Decoding host-microbiome interactions through co-expression network analysis within the non-human primate intestine

Mika Uehara[1], Takashi Inoue[2], Sumitaka Hase[1], Erika Sasaki[2, 4], Atsushi Toyoda[3], Yasubumi Sakakibara[1*]

1) Department of Biosciences and Informatics, Keio University, Yokohama, Kanagawa 223-8522, Japan

2) Department of Marmoset Biology and Medicine, Central Institute for Experimental Animals, Kawasaki, Kanagawa 210-0821, Japan

3) Department of Genomics and Evolutionary Biology, National Institute of Genetics, Mishima, Shizuoka 411-8540, Japan

4) Laboratory for Marmoset Neural Architecture, RIKEN Center for Brain Science, 2-1 Hirosawa, Wako-shi, Saitama 351-0198, Japan.

*Corresponding Author

Yasubumi Sakakibara

3-14-1 Hiyoshi, Kohoku-ku, Yokohama, 223-8522, Japan

Phone/Fax: +81-45-566-1791

E-mail: yasu@bio.keio.ac.jp

The authors declare no competing interests.

ABSTRACT

The gut microbiome affects the health status of the host through complex interactions with the host's intestinal wall. These host-microbiome interactions spatially vary along the physical and chemical environment of the intestine, but these changes remain unknown. This study investigated these intricate relationships through a gene co-expression network analysis based on dual transcriptome profiling of different intestinal sites (cecum, transverse colon, rectum) of the primate common marmoset. We proposed a gene module extraction algorithm based on graph theory to find tightly interacting gene modules of the host and the microbiome from a vast co-expression network. The 27 gene modules identified by this method, which include both host and microbiome genes, not only produced results consistent with previous studies regarding the host-microbiome relationships, but also provided new insights into microbiome genes acting as potential mediators in host-microbiome interplays. Specifically, we discovered associations between the host gene FBP1, a cancer marker, and polysaccharide degradation-related genes (*pfkA*, *fucI*) coded by *Bacteroides vulgatus*, as well as relationships between host genes CD19, CD22, CD79B, PTPN6 involved in B-cell immune responses and a tryptophan synthesis gene (*trpB*) coded by *Parabacteroides distasonis*. Furthermore, our proposed module extraction algorithm surpassed existing approaches by successfully defining more functionally related gene modules, providing insights for understanding the complex relationship between the host and the microbiome.

INTRODUCTION

The mutual interaction between the intestinal microbiome and the host constitutes a complex biological process, with intertwined dependencies essential for survival and health of both. The

interactions between the intestinal microbiome and the host are formed by various factors such as the effects of metabolites produced by bacteria on the host, the binding of bacterial-derived proteins and host receptors, and the binding of bacterial cell surface structures and host receptors. Studies on host-microbiome interaction through bacterial-derived proteins have discovered that a novel protein P9 secreted by *Akkermansia muciniphila* binds to host ICAM-2, thereby increasing the secretion of glucagon-like peptide 1 (GLP-1), suggesting that the interaction between P9 and ICAM-2 could be a potential therapeutic target in type 2 diabetes patients [1]. It has also been reported that bacterial lipopolysaccharides (LPS), cell wall components of gram-negative bacteria, are structurally recognized by the Toll-like receptor 4 (TLR4) complex, inducing innate immune responses [2]. Among these, the host-microbiome interaction mediated by metabolites is particularly well-studied. For instance, primary bile acids, produced in the host's liver cells, are converted into secondary bile acids by the microbiome in the colon. This microbiome-derived secondary bile acid is known to regulate the homeostasis of regulatory T cells, immune suppressive cells in the host [3]. In this way, the intestinal microbiome and host mutually interact and establish a symbiotic relationship while forming a metabolic pathway or physiological response.

To understand such complex interactions between the microbiome and the host, it is beneficial to analyze combined profiles of both the microbiome and the host. A cohort study of patients with colorectal cancer, irritable bowel syndrome, and inflammatory bowel disease investigated host gene expression levels and microbiota composition [4]. While the relationship between disease and microbiome, or disease and host genes, has been investigated individually, these approach combining profiles of both the microbiome and the host have been a significant clue to unravel the relationship among the microbiome, host genes, and diseases.

