Metatranscriptomic analysis indicates prebiotic effect of Isomalto/malto-polysaccharides on human colonic microbiota *in-vitro*

Klaudyna Borewicz[1, §, ^], Bastian Hornung[1,2, §,#,*], Fangjie Gu[3,$], Pieter H. van der Zaal[4,&], Henk A. Schols[3], Peter J. Schaap[2], Hauke Smidt[1]

[1]Laboratory of Microbiology, Wageningen University & Research, Stippeneng 4, 6708 WE Wageningen, The Netherlands

[2]Laboratory of Systems and Synthetic Biology, Wageningen University & Research, Stippeneng 4, 6708 WE, Wageningen, The Netherlands

[3]Laboratory of Food Chemistry, Wageningen University & Research, Bornse Weilanden 9, 6708 WG Wageningen, The Netherlands

[4]Biobased Chemistry and Technology, Wageningen University & Research, Bornse Weilanden 9, 6708 WG Wageningen, The Netherlands

[§]These authors contributed equally to this work

[^] Present address: Mead Johnson, Middenkampweg 2, 6545 CJ Nijmegen, The Netherlands

[#] Present address: CBG-MEB, Graadt van Roggenweg 500, 3531AH Utrecht, Netherlands

[$] Present address : TUMCREATE, 1 CREATE Way, CREATE Tower, #10-02 Singapore 138602, Singapore

[&] Present address : IFF, Willem Einthovenstraat 4, 2342 BH, Oegstgeest, The Netherlands

[*]Corresponding author:

Bastian Hornung

Bastian.hornung@gmx.de

CBG-MEB, Graadt van Roggenweg 500, 3531AH Utrecht, Netherlands

Phone number : +31646806979

Abstract

Isomalto/malto-polysaccharides (IMMPs) are a novel type of soluble dietary fibres with a prebiotic potential promoting growth of beneficial microbes in the gut. However, the mode of action of IMMPs remains unknown. Previous studies on IMMPs showed an increase in total bacteria, especially lactobacilli, and higher production of short chain fatty acids (SCFA) when IMMPs were fed to rats or used during *in vitro* fermentation. Here we investigated with metatranscriptomics how IMMPs with different amounts of α-(1→6) glycosidic linkages affected microbial function during incubation with human faecal inoculum. We showed that active microbial community dynamics during fermentation varied depending on the type of IMMP used and that the observed changes were reflected in the community gene expression profiles. Based on metatranscriptome analysis, members of *Bacteroides*, *Lactobacillus* and *Bifidobacterium* were the predominant degraders of IMMPs, and the increased gene expression in these bacteria correlated with high amounts of α-(1→6) glycosidic linkages. We also noted an increase in relative abundance of these bacteria and an activation of pathways involved in SCFA synthesis. Our findings could provide a baseline for more targeted approaches in designing prebiotics for specific bacteria and to achieve more controlled modulation of microbial activity towards desired health outcomes.

Introduction

The human gut is home to a diverse ecosystem inhabited by bacteria, archaea, viruses and eukaryotes, which play an important role in their host's health and well-being [1-3]. These organisms interact with each other and with the host via a complex network of relationships, and knowing the mechanisms of these interactions and how to influence them might provide a useful tool for refining the function of this ecosystem to promote homeostasis and to strengthen host's immunity against infections [4]. Currently ways to manipulate the composition and function of gut microbiota range from mild measures, such as use of dietary supplements (especially pro- and prebiotics [5]) or implementation of dietary regimes, to more extreme measures, such as the use of antibiotics [6] or faecal transplantations [7]. Prebiotics are complex carbohydrates, often soluble dietary fibres, that cannot be digested by human enzymes but are readily used by the colonic microbiota and provide a health benefit for the host [8]. A range of different prebiotics may stimulate growth and activity of specific microbial groups (e.g. butyrogenic bacteria [9]), leading to the production of different metabolites with health-supporting effects. However, the exact mode of action and the specific impact on microbial interactions needs to be investigated.

Isomalto/malto-polysaccharides (IMMPs) comprise a novel class of soluble dietary fibres with prebiotic potential. IMMPs are synthetized from starch by enzymatic conversion of α-(1→4) glycosidic linkages into α-(1→6) glycosidic linkages by 4,6-α-glucanotransferase (GTFB) from *Lactobacillus reuteri* 121 [10]. The resulting α-(1→6) linkages make IMMPs resistant to digestion by human enzymes in the small intestine. As such, IMMPs can pass into the large intestine where they are fermented by the resident microbes capable of breaking down the α-(1→6) glycosidic linkages. This property of IMMPs makes them interesting as a prebiotic food ingredient. A previous *in vitro* study has reported an increased production of short chain fatty acids (SCFA), especially acetate and propionate, when IMMPs were used as a carbon source for fermentation with human faecal inoculum as the microbial source [10]. Here we investigated the effects of three different IMMPs on microbial composition and function during *in vitro* batch fermentations with faecal inoculum from healthy human adults. Additionally, we used a fourth substrate, an IMMP preparation after removal of remaining digestible starch segments consisting mainly of α-(1→4) glucose moieties, to investigate the direct impact of having mainly α-(1→6) glycosidic linkages available to the gut microbiota.

We showed that specific changes of the microbiota, such as growth of *Bifidobacterium* and *Lactobacillus* could be attributed to the IMMPs, and that these changes were also reflected at the transcriptomic level, i.e. upregulation of specific gene groups, as well as in enzymatic activity and increase in production of SCFA.

Results

We performed two *in vitro* batch fermentation experiments to understand how the IMMPs containing different amounts of α-(1→6) glycosidic linkages were broken down by human faecal microorganisms over time, and how the chemical structure of these compounds affected the functional dynamics of the microbial community during fermentation. Experiment A included fermentation of IMMPs of varying percentage of α-(1→6) glycosidic linkages (27%, IMMP-27; 94%, IMMP-94) at three different time points. This was complemented by experiment B that was performed with IMMP with 96% α-(1→6) linkages (IMMP-96) and IMMP-27 after treatment with α-amylase and amyloglucosidase (IMMP-dig27). Furthermore, in experiment B an additional set of time points was evaluated to provide a more detailed understanding of microbial community dynamics. In both experiments a control blank without any IMMP substrate was included. We then performed metatranscriptome sequencing of all samples and assembled the resulting data into one reference metatranscriptome. Afterwards, machine learning techniques were applied to identify groups of similarly behaving bacteria and to identify consistent dynamic patterns in gene expression.

Quality control and statistics

The metatranscriptomes were sequenced and subjected to a quality control before the data was further analysed (Figure S1). As a result, 320 million reads (89% of the raw reads and 54% of all bases) passed the quality check and were used for contig assembly. In experiment A, the assembly yielded over 140,000 contigs, with more than 200,000 protein coding genes, and contained, on average, 81% of the input reads (range 71% - 85%) per sample. Read counts for experiment B were acquired by mapping to the same assembly obtained from experiment A (Table S1), and showed the same average mapping rate (81%, range 71% - 89%). After mapping, the biological replicates within each experiment showed a spearman correlation of on average 0.86 (range 0.78 - 0.93), indicating good reproducibility within sets of samples from the same treatment group (Table S2). The six samples from time point 0 showed a correlation of at least 0.79 (max 0.86), and all samples taken from batches incubated in the presence of prebiotics showed a correlation of at least 0.7 for time point 24h, and all (except two) showed a correlation of at least 0.6 at time point 48h, indicating similar development over all cultures.

Of the ~200.000 protein coding genes, ~36.000 were full length genes. To ~144.000 protein sequences at least one Interpro domain (excluding “Coils” domain) could be identified. These included more than 24.000 genes with at least one (partial) EC number.

Community structure and expression patterns

Taxonomic classification to at least the superkingdom of Bacteria was assigned to 190,000 of the 200,000 genes obtained from the RNA-assembly. Less than 3,000 genes were assigned to eukaryotes and less than 2,000 to Archaea. Of the bacterial groups, most genes were assigned to the orders Bacteroidales (>67,000), Clostridiales (>40,000), Lactobacillales (27,000) and Enterobacteriales (>14,000). The genus with the highest number of assigned genes was the genus *Bacteroides* (>54.000). Figure 1 shows the relative abundance of transcripts per assigned genus.

To identify bacterial gene expression patterns, we focused on RNA reads for which KEGG Orthology (KO) or EC identifiers could be assigned. The percentage of reads with defined KO or EC ranged from 42% to 83% for different samples. Most of the data with assigned KO or EC identifiers came from 22 bacterial groups, of which 12 could be assigned to a known genus, and only a small number of genes was assigned to minor groups (3%), unclassifiable sequences (3%), and sequences not classifiable beyond the superkingdom Bacteria (3.5%). In the activated inoculum at the start of the incubation (t0), unclassified Enterobacteriaceae were the most active group (Figure 1), probably due to residual oxygen during the activation. However, once the incubation had started, the relative expression of *Bacteroides* increased in all treatment groups. In all samples combined across all treatments and time points, 39% of all expression data came from the genus *Bacteroides* and 27% from unclassified Enterobacteriaceae. Overall, the relative abundance of different bacterial groups based on the metatranscriptome data corresponded to the pattern in the relative abundance of different taxa based on the 16S rRNA gene analysis described previously by Gu *et al.* [11] (Figure 2).

Global and IMMP specific co-occurrence of taxa

It is known that in microbial ecosystems bacterial taxa can occupy different niches and co-exist forming a complex network of co-dependencies. We wanted to assess whether, based on the metatranscriptome data, we could identify bacterial groups which co-occurred in our samples, with particular attention to the effect of specific IMMPs. We performed clustering analysis based on mRNA reads from all samples in our dataset to test for global co-occurrence patterns. We showed that clustering into nine groups was most stable, as it could be reproduced in multiple rounds of clustering. These resulting nine different groups showed different behaviours over all investigated conditions. An overview of organism assignment per cluster, with number of assigned genes and differentially expressed genes is provided in Table S3. One of these clusters was present in all t0 samples but decreased or was absent at all other time points. This cluster consisted mostly of reads assigned to *Ruminococcus* and *Lactococcus* as well as reads that could be largely classified as contamination from the inoculum/preparation (e.g. *Homo*, *Mus*, *Bos*, unclassified Mammalia). The second and third cluster consisted mainly of genera including many genera known to contain many probiotic organism candidates, i.e. *Bifidobacterium*, *Lactobacillus*, and *Enterococcus*, and sequences, which could not be classified beyond a related higher taxon (e.g. unclassified Bifidobacteriaceae, unclassified Lactobacillaceae). These clusters also contained a related phage group (*Myoviridae*, mainly *Lactobacillus* phages), and an unrelated genus (*Fusobacterium*). The identified genera in cluster two and three showed increasing relative transcript abundance in all cultures supplied with IMMP substrates, whereas relative transcript abundance was decreased or undetected in the control cultures. The fourth cluster was predominated by *E. coli* and related higher order classifications (e.g. unclassified Enterobacteriaceae), together with other enterobacteria such as *Enterobacter, Citrobacter* and *Klebsiella*, and the unrelated Gram-positive genus *Eubacterium*. This cluster was mainly present in the samples without IMMP and declined in the samples with IMMP. The fifth cluster was predominated by *Bacteroides* and showed an increase with time in all incubations. This cluster also included *Parabacteroides, Prevotella, Flavobacterium,* and *Desulfosporosinus*. The sixth cluster consisted only of *Clostridium*/unclassified Clostridia, which showed some increase with time in all incubations. The seventh cluster contained *Anaerostipes* and related higher level taxon classifications (unclassified Clostridiales, unclassified Lachnospiraceae) and showed a similar pattern as cluster six. No clear pattern was seen for the eighth cluster consisting of *Corynebacterium, Ethanoligenes, Odoribacter,* and *Sutterella*. Finally, the ninth group consisted of different bacterial genera, some related to non-carbohydrate metabolizing bacteria (*Acidaminococcus, Bilophila, Phascolarctobacterium*), and some known gut symbionts like *Veillonella* and *Megasphaera*. This group was common in samples of incubations without any prebiotics at 48 h and was nearly absent in all the other samples.

