the pectin-degrading intestinal *Monoglobus pectinolyticus* [65,66], which does not encode the corresponding rhamnosidase. Below a 65 PID cutoff, mainly species from the genera *Parabacteroides* and *Bacteroides* and the species *Alistipes onderdonkii* (approximately 45 to 50 PID to MGR2) were found to be highly abundant (Supplementary Figure S4). Several distinct rhamnosidase families can be distinguished phylogenetically, of which the largest cluster does not include characterized ones and contains mainly sequence hits from the *Bacteroides* and *Parabacteroides* genomes. Interestingly, most characterized rhamnosidases do not belong to phylogenetic nodes that mainly represent the corresponding species: *B. dentium* is assigned to a cluster containing mostly *B.*

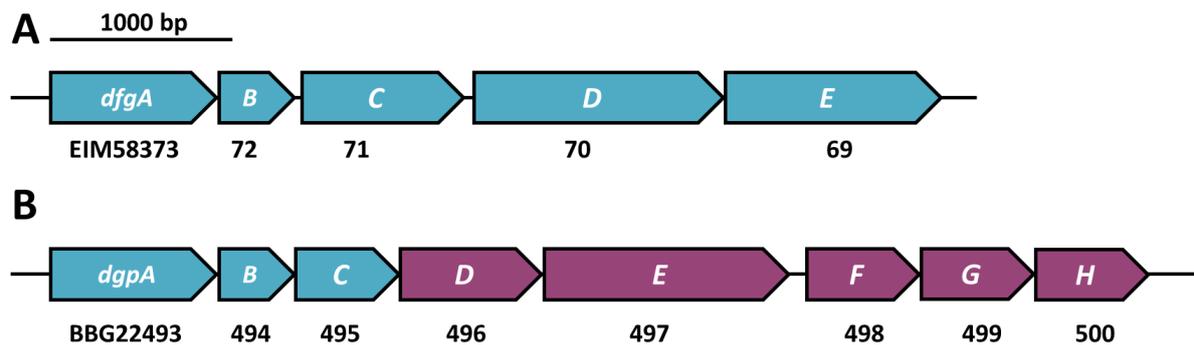
*pseudocatenulatum* genomes, *L. acidophilus* belongs to a large node containing mostly *Fusicatenibacter saccharivorans*, the node including the characterized *Bifidobacterium breve* rhamnosidase is dominated by *Parabacteroides merdae*, while the LpR2 mainly produced hits from *Bacteroides dorei* and *Parabacteroides species* (Supplementary Figure S5). The *B. thetaiotaomicron* rhamnosidase and LpR1 cluster together in a large node dominated by *Bacteroides dorei*. These results show that the characterized isolate species often do not represent the species that are mainly responsible for the de-rhamnosylating activity in the human gut.



**Figure 2.** Rhamnosidase homologs in human gut MAGs. A PID threshold of 65 was chosen for the queries shown in the color code (for abbreviations and details, see Table 1; hits with a lower PID are shown in Supplementary Figure S4). Hits were filtered for at least 20 occurrences, so that each bubble represents a number of redundant sequences ranging from 20 to 2176.

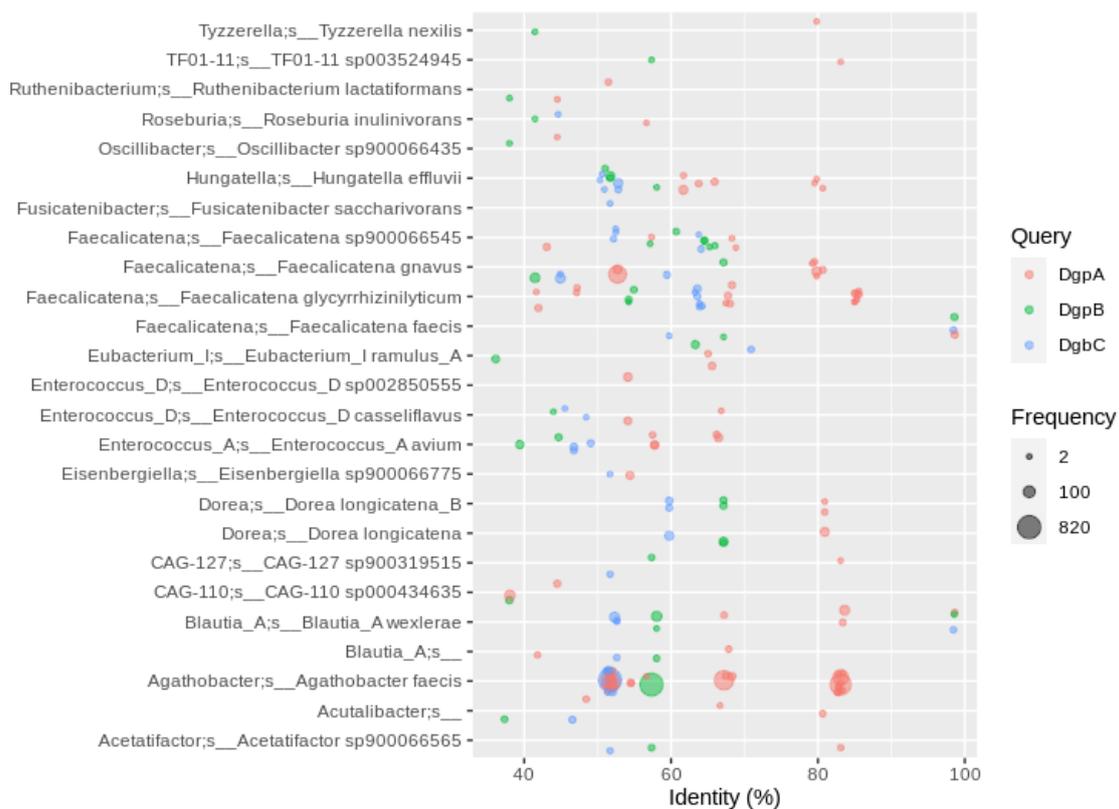
### 2.5. C-Deglycosylation of Flavonoids

The current gene clusters reported to play a role in C-deglycosylation are depicted in Figure 3. Of the *Eubacterium cellulosolvens* *dfgABCDE* gene cluster (Figure 3A, [17]), all sequences were considered as queries in this study because the question as to which gene products are involved in catalysis has not been completely resolved. Of the *dgp* gene cluster of strain PUE (Figure 3B), we used only the first three genes (*dgpABC*) as queries, since only these genes were shown to catalyze reactions in the C-deglycosylation pathway [25,26].



**Figure 3.** Two characterized C-deglycosylation gene clusters. (A) *Eubacterium cellulosolvens* with five genes involved in C-deglycosylation [17] and (B) strain PUE [25,26], with *dgpA* encoding an oxidoreductase and *dgpBC* as oxo-puerarin-deglycosylating enzymes, of which all three are involved in C-deglycosylation. Genes encoding deglycosylating enzymes are shown in turquoise, accessory genes in violet. The full GenBank accession number is given for the first gene and only variable digits are given for the downstream genes.

To respond to the involvement of more than one enzyme in this reaction, we filtered for genomes encoding all three catalytic subunits *dgpABC* or at least three of *dfgABCDE*. The latter resulted in only a few similar hits with a PID higher than 40 (less than 100, Supplementary Figure S6), most of them in the *Faecalicatena* genus (Supplementary Figure S6). Since *dfgD* was not found in the majority of MAGs containing *dfgABCE*, it is likely that DfgD is not involved in C-deglycosylation. The effect of a lack of *dfgD* in the *dfgABCDE* gene cluster in C-deglycosylation was not specifically investigated [17]. A higher number of hits to DfgA (more than 300) with a PID lower than 40 was observed in *Faecalicatena* species, but these were single genes that were not part of a *dfgABCDE* cluster (Supplementary Figure S6). The *dgpABC* cluster was much more prominent in the human gut MAGs than the *dfgABCE* cluster, with most hits in *Agathobacter faecis* (formerly *Roseburia faecis*, PID of 50 to 85%) plus several dozen in *Blautia*, *Dorea*, *Enterococcus* and *Faecalicatena* MAGs (Figure 4). However, only a low percentage of *A. faecis* MAGs carried the *dgpABC* cluster, i.e., 804 of the 2819 MAGs assigned to that species (~30%), so the occurrence of *A. faecis* cannot necessarily be linked to the C-deglycosylation of flavonoids. The occurrence of C-deglycosylating genes in *Faecalicatena gnavus* seems even less specific, with only 84 MAGs of 1197 encoding DgpABC (Supplementary Table S1).