Recently, by profiling the gene expression of both the host and the microbiome, the functional involvement of the microbiome in the host response has been revealed, extending beyond just the relationship between microbiome members and the host [5, 6, 7]. Gene expression network analysis is suitable for comprehensively deciphering the complex interactions between the host's genes and the diverse microbiome's genes. Some studies have adopted network analysis to highlight the characteristics of host-microbiome interactions, identifying functional gene modules related to various physiological and pathological processes in both the host and microbiome. For instance, a study [5]characterizing the host-microbiome interaction during influenza virus infection identified modules from the host's gene co-expression network. It discovered correlations between host gene modules involved in interferon signal transduction and the sulfonamide antibiotic resistance genes of the microbiome. In another study [6] applying network analysis to the gene expression profiles of the host and microbiome in the nasopharyngeal airway during bronchiolitis, networks were built for both the host and microbiome, then modules were identified, revealing host modules involved in host-T-cell regulation and microbiome modules involved in microbial-BCAA metabolism and oxidative stress response were associated with disease severity. Thus, considering network topology, such as modularity, helps interpret mutual biological processes between the microbiome and host from complex and large networks.

Despite such progress, there remain several limitations to the understanding of the interactions between the host and the intestinal microbiome. First, there's a lack of understanding about the spatial variation in host-microbiome interactions. As the function of the intestinal tract varies by location and the physicochemical environment such as nutrients, oxygen, and pH also

differ [8, 9], the mechanisms of interaction between host gut cells and the microbiome could vary as well. Indeed, studies using animal models have reported that the composition and functional fluctuation of the microbiome varies due to environmental changes depending on intestinal location (biogeographical location) [10, 11, 12, 13]. However, these studies focused solely on the microbiome, and the coordinated spatial variation of both the microbiome and host gut cells along the intestinal tract has yet to be clarified.

Second, there's a lack of computational methods for extracting significant modules (subgraphs) from the host-microbiome gene expression network. Analyzing a massive gene expression network integrating the gene expression profiles of both the host and microbiome helps interpret the interactions between the genes of the host and the microbiome. Especially in a huge gene expression network, it is crucial to extract densely connected subgraphs (modules) to divide the network into understandable units; however, existing algorithms for extracting modules from the network are insufficient for accurately capturing the interaction between the host and microbiome because they were not designed with the purpose of understanding interactions between organisms [14, 15].

Lastly, there are concerns regarding the applicability of the results of studies using animal models to humans. Mouse models have been the basis of biological research for a long time due to ethical constraints, experiments in controlled environments, and considerations of reproducibility, but there are several issues and limitations to their use. The physiology, anatomy, and pharmacology of mice differ from those of humans, and there are differences in the composition and function of the intestinal microbiome between mice and humans [13, 16]. Particularly considering the complexity of host-microbiome interactions, it is challenging to precisely replicate the relationship between humans and their intestinal microbiome. In contrast,

116 the common marmoset, a small New World primate, is considered a useful model in preclinical
117 trials because it is genetically close to humans and has common physiological and anatomical
118 characteristics [17]. Moreover, the common marmoset is the only non-human primate in which a
119 germ-free state has been successfully achieved [18], and it has a fecal microbiome similar to that
120 of humans compared to other animal models, including mice [13], potentially broadening the
121 scope of intestinal microbiome research.