Detection of specific gene expression patterns

Besides the co-occurrence of bacterial groups, the specific gene expression patterns within these groups were investigated based on the optimal gene clustering for all bacterial groups using DBSCAN. The clustering with the optimal tau was chosen for all bacterial groups, except for the genus *Enterococcus*, for which a suboptimal tau led to better cluster separation. As a result, the DBSCAN gene clustering analysis revealed the presence of three main patterns in the expression in nearly all observed bacterial groups (Figure S3). These three patterns comprised in all cases at least 80% of all investigated genes, which were not considered noise. The first pattern was present in all incubations and was characterized by genes which were expressed only at t0, and not expressed at any of the later time points. The second pattern was found only in the control group and only at 48 h. The third and most common pattern found in all experimental groups included genes that were not expressed at t0 but showed upregulation at the later time points during incubation. This pattern was characteristic for genes assigned to the genera *Enterococcus* and *Bacteroides*, for which respectively 40% and 99% of differentially expressed genes were increasingly expressed over time in all treatment groups including the control group. *Bifidobacterium/Lactobacillus* and *Clostridium* also showed the same pattern, but only in the groups where IMMPs were present. *Anaerobutyricum hallii* (formerly *Eubacterium hallii*) showed the same gene expression pattern, but only in the group supplemented with IMMP-27 (Figure S3).

Considering the experimental set-up chosen for this study, the expression levels of genes assigned to a specific bacterial group indicate the contribution of this group to the utilization of the specified substrate, or its by-products. The high overall relative expression of genes assigned to bifidobacteria (and unclassified Bifidobacteriaceae), lactobacilli, enterococci, and unclassified Actinobacteria was positively associated with the presence of IMMPs (Figure 1). Conversely, the expression of genes assigned to unclassified

Proteobacteria, *Prevotella*, *Sutterella*, *Acinetobacter*, *Eggerthella*, *Acidaminococcus*, *Streptococcus*, *Phascolarctobacterium*, and *Bilophila* was negatively associated with the presence of IMMPs, as compared to the control group.

General metabolic effects of IMMP

We wanted to further investigate the expression of genes and corresponding bacterial groups associated with the fermentation of different IMMPs. Our analysis of the metabolic clusters revealed that five bacterial groups found in the faecal inoculum, namely *Bifidobacterium/Lactobacillus, Enterococcus, Bacteroides, Clostridium,* and *Anaerobutyricum hallii,* showed a considerable upregulation of general metabolic pathways like glycolysis, nucleic acid or fatty acid biosynthesis, as compared to the gene expression at t0. When we compared metabolic patterns between different bacterial groups, the groups exhibited overall different metabolic patterns. Members of the genus *Bacteroides* showed at first a unique partial upregulation of Vitamin B12 metabolism. An investigation of the cofactor requirements showed that Vitamin B12 in *Bacteroides* is essential for methionine synthase and methylmalonyl-CoA mutase, the latter of which produces methylmalonyl-CoA from succinyl-CoA (Figure 3).

Methylmalonyl-CoA mutase is involved in propionate biosynthesis, and our data showed that the whole pathway for propionate biosynthesis was, in fact, upregulated. The data further showed that many genes coding for proteins involved in iron scavenging were also upregulated (e.g. FecR). One of the genes coding for an enzyme with iron requirements was that encoding succinate dehydrogenase, which converts succinate into fumarate. This function, as well as all others in the TCA cycle, showed upregulation in all samples tested. The genus *Clostridium* also showed an upregulation of genes involved in Vitamin B12 production, but the biosynthesis occurred via glutamate, whereas in the *Bacteroides* group it was produced via succinate. The genes in the pathway for propionate production were overall upregulated (production via acetyl-CoA, not succinyl-CoA), similar to the genes in lactate and butyrate production pathways. The only other enzyme requiring Vitamin B12 in the microbiome was a multimer of propanediol dehydratase or glycerol dehydratase (ambiguous functional assignment), both involved in the breakdown of glycerol/glycerone phosphate to propanol/propionate/1,3-propanediol. However, a full upregulation of either pathway was not observed. The *Bifidobacterium*/*Lactobacillus* group and the *Enterococcus* group showed upregulation of genes related to production of lactate from pyruvate, and the *Clostridium* group also showed upregulation of genes encoding proteins involved in butyrate production, but it is unclear if butyrate would be directly produced from pyruvate, or derived from external acetate. *Anaerobutyricum hallii*, on the other hand, showed high gene expression related to converting lactate into butyrate, as also shown previously [12]. In addition, our data indicated that formate might have been produced by proteins encoded by genes belonging to the *Enterococcus* and *Bacteroides* populations.

Microbial groups directly involved in the degradation of the IMMPs

In order to gain insight into which bacterial groups were directly involved in degradation of different IMMPs, we used the KEGG reference pathway for starch and sucrose metabolism [13]. We surveyed our data for the expression of the genes encoding enzymes that are known to be involved in sucrose and starch metabolism. More specifically we focused on genes encoding enzymes from glycoside hydrolase family 13 (http://www.cazy.org/GH13_bacteria.html), as this family includes a number of bacterial proteins shown to be essential in degradation of similar compounds, such as isomaltooligosaccharides (IMOs) [14]. The majority of genes listed in the KEGG starch and sucrose metabolism pathway were detected in our transcriptome data, as well as some additional genes in glycoside hydrolase family 13 (EC 3.2.1.135, 3.2.1.68 and 3.2.1.11), which were not listed in the KEGG pathway, but which have been shown to be activated during the degradation of pullulan and dextran [15-18]. It is interesting to note that the relative contribution of these starch and sucrose metabolism genes to the total number of genes from each sample did not correlate with the presence or absence of IMMPs in the samples. The only exception was incubation with IMMP-dig27, in which starch and sucrose metabolism genes reached an expression of 10% at 12 h and about 12% at 48 h, whereas in other groups they ranged between 4 to 5% (Figure S4). Despite of the similarities in the overall expression of the starch and sucrose metabolism genes in all samples, we could see differences in the relative abundance of genes coding for specific enzymes depending on the IMMP used, and the duration of the fermentation (Figure S5).

One of the aims of this study was to better understand the functional dynamics of the bacterial communities during IMMP degradation. Previously reported HPAEC and HPSEC analyses [11] showed that the degradation of IMMP-94 and IMMP-96 occurred between 12 h and 24 h of the incubation. At 24 h and 48 h we noted an increase in the expression of genes coding for enzymes that might be directly involved in the hydrolysis of α-(1→6) glycosidic linkages, namely EC 3.2.1.10 – oligo-1,6-glucosidase, EC 3.2.1.11 – dextranase, and EC 3.2.1.33 – amylo-α-1,6- glucosidase (Figure S6a,b). There was also an increase in the expression of genes coding for enzymes that can hydrolyse α-(1→4) glycosidic linkages, mainly EC 3.2.1.1 – α-amylase, EC 3.2.1.20 – α–glucosidase 4-α-glucanotransferase, and EC 2.4.1.25 – 4-α-glucanotransferase. Since IMMP-27 contains lower amounts of α-(1→6) linkages, its degradation also involves the activation of the same genes, however, the expression levels of the genes encoding enzymes which hydrolyse α-(1→6) linkages were much lower (Figure S6a,b). Bacterial groups that contributed the most expression to the primary degradation of IMMP's α-(1→6) linkages were *Lactobacillus*, *Bifidobacterium* and *Bacteroides*, all expressing the genes encoding EC 3.2.1.10 oligo-1,6-glucosidase and EC 3.2.1.11 dextranase. On the other hand, the metatranscriptomic data suggested that α-(1→4) linkages were hydrolysed mainly by proteins encoded by *Bacteroides*, unclassified Bacteroidales, unclassified Enterobacteriaceae, *Lactobacillus* and *Bifidobacterium* via EC 3.2.1.1 alpha-amylase and EC 2.4.1.1 glycogen/amylophosphorylase (Figure S7). Based on the transcript data, genes belonging to the genera *Bifidobacterium* and *Lactobacillus* were mainly expressed in the degradation of IMMP-94 and IMMP-96 at 24 h (Figure 4, and Figure S7). These genera showed also expression during degradation of IMMP-27 and the IMMP-dig27, but their relative contributions were much lower (Figure 4, and Figure S7). The breakdown of IMMPs at 24 h and 48 h was otherwise dominated by expression from *Bacteroides*, with the exception of IMMP-dig27 at 48 h, which showed a high level of expression of genes assigned to unclassified Enterobacteriaceae. Figure 4 summarizes our model of IMMP degradation and shows the proposed specialized role of lactobacilli and bifidobacteria in hydrolysis of α-(1→6) linkages. It also reveals the important contribution of *Bacteroides* as both primary and secondary degraders of IMMPs and their by-products.

Comparison to other metatranscriptomic projects

We identified few limitations in our study, such as the fact that our starting culture was a mix of faecal samples from several donors for which we lack the initial gut microbiota profiles. Another limitation was the use of *in-vitro* batch fermentation system which does not allow to control for nutrient depletion and by-products accumulation. Therefore, to validate our findings we compared our results with other *in-vivo* and *in-vitro* gut metatranscriptome projects. We selected all available Illumina single-end metatranscriptomic bioprojects at the NCBI, as well as two relevant bioprojects with paired-end Illumina data. For each project we profiled taxonomic composition and compared the projects using PCA on the inverted Bray-

Curtis distances. As it can be seen in Figure 5, our data showed stronger clustering than a dysbiotic microbiome project (PRJNA396964, from graft-vs-host disease patients [19]). Our project data was also closely positioned with clusters of PRJNA416988 project [20], another *in-vitro* study, where samples were derived from a single person and RNA was extracted after 7 days in a bioreactor and 24h in a Hungate tube. The second most closely positioned project was PRJNA509512, where samples were derived from healthy volunteers[21]. Another *in-vitro* project, PRJNA593787[22], with also a roughly comparable setup, was displayed further away, indicating a variation in the *in-vitro* setups comparable to *in-vivo* data.

Discussion

Prebiotic food components should be resistant to host's gastric enzymes, fermentable by the host's intestinal microbiota and capable of promoting growth and activity of bacterial groups associated with health [8]. The IMMPs seem to fulfil all these criteria [23-25]. Earlier studies demonstrated that (reduced) high DP IMOs are not or poorly digestible by rat gastric enzymes [25], and that diets containing IMOs are associated with higher numbers of lactobacilli, and an overall increase in the number of intestinal bacteria [26]. Moreover, a recent study with human inoculum reported that IMMPs can be fermented by human large intestinal microbiota and that SCFAs, in particular acetate and propionate, are produced, indicating that IMMPs may stimulate activity of probiotic groups [10]. This is in accordance with earlier findings from a small human trial that showed an increased level of bifidobacteria in subjects who received IMOs in their diets [23].