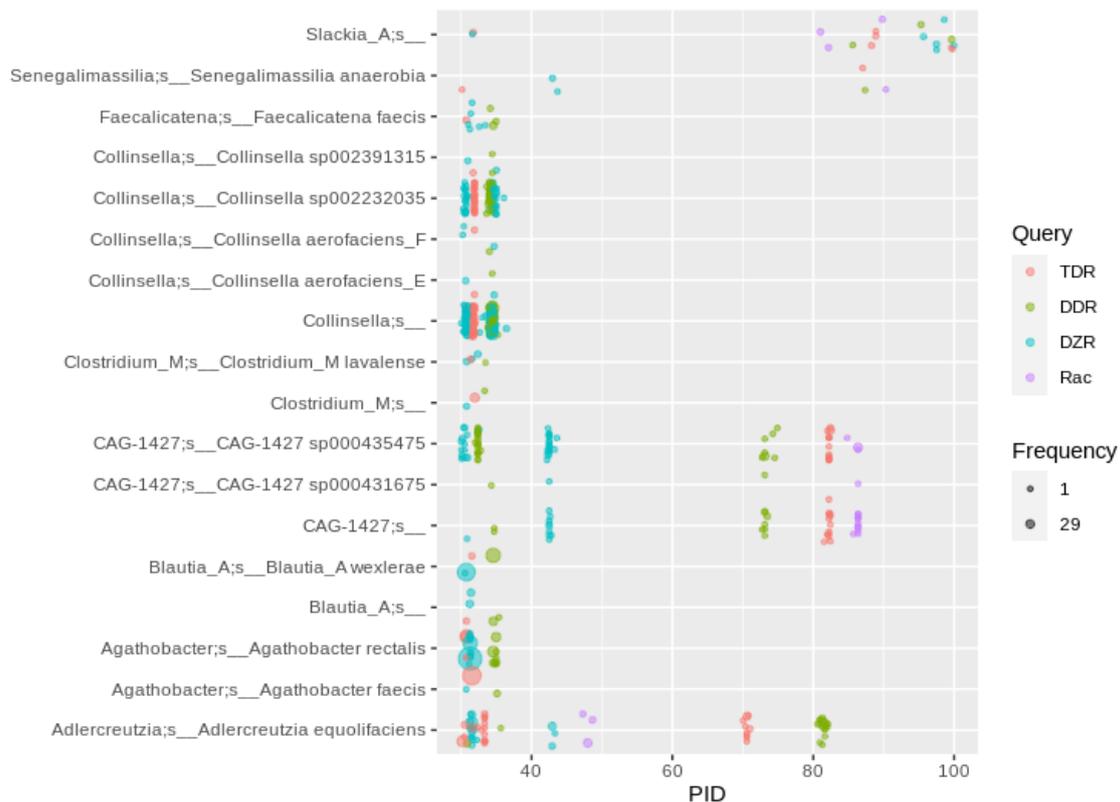


**Figure 4.** Homologs to enzymes involved in flavonoid C-deglycosylation using sequences of strain PUE. Hits were filtered for the co-occurrence of all three genes required for C-deglycosylation (*dgpABC*) in the same MAG. Hits were filtered for at least 2 occurrences, so that each bubble represents a number of redundant sequences ranging from 2 to 820.

## 2.6. Daidzein-to-Equol Conversion

The three enzymes responsible for daidzein-to-equol conversion (DZR, DDR and TDR) plus the optional dihydrodaidzein racemase encoded by *Lactococcus garviae* [30] are encoded in a single cluster, which is very similar in *L. garviae*, *A. equolifaciens* and *S. isoflavoniconvertens*. Due to the high sequence similarities of the enzymes from the three species, we used only the *S. isoflavoniconvertens* sequences (plus the *L. garviae* racemase sequence) as queries. Up- and downstream genes were not considered in this study, since they were suggested to play accessory roles [29]. Hits in *Slackia*, *Lactococcus* or *Eggerthella* were scarce, which contradicts reports in the literature on a widely distributed daidzein-to-equol conversion by these species in the human gut. Four hits of *Slackia* MAGs containing the daidzein-to-equol gene cluster were assigned to the same species represented by GUT\_GENOME145587. A comparison of this representative genome to the available *S. isoflavoniconvertens* DSM22006 genome (GenBank assembly number GCA\_003725955.1) using ANI calculator [67] resulted in an ANI value of 95%. Therefore, the classification of these MAGs to *S. isoflavoniconvertens* is justified. Since the other 261 MAGs representing this species do not encode the daidzein-to-equol pathway enzymes, *S. isoflavoniconvertens* appears not to be a typical equol producer in the human gut. Therefore, conclusions of equol production in an individual should not be based on 16S rRNA analyses, which obviously cannot resolve the daidzein-to-equol transforming potential of *Slackia* spp. Another bacterial species described to transform daidzein to equol, *Adlercreutzia equolifaciens*, was more abundant, with 33 occurrences of the *ddr* gene, but also here, the percentage of MAGs carrying the daidzein-to-equol gene cluster was relatively low with less than ten percent (a total of 271 *A. equolifaciens* MAGs). This observation is strengthened by a report of a human gut *A. equolifaciens* isolate that lacks the cluster and is not able to transform daidzein to equol (Valezquez et al., 2020). Another species whose MAGs frequently carried daidzein-to-equol cluster genes was CAG-1427 sp000435475 (*Eggerthellaceae*, Figure

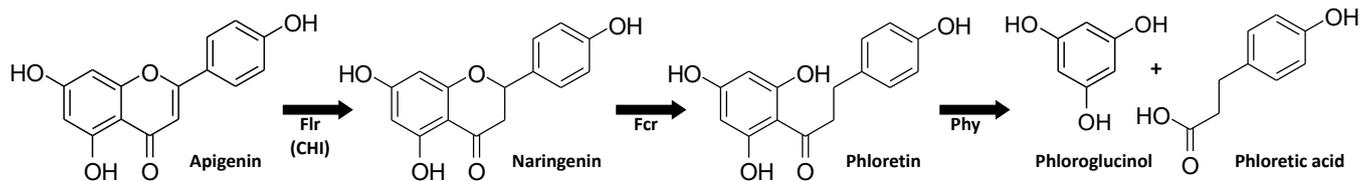
5), of which nearly 50% (a number of 35) of the MAGs (Supplementary Table S1) carry the daidzein-to-equol gene cluster. Therefore, the occurrence of these bacterial species cannot give a reliable conclusion on the main equol producers in the human gut. Specifically designed primers in a PCR-based method revealed that three of 17 arbitrarily selected individuals (of which nine were “equol-producers”) carried equol-producing *Slackia*, and another four carried *Adlercreutzia* [52], which together with the present study hints towards an at least equal or higher importance of *A. equolifaciens* compared to *S. equolifaciens* in equol production. However, the higher percentage of equol-producing individuals is in contrast to the low number of MAGs from potential equol-producing bacteria in the present study. This might be explained by a low abundance of equol-producers in most individuals, preventing an efficient assembly of MAGs from metagenomes, which of course is much less sensitive compared to sensitive HPLC-based equol detection or the highly specific PCR-based detection of the corresponding genes. Since two of the equol-producing individuals of the mentioned study did not result in a PCR product with the employed primer pairs, undiscovered equol-forming bacteria exist in the human gut. A high number of hits with a PID of 30 to 40 was identified in *Agathobacter rectalis* (previously *Eubacterium rectale*) and several, mostly uncharacterized, *Collinsella* species (Figure 5). The occurrence of this latter genus in human gut microbiomes correlated with the ability to transform daidzein to equol in several studies [68–70]. However, since the PID to the characterized daidzein-transforming enzymes is rather low, with values of only about 30 to 35, and the corresponding genes are not clustered, but spread out in the genome, this assumption warrants further research. A relatively high number of genomes showing only hits to the daidzein reductase sequence might be responsible for the large number of phenotypes transforming daidzein only to O-Desmethylangolensin. In particular, bacteria assigned to *Erysipelatoclostridium* and *Clostridium* M genera carry putative *dzt* genes (Supplementary Figure S7, Supplementary Excel File).



**Figure 5.** Homologs to enzymes involved in daidzein-to-equol transformation in human gut MAGs. Hits were filtered for the co-occurrence of all three genes required for daidzein-to-equol conversion (*dzt*, *ddr*, *tdr*) in a single MAG. For a plot showing the hits to individually occurring genes (mainly *dzt*), see Supplementary Figure S7.