122 To address these challenges, this study aimed to reveal the interactions between the host
123 and the microbiome along multiple sites of the common marmoset colon (cecum, transverse
124 colon, rectum) by developing a module extraction method applicable to the gene expression
125 network that represents the complex interactions between the host and the microbiome. Through
126 simultaneous profiling of host and microbiome transcripts in the intestinal sites of the common
127 marmoset, we constructed an integrated gene expression network, incorporating both host and
128 microbiome genes. From this network, we further identified densely covariant subgraphs (gene
129 modules) to interpret the interactions between genes. As conventional methods for extracting
130 gene modules from the network were insufficient for interpreting the interactions between
131 organisms examined in this study, we proposed a module extraction algorithm based on graph
132 theory. Our proposed method successfully extracted more modules containing functionally
133 related genes of both the host and microbiome compared to existing algorithms, underlining its
134 suitability for predicting host-microbiome interactions. This method extracted 27 gene modules
135 containing both host and microbiome, indicating that the gut microbiome is functionally closely
136 involved in the variation of host gene expression along the intestinal tract.

137

138 **RESULTS**

139 ***Overview of the analysis pipeline integrating host and microbiome transcriptomes:***

140 ***Identification of closely interacting gene modules from gene expression networks***

141 To investigate the host and microbiome interactions along the intestinal tract, we proposed an
142 analysis method integrating the host transcriptome and the microbiome metatranscriptome
143 (Figure 1). First, we computed gene expression profiles from the host transcriptome and
144 microbiome metatranscriptome obtained via dual RNA-seq and created three types of gene
145 expression networks: host-host, host-microbiome, and microbiome-microbiome. By unifying
146 these three networks, we constructed a host and microbiome gene co-expression network, with
147 the genes of the host and microbiome serving as nodes, and covariant genes interconnected.
148 Second, we unified gene modules containing closely interacting genes from both the host and
149 microbiome from the gene co-expression network. Conventional algorithms for extracting
150 subgraphs from the network, such as modularity maximization, may not be suitable for
151 accurately capturing interactions between the host and microbiome, as they were not designed
152 with the purpose of understanding inter-organism interactions. Therefore, we proposed a clique-
153 search-based module extraction method that takes into account the structure of three networks:
154 host-host, host-microbiome, and microbiome-microbiome. This method begins by exploring
155 cliques from the gene co-expression network. For overlapping cliques, it calculates the
156 cliqueness in each of the original three networks and merges them if the cliqueness of all
157 networks exceeds a certain threshold, thereby obtaining the final gene modules (refer to the
158 Methods section for details).

159

160 ***Computing gene expression profiles of the host and the microbiome from dual RNA-seq data***

161 Firstly, we extracted high-quality RNA (RNA integrity number (RIN) scores ≥8.6;
162 Supplementary Table S1) from the intestinal sites (cecum, transverse colon, rectum) of five
163 common marmosets and carried out dual RNA-seq with NovaSeq 6000 v1.5 (100 bp PE). The
164 obtained RNA read data amounted to a total of 4,724 M reads (mean 157,450,033 paired-end
165 reads/sample; Supplementary Table S2).

166 Next, as the RNA read data obtained by dual RNA-seq contains a mix of reads derived
167 from both the host and the microbiome, it is necessary to separate these two read data in silico
168 (Supplementary Table S3). The pipeline established in this study separates the host-derived read
169 data and the microbiome-derived read data by aligning the RNA reads to the host reference
170 genome (Fig. 1B). To avoid cross-contamination of read data between organisms at this stage, we
171 adjusted the pipeline parameters using simulated RNA read data from the host and the
172 microbiome. As a result, we achieved a precision of 1.000 in separating the read data of the host
173 and the microbiome, and our established pipeline completely separated the reads derived from
174 the host and those from the microbiome (Supplementary Table S4, Supplementary Note Section
175 3).

176 After the above preprocessing, we calculated the gene expression levels for both the host
177 and the microbiome. As a result, we detected 16,578 host expressed genes and 8,533 microbiome
178 expressed genes in total (Supplementary Table S5 and S6, Supplementary Note Section 4 and 5).
179 To examine the expression patterns between different sites and between different individuals for
180 both host genes and microbiome genes, we applied Principal Component Analysis (PCA) to the
181 gene expression profiles for visualization. In the host profiles, distinct clusters were formed for
182 each site, while the clusters for each individual overlapped (Fig. 2A and B). In the microbiome
183 profiles, clusters were separated for the cecum and rectum, which are the beginning and end of