In our study, we confirmed the prebiotic character of the IMMPs and showed that the specific effect of different IMMPs on human faecal microbiota composition and gene expression varied during *in vitro* fermentation, depending on the relative amount of α-$(1\rightarrow6)$ glycosidic linkages present in the substrate (Figure 4). When IMMP-94 and IMMP-96 were used as a carbon source, we observed a strong upregulation of genes in the probiotic cluster, specifically genes assigned to bifidobacteria and lactobacilli. Furthermore, high relative gene expression of these bacteria corresponded with an increase in their relative abundance as estimated by rRNA gene sequencing [11]. In contrast, when the IMMP-dig27 also containing mainly α-$(1\rightarrow6)$ glycosyl units was used as a substrate, the relative expression of genes assigned to bifidobacteria and lactobacilli was lower. These bacteria showed lower expression in the control group, and their highest gene expression in the presence of IMMP-27 was delayed to 48 h. Interestingly, all IMMP treatment groups showed a time lag between the maximum relative expression, and the increase in the corresponding bacterial relative abundance as measured by rRNA gene-targeted community analysis [11]. For example, the maximum expression of bifidobacterial genes was observed at 24 h of incubation when IMMP-94 and IMMP-96 were used as substrates. Yet, bifidobacteria reached their highest relative abundance only at 48 h when relative expression of their genes had already decreased. The relative gene expression of lactobacilli followed a pattern similar to that of bifidobacteria in all treatment groups, except for incubations with IMMP-94 where genes assigned to lactobacilli showed maximum relative expression at 12 h, whereas bifidobacterial gene expression peaked at 24 h. Relative expression of genes assigned to *Bacteroides* was very high in all groups, regardless of the incubation time, or presence and type of the IMMP used. *Bacteroides* spp. are known to be generalists that are able to break down a wide array of carbon sources [27]. Bifidobacteria and lactobacilli are often more specialized and can grow on substrates that are not necessarily accessible to other bacteria in the microbial ecosystem (e.g. human milk oligosaccharides [28,29], galactooligosaccharides [30], other complex carbohydrates [31]) . This may be the reason that these groups show delayed gene expression in relation to *Bacteroides*, as only after the depletion of the easily accessible IMMP fractions containing the α-(1→4) glycosidic linkages the bacterial groups capable of utilizing the α-(1→6) glycosidic linkages gained a competitive advantage.

While relative abundance and gene expression of some of the beneficial bacteria increased with the presence of IMMPs, we also noted that the exclusive use of these prebiotics put a selective pressure on other beneficial microbes. For example, *Lactococcus lactis* and *Ruminococcus bromii* - two specialized beneficial degraders, did not show any survival in our samples [11]. This can be explained by the lack of suitable substrate for both species, given that neither any simple mono- or disaccharides (for *Lactococcus* [32]) nor type II or III resistant starch (for *Ruminococcus* [33]) were present in this experiment. Both organisms can be considered beneficial, but were not stimulated in the particular prebiotic environment tested here, enforcing the notion that prebiotics can selectively stimulate gene expression and growth of specific groups, whereas in general, a diverse diet may be necessary to comprehensively support a stable community of commensal microbes. An alternative explanation, which just recently got discovered, is that at least *L. lactis* could be a contamination from the medium, which we can as well not rule out [34], but this has not yet been shown for *R. bromii*.

In our study we also observed a clear effect of having no carbohydrate source in the control samples. With the absence of the prebiotics, there was a switch of the community from processing carbohydrates to utilizing amino acids [35], as indicated by the increase of relative abundance of *Acidaminococcus* [11]. In addition, there was an increase of *Bilophila* in the control samples, which is an organism previously associated with gut dysbiosis [36], and which previously has been shown to decrease during the use of other prebiotics [37].

A total of 130 families of glycoside hydrolases, 22 families of polysaccharide lyases, and 16 families of carbohydrate esterases have been described, and many of these enzymes are encoded only by the genomes of microbes (www.cazy.org) [38]. We surveyed our data for the presence of genes encoding the enzymes that are known to be involved in sucrose and starch metabolism, mostly genes from glycoside hydrolase family 13. Few studies up to date looked at the genetics and enzymology of degradation of IMMPs mainly in lactobacilli [16,39,40], bifidobacteria [41] and *Bacteroides*. However, microbial species in the gut do not act in isolation, but rather can interact with each other through a network of syntrophic interactions often making the utilization of the substrate more effective [42]. Metabolic potential and fermentation efficiency vary between different species, and complete degradation of IMMPs in the gut is a result of different bacterial groups working together in a complementary fashion, likely leading to the formation of microbial food chains [42,43]. Certain bacterial groups may show a higher expression of specific genes coding for enzymes required to catalyze given reactions at specific degradation steps. This was also visible in our experiments. The expression of oligo-1-6-glucosidase encoding genes was dominated by lactobacilli and bifidobacteria when IMMP-94 and IMMP-96 were used as a substrate, whereas *Bacteroides* and unclassified Bacteroidales showed also high gene expression with IMMP-27 or IMMP-dig27 substrate. Similar patterns could be observed in expression of other genes that code for enzymes involved in sucrose and starch metabolism (Suppl. Figure S7). While some of the carbohydrate breakdown steps were dominated by gene expression of known probiotic genera, many of the primary and secondary degradation processes were also performed by members of *Bacteroides*. Our data showed that once the fermentation started, one of the very specialized enzymes, dextranase, was expressed only by *Bacteroides*. Other processes were found reliant on multiple genera as based on the gene expression data. For example, the breakdown of IMMP27 and IMMP-dig27 to maltose and maltotriose by α–amylases was dominated by gene expression of the genus *Bacteroides*, whereas further metabolisation was performed also by bifidobacteria and lactobacilli. Furthermore, other groups such as enterobacteria or *Parabacteroides* were not involved in most of these breakdown processes, but still constituted viable populations in the communities. Their functional role in the community is, however, not clear.

Previously published experimental results of this experiment showed that the administration of IMMPs lead to an increased production of different SCFAs, mainly acetate and succinate [11]. While succinate normally does not accumulate in this medium [44], the excess of substrate [45], high CO_2 levels, and the upregulation of all the necessary steps [46] in our metabolic mapping, including the necessity for iron, could explain such accumulation. In addition, previous studies showed that succinate accumulation is associated with oversupply of complex substrates [47], such as prebiotics, or in our case IMMPs or when further metabolization of succinate is unnecessary [44]. It is also possible that the lack of Vitamin B12, which is necessary for propionate production [46], resulted in the accumulation of succinate instead of propionate. We also indeed observed an upregulation of genes involved in Vitamin B12 production. However, we were unable to unequivocally determine the exact reason based on our data.

A further limitation of our study is that we do not have the gut microbiota profile of the initial samples available, that our starting culture was a mix of seven different faecal samples, and that the overall setup of this type of *in-vitro* experiments didn’t allow for continious monitoring and controlling. We performed pH measurements [11], but the pH dropped as a result of SCFA production, which might have impacted the composition of the microbiota. In order to better understand the validity of our findings and how our sequencing results compare with other metatranscriptomic data from other projects (despite that differing sequencing runs have the biggest impact on the sequencing results), we compared our data with other in-vivo and *in-vitro* human metatranscriptomic studies. The results showed that our data is comparable to other metatranscriptomic projects (Figure 5), and is closely positioned to data derived from healthy volunteers and from other *in-vitro* projects. And while the exact composition of our *in-vitro* does not directly reflect the human gut (e.g. loss of many Firmicutes), we could still identify many of the known human gut microbes (Figure 1). We therefore consider this experiment as being in general similar to relevant physiological conditions.

In conclusion, dietary fibres, including modified starches such as IMMPs offer a promising, non-invasive way to intentionally manipulate gut microbiota composition.

Investigations of whole bacterial communities and understanding of the mechanisms by which microorganisms interact to degrade different dietary carbohydrates are essential for our ability to manipulate gut microbiota to benefit our health. We showed how IMMPs can increase the relative abundance and gene expression of beneficial bacteria, making these novel prebiotics potentially useful in improving host's health from the aspect of nutrition, to achieve prevention or even alleviation of diseases.

Materials and Methods

In vitro fermentation; design and sampling

The faecal inoculum stock was prepared at TNO (Zeist, The Netherlands) from fresh faeces of seven healthy adult donors, to reduce variation within the inoculum [48]. The stock was mixed, aliquoted and stored anaerobically at -80 °C [48]. Sterile 20 mL anaerobic serum bottles were filled with 10 mL of the Standard Ileal Efflux Medium (SIEM; Tritium Microbiology, Eindhoven, The Netherlands). The SIEM was prepared according to Rösch et *al.* [49], but omitting the carbon source and Tween 80. The modified SIEM medium contained 40% (v/v) BCO medium, 1.6% (v/v) salt solution, 0.8% (v/v) MgSO4 (50 g/L), 0.4% (v/v) cysteine hydrochloride (40 g/L), 0.08% (v/v) vitamin solution and 10% (v/v) MES buffer (1 M, pH 6.0) in water. Before inoculation, a faeces stock aliquot was mixed with SIEM at 1:10 v/v and incubated overnight at 37 °C. The activated inoculum was then added to the fermentation bottles at 1% (v/v) final concentration. Three different IMMP fibres were tested, with 27% (IMMP-27), 94% (IMMP-94) and 96% (IMMP-96) of α-(1→6) glycosidic linkages as compared to the total amount of glycosidic linkages. In addition, a pre-treated IMMP-27 (IMMP-dig27) sample was included after it had been digested with α-amylase and amyloglucosidase to imitate passage through the small intestine [11]. Samples were prepared and processed in duplicate with fibres added to individual fermentation bottles at a final concentration of 10 mg/mL. Flasks were incubated at 37 °C in an anaerobic chamber with 81% N2, 15% CO2,and 4%H2. 0.5 to 2 mL of each culture was removed at different sampling time points, depending on the experiment.

In experiment A, cultures supplied with two different prebiotic fibres (IMMP-27 and IMMP-94) and one control culture without any substrate (IMMP blank, with only the modified SIEM medium, without any carbon source) were monitored over 48 h (in duplicate), and aliquots were removed at time points 0 (up to 15 min after addition of the prebiotic), 24 h and 48 h. In experiment B, cultures were supplied with two other prebiotics, IMMP-dig27 and IMMP-96, and one culture was left with no prebiotic (IMMP blank). Experiment B was monitored for 48 hours, and samples were taken at 6 h, 12 h, 24 h and 48 h (in duplicate). An aliquot of the activated blank inoculum was taken at time point 0, just before the addition of the IMMP. All samples (18) from experiment A were subjected to metatranscriptomic sequencing. In experiment B the metatranscriptomics sequencing was done for the activated inoculum at time point t0 and for the treatment groups at all time points (17). Samples for metatranscriptomics were harvested and immediately stabilized in RNAprotect (Qiagen, Hilden, Germany) following the manufacturer's instructions, and bacterial pellets were stored at -80 °C for up to three weeks before further processing.