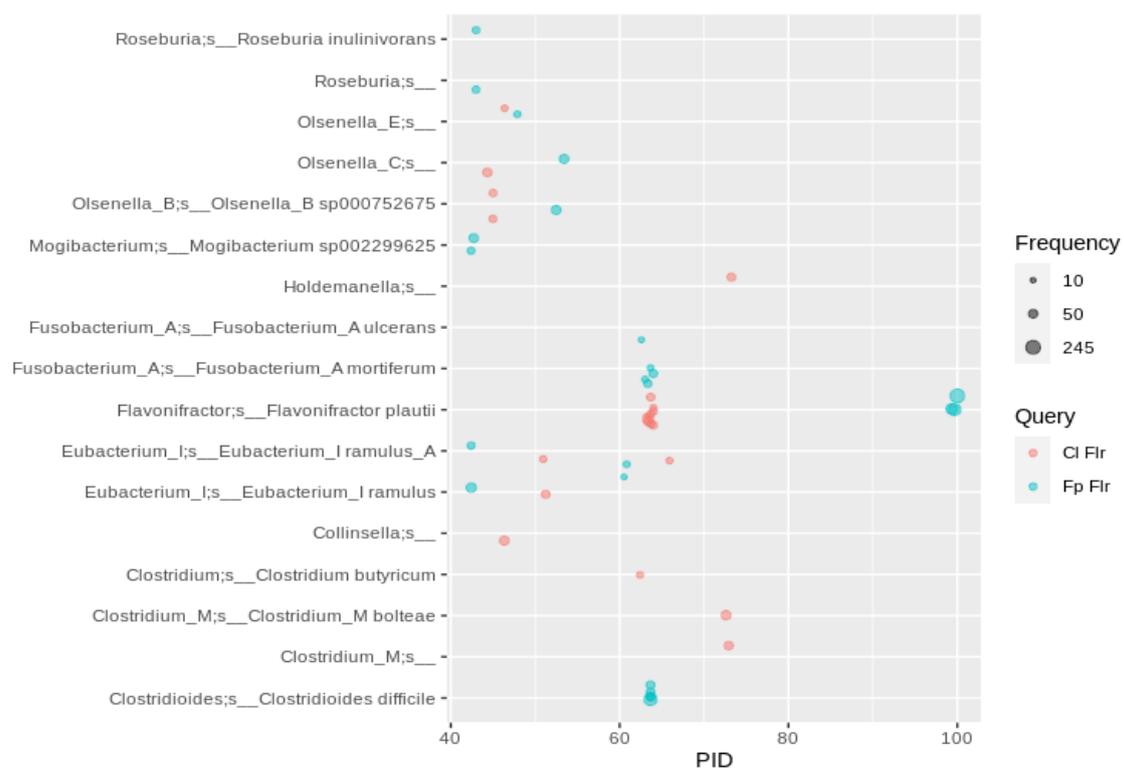
### 2.7. Flavone/Flavonol Reduction, Flavanone/Flavanonol Ring Cleavage, Chalcone Isomerization and Phloretin Hydrolysis

Since these four reactions, the reduction of flavons and flavonols to flavanones and flavanonols, respectively (catalyzed by Flr), the C-ring cleavage of these products (via Fcr), isomerization of chalcones and auronols (via CHI), and the hydrolysis of the dihydrochalcone phloretin (via Phy) can be viewed as steps in the same pathway of flavone/flavonol degradation (e.g., of apigenin via naringenin, Figure 6), the results are compiled in one subchapter.



**Figure 6.** Reductive degradation of flavonoids depicted with apigenin as a flavone example.

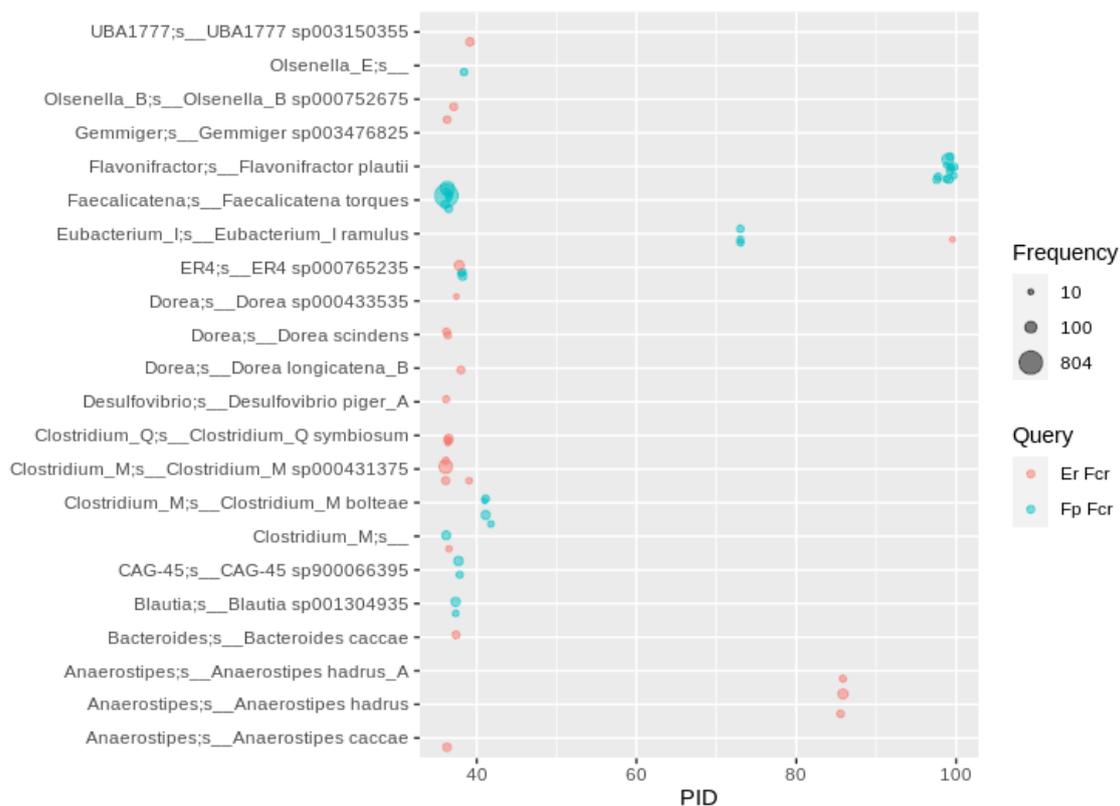
The very recently characterized Flrs from *Flavonifractor plautii* and *Clostridium ljungdahlii* [50] resulted in a large number of hits in *F. plautii* and *Clostridioides difficile* MAGs (the latter interestingly with a higher PID to the sequence of *F. plautii* than to the sequence of *C. ljungdahlii*), and, with a lower PID, in MAGs of the *Olsenella*, *Eubacterium* and *Fusobacterium* species (Figure 7). Most of the *F. plautii* MAGs (and also the strain type DSM 6740) harbor two copies of the *flr* gene, with one of them more closely related (PID of ~ 60) to that of *Clostridium ljungdahlii* than to the “main” Flr of *F. plautii*. This second putative Flr of *F. plautii* did not show any apigenin-reducing activity [50], but might reduce other flavonoid classes, which were not tested in the corresponding study.



**Figure 7.** Flr sequence hits of human gut MAGs. A PID threshold of 40 was chosen to the queries shown in the color code (for abbreviations and details see Table 1). Hits were filtered for at least 10 occurrences, so that each bubble represents a number of redundant sequences ranging from 10 to 245.

The BLAST analysis with the characterized flavanone- and flavanone-cleaving reductase (Fcr) from *Eubacterium ramulus* resulted in hits identified in *F. plautii* MAGs (523 overall occurrences, Supplementary Excel File), while *E. ramulus* MAGs revealed a lower number of 188 occurrences. Similar to the situation with Flr, several *E. ramulus* genomes contained two *fcr* genes, one of which had a higher PID to the *F. plautii* gene, which might reflect a different substrate spectrum of the encoded Fcrs. Besides *F. plautii* and *E. ramulus*, MAGs assigned to *Anaerostipes hadrus* frequently (138 MAGs) encoded an Fcr-like enzyme with approximately 85% identity to Fcr from *E. ramulus*. Besides these three main species, a number of MAGs from *Clostridium boltae* encoded a putative Fcr with a PID of approximately 42 to the Fcr from *F. plautii*. Most hits to Fcr (several hundred up to nearly 1000) showed around 35 PID (Figure 8) and frequently occurred, in particular, in *Faecalicatena* spp. Whether these putative enzymes have a similar function and whether *Faecalicatena* and other species contribute to flavonoid degradation warrants further investigation.