184 the colon, respectively, with the transverse colon positioned between the cecum and rectum
185 overlapping between the two clusters (Fig. 2C). The clusters for each individual also overlapped
186 in the microbiome profiles (Fig. 2D). These results suggest that both host and microbiome gene
187 expression profiles are more variable across different sites compared to those across individuals.
188

189 ***Spatial variations in host gene expression among intestinal sites***

190 Of the 16,578 host genes detected in three regions of the common marmoset's intestinal tract
191 (cecum, transverse colon, and rectum), a total of 2,615 genes showed significant expression
192 variations among the sites (FDR<0.01). Furthermore, upon conducting functional pathway
193 analysis, 40 pathways were identified, all of which were significantly upregulated in the cecum
194 compared to the transverse colon (normalized enrichment score >1.5, FDR<0.1; Supplementary
195 Table S7). In particular, pathways with high normalized enrichment scores, such as Fc epsilon RI
196 signaling pathway, T cell receptor signaling pathway, leukocyte transendothelial migration,
197 Yersinia infection, and B cell receptor signaling pathway were detected, indicating that pathways
198 related to immune functions were upregulated in the cecum relative to the transverse colon (Fig.
199 3A). Additionally, a substantial number of genes involved in these functional pathways were
200 found to be significantly upregulated in the cecum (FDR<0.01). For example, among the genes
201 specifically upregulated in the cecum, we found genes related to B cell receptor signaling. These
202 include CD19, which acts as a coreceptor for the B cell antigen receptor complex and augments
203 B cell activation [19]; CD79A and CD79B, components of the B cell antigen receptor that
204 generate signals upon antigen recognition [20]; and CD22 and CD72, which inhibit B cell
205 antigen receptor signal transduction [21, 22]. In addition, genes associated with T cell receptor
206 signaling were also found as being highly expressed specifically in the cecum. These include

207 CD28, a co-stimulatory receptor for T cell receptors that activates T cells [23]; CD3D, a
208 component of CD3 that plays a role in signal transduction during T cell activation [24];
209 CD40LG, which is induced on the T cell surface following T cell receptor (TCR) activation [25];
210 and CTLA4, which attenuates T cell activation and fine-tunes immune responses [26] (Fig. 3B).
211

212 ***Interpreting host-microbiome interactions through spatial co-expression gene network***
213 ***analysis within the intestine***

214 We constructed the unified gene expression network composed of 3,219 gene nodes, which
215 embodies co-expression information on host-host, host-microbiome, and microbiome-
216 microbiome interactions. Subsequently, to identify densely co-expressed gene sets within this
217 network, we implemented the clique-based module extraction algorithm proposed in this study.
218 In graph theory, a "clique" refers to a subgraph within the network where all nodes are fully
219 connected. Our algorithm discerned modules of interacting genes in the gene co-expression
220 network as aggregates of overlapping cliques. As a result, we then identified 27 gene modules
221 containing both host and microbiome genes (Supplementary Table S8). Furthermore, upon
222 conducting a functional pathway enrichment analysis on these gene modules, we found that 19 of
223 them, representing 70.4%, could be functionally characterized (Supplementary Note Section 6).
224 The modules we identified were characterized by functional pathways including the B
225 cell receptor signaling pathway, fructose and mannose metabolism, oxidative phosphorylation,
226 and biosynthesis of amino acids (Fig. 4). In addition, we provided visualizations of the bacterial
227 species that contribute to microbiome gene expression within these identified modules, together
228 with host and microbiome genes (Fig.5, 6, 7 and S2). The results obtained from these identified
229 modules were not only in agreement with previous research examining host genes, microbiome

genes, and bacterial species but also revealed novel interactions between host genes and the microbiome, including the identification of microbiome genes that serve as intermediaries between host genes and the microbiome.