RNA extraction and Illumina sequencing

Total RNA was extracted by using the beat beating - TRIzol - column method modified from Kang et al. [50]. Briefly, bacterial pellets were re-suspended in 100 µL TE buffer (30 mM Tris-HCl, 1 mM EDTA, pH = 8.0) containing 15 mg/mL Lysozyme, 10 U/mL of Mutanolysin and 100 µg/mL of Proteinase K. Samples were vortexed for 10 s and incubated at room temperature for 10 min, and 400 µL of RLT buffer (Qiagen, Hilden, Germany) containing 4 µL of β-mercaptoethanol was added. Samples were then vortexed, mixed with 500 µL of TRIzol Max reagent (Invitrogen, Carlsbad, CA, USA) and homogenized with 0.8 g of sterilized 0.1 mm zirconia beads for three min (3 × 1 min with cooling in between) at 5.5 ms using a bead beater (Precellys 24, Bertin Technologies). Following the beating step, samples were cooled on ice, gently mixed by inverting the tube with 200 µL of ice cold chloroform for 15 s and centrifuged for 15 min at 4 °C at 12,000 × g. The aqueous phase containing total RNA was transferred to fresh tubes and mixed with an equal volume of 70% ethanol. The mixture was placed on a Qiagen RNeasy mini column (RNeasy Mini Kit, Qiagen, Hilden, Germany) and centrifuged at 8,000 × g for 15 s to bind RNA into the column. Filtrate was discarded, and the RNA binding step was repeated until the complete sample was filtered through the column. The columns were rinsed with 350 µL of RW1 buffer (RNeasy Mini Kit, Qiagen, Hilden, Germany), and 80 µL DNAse I solution (Roche, Manheim, Germany) was applied to the column and incubated for 15 min at RT to digest DNA. The columns were rinsed twice with 350 µL RW1 buffer, and twice with 700 µL of RPE buffer (RNeasy Mini Kit, Qiagen, Hilden, Germany), following with a final wash with 80% ethanol. Columns were dried by a 2 min centrifugation at maximum speed, and total RNA was eluted with 30 µL of DNAse/RNAse free water. The total RNA concentrations were measured spectrophotometrically with an ND-1000 spectrophotometer (NanoDrop® Technologies, Wilmington, DE, USA), and residual DNA concentrations were measured with the Qubit® dsDNA BR Assay Kit (Life Technologies, Leusden, the Netherlands). Samples which contained over 10 ng/µL DNA contamination were treated with the Turbo DNAfree® Kit (Ambion, Bleiswijk, Netherlands) following manufacturer's instructions and purified using the RNeasy Mini Kit. Total RNA quality was evaluated using the Experion RNA StdSens kit (Biorad Laboratories INC, USA), total RNA concentrations were measured with NanoDrop® and DNA contamination concentrations were measured with the Qubit®dsDNA BR Assay Kit. Between 3-5 µg of total RNA from each sample was used for mRNA enrichment with the RiboZero Bacterial rRNA Removal Kit (Illumina, San Diego, CA, USA), and the quality and quantity of enriched mRNA was assessed as described above for total RNA. Between 200-500 ng of enriched mRNA was used for cDNA production using the ScriptSeq®v2RNA-Seq Library Preparation Kit (Epicentre, Madison, WI, USA), FailSafe®PCR Enzyme Mix (Epicentre, Madison, WI, USA) and ScriptSeq®Index PCR Primers (Epicentre, Madison, WI, USA) for amplification and barcoding of di-tagged cDNA. The PCR product presence was confirmed with gel electrophoresis using the FlashGel® System (Lonza, Rockland, ME, USA). PCR products were then purified with the HighPrep® PCR kit (MagBio Genomics, Gaithersburg, MD, USA) and concentrations of indexed cDNA were measured using the Qubit®dsDNA BR Assay Kit (Invitrogen, Carlsbad, CA, USA). Approximately 28 ng of DNA from each sample was added to a pool, and final volume of each library was adjusted to 25 µL using the HighPrep® PCR kit. Two libraries were prepared containing either 17 or 18 samples, with final concentrations of 20 ng/µL in each library. Libraries were sent for single end 150 bp Illumina HiSeq2000 sequencing (GATC, Konstanz, Germany).

Bioinformatic processing, read assembly and annotation

The bioinformatics workflow was adapted from Davids et *al.* [51]. SortMeRNA v1.9 [52] software was used to screen the metatranscriptome data against all databases deployed with the program and to remove rRNA reads. Adapters were trimmed with cutadapt v1.2.1 [53] using default settings. Quality trimming was performed with PRINSEQ Lite v0.20.0 [54] with a minimum sequence length of 40 bp and a minimum quality of 30 on both ends of the read, and as mean quality. All reads containing more than three Ns or non-IUPAC characters were discarded.

Reads from experiment A (Suppl. Figure S1) were pooled and assembled with IDBA_UD version 1.1.1 [55] using two rounds of assembly; firstly, with the options –min_count 200 and – min_support 5, and secondly, the reads, which could not be mapped to this assembly with bowtie2 v2.0.6 [56], standard parameters, were extracted, and assembled with standard options, but with the output from the previous run provided as long reads. Contigs with an A/T content of >80% were removed from the final assembly. Due to hardware restrictions, we did not include reads from experiment B in the assembly, but rather mapped reads to the assembly generated from reads obtained from experiment A as described below. Since mapping rate was comparable (table S1), it was concluded that this approach is appropriate. Prodigal v2.5 was used for prediction of protein coding DNA sequences with the option for meta samples [57]. Protein sequences were annotated with InterProScan 5.4-47.0 [58] on the Dutch science grid (offered by the Dutch National Grid Initiative via SurfSara), and enriched by adding EC numbers using PRIAM version March 06, 2013 [59]. Carbohydrate active enzymes were predicted with dbCAN release 3.0 [60]. Further enrichment for EC numbers was obtained by matching all InterProScan derived domain names against the BRENDA database (download 13.06.13) [61] and using a text mining algorithm that included removal of the non-alphanumerical characters (colons, commas, brackets, etc.), partial and generic terms (type, terminal, subunit, domain, enzyme, like, etc.), as well as other smaller modifications. Details are provided in Supplementary Materials and Methods.

Read counts from experiment A and B (Figure S1) were obtained with Bowtie2 v2.0.6 [56] using default settings. BAM files were converted with SAMtools v0.1.18 [62], and gene coverage was calculated with subread version 1.4.6 [63]. Read mappings to the RNA-assemblies were inspected with Tablet [64].

Taxonomic assignments

RNA sequences from the metatranscriptome assembly were compared with Blast 2.2.29 [65] against the NCBI NT database (download 22.01.2014) using standard parameters, besides an E-value of 0.0001, to the human microbiome (download 08.05.2014), NCBI bacterial draft genomes (download 23.01.2014), NCBI protozoa genomes (download 08.05.2014), and the human genome (download 30.12.2013, release 08.08.2013, NCBI Homo sapiens annotation release 105). Taxonomy was estimated with a custom version of the LCA algorithm as implemented in MEGAN [66], but with the following changes: only hits, which exceeded a bit-score of 50 were considered, and of these, only hits with a length of more than 100 nucleotides and which did not deviate more than 10% from the longest hit were accepted.

From all sequences from the assembly that did not have a match in any of the former blast analyses, another run with the – blastn option was performed against the same databases. In case this did not yield any results, a blastp of the predicted proteins was performed against a custom version of the KEGG Orthology database (http://www.genome.jp/kegg/ko.html, download 25.04.2014). Taxonomic assignment was again performed with the LCA algorithm, and for the blastp run only hits which did not deviate by more than 10% from the hit with the maximum identity were considered.

Differential expression

Differential expression analysis was performed separately at genus level per genus in R version 3.1.1 [67] with the TCC package release 1.6.5 [68], with 36 iterations and the combination of tmm normalization and edgeR, with an FDR=0.1. Only genes with a q-value (multitest corrected p-value) of less than 0.01 in any of the relevant comparisons were considered to be significantly differentially expressed, unless otherwise mentioned. A volcano plot can be seen in Figure S8.

Metabolic mapping

Two rounds of clustering were performed to detect patterns in the expressed genes (Figure S2). All genera, which either had an average read count of >=10 per gene, or which exceeded 1% of all reads in any given condition, were clustered into groups based on relative counts per group using the k-means algorithm in Scipy version 1.6.1 [69]. To determine the stability of the clustering, 50 iterations with a clustering between 1 and 20 clusters were performed, with the option “iter” set to 100.000. Afterwards the average cluster support per amount of clusters over all the iterations was computed, and additionally, the clustering was investigated with a custom python implementation of clustergrams [70]. Within the clustered genera, genes with similar expression patterns were identified with the DBSCAN algorithm [71]. Clustering on expression patterns was performed with ELKI 0.7.0~20150828 [72], the –minpts parameter was fixed to 3 and the epsilon parameter was varied in percentages. Final clustering was evaluated using the Tau index as implemented in ELKI, and the clustering result with the best Tau was chosen, unless a lower Tau led to better cluster separation.

Only genes which were differentially expressed in at least one sampling time point in any of the incubations (i.e. Ino.BL, IMMP-27, IMMP-94, IMMP-96, IMMP-dig27), were considered in the clustering analysis. Genes were normalized per row before the clustering. All derived EC numbers were mapped with custom scripts onto the KEGG database [13] and visualized with Python Scipy version 1.6.1 and NumPy version 0.9.0 [69]. Correlations were calculated with the mentioned versions of Scipy/NumPy. Differentially expressed genes were mapped separately for groups of interest, and changed functions were derived from visual inspections. Cofactor requirements were investigated with the Expasy database [73].

Comparison with other projects

The NCBI was searched (June 20, 2021) for all bioprojects containing single-end Illumina sequencing data derived from the human gut microbiota, as well as relevant projects with paired-end Illumina sequencing. In total 10 projects could be identified (PRJEB30770, PRJNA396964, PRJNA416988, PRJNA476770, PRJNA509512, PRJNA510123, PRJNA515307, PRJNA526861, PRJNA593787, PRJNA597902), with in total 457 samples. Of the 457 samples one sequencing run, SRR8567581, was excluded, due to not containing enough reads. All samples were downloaded with fasterq-dump from the sra-toolkit 2.11.0 (with option –concatenate-reads),and were classified with Kraken2 v2.1.1 with the PlusPF database, https://benlangmead.github.io/aws-indexes/k2 [74]. Bray-Curtis dissimilarity was calculated with the beta-diversity script from KrakenTools v1.2, https://github.com/jenniferlu717/KrakenTools/. The beta-diversity measures were inverted and a PCA was performed with the pca python package, https://github.com/erdogant/pca/ .

Ethical approval and informed consent

Faeces collection in the Netherlands does not require medical ethical approval if volunteers have not undergone any form of intervention, such as in this study. However, an informed consent was signed by all donors upon providing the faecal samples.

Data accessibility

The datasets generated and analyzed during the current study are available in the ENA repository, under accession number PRJEB13209, https://www.ebi.ac.uk/ena/data/view/PRJEB13009 .

References

1 Dethlefsen, L., McFall-Ngai, M. & Relman, D. A. An ecological and evolutionary perspective on human-microbe mutualism and disease. *Nature* **449**, 811-818, doi:10.1038/nature06245 (2007).

2 Flint, H. J., Duncan, S. H., Scott, K. P. & Louis, P. Interactions and competition within the microbial community of the human colon: links between diet and health. *Environmental microbiology* **9**, 1101-1111, doi:10.1111/j.1462-2920.2007.01281.x (2007).

3 Cani, P. D. & Everard, A. Talking microbes: When gut bacteria interact with diet and host organs. *Molecular Nutrition and Food Research* **60**, 58-66, doi:10.1002/mnfr.201500406 (2016).

4 Garrett, W. S., Gordon, J. I. & Glimcher, L. H. Homeostasis and inflammation in the intestine. *Cell* **140**, 859-870, doi:10.1016/j.cell.2010.01.023 (2010).

5 Hood, L., Balling, R. & Auffray, C. Revolutionizing medicine in the 21st century through systems approaches. *Biotechnology journal* **7**, 992-1001, doi:10.1002/biot.201100306 (2012).

6 Becattini, S., Taur, Y. & Pamer, E. G. Antibiotic-Induced Changes in the Intestinal Microbiota and Disease. *Trends Mol Med* **22**, 458-478, doi:10.1016/j.molmed.2016.04.003 (2016).

7 Bowman, K. A., Broussard, E. K. & Surawicz, C. M. Fecal microbiota transplantation: current clinical efficacy and future prospects. *Clin Exp Gastroenterol* **8**, 285-291, doi:10.2147/CEG.S61305 (2015).

8 Gibson, G. R. *et al.* Expert consensus document: The International Scientific Association for Probiotics and Prebiotics (ISAPP) consensus statement on the definition and scope of prebiotics. *Nature reviews. Gastroenterology & hepatology* **14**, 491-502, doi:10.1038/nrgastro.2017.75 (2017).

9 Scott, K. P., Martin, J. C., Duncan, S. H. & Flint, H. J. Prebiotic stimulation of human colonic butyrate-producing bacteria and bifidobacteria, in vitro. *FEMS microbiology ecology* **87**, 30-40, doi:10.1111/1574-6941.12186 (2014).