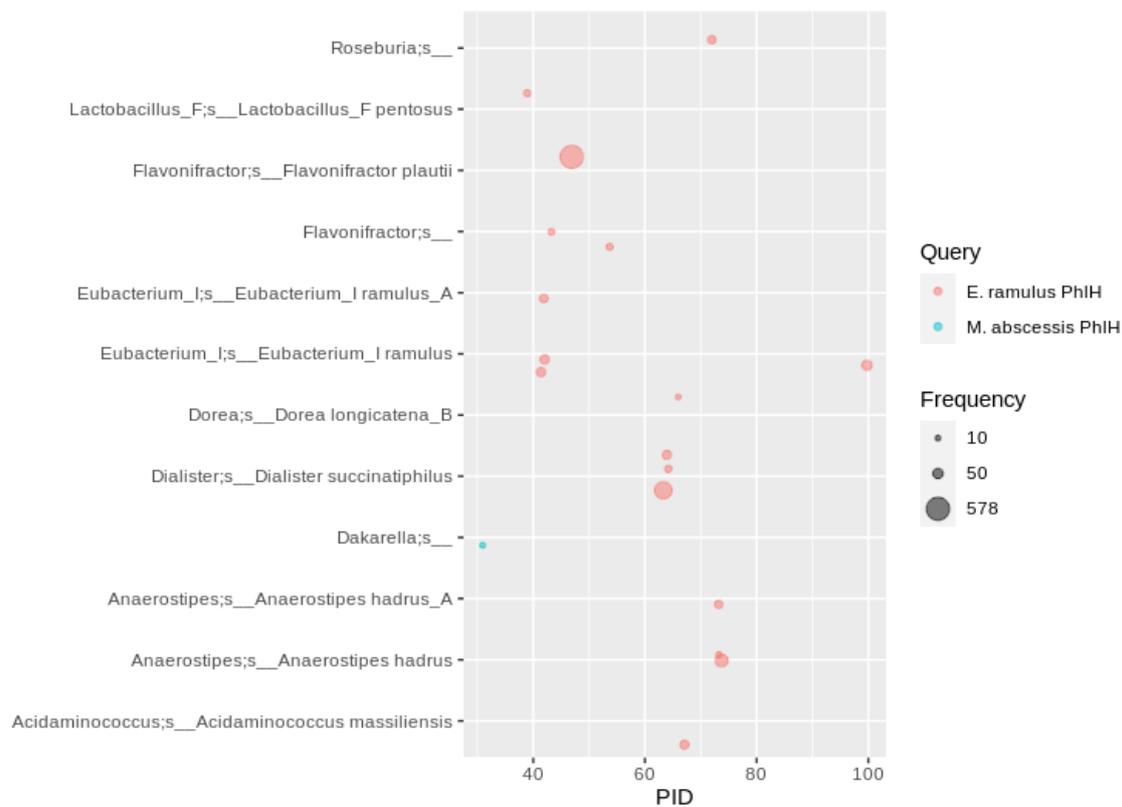
Genes similar to that encoding CHI (Supplementary Figure S8) were highly abundant in MAGs assigned to *Flavonifractor plautii* (593 overall occurrences) and *Clostridoides difficile* (375) and, to a lesser extent, were present in *Eubacterium ramulus* (137), in which it was initially described. Below a PID of 40, only MAGs from *Fusobacterium mortiferum* were detected in a higher number (ca. 100).



**Figure 8.** Fcr-like enzymes in human gut MAGs. Hits were filtered for at least 10 occurrences, so that each bubble represents a number of redundant sequences ranging from 10 to 804.

The BLAST search using the phloretin hydrolase (Phy) sequence of *E. ramulus* resulted in several hits with a PID of more than 90, assigned to 159 *E. ramulus* MAGs (Figure 9). The most abundant Phy homologs with a PID of 46 (579 occurrences, Supplementary Excel file) were found in MAGs assigned to *Flavonifractor plautii*, which was described to hydrolyze phloretin [71]. Genes of two species never reported to transform flavonoids, *Dialister succinatiphilus* (322 occurrences overall) and *Anaerostipes hadrus* (119 occurrences), were abundant with a PID of approximately 60 and 75 to the Phy query. Opposed to the

Phy of *E. ramulus*, the BLAST search with the Phy of *M. abscessis* resulted in only few hits below 40 PID in a *Dakarella* species.



**Figure 9.** Distribution of Phy homologs in human gut MAGs. Depicted are MAGs with at least five identical amino acid sequences. Hits were filtered for at least ten occurrences, so that each bubble represents a number of redundant sequences ranging from 10 to 578.

The co-occurrence of hits to all enzyme sequences (Flr, Fcr, CHI and Phy) was observed in most of the *F. plautii* and *E. ramulus* MAGs, emphasizing the co-functionality of these enzymes in a single pathway of flavone degradation. *F. plautii* was by far the most abundant species in this analysis, pointing toward its major role in flavone degradation in the human gut. Of the 610 MAGs assigned to *F. plautii*, most (509 to 610) contained at least one of the genes. In contrast, *A. hadrus*, containing only Fcr and Phy, showed only a minor amount of MAGs carrying flavonoid-degrading enzymes (less than 5% of the overall MAGs).

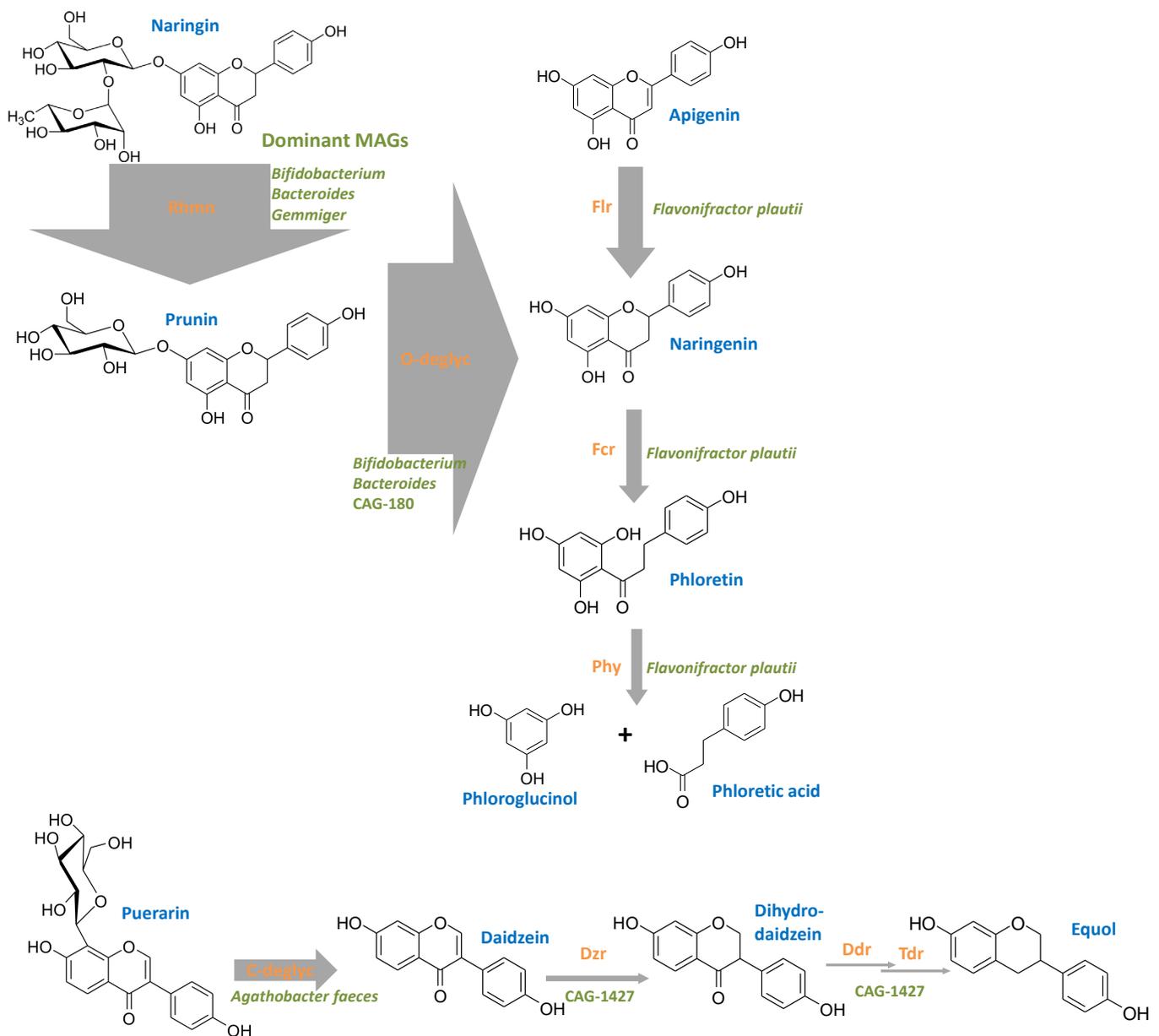
### 2.8. O-Demethylation of Flavonoids

Similar to C-deglycosylation, O-demethylation requires several proteins (MT1 and MT2, CP and AE) acting together in a complex reaction pathway. Since no specific flavonoid O-demethylation system has been characterized, the four enzyme sequences of the O-demethylase system from *Eubacterium limosum*, which acts on the polyphenolic lignan secoisolariciresinol [32], were used as queries. Above 70 PID, only a few dozen hits from the *Eubacterium* genus, notably from the species *E. callenderi* and to a lesser extent *E. limosum* (characterized for flavonoid O-demethylation), were observed (Supplementary Figure S9). The highest number of similar O-demethylase sequences was detected in MAGs of *Intestinibacter bartlettii* (PIDs of 60 for MT1, ca. 40 to 50 for MT2 and CP; ca. 33 for AE). The high PID of MT1, which defines the substrate specificity of O-demethylation [34], suggests that *I. bartlettii* (never tested for O-demethylation of flavonoids) has a similar substrate spectrum for O-demethylation as *E. limosum*. The MTI gene is part of a complete O-demethylase gene cluster in *I. bartlettii* (Supplementary Figure S10), including two CP

genes and *cobW*, which encodes a cobalt-binding protein. The high number of MAGs with hits to the CP of the *O*-demethylase operon sequences (Supplementary Excel File) can be explained by the high abundance of proteins homologous to CP due to the common B<sub>12</sub>-binding and Fe-S cluster domains. In fact, only a low number of MAGs were found to contain the complete operon with a PID of above 40, with *I. bartlettii* MAGs being the most numerous (69 occurrences, Supplementary Excel File).