Module_06, as shown in Figure 5, was identified as a module with a positive correlation between the host gene Fructose-1,6-bisphosphatase 1 (FBP1), which is highly expressed in the transverse colon, and 40 microbiome genes. These 40 gut microbiome genes included those associated with fructose-mannose metabolism, such as K01840 (manB), K00850 (pfkA), and K01818 (fucI), as well as those related to oxidative phosphorylation, such as K00241 (sdhC, frdC), K00426 (cydB), K02112 (atpD), and K02121 (ntpE, atpE). The most substantial contributor to the expression of these microbiome genes was *Bacteroides vulgatus*. FBP1 is known to have reduced expression in various tumor tissues, including the esophagus [27], liver [28], lung [29], and ovarian cancer [30]. It has been confirmed that its expression increases with fructose treatment in intestinal organoids [31]. Previous studies have identified FBP1 as a differentially expressed gene in the small intestine of germ-free mice, suggesting that the intestinal microbiome influences the expression of the host's FBP1 gene [32], but the bacterial species and genes affecting the expression of the FBP1 gene have not been identified. *Bacteroides vulgatus*, which contributed to the expression of fructose-mannose metabolism-related microbiome genes observed in Module_06, has been reported to release fructose by decomposing polysaccharides [33, 34]. In this module, the gene K00850 (*pfkA*) involved in the fructo-oligosaccharide degradation and the gene K01818 (*fucI*) involved in the 2'-fucosyllactose degradation were detected, revealing a novel association between polysaccharide degradation genes encoded by *Bacteroides vulgatus* and the host gene FBP1.

252 In Module_01, the B cell receptor signaling pathway was enriched (Fig. 6). This module
253 featured covariation between host genes involved in B cell immune responses—including CD19,
254 CD22, CD79B, PTPN6, which were highly expressed in the cecum—and the tryptophan
255 synthesis gene K01696 (*trpB*). This tryptophan synthesis gene K01696 (*trpB*) expression was
256 contributed to by *Parabacteroides distasonis*. Tryptophan is an essential amino acid that, upon
257 uptake by intestinal cells and subsequent metabolism, can activate the Aryl hydrocarbon receptor
258 and influence B cell immune responses [35]. Tryptophan is also known to ameliorate
259 autoimmune diseases, such as inflammatory bowel disease [36]. *Parabacteroides distasonis* has
260 been suggested to be involved in various autoimmune conditions in several studies [37, 38]. For
261 instance, *Parabacteroides distasonis* levels have been shown to be lower in multiple sclerosis
262 (MS) patients compared to healthy controls, and transplanting the intestinal microbiome from
263 MS patients to germ-free mice resulted in increased severity of autoimmune encephalomyelitis
264 symptoms [37]. While the associations between tryptophan, B-cell immune responses, and
265 autoimmune diseases, as well as between *Parabacteroides distasonis* and autoimmune diseases,
266 are already reported, our study provides additional insight. We suggest that tryptophan
267 synthesized by *Parabacteroides distasonis* could exert an influence on the upregulation of B-cell
268 immune response-related genes in the cecum, including CD19, CD22, CD79B, and PTPN6.
269 In Module_19, a group consisting of three host genes and 40 microbiome genes was
270 identified. The three host genes included MMP10, two uncharacterized genes LOC103795407
271 and CCDC162P, all of which showed high expression in the cecum (Fig. 7). The microbiome
272 genes that showed a positive correlation with these genes included K01491 (folD), K00605
273 (gcvT), K01803 (*tpiA*), K00831 (*serC*), K15633 (*gpmI*), K00024 (*mdh*), all of which are
274 involved in carbon metabolism pathways (Supplementary Fig. S1). *Bacteroides vulgatus* was