10 Leemhuis, H. *et al.* Isomalto/malto-polysaccharide, a novel soluble dietary fiber made via enzymatic conversion of starch. *Journal of agricultural and food chemistry* **62**, 12034-12044, doi:10.1021/jf503970a (2014).

11 Gu, F. *et al.* In vitro fermentation behaviour of isomalto/malto-polysaccharides using human faecal inoculum indicates prebiotic potential. *Molecular Nutrition and Food Research* **62**, 1800232 (2018).

12 Duncan, S. H., Louis, P. & Flint, H. J. Lactate-utilizing bacteria, isolated from human feces, that produce butyrate as a major fermentation product. *Applied and environmental microbiology* **70**, 5810-5817, doi:10.1128/AEM.70.10.5810-5817.2004 (2004).

13 Kanehisa, M., Goto, S., Sato, Y., Furumichi, M. & Tanabe, M. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic acids research* **40**, D109-114, doi:10.1093/nar/gkr988 (2012).

14 Abou Hachem, M. *et al.* A Snapshot into the Metabolism of Isomalto-oligosaccharides in Probiotic Bacteria. *Journal of Applied Glycoscience* **60**, 95-100, doi:10.5458/jag.jag.JAG-2012_022 (2013).

15 Kuchtová, A. & Janeček, Š. Domain evolution in enzymes of the neopullulanase subfamily. *Microbiology* **162**, 2099-2115, doi:doi:10.1099/mic.0.000390 (2016).

16 Ganzle, M. G. & Follador, R. Metabolism of oligosaccharides and starch in lactobacilli: a review. *Frontiers in microbiology* **3**, 340, doi:10.3389/fmicb.2012.00340 (2012).

17 Bailey, R. W. & Clarke, R. T. J. A bacterial dextranase. *Biochemical Journal* **72**, 49-54 (1959).

18 Khalikova, E., Susi, P. & Korpela, T. Microbial Dextran-Hydrolyzing Enzymes: Fundamentals and Applications. *Microbiology and Molecular Biology Reviews* **69**, 306-325, doi:10.1128/mmbr.69.2.306-325.2005 (2005).

19 Legoff, J. *et al.* The eukaryotic gut virome in hematopoietic stem cell transplantation: new clues in enteric graft-versus-host disease. *Nature Medicine* **23**, 1080-1085, doi:10.1038/nm.4380 (2017).

20 Defois, C. *et al.* Food Chemicals Disrupt Human Gut Microbiota Activity And Impact Intestinal Homeostasis As Revealed By In Vitro Systems. *Scientific Reports* **8**, 11006, doi:10.1038/s41598-018-29376-9 (2018).

21 Langelier, C. *et al.* Microbiome and Antimicrobial Resistance Gene Dynamics in International Travelers. *Emerging Infectious Disease journal* **25**, 1380, doi:10.3201/eid2507.181492 (2019).

22 Hanson, B. T. *et al.* Sulfoquinovose is a select nutrient of prominent bacteria and a source of hydrogen sulfide in the human gut. *The ISME Journal* **15**, 2779-2791, doi:10.1038/s41396-021-00968-0 (2021).

23 Kohmoto, T. *et al.* Effect of Isomalto-oligosaccharides on Human Fecal Flora. *Bifidobacteria and Microflora* **7**, 61-69 (1988).

24 Roberfroid, M. Prebiotics: The Concept Revisited. *The Journal of nutrition* **137**, 830S-837S (2007).

25 Kaneko, T., Yokoyama, A. & Suzuki, M. Digestibility Characteristics of Isomaltooligosaccharides in Comparison with Several Saccharides Using the Rat Jejunum Loop Method. *Bioscience, Biotechnology, and Biochemistry* **59**, 1190-1194 (1995).

26 Ketabi, A., Dieleman, L. A. & Ganzle, M. G. Influence of isomalto-oligosaccharides on intestinal microbiota in rats. *Journal of applied microbiology* **110**, 1297-1306, doi:10.1111/j.1365-2672.2011.04984.x (2011).

27 Cockburn, D. W. & Koropatkin, N. M. Polysaccharide Degradation by the Intestinal Microbiota and Its Influence on Human Health and Disease. *Journal of Molecular Biology* **428**, 3230-3252, doi:https://doi.org/10.1016/j.jmb.2016.06.021 (2016).

28 Marcobal, A. *et al.* Consumption of human milk oligosaccharides by gut-related microbes. *Journal of agricultural and food chemistry* **58**, 5334-5340, doi:10.1021/jf9044205 (2010).

29 Thongaram, T., Hoeflinger, J. L., Chow, J. & Miller, M. J. Human milk oligosaccharide consumption by probiotic and human-associated bifidobacteria and lactobacilli. *Journal of dairy science* **100**, 7825-7833, doi:10.3168/jds.2017-12753 (2017).

30 Thongaram, T., Hoeflinger, J. L., Chow, J. & Miller, M. J. Prebiotic Galactooligosaccharide Metabolism by Probiotic Lactobacilli and Bifidobacteria. *Journal of agricultural and food chemistry* **65**, 4184-4192, doi:10.1021/acs.jafc.7b00851 (2017).

31 Crociani, F., Alessandrini, A., Mucci, M. M. & Biavati, B. Degradation of complex carbohydrates by *Bifidobacterium* spp. *International journal of food microbiology* **24**, 199-210 (1994).

32 De Vos, W. M. Metabolic engineering of sugar catabolism in lactic acid bacteria. *Antonie van Leeuwenhoek* **70**, 223-242 (1996).

33 Ze, X., Duncan, S. H., Louis, P. & Flint, H. J. *Ruminococcus bromii* is a keystone species for the degradation of resistant starch in the human colon. *The ISME journal* **6**, 1535-1543, doi:10.1038/ismej.2012.4 (2012).

34 Bisanz, J. E., Upadhyay, V., Turnbaugh, J. A., Ly, K. & Turnbaugh, P. J. Meta-Analysis Reveals Reproducible Gut Microbiome Alterations in Response to a High-Fat Diet. *Cell host & microbe*, doi:10.1016/j.chom.2019.06.013 (2019).

35 Rogosa, M. *Acidaminococcus* gen. n., *Acidaminococcus fermentans* sp. n., Anaerobic Gram-negative Diplococci Using Amino Acids as the Sole Energy Source for Growth. *Journal of bacteriology* **92**, 756-766 (1969).

36 Baron, E. J. Bilophila wadsworthia: a Unique Gram-negative Anaerobic Rod. *Anaerobe* **3**, 83-86, doi:https://doi.org/10.1006/anae.1997.0075 (1997).

37 Vandeputte, D. *et al.* Prebiotic inulin-type fructans induce specific changes in the human gut microbiota. *Gut* **66**, 1968-1974, doi:10.1136/gutjnl-2016-313271 (2017).

38 Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M. & Henrissat, B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic acids research* **42**, D490-495, doi:10.1093/nar/gkt1178 (2014).

39 Moller, M. S. *et al.* Enzymology and structure of the GH13_31 glucan 1,6-alpha-glucosidase that confers isomaltooligosaccharide utilization in the probiotic Lactobacillus acidophilus NCFM. *J Bacteriol* **194**, 4249-4259, doi:10.1128/JB.00622-12 (2012).

40 Hu, Y., Ketabi, A., Buchko, A. & Ganzle, M. G. Metabolism of isomalto-oligosaccharides by *Lactobacillus reuteri* and bifidobacteria. *Letters in applied microbiology* **57**, 108-114, doi:10.1111/lam.12076 (2013).

41 Liu, S. *et al.* Starch and starch hydrolysates are favorable carbon sources for bifidobacteria in the human gut. *BMC microbiology* **15**, 54, doi:10.1186/s12866-015-0362-3 (2015).

42 Koropatkin, N. M., Cameron, E. A. & Martens, E. C. How glycan metabolism shapes the human gut microbiota. *Nature reviews. Microbiology* **10**, 323-335, doi:10.1038/nrmicro2746 (2012).

43 Fischbach, M. A. & Sonnenburg, J. L. Eating for two: how metabolism establishes interspecies interactions in the gut. *Cell host & microbe* **10**, 336-347, doi:10.1016/j.chom.2011.10.002 (2011).

44 Gibson, G. R., Cummings, J. H. & MacFarlane, G. T. Use of a Three-Stage Continuous Culture System to Study the Effect of Mucin on Dissimilatory Sulfate Reduction and Methanogenesis by Mixed Populations of Human Gut Bacteria. *Applied and environmental microbiology* **54**, 2750-2755 (1988).

45 Macfarlane, S. & Macfarlane, G. T. Regulation of short-chain fatty acid production. *The Proceedings of the Nutrition Society* **62**, 67-72, doi:10.1079/PNS2002207 (2003).

46 Macy, J. M. & Probst, I. The biology of gastrointestinal bacteroides. *Annual Reviews in Microbiology* **33**, 561-591 (1979).

47 Sakata, T., Kojima, T., Fujieda, M., Takahashi, M. & Michibata, T. Influences of probiotic bacteria on organic acid production by pig caecal bacteria in vitro. *The Proceedings of the Nutrition Society* **62**, 73-80, doi:10.1079/PNS2002211 (2003).

48 Aguirre, M., Ramiro-Garcia, J., Koenen, M. E. & Venema, K. To pool or not to pool? Impact of the use of individual and pooled fecal samples for in vitro fermentation studies. *Journal of microbiological methods* **107**, 1-7, doi:10.1016/j.mimet.2014.08.022 (2014).

49 Rösch, C., Venema, K., Gruppen, H. & Schols, H. A. Characterisation and in vitro fermentation of resistant maltodextrins using human faecal inoculum and analysis of bacterial enzymes present. *Bioactive Carbohydrates and Dietary Fibre* **6**, 46-53, doi:10.1016/j.bcdf.2015.07.002 (2015).

50 Kang, S., Denman, S. E., Morrison, M., Yu, Z. & McSweeney, C. S. An efficient RNA extraction method for estimating gut microbial diversity by polymerase chain reaction. *Current microbiology* **58**, 464-471, doi:10.1007/s00284-008-9345-z (2009).

51 Davids, M. *et al.* Functional Profiling of Unfamiliar Microbial Communities Using a Validated De Novo Assembly Metatranscriptome Pipeline. *PLOS ONE* **11**, e0146423 (2016).

52 Kopylova, E., Noe, L. & Touzet, H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* **28**, 3211-3217, doi:10.1093/bioinformatics/bts611 (2012).

53 Martin, M. Cutadapt removes adapter sequences from high-througput sequencing reads. *EMBnet.journal* **17**, 10-12 (2011).

54 Schmieder, R. & Edwards, R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* **27**, 863-864, doi:10.1093/bioinformatics/btr026 (2011).

55 Peng, Y., Leung, H. C., Yiu, S. M. & Chin, F. Y. IDBA-UD: A de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics* **28**, 1420-1428 (2012).

56 Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nature methods* **9**, 357-359, doi:10.1038/nmeth.1923 (2012).

57 Hyatt, D. *et al.* Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC bioinformatics* **11**, 119, doi:10.1186/1471-2105-11-119 (2010).

58 Hunter, S. *et al.* InterPro in 2011: new developments in the family and domain prediction database. *Nucleic acids research* **40**, D306-312, doi:10.1093/nar/gkr948 (2012).

59 Claudel-Renard, C., Chevalet, C., Faraut, T. & Kahn, D. Enzyme-specific profiles for genome annotation: PRIAM. *Nucleic acids research* **31**, 6633-6639 (2003).

60 Yin, Y. *et al.* dbCAN: a web resource for automated carbohydrate-active enzyme annotation. *Nucleic acids research* **40**, W445-451, doi:10.1093/nar/gks479 (2012).

61 Chang, A. *et al.* BRENDA in 2015: exciting developments in its 25th year of existence. *Nucleic acids research* **43**, D439-446, doi:10.1093/nar/gku1068 (2015).