### 3. Conclusions

The in silico analysis of potential flavonoid-modifying enzymes of human gut bacteria presented here is a first step to uncover the complete picture of flavonoid transformation in the human intestine. The findings constitute a basis for targeted testing of the detected potential novel enzymes and their bacterial sources. Figure 10 summarizes some representative flavonoid-transformations, the prevalence of MAGs encoding the corresponding flavonoid-transforming enzymes and the main bacterial species involved in these reactions. In addition to the already characterized gut bacteria involved in flavonoid transformation, potential species and genera so far not associated with the modification of these polyphenols in the human gut were identified. Of these novel flavonoid-modifying candidates, several species are as-yet not isolated. These include, for example, *Gemmiger* sp003476825, which is potentially involved in *O*-deglycosylation and derhamnosylation, or the CAG-1427 genus, which could be involved in *O*-deglycosylation. Heterologous expression of the potentially involved enzymes could reveal their flavonoid deglycosylation potential. Some other bacteria were already known to play a crucial metabolic role in the human gut (e.g., production of short-chain fatty acids or polysaccharide degradation), but their ability to convert flavonoids may have been overlooked. An example is *Anaerostipes hadrus*, so far mainly known for butyrate production [72]. An (albeit minor) portion of MAGs of this highly abundant human gut bacterium was observed to encode several putative flavonoid-transforming enzymes including a Phy and an Fcr. A major flavonoid-degrading bacterium seems to be *Flavonifractor plautii*. Most of its MAGs contain the Flr, Fcr, Phy and CHI genes, and it is relatively abundant. Thus, *F. plautii* might serve as a good indicator of the flavonoid degradation potential in 16S rRNA gene-based gut microbiota analyses. On the other hand, our quantitative results on flavonoid-modifying bacteria imply that several key flavonoid modification processes in the human gut, which might have been so far assigned to bacteria, are neither representative nor the main flavonoid-transforming bacteria in this ecosystem. For example, equol formation from daidzein appears to mainly be catalyzed in the human intestine by *Eggerthella*, and it may be catalyzed by *Collinsella* species more frequently than by the thoroughly studied *Slackia isoflavoniconvertens* and *Adlercreutzia equolifaciens*.



**Figure 10.** Overview of the flavonoid conversion pathways in the human gut based on the MAGs study. Thickness of arrows reflects the abundance of observed hits to the sequences of the characterized flavonoid-transforming enzymes. Color code: blue, flavonoid trivial names; orange, flavonoid-converting enzymes; green, most abundant bacterial species carrying out the corresponding reaction in this analysis. For abbreviations, see text.

The observation of the wide array of uncharacterized, putative flavonoid-transforming enzymes in human gut bacteria presented here might be a valuable source for identification of novel enzymes of biotechnological interest, especially for the production of unusual or difficult-to-produce flavonoids. Several bacterial species seem to be specialized in metabolizing flavonoids and possibly other polyphenols as well. For example, nearly all *Flavonifractor plautii* MAGs harbor the three enzymes Phy, Fcr and CHI, and therefore seem to be the main flavanone and flavanone-degrading bacteria in the human gut. While further research on the effect of this degradation on human health is necessary, it might be one of the reasons why the presence of *F. plautii* was associated with the development of colon cancer in an Indian population [73]. Finally, the outcome of the in silico analysis presented here could contribute to the clarification of the impact of gut bacteria on flavonoid-mediated health effects in humans.

## 4. Methods

### 4.1. Retrieval of Query Enzyme Sequences, MAGs Data and Set-Up of UHGP BLAST Database

Besides extracting the most recent review on flavonoid modification [11], a literature search on the databases from Web of Science and Pubmed was performed. In addition, a Google scholar search was employed. The keywords used were “flavonoid(s)” AND “modification” OR “transformation” OR “deglycosylation” OR “demethylation” OR “cleavage”. Furthermore, screening in the citation network of the retrieved publications was performed and manually reviewed before Genbank accession numbers and the corresponding amino acid sequences were retrieved from the NCBI Genbank database (Table 1).

A BLAST-searchable database including all non-redundant amino acid sequences from the UHGP catalog was set up as follows: The UHGP-100 version 1.0 non-redundant protein fasta file (uhgp-100.faa) from 286,997 MAGs assembled from gut metagenomes [61] was downloaded as a packed uhgp-100.tar.gz file from the MGnify ftp server ([http://ftp.ebi.ac.uk/pub/databases/metagenomics/mgnify\\_genomes/human-gut/v1.0/](http://ftp.ebi.ac.uk/pub/databases/metagenomics/mgnify_genomes/human-gut/v1.0/), accessed on 1 July 2020) and subsequently extracted. With BLAST+ (stand-alone, version 2.9.0 [74]) installed on Linux Ubuntu 18.04 and the `makeblastdb` command, a BLAST database was created from the uhgp-100 file. For quantification purposes, the uhgp-100.tsv table, including all of the redundant amino acid sequence IDs (all assigned to a non-redundant “representative” sequence ID corresponding to the identical amino acid sequence in other MAGs), was extracted from the same packed file (uhgp-100.tar.gz). The UHGP metadata, which included MAGs taxonomy assigned via the genome taxonomy database (GTDB, release 89), genome quality values, study and sample information and others, were downloaded from the same mgnify ftp-address (genomes-all\_metadata.tsv file). Please note that there might be differences in the taxonomy, since the current release of the GTDB is no. 202 and the metadata file was not updated.

### 4.2. BLAST Searches, Data Mangling and Quantification

The queries were searched against the UHGP-100 BLAST database using the BLASTP search function and an e-value cutoff of  $1e-25$  for pathways including short sequences (150 to 350 amino acids; daidzein-to-equol, CHI, O-demethylation, Phy, Flr, DfgCD) or  $1e-60$  for longer sequences (>350 amino acids; Fcr, O-glycosidases, rhamnosidases) and  $1e-20$  for the C-deglycosylation enzymes, which contained two sequences below 150 amino acids. The maximal number of target hits was set to 100,000 via the `max_target_seqs` argument. Blast table output format 6 was chosen with the specified columns (the Linux BASH commands are given in the Supplementary txt-file). The output files were concatenated and filtered for a PID of at least 30 and a query coverage value of at least 75%. To avoid redundant hits to more than one query, only the hits with the best bitscore were kept. To retrieve the redundant amino acid sequence IDs for quantification of the hits, the `grep` command was used on the uhgp-100.tsv file with a list of the non-redundant sequence IDs as input. The resulting table was merged with the metadata and subsequently filtered for MAGs only (isolate genomes also included in the uhgp catalog were discarded) and the occurrences of the redundant amino acid sequences were counted; the values accounting for the frequency in a given non-redundant sequence group were placed into a new column (designated Freq). For further filtering and data visualization, R (v 3.6.3) was used via RStudio (v 1.3), employing the package `ggplot2`, v.3.3.0 [75].

### 4.3. MAGs Gene Cluster Analysis

To compare and analyze gene clusters among MAGs or isolate genomes, the corresponding genomes were downloaded from the NCBI database or the UHGP MAGs folder and uploaded to the rapid annotation subsystem technology (RAST) server (with their original annotation preserved if possible). The sequence comparison tool was used to compare up to ten genomes.