identified as the bacteria contributing to the expression of these microbiome genes. The MMP10 gene plays a crucial role in maintaining epithelial barrier function and is necessary for resolving colonic epithelial damage [39]. A study examining the host transcriptome and microbiota composition of colorectal cancer tumors found high expression of MMP10 in tumor tissues and a significant increase in the genus *Bacteroides*, indicating a link between the host gene MMP10 and the genus *Bacteroides* [40]. These findings are consistent with our observations. Furthermore, in the identified Module_19, not only the relationship between the host gene MMP10 and *Bacteroides vulgatus*, but also a group of microbiome genes related to carbon metabolism pathways showing a positive correlation with the host gene MMP10 was identified. One hypothesis that suggests a positive correlation between these carbon metabolism-related genes coded by *Bacteroides vulgatus* and the host gene MMP10 is that the microbiome might form colonies in the host's mucosal layer. Mucin, the main component of mucus generated by the gastrointestinal mucosal epithelium, contains various sugars such as glucose and fructose [41]. In addition to promoting bacterial settlement and adhesion, it also serves as a nutrient source for bacteria. In this Module_19, it was observed that the microbiome gene group, including K01803 (*tpiA*), K00831 (*serC*), K15633 (*gpmI*) related to the glycolysis of glucose, fructose, and galactose, positively correlated with the host gene MMP10, which is upregulated specifically in the cecum. This suggests the possibility that in the cecum, the microbiome, including *Bacteroides vulgatus*, forms colonies using mucin as a nutrient source, and the high expression of the host's MMP10 gene protects the epithelium from the colonization of the microbiome. In this Module_19, we observed that a group of peptide genes such as K01803 (*tpiA*), K00831 (*serC*), and K15633 (*gpmI*) related to glycolysis of glucose, fructose, and galactose were positively correlated with the host gene MMP10, which was up-regulated specifically in the

298 cecum. This observation suggests that in the cecum, microbiome including *Bacteroides vulgatus*
299 colonize the mucin, the main mucus component, as a nutrient source, and that the host MMP10
300 gene is highly expressed to protect the epithelium from microbiome colony formation.

301 In Module_16, three host genes and 54 microbiome genes were identified, and
302 furthermore, the expression of these microbiome genes was influenced by five bacterial species.
303 The relationships between these host genes and microbiome genes, as well as between the host
304 genes and bacterial species, were consistent with existing knowledge about bile acid synthesis
305 and degradation (Fig. 8). This module included the host gene CYP27A1, which is involved in
306 bile acid production [42], and NR1H3, a host gene that codes for a bile acid receptor [43]. These
307 genes were upregulated in the cecum and transverse colon compared to the rectum. The
308 microbiome genes interacting with these host genes included K03664 (*smpB*), a gene essential
309 for bacterial survival when upregulated in the presence of secondary bile acid deoxycholic acid
310 (DCA), K02835 (*prfA*) which positively regulates the expression of a gene coding for bile acid
311 hydrolase, and K03527 (*ispH*, *lytB*), a gene involved in bile acid sensitivity. Additionally,
312 *Bacteroides vulgatus* contributed to the expression of all these microbiome genes, which aligns
313 with previous studies showing an increase in *Bacteroides* spp. correlating with the expression of
314 the host gene NR1H3 [44].

315

316 **DISCUSSION**

317 In this study, we investigated the interactions between the host and intestinal microbiome in the
318 common marmoset's intestinal sites (cecum, transverse colon, and rectum) by utilizing dual
319 transcriptome profiling and subsequent gene co-expression network analysis. To discern modules
320 of tightly interacting host and microbiome genes from a vast gene co-expression network with

321 nodes representing both host and microbiome genes, we proposed a clique-based algorithm that
322 considers the interconnectedness of nodes (cliqueness), in three different networks: host-host,
323 host-microbiome, and microbiome-microbiome. As a result, we identified 27 co-expression
324 modules composed of both host and microbiome genes, with 70.4% of them, or 19 modules,
325 being functionally characterized by enrichment analysis (Fisher's exact test, FDR<0.05;
326 Supplementary Table S9).

327 To compare our method with conventional module detection methods, we applied the
328 Newman [14] and Louvain [15] algorithms, both of which maximize modularity, to the same
329 gene co-expression network. Upon conducting pathway enrichment analyses on the modules
330 extracted by these conventional methods, the proportion of functionally characterized modules
331 was only 22.2% and 50.0% respectively, which is lower than what we achieved with our method
332 (Supplementary Table S10). These results underscore the success of our method in defining a
333 larger number of functionally interpretable gene modules compared to conventional approaches
334 and emphasize its suitability for predicting host-microbiome interactions, over conventional
335 modularity-maximizing approaches.