62 Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079, doi:10.1093/bioinformatics/btp352 (2009).

63 Liao, Y., Smyth, G. K. & Shi, W. The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic acids research* **41** (2013).

64 Milne, I. *et al.* Using Tablet for visual exploration of second-generation sequencing data. *Briefings in bioinformatics* **14**, 193-202, doi:10.1093/bib/bbs012 (2013).

65 Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic Local Alignment Search Tool. *Journal of Molecular Biology* **215**, 403-410 (1990).

66 Huson, D. H., Mitra, S., Ruscheweyh, H. J., Weber, N. & Schuster, S. C. Integrative analysis of environmental sequences using MEGAN4. *Genome research* **21**, 1552-1560, doi:10.1101/gr.120618.111 (2011).

67 R Core Team. *R: A language and environment for statistical computing.*, (R Foundation for Statistical Computing, 2012).

68 Sun, J., Nishiyama, T., Shimizu, K. & Kadota, K. TCC: an R package for comparing tag count data with robust normalization strategies. *BMC bioinformatics* **14**, 219 (2013).

69 van der Walt, S., Colbert, C. & Varoquaux, G. The NumPy Array: A structure for Efficient Numerical Computation. *Computing In Science & Engineering* **13**, 22-30 (2011).

70 Schonlau, M. Visualizing non-hierarchical and hierarchical cluster analyses with clustergrams. *Computational Statistics* **19**, 95-111 (2004).

71 Ester, M., Kriegel, H.-P., Sander, J. & Xu, X. in *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining.* 226-231.

72 Schubert, E. *et al.* A Framework for Clustering Uncertain Data. *Proceedings of the VLDB Endowment* **18**, 1976-1979 (2015).

73 Bairoch, A. The ENZYME database in 2000. *Nucleic acids research* **28**, 304-305 (2000).

74 Wood, D. E., Lu, J. & Langmead, B. Improved metagenomic analysis with Kraken 2. *Genome Biology* **20**, 257, doi:10.1186/s13059-019-1891-0 (2019).

Acknowledgements

The authors thank Jasper Koehorst (Wageningen University & Research, Laboratory of Systems and Synthetic Biology) for his help with the transcriptome annotation, and Maria Suarez-Diez and Edoardo Saccenti (Wageningen University & Research, Laboratory of Systems and Synthetic Biology) for helpful discussions. The authors want to thank Bartholomeus van den Bogert (Wageningen University & Research, Laboratory of Microbiology) for the design of Figure S1.

This research was performed in the framework of the public-private partnership CarboHealth coordinated by the Carbohydrate Competence Center (CCC, www.cccresearch.nl) and financed by participating partners and allowances of the TKI Agri&Food program, Ministry of Economic Affairs. This work was furthermore supported by Wageningen University and the Wageningen Institute for Environment and Climate Research (WIMEK) to BH through the IP/OP program Systems Biology [project KB-17-003.02-023].

This work was carried out on the Dutch national e-infrastructure with the support of SURF Foundation.

Author contributions

KB and FG conducted the experiments. BH performed the bioinformatics processing. KB and BH analysed the data, designed figures and wrote the manuscript. HAS and HS conceived the study. PS advised on bioinformatics methods. PZ provided the IMMP substrates and provided biochemical assistance. All authors reviewed the manuscript. KB and BH contributed equally to this manuscript.

Competing interests

KB is currently working for Mead Johnson. FG is currently working for TUMCREATE Ltd. PZ is currently working for IFF. None of the companies were involved in this study.

Figure legends

Figure 1. Average relative transcript expression of different genus level taxa in incubations sampled at time points 0 h, 6 h, 12 h, 24 h and 48 h. When the taxonomic assignment could not be made at genus level, the lowest classifiable taxonomy assignment was used for display. Low abundance genera are summarized as “Other taxonomic assignments” for display purposes.

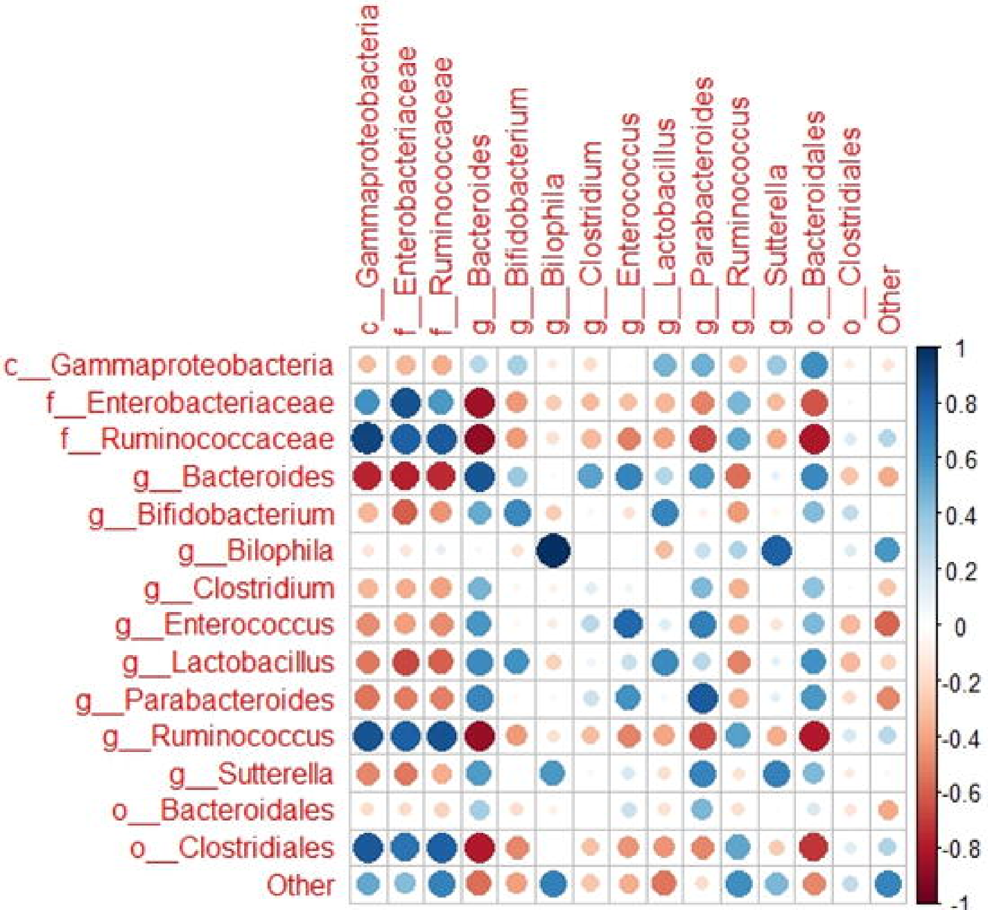
Figure 2. Correlation between the relative activity of the main bacterial groups based on metatranscriptome data, and their relative abundance based on 16S rRNA gene sequencing data [11]. In case when genus level assignment was ambiguous, unclassified fraction within the next higher taxonomic level was used.

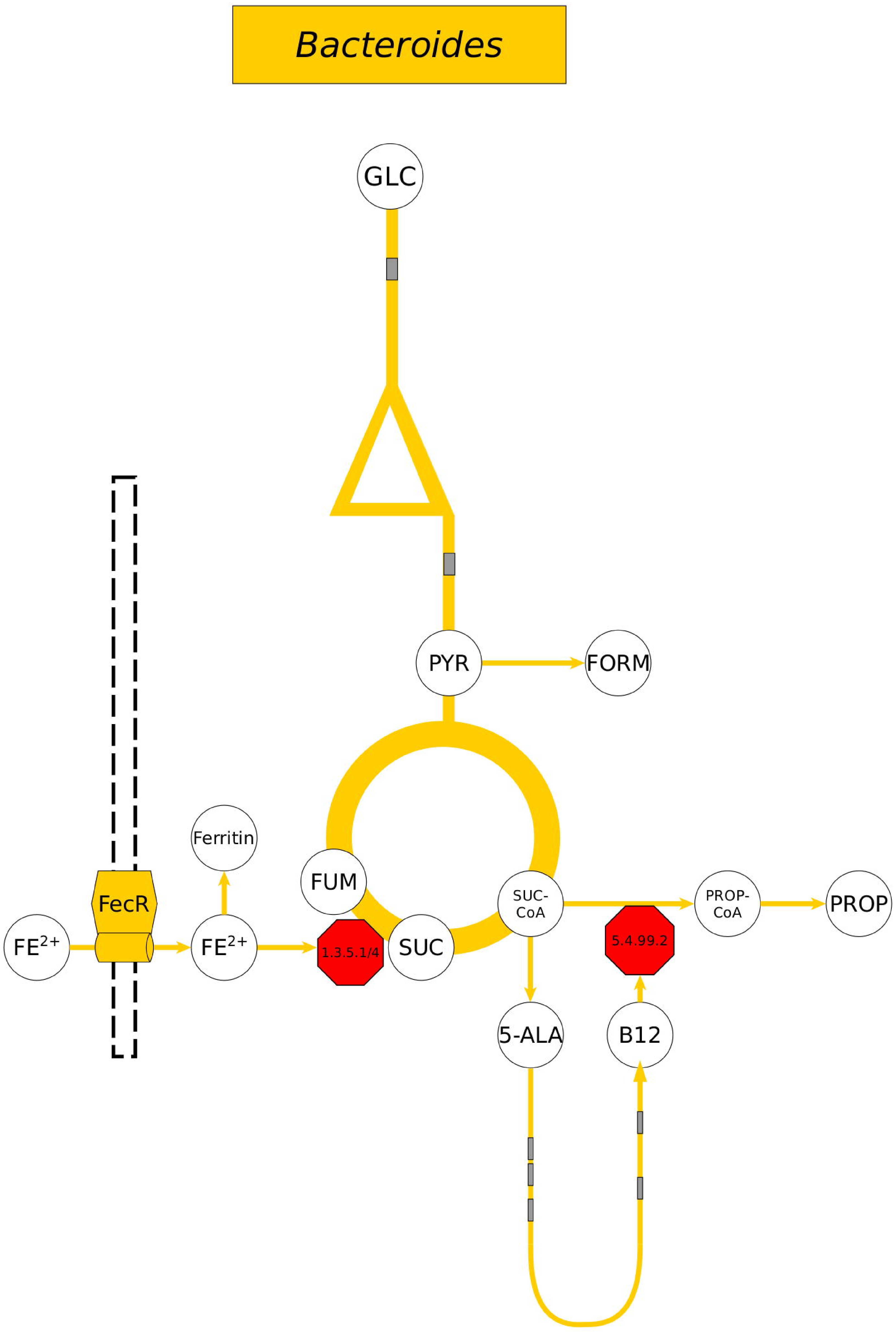
Figure 3. Overview of the metabolism of specific microbial groups observed in the samples taken during *in vitro* fermentation of different IMMPs by human faecal inoculum All samples show in general the same patterns for all organisms, besides for *Anaerobutyricum hallii*, which only showed expression in the samples with IMMP-dig27. The genus *Enterococcus* showed the same pattern as *Bifidobacterium/Lactobacillus*, but at lower relative transcript abundance. Grey indicates that certain genes were not differentially expressed within a pathway. 5-ALA = 5-Aminolevulinate, AC = Acetate, Ac-CoA = Acetyl-CoA, BUT = Butyrate, FORM = Formate, FUM = Fumarate , GLC = Glucose, LAC = Lactate, PROP = Propionate, PROP-CoA = Propanoyl-CoA, PYR = Pyruvate, SUC = Succinate, SUC-CoA = Succinyl-CoA