#### 4.4. Construction of Phylogenetic Trees

Phylogenetic trees of the rhamnosidase and O-glycosidase sequences were constructed after clustering of the sequences with CD-hit (v 4.8.1 [76]) using standard settings (90% PID clustering threshold) and sequence alignment with ClustalOmega (v 1.2.4, [77]) installed on Linux Ubuntu. The alignment was refined with Clipkit (v.1.1.2, [78]) before the tree was generated with Fasttree (v 2.1.10, standard settings [79]). Tree visualization was performed via iTol (v 6, [80]) and manually refined in Inkscape for Windows (v1.01). The size of the node triangles was drawn to scale using the number of CD-hit clustered sequences as well as the redundant sequence IDs of the uhgp-100.tsv file.

**Supplementary Materials:** The following are available online at [www.mdpi.com/2072-6643/13/8/2688/s1](http://www.mdpi.com/2072-6643/13/8/2688/s1), Figure S1: Hits of human gut MAGs to flavonoid O-deglycosidases; Figure S2: Phylogenetic tree of O-deglycosidase sequences; Figure S3: Hits to sequences of the flavonoid O-deglycosidase DfgCD from *E. ramulus*; Figure S4: Hits with 40 to 65 PID to flavonoid rhamnosidase sequences; Figure S5: Phylogenetic tree of rhamnosidase sequences; Figure S6: Hits to the enzyme query sequences of the C-deglycosylation pathway of *E. cellulosolvens*; Figure S7: Hits to daidzein-to-equol enzyme query sequences; Figure S8: CHI-like enzymes in human gut bacteria; Figure S9: Putative O-demethylase systems in human gut bacteria; Figure S10: O-Demethylase gene cluster of *Intestinibacter bartlettii*; Table S1: Distribution of BLAST hits (PID > 40) to flavonoid-modifying enzymes across all MAGs of species discussed in the text; Table S2: Potential flavonoid-modifying bacteria from a screening of the BRENDA enzyme database performed in a previous study.

**Author Contributions:** Conceptualization, formal analysis, investigation, writing—original draft preparation, data curation, visualization T.G., methodology, writing—review and editing, T.G.; A.B.; R.R.C.C. project administration, funding acquisition, A.B. All authors have read and agreed to the published version of the manuscript.

**Funding:** This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement number 814650 (SynBio4Flav).

**Data Availability Statement:** Data are supplied as supplementary material.

**Conflicts of Interest:** The authors have no conflicting interests to declare.

## References

1. Vargas, F.; Romecin, P.; Garcia-Guillen, A.I.; Wangesteen, R.; Vargas-Tendero, P.; Paredes, M.D.; Atucha, N.M.; Garcia-Estan, J. Flavonoids in Kidney Health and Disease. *Front. Physiol.* **2018**, *9*, 394.
2. Panche, A.N.; Diwan, A.D.; Chandra, S.R. Flavonoids: An overview. *J. Nutr. Sci.* **2016**, *5*, e47.
3. Pei, R.; Liu, X.; Bolling, B. Flavonoids and gut health. *Curr. Opin. Biotechnol.* **2020**, *61*, 153–159.
4. Farzaei, M.H.; Singh, A.K.; Kumar, R.; Croley, C.R.; Pandey, A.K.; Coy-Barrera, E.; Kumar Patra, J.; Das, G.; Kerry, R.G.; Annunziata, G.; et al. Targeting Inflammation by Flavonoids: Novel Therapeutic Strategy for Metabolic Disorders. *Int. J. Mol. Sci.* **2019**, *20*, 4957.
5. Caro-Ordieres, T.; Marin-Royo, G.; Opazo-Rios, L.; Jimenez-Castilla, L.; Moreno, J.A.; Gomez-Guerrero, C.; Egido, J. The Coming Age of Flavonoids in the Treatment of Diabetic Complications. *J. Clin. Med.* **2020**, *9*, 346.
6. Bondonno, N.P.; Dalgaard, F.; Kyro, C.; Murray, K.; Bondonno, C.P.; Lewis, J.R.; Croft, K.D.; Gislason, G.; Scalbert, A.; Cassidy, A.; et al. Flavonoid intake is associated with lower mortality in the Danish Diet Cancer and Health Cohort. *Nat. Commun.* **2019**, *10*, 1–10.
7. Bondonno, N.P.; Lewis, J.R.; Blekkenhorst, L.C.; Bondonno, C.P.; Shin, J.H.; Croft, K.D.; Woodman, R.J.; Wong, G.; Lim, W.H.; Gopinath, B.; et al. Association of flavonoids and flavonoid-rich foods with all-cause mortality: The Blue Mountains Eye Study. *Clin. Nutr.* **2020**, *39*, 141–150.
8. Cassidy, A.; Minihane, A.M. The role of metabolism (and the microbiome) in defining the clinical efficacy of dietary flavonoids. *Am. J. Clin. Nutr.* **2017**, *105*, 10–22.
9. Chen, Z.; Zheng, S.; Li, L.; Jiang, H. Metabolism of flavonoids in human: A comprehensive review. *Curr. Drug. Metab.* **2014**, *15*, 48–61.
10. Rodriguez-Mateos, A.; Vauzour, D.; Krueger, C.G.; Shanmuganayagam, D.; Reed, J.; Calani, L.; Mena, P.; Del Rio, D.; Crozier, A. Bioavailability, bioactivity and impact on health of dietary flavonoids and related compounds: An update. *Arch. Toxicol.* **2014**, *88*, 1803–1853.
11. Braune, A.; Blaut, M. Bacterial species involved in the conversion of dietary flavonoids in the human gut. *Gut Microbes* **2016**, *7*, 216–234.