336 The gene modules extracted, which encompassed both host and microbiome genes,
337 frequently exhibited relationships where a small number of genes regulated a larger number of
338 genes between the host and microbiome genes. These observations are thought to reflect the
339 characteristics of the interactions between the host and the microbiome. Several studies have
340 reported the impact of host gene mutations on the intestinal microbiome. For instance, it has
341 been reported that the host gene VDR, which codes for the Vitamin D receptor, influences the
342 colonization of *Parabacteroides* [45], and that the host gene ABO is associated with
343 *Faecalicatena lactaris* [46]. In another study investigating the relationship between the intestinal

344 microbiome and host gene loci through quantitative trait locus analysis, it was suggested that
345 *Akkermansia muciniphila* regulates the expression of ATF3, which codes for a transcription
346 factor that plays a crucial role in metabolic and immune regulation, and this bacterium is related
347 to only four genes [47]. These studies highlight the fact that a single or a small number of host
348 genes are controlling, or are controlled by, the intestinal microbiome. Conversely, a small
349 number of microbiome genes can affect the host's metabolic [48], fitness [49], and longevity [50]
350 phenotypes. For example, structural mutations in the gene related to butyrate transport in
351 *Anaerostipes hadrus* are strongly associated with a reduction in the host's metabolic risk [48],
352 and the intestinal microbiome's thiamine biosynthesis genes, transferred by horizontal
353 transmission, influence the host's fitness [49]. Considering these previous studies and our results,
354 it is suggested that the interaction between the host and the microbiome involves widespread
355 influence of one's genes on the activity of the other.

356 The common marmoset used in this study bears anatomical and pharmacological
357 similarities to humans [17], and the fecal microbiome of the common marmoset is more similar
358 to that of humans compared to other animal models [13]. While the common marmoset's
359 cecum—one of the study's targeted intestinal sites—does not possess an appendage as
360 pronounced as the human appendix, it is known to share significant similarities with the human
361 appendix in terms of microanatomical structure and immune reactivity [51, 52]. The results of
362 this study using common marmosets could therefore hold higher applicability to humans
363 compared to other animal models.

364 In conclusion, we union host transcriptome and microbiome metatranscriptome profiles
365 of the common marmoset and performed gene co-expression network analysis. The algorithm we
366 proposed for extracting gene modules in the gene co-expression network enabled the

367 identification of dense, functionally related host-microbiome co-expression modules, which was
368 challenging using the conventional module maximization algorithms. The identified gene
369 modules provided new insights into relationships between host genes and intestinal microbiome
370 and microbial genes bridging interactions between the host and bacteria. Our findings underscore
371 that interactions between the host and microbiome are not uniform throughout the intestine but
372 are localized, indicating spatial variation in host-microbiome interactions along the gut. These
373 insights could be crucial in elucidating mechanisms of localized gut diseases, such as colorectal
374 cancer and inflammatory bowel disease, and in developing localized therapies.
375

376 **MATERIALS AND METHODS**

377 ***Animal experiments***

378 Common marmosets were housed at the Central Institute for Experimental Animals (Kawasaki,
379 Japan). Five marmosets were euthanized with intravenous administration of pentobarbital
380 overdose, and the digestive tract was isolated (Supplementary Table S11, Supplementary Note
381 Section 1). Three gastrointestinal site samples (cecum, transverse colon and rectum), including
382 the mucosal layer, were used for dual RNA-seq. These three sites were targeted at the beginning,
383 middle and end of the colon, which has an abundant microbiome. The animal experiment
384 protocol was approved by the CIEA Institutional Animal Care and Use Committee (approval no.
385 17031, 18032, and 21084). The study was conducted in accordance with the guidelines of CIEA
386 that comply with the Guidelines for Proper Conduct of Animal Experiments published by the
387 Science Council of Japan. Animal care was conducted in accordance with the Guide for the Care
388 and Use of