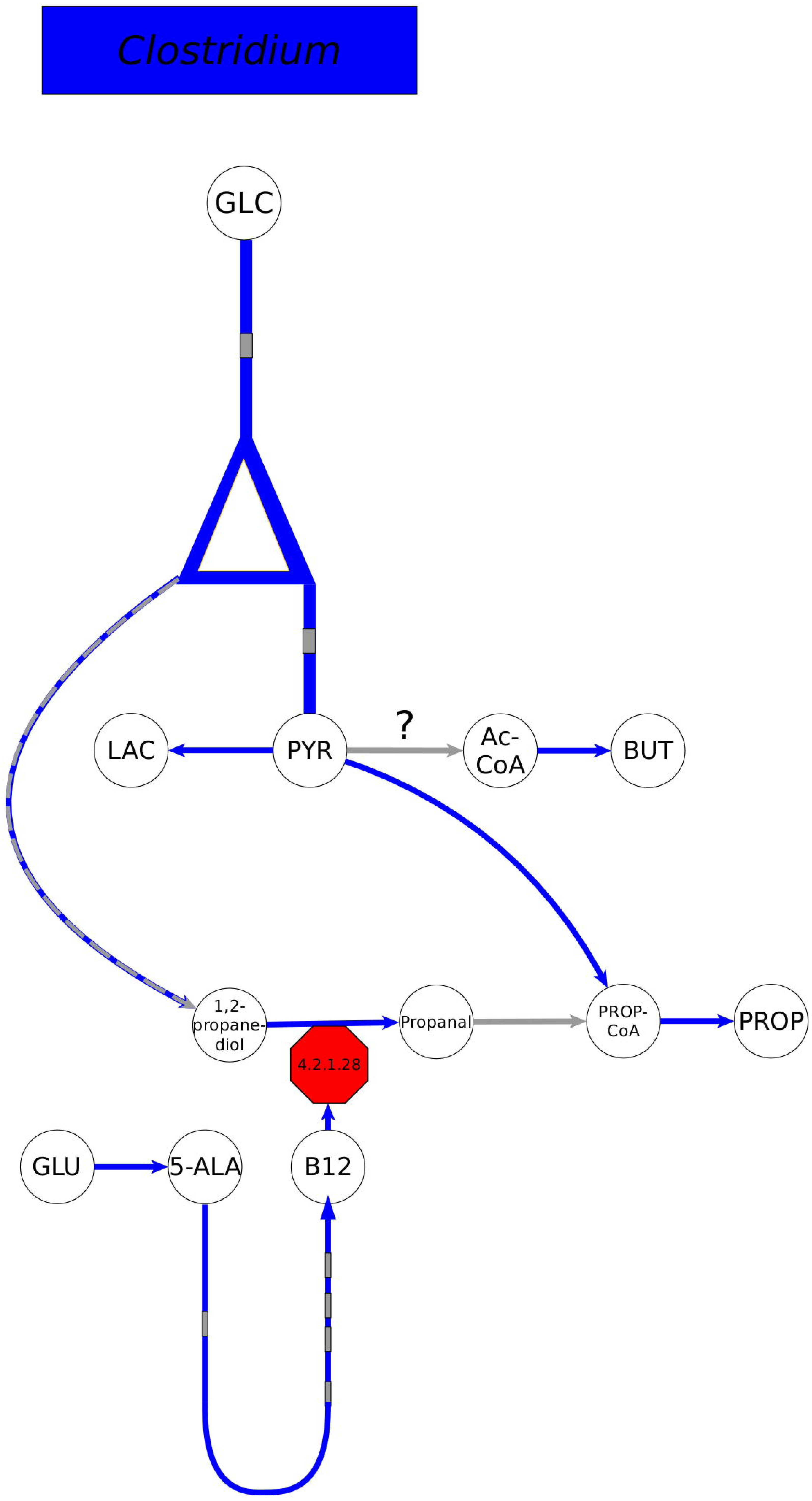
Figure 4. Overview over the main degradation pathways starting from dextran. Colours in the top panel indicate the main contributors to a reaction. The bottom panel shows the overall expression in reads per kilobase per million (RPKM) per organism at time points 24 and 48h for all conditions. If an enzyme could not be identified by its associated EC number, then the KO or CAZy identifier used for identification are given in brackets. Only enzymes, which were significantly differentially expressed over at least one time series, are displayed. Please note the different scales between the first and the last enzymes in the bottom panels, due to differences in expression scale.

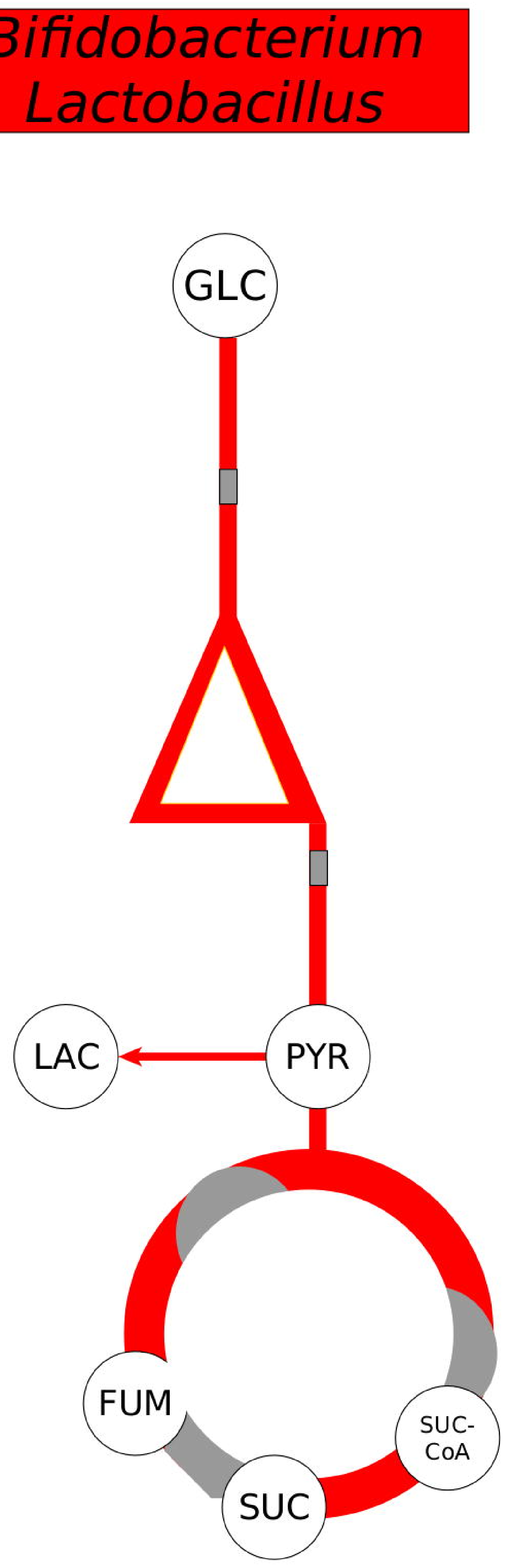
Figure 5. PCA based on Bray-Curtis distance between samples from different bioprojects. The current study (light blue) groups with samples from two other studies (green, purple), which are derived from either an in-vitro model based on a single volunteer or direct sequencing of healthy volunteer samples, respectively. The samples displayed in orange are derived from a graft-vs-host disease study and include disturbed microbiomes.

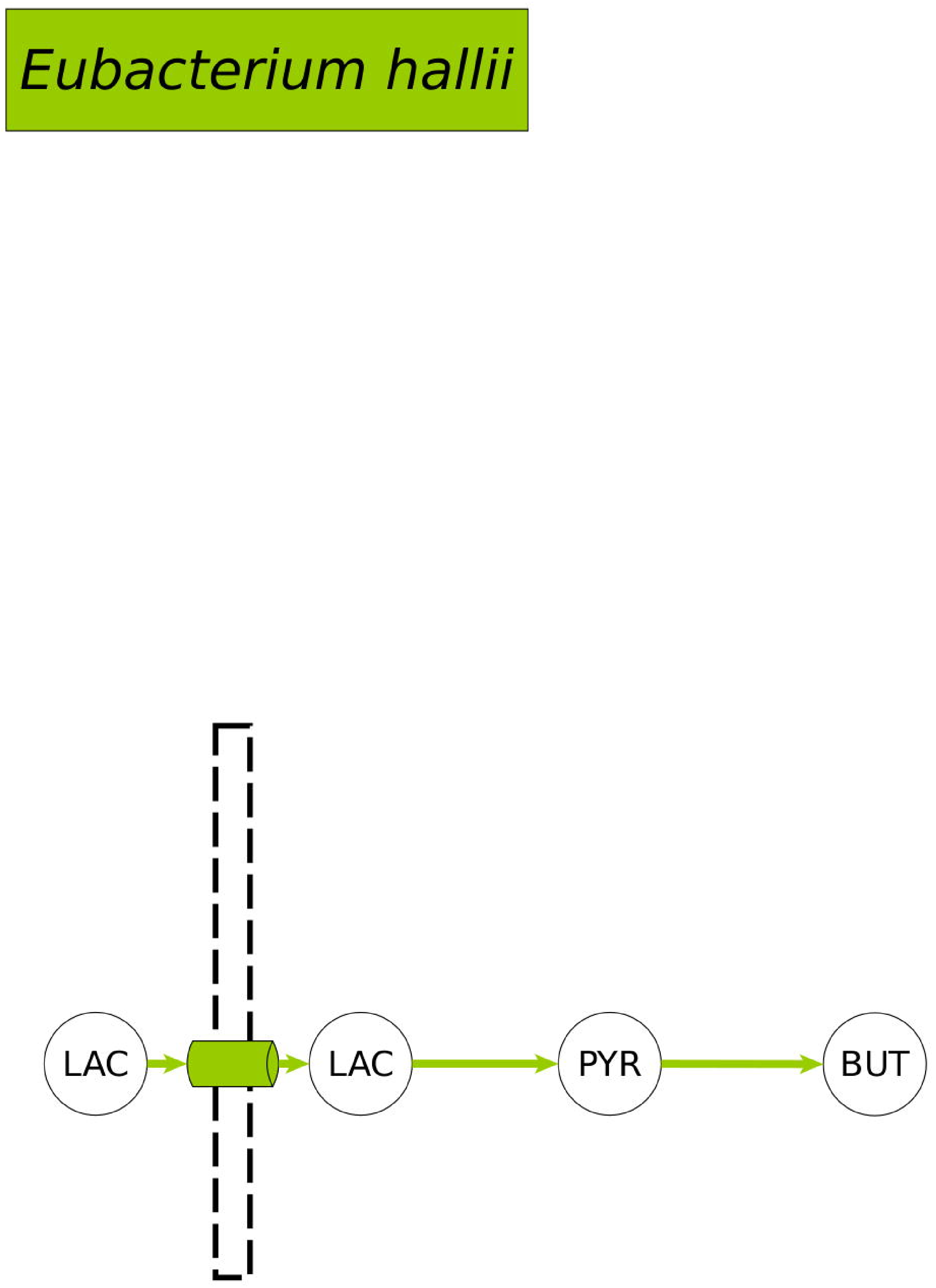
Percent
100
80
60
40
20
0
Ino.BL.0h
Ino.BL.24h
Ino.BL.48
IMMP-27.0h
IMMP-27.24h
IMMP-27.48h
IMMP-94.0h
IMMP-94.24h
IMMP-94.48h
InoBL.B2.0h
IMMP-96.6h
IMMP-96.12h
IMMP-96.24h
IMMP-96.48h
IMMP-dig27.6h
IMMP-dig27.12h
IMMP-dig27.24h
IMMP-dig27.48h
Sample
no classification assigned
other taxonomic assignments
unclassified Bacteria
unclassified Bacteroidetes
unclassified Bacteroidales
Bacteroides
unclassified Firmicutes
unclassified Clostridiales
Clostridium
Lactobacillus
Enterococcus
Ruminococcus
Lactococcus
unclassified Gammaproteobacteria
unclassified Enterobacteriaceae
Escherichia
Sutterella
Bilophila
Phascolarctobacterium
unclassified Actinobacteria
Bifidobacterium