12. Duda-Chodak, A.; Tarko, T.; Satora, P.; Sroka, P. Interaction of dietary compounds, especially polyphenols, with the intestinal microbiota: A review. *Eur. J. Nutr.* **2015**, *54*, 325–341.
13. Rafii, F. The role of colonic bacteria in the metabolism of the natural isoflavone daidzein to equol. *Metabolites* **2015**, *5*, 56–73.
14. Mayo, B.; Vazquez, L.; Florez, A.B. Equol: A Bacterial Metabolite from The Daidzein Isoflavone and Its Presumed Beneficial Health Effects. *Nutrients* **2019**, *11*, 2231.
15. Guadamuro, L.; Florez, A.B.; Alegria, A.; Vazquez, L.; Mayo, B. Characterization of four beta-glucosidases acting on isoflavone-glycosides from *Bifidobacterium pseudocatenulatum* IPLA 36007. *Food Res. Int.* **2017**, *100*, 522–528.
16. Youn, S.Y.; Park, M.S.; Ji, G.E. Identification of the beta-glucosidase gene from *Bifidobacterium animalis* subsp. lactis and its expression in *B. bifidum* BGN4. *J. Microbiol. Biotechnol.* **2012**, *22*, 1714–1723.
17. Byun, D.H.; Choi, H.J.; Lee, H.W.; Jeon, H.Y.; Choung, W.J.; Shim, J.H. Properties and applications of O-glycosidase from *Bacteroides thetaiotaomicron* that specifically hydrolyses isoflavone glycosides. *Int. J. Food Sci. Technol.* **2015**, *50*, 1405–1412.
18. Yadav, V.; Yadav, P.K.; Yadav, S.; Yadav, K.D.S. alpha-L-Rhamnosidase: A review. *Process. Biochem.* **2010**, *45*, 1226–1235.
19. Bang, S.H.; Hyun, Y.J.; Shim, J.; Hong, S.W.; Kim, D.H. Metabolism of rutin and poncirin by human intestinal microbiota and cloning of their metabolizing alpha-L-rhamnosidase from *Bifidobacterium dentium*. *J. Microbiol. Biotechnol.* **2015**, *25*, 18–25.
20. Zhang, R.; Zhang, B.L.; Xie, T.; Li, G.C.; Tuo, Y.; Xiang, Y.T. Biotransformation of rutin to isoquercitrin using recombinant alpha-L-rhamnosidase from *Bifidobacterium breve*. *Biotechnol. Lett.* **2015**, *37*, 1257–1264.
21. Wu, T.; Pei, J.; Ge, L.; Wang, Z.; Ding, G.; Xiao, W.; Zhao, L. Characterization of a alpha-L-rhamnosidase from *Bacteroides thetaiotaomicron* with high catalytic efficiency of epimedin C. *Bioorg. Chem.* **2018**, *81*, 461–467.
22. Beekwilder, J.; Marcozzi, D.; Vecchi, S.; de Vos, R.; Janssen, P.; Francke, C.; van Hylckama Vlieg, J.; Hall, R.D. Characterization of Rhamnosidases from *Lactobacillus plantarum* and *Lactobacillus acidophilus*. *Appl. Environ. Microbiol.* **2009**, *75*, 3447–3454.
23. Avila, M.; Jaquet, M.; Moine, D.; Requena, T.; Pelaez, C.; Arigoni, F.; Jankovic, I. Physiological and biochemical characterization of the two alpha-L-rhamnosidases of *Lactobacillus plantarum* NCC245. *Microbiology* **2009**, *155*, 2739–2749.
24. Li, B.C.; Zhang, T.; Li, Y.Q.; Ding, G.B. Target Discovery of Novel alpha-L-Rhamnosidases from Human Fecal Metagenome and Application for Biotransformation of Natural Flavonoid Glycosides. *Appl. Biochem. Biotechnol.* **2019**, *189*, 1245–1261.
25. Braune, A.; Engst, W.; Blaut, M. Identification and functional expression of genes encoding flavonoid O- and C-glycosidases in intestinal bacteria. *Environ. Microbiol.* **2016**, *18*, 2117–2129.
26. Nakamura, K.; Zhu, S.; Komatsu, K.; Hattori, M.; Iwashima, M. Expression and Characterization of the Human Intestinal Bacterial Enzyme Which Cleaves the C-Glycosidic Bond in 3'-Oxo-puerarin. *Biol. Pharm. Bull.* **2019**, *42*, 417–423.
27. Nakamura, K.; Zhu, S.; Komatsu, K.; Hattori, M.; Iwashima, M. Deglycosylation of the Isoflavone C-Glucoside Puerarin by a Combination of Two Recombinant Bacterial Enzymes and 3-Oxo-Glucose. *Appl. Environ. Microbiol.* **2020**, *86*, e00607-20.
28. Han, J.T.; Zhang, S.P.; Jia, W.J.; Zhang, Z.; Wang, Y.; He, Y.X. Discovery and structural analysis of a phloretin hydrolase from the opportunistic human pathogen *Mycobacterium abscessus*. *FEBS J.* **2019**, *286*, 1959–1971.
29. Schoefer, L.; Braune, A.; Blaut, M. Cloning and expression of a phloretin hydrolase gene from *Eubacterium ramulus* and characterization of the recombinant enzyme. *Appl. Environ. Microbiol.* **2004**, *70*, 6131–6137.
30. Schröder, C.; Matthies, A.; Engst, W.; Blaut, M.; Braune, A. Identification and expression of genes involved in the conversion of daidzein and genistein by the equol-forming bacterium *Slackia isoflavoniconvertens*. *Appl. Environ. Microbiol.* **2013**, *79*, 3494–3502.
31. Shimada, Y.; Takahashi, M.; Miyazawa, N.; Abiru, Y.; Uchiyama, S.; Hishigaki, H. Identification of a novel dihydrodaidzein racemase essential for biosynthesis of equol from daidzein in *Lactococcus* sp. strain 20-92. *Appl. Environ. Microbiol.* **2012**, *78*, 4902–4907.
32. Tsuji, H.; Moriyama, K.; Nomoto, K.; Akaza, H. Identification of an enzyme system for daidzein-to-equol conversion in *Slackia* sp. strain NATTS. *Appl. Environ. Microbiol.* **2012**, *78*, 1228–1236.
33. Chiu, H.H.; Hsieh, Y.C.; Chen, Y.H.; Wang, H.Y.; Lu, C.Y.; Chen, C.J.; Li, Y.K. Three important amino acids control the regioselectivity of flavonoid glucosidation in glycosyltransferase-1 from *Bacillus cereus*. *Appl. Microbiol. Biotechnol.* **2016**, *100*, 8411–8424.
34. Engelmann, T.; Kaufmann, F.; Diekert, G. Isolation and characterization of a veratrol:corrinoid protein methyl transferase from *Acetobacterium dehalogenans*. *Arch. Microbiol.* **2001**, *175*, 376–383.
35. Studenik, S.; Vogel, M.; Diekert, G. Characterization of an O-demethylase of *Desulfotobacterium hafniense* DCB-2. *J. Bacteriol.* **2012**, *194*, 3317–3326.
36. Paraiso, I.L.; Plagmann, L.S.; Yang, L.; Zielke, R.; Gombart, A.F.; Maier, C.S.; Sikora, A.E.; Blakemore, P.R.; Stevens, J.F. Reductive Metabolism of Xanthohumol and 8-Prenylnaringenin by the Intestinal Bacterium *Eubacterium ramulus*. *Mol. Nutr. Food Res.* **2019**, *63*, e1800923.
37. Possemiers, S.; Rabot, S.; Espin, J.C.; Bruneau, A.; Philippe, C.; Gonzalez-Sarrias, A.; Heyerick, A.; Tomas-Barberan, F.A.; De Keukeleire, D.; Verstraete, W. *Eubacterium limosum* activates isoxanthohumol from hops (*Humulus lupulus* L.) into the potent phytoestrogen 8-prenylnaringenin in vitro and in rat intestine. *J. Nutr.* **2008**, *138*, 1310–1316.
38. Kawada, Y.; Yokoyama, S.; Yanase, E.; Niwa, T.; Suzuki, T. The production of S-equol from daidzein is associated with a cluster of three genes in *Eggerthella* sp. YY7918. *Biosci. Microbiota Food Health* **2016**, *35*, 113–121.
39. Shimada, Y.; Yasuda, S.; Takahashi, M.; Hayashi, T.; Miyazawa, N.; Sato, I.; Abiru, Y.; Uchiyama, S.; Hishigaki, H. Cloning and expression of a novel NADP(H)-dependent daidzein reductase, an enzyme involved in the metabolism of daidzein, from equol-producing *Lactococcus* strain 20-92. *Appl. Environ. Microbiol.* **2010**, *76*, 5892–5901.