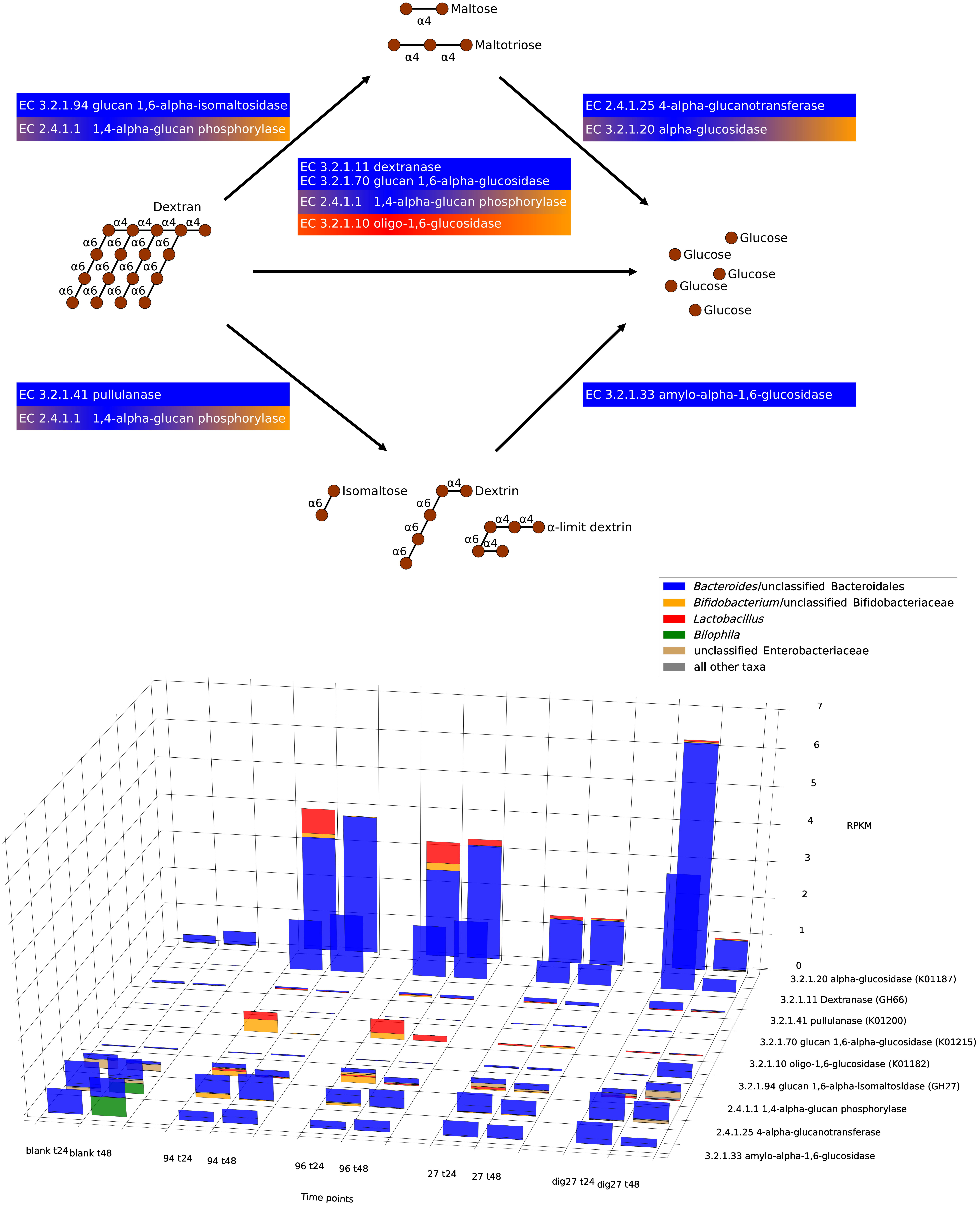

c__Gammaproteobacteria
f__Enterobacteriaceae
f__Ruminococcaceae
g__Bacteroides
g__Bifidobacterium
g__Bilophila
g__Clostridium
g__Enterococcus
g__Lactobacillus
g__Parabacteroides
g__Ruminococcus
g__Sutterella
o__Bacteroidales
o__Clostridiales
Other
c__Gammaproteobacteria
f__Enterobacteriaceae
f__Ruminococcaceae
g__Bacteroides
g__Bifidobacterium
g__Bilophila
g__Clostridium
g__Enterococcus
g__Lactobacillus
g__Parabacteroides
g__Ruminococcus
g__Sutterella
o__Bacteroidales
o__Clostridiales
Other
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1

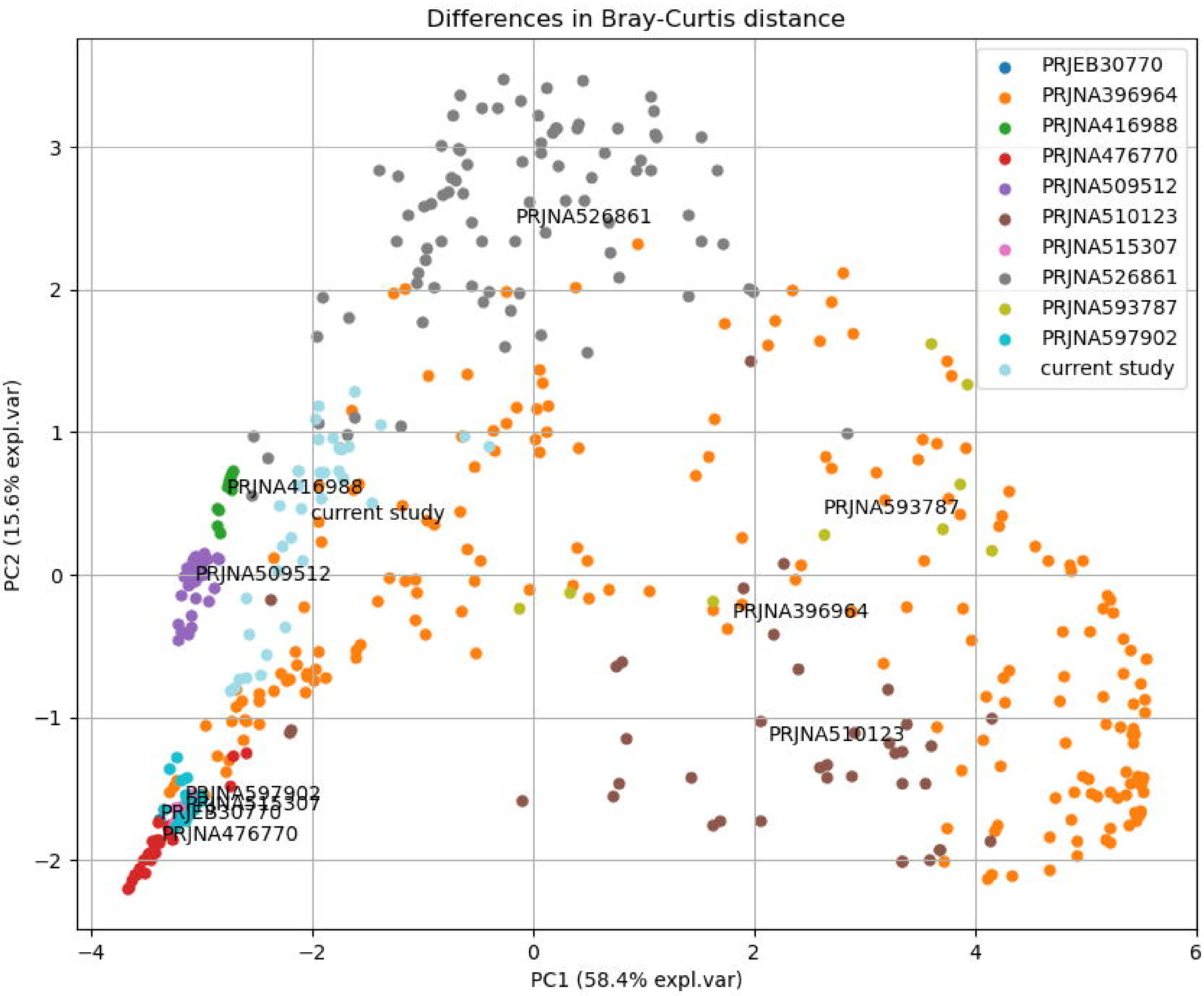

Bacteroides
GLC
PYR
FORM
FUM
SUC
SUC-CoA
PROP-CoA
PROP
Ferritin
FecR
FE^{2+}
FE^{2+}
1.3.5.1/4
5.4.99.2
5-ALA
B12


Clostridium
GLC
LAC
PYR
?
Ac-CoA
BUT
1,2-propanediol
Propanal
PROP-CoA
PROP
4.2.1.28
GLU
5-ALA
B12


Bifidobacterium
Lactobacillus
GLC
LAC
PYR
FUM
SUC
SUC-CoA


Eubacterium hallii
LAC
LAC
PYR
BUT


Maltose
α4
Maltotriose
α4 α4
EC 3.2.1.94 glucan 1,6-alpha-isomaltosidase
EC 2.4.1.1 1,4-alpha-glucan phosphorylase
EC 2.4.1.25 4-alpha-glucanotransferase
EC 3.2.1.20 alpha-glucosidase
EC 3.2.1.11 dextranase
EC 3.2.1.70 glucan 1,6-alpha-glucosidase
EC 2.4.1.1 1,4-alpha-glucan phosphorylase
EC 3.2.1.10 oligo-1,6-glucosidase
Dextran
α4 α4 α4 α4
α6 α6 α6 α6
α6 α6 α6 α6
α6 α6 α6 α6
Glucose
Glucose
Glucose
Glucose
Glucose
EC 3.2.1.41 pullulanase
EC 2.4.1.1 1,4-alpha-glucan phosphorylase
EC 3.2.1.33 amylo-alpha-1,6-glucosidase
Isomaltose
α6
α4
Dextrin
α6
α6
α6
α4 α4
α-limit dextrin
α6
α4
Bacteroides/unclassified Bacteroidales
Bifidobacterium/unclassified Bifidobacteriaceae
Lactobacillus
Bilophila
unclassified Enterobacteriaceae
all other taxa
7
6
5
4
3
2
1
0
RPKM
3.2.1.20 alpha-glucosidase (K01187)
3.2.1.11 Dextranase (GH66)
3.2.1.41 pullulanase (K01200)
3.2.1.70 glucan 1,6-alpha-glucosidase (K01215)
3.2.1.10 oligo-1,6-glucosidase (K01182)
3.2.1.94 glucan 1,6-alpha-isomaltosidase (GH27)
2.4.1.1 1,4-alpha-glucan phosphorylase
2.4.1.25 4-alpha-glucanotransferase
3.2.1.33 amylo-alpha-1,6-glucosidase
blank t24
blank t48
94 t24
94 t48
96 t24
96 t48
27 t24
27 t48
dig27 t24
dig27 t48
Time points


Differences in Bray-Curtis distance
PC2 (15.6% expl.var)
PC1 (58.4% expl.var)
3
2
1
0
−1
−2
−4
−2
0
2
4
6
PRJEB30770
PRJNA396964
PRJNA416988
PRJNA476770
PRJNA509512
PRJNA510123
PRJNA515307
PRJNA526861
PRJNA593787
PRJNA597902
current study
PRJNA526861
PRJNA416988
current study
PRJNA509512
PRJNA593787
PRJNA396964
PRJNA510123
PRJNA597902
PRJNA515307
PRJEB30770
PRJNA476770