40. Iino, C.; Shimoyama, T.; Iino, K.; Yokoyama, Y.; Chinda, D.; Sakuraba, H.; Fukuda, S.; Nakaji, S. Daidzein Intake Is Associated with Equol Producing Status through an Increase in the Intestinal Bacteria Responsible for Equol Production. *Nutrients* **2019**, *11*, 433.
41. Song, K.B.; Atkinson, C.; Frankenfeld, C.L.; Jokela, T.; Wahala, K.; Thomas, W.K.; Lampe, J.W. Prevalence of daidzein-metabolizing phenotypes differs between Caucasian and Korean American women and girls. *J. Nutr.* **2006**, *136*, 1347–1351.
42. Frankenfeld, C.L.; Atkinson, C.; Thomas, W.K.; Gonzalez, A.; Jokela, T.; Wahala, K.; Schwartz, S.M.; Li, S.S.; Lampe, J.W. High concordance of daidzein-metabolizing phenotypes in individuals measured 1 to 3 years apart. *Br. J. Nutr.* **2005**, *94*, 873–876.
43. Frankenfeld, C.L. O-desmethylangolensin: The importance of equol's lesser known cousin to human health. *Adv. Nutr.* **2011**, *2*, 317–324.
44. Schneider, H.; Blaut, M. Anaerobic degradation of flavonoids by *Eubacterium ramulus*. *Arch. Microbiol.* **2000**, *173*, 71–75.
45. Braune, A.; Gutschow, M.; Blaut, M. An NADH-Dependent Reductase from *Eubacterium ramulus* Catalyzes the Stereospecific Heteroring Cleavage of Flavanones and Flavanonols. *Appl. Environ. Microbiol.* **2019**, *85*, e01233–219.
46. Gall, M.; Thomsen, M.; Peters, C.; Pavlidis, I.V.; Jonczyk, P.; Grunert, P.P.; Beutel, S.; Scheper, T.; Gross, E.; Backes, M.; et al.: Enzymatic conversion of flavonoids using bacterial chalcone isomerase and enoate reductase. *Angew. Chem. Int. Ed. Engl.* **2014**, *53*, 1439–1442.
47. Braune, A.; Engst, W.; Elsinghorst, P.W.; Furtmann, N.; Bajorath, J.; Gutschow, M.; Blaut, M. Chalcone Isomerase from *Eubacterium ramulus* Catalyzes the Ring Contraction of Flavanonols. *J. Bacteriol.* **2016**, *198*, 2965–2974.
48. Herles, C.; Braune, A.; Blaut, M. First bacterial chalcone isomerase isolated from *Eubacterium ramulus*. *Arch. Microbiol.* **2004**, *181*, 428–434.
49. Thomsen, M.; Tuukkanen, A.; Dickerhoff, J.; Palm, G.J.; Kratzat, H.; Svergun, D.I.; Weisz, K.; Bornscheuer, U.T.; Hinrichs, W. Structure and catalytic mechanism of the evolutionarily unique bacterial chalcone isomerase. *Acta Crystallogr. D Biol. Crystallogr.* **2015**, *71*, 907–917.
50. Yang, G.; Hong, S.; Yang, P.; Sun, Y.; Wang, Y.; Zhang, P.; Jiang, W.; Gu, Y. Discovery of an ene-reductase for initiating flavone and flavonol catabolism in gut bacteria. *Nat. Commun.* **2021**, *12*, 790.
51. Shimada, Y.; Takahashi, M.; Miyazawa, N.; Ohtani, T.; Abiru, Y.; Uchiyama, S.; Hishigaki, H. Identification of two novel reductases involved in equol biosynthesis in *Lactococcus* strain 20-92. *J. Mol. Microbiol. Biotechnol.* **2011**, *21*, 160–172.
52. Braune, A.; Blaut, M. Evaluation of inter-individual differences in gut bacterial isoflavone bioactivation in humans by PCR-based targeting of genes involved in equol formation. *J. Appl. Microbiol.* **2018**, *124*, 220–231.
53. Qin, J.; Li, R.; Raes, J.; Arumugam, M.; Burgdorf, K.S.; Manichanh, C.; Nielsen, T.; Pons, N.; Levenez, F.; Yamada, T.; et al. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **2010**, *464*, 59–65.
54. Nielsen, H.B.; Almeida, M.; Juncker, A.S.; Rasmussen, S.; Li, J.H.; Sunagawa, S.; Plichta, D.R.; Gautier, L.; Pedersen, A.G.; Le Chatelier, E.; et al.: Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes. *Nat. Biotechnol.* **2014**, *32*, 822–828.
55. Almeida, A.; Mitchell, A.L.; Boland, M.; Forster, S.C.; Gloor, G.B.; Tarkowska, A.; Lawley, T.D.; Finn, R.D. A new genomic blueprint of the human gut microbiota. *Nature* **2019**, *568*, 499–504.
56. Nayfach, S.; Shi, Z.J.; Seshadri, R.; Pollard, K.S.; Kyrpides, N.C. New insights from uncultivated genomes of the global human gut microbiome. *Nature* **2019**, *568*, 505–510.
57. Pasolli, E.; Asnicar, F.; Manara, S.; Zolfo, M.; Karcher, N.; Armanini, F.; Beghini, F.; Manghi, P.; Tett, A.; Ghensi, P.; et al. Extensive Unexplored Human Microbiome Diversity Revealed by Over 150,000 Genomes from Metagenomes Spanning Age, Geography, and Lifestyle. *Cell* **2019**, *176*, 649–662.e620.
58. Hur, H.G.; Lay, J.O., Jr.; Beger, R.D.; Freeman, J.P.; Rafii, F. Isolation of human intestinal bacteria metabolizing the natural isoflavone glycosides daidzin and genistin. *Arch. Microbiol.* **2000**, *174*, 422–428.
59. Mattarelli, P.; Bonaparte, C.; Pot, B.; Biavati, B. Proposal to reclassify the three biotypes of *Bifidobacterium longum* as three subspecies: *Bifidobacterium longum* subsp. *longum* subsp. nov., *Bifidobacterium longum* subsp. *infantis* comb. nov. and *Bifidobacterium longum* subsp. *suis* comb. nov. *Int. J. Syst. Evol. Microbiol.* **2008**, *58*, 767–772.
60. Ivey, K.L.; Chan, A.T.; Izard, J.; Cassidy, A.; Rogers, G.B.; Rimm, E.B. Role of Dietary Flavonoid Compounds in Driving Patterns of Microbial Community Assembly. *mBio* **2019**, *10*, e01205–e01219.
61. Almeida, A.; Nayfach, S.; Boland, M.; Strozzi, F.; Beracochea, M.; Shi, Z.J.; Pollard, K.S.; Sakharova, E.; Parks, D.H.; Hugenholtz, P.; et al. A unified catalog of 204,938 reference genomes from the human gut microbiome. *Nat. Biotechnol.* **2021**, *39*, 105–114.
62. Shimojo, Y.; Ozawa, Y.; Toda, T.; Igami, K.; Shimizu, T. Probiotic *Lactobacillus paracasei* A221 improves the functionality and bioavailability of kaempferol-glucoside in kale by its glucosidase activity. *Sci. Rep.* **2018**, *8*, 9239.
63. Gaya, P.; Peiroten, A.; Alvarez, I.; Medina, M.; Landete, J.M. Production of the bioactive isoflavone O-desmethylangolensin by *Enterococcus faecium* INIA P553 with high efficiency. *J. Funct. Foods* **2018**, *40*, 180–186.
64. Gaya, P.; Peiroten, A.; Medina, M.; Landete, J.M. Isoflavone metabolism by a collection of lactic acid bacteria and bifidobacteria with biotechnological interest. *Int. J. Food Sci. Nutr.* **2016**, *67*, 117–124.
65. Kim, C.C.; Healey, G.R.; Kelly, W.J.; Patchett, M.L.; Jordens, Z.; Tannock, G.W.; Sims, I.M.; Bell, T.J.; Hedderley, D.; Henrissat, B.; et al. Genomic insights from *Monoglobus pectinilyticus*: A pectin-degrading specialist bacterium in the human colon. *ISME J.* **2019**, *13*, 1437–1456.

66. Kim, C.C.; Kelly, W.J.; Patchett, M.L.; Tannock, G.W.; Jordens, Z.; Stoklosinski, H.M.; Taylor, J.W.; Sims, I.M.; Bell, T.J.; Rosendale, D.I. *Monoglobus pectinilyticus* gen. nov., sp. nov., a pectinolytic bacterium isolated from human faeces. *Int. J. Syst. Evol. Microbiol.* **2017**, *67*, 4992–4998.
67. Goris, J.; Konstantinidis, K.T.; Klappenbach, J.A.; Coenye, T.; Vandamme, P.; Tiedje, J.M. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int. J. Syst. Evol. Microbiol.* **2007**, *57*, 81–91.
68. Nakatsu, C.H.; Armstrong, A.; Clavijo, A.P.; Martin, B.R.; Barnes, S.; Weaver, C.M. Fecal bacterial community changes associated with isoflavone metabolites in postmenopausal women after soy bar consumption. *PLoS ONE* **2014**, *9*, e108924.
69. Guadamuro, L.; Dohrmann, A.B.; Tebbe, C.C.; Mayo, B.; Delgado, S. Bacterial communities and metabolic activity of faecal cultures from equol producer and non-producer menopausal women under treatment with soy isoflavones. *BMC Microbiol.* **2017**, *17*, 93.
70. Guadamuro, L.; Azcarate-Peril, M.A.; Tojo, R.; Mayo, B.; Delgado, S. Use of high throughput amplicon sequencing and ethidium monoazide dye to track microbiota changes in an equol-producing menopausal woman receiving a long-term isoflavones treatment. *AIMS Microbiol.* **2019**, *5*, 102–116.
71. Schoefer, L.; Mohan, R.; Schwiertz, A.; Braune, A.; Blaut, M. Anaerobic degradation of flavonoids by *Clostridium orbiscindens*. *Appl. Environ. Microbiol.* **2003**, *69*, 5849–5854.
72. Allen-Vercoe, E.; Daigneault, M.; White, A.; Panaccione, R.; Duncan, S.H.; Flint, H.J.; O’Neal, L.; Lawson, P.A. *Anaerostipes hadrus* comb. nov., a dominant species within the human colonic microbiota; reclassification of *Eubacterium hadrum* Moore et al. 1976. *Anaerobe* **2012**, *18*, 523–529.
73. Gupta, A.; Dhakan, D.B.; Maji, A.; Saxena, R.; Prasoodanan, P.K.V.; Mahajan, S.; Pulikkan, J.; Kurian, J.; Gomez, A.M.; Scaria, J.; et al. Association of *Flavonifractor plautii*, a Flavonoid-Degrading Bacterium, with the Gut Microbiome of Colorectal Cancer Patients in India. *mSystems* **2019**, *4*, e00438-19.
74. Camacho, C.; Coulouris, G.; Avagyan, V.; Ma, N.; Papadopoulos, J.; Bealer, K.; Madden, T.L. BLAST+: Architecture and applications. *BMC Bioinform.* **2009**, *10*, 421.
75. Wickham, H. *Ggplot2: Elegant Graphics for Data Analysis*; Springer-Verlag: New York, NY, USA. Available online: <https://ggplot2.tidyverse.org> (accessed on 1 June 2021).
76. Fu, L.; Niu, B.; Zhu, Z.; Wu, S.; Li, W. CD-HIT: Accelerated for clustering the next-generation sequencing data. *Bioinformatics* **2012**, *28*, 3150–3152.
77. Sievers, F.; Higgins, D.G. Clustal Omega for making accurate alignments of many protein sequences. *Protein. Sci.* **2018**, *27*, 135–145.
78. Steenwyk, J.L.; Buida, T.J., 3rd; Li, Y.; Shen, X.X.; Rokas, A. ClipKIT: A multiple sequence alignment trimming software for accurate phylogenomic inference. *PLoS Biol.* **2020**, *18*, e3001007.
79. Price, M.N.; Dehal, P.S.; Arkin, A.P. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS ONE* **2010**, *5*, e9490.
80. Letunic, I.; Bork, P. Interactive Tree Of Life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **2021**, *49*, 293–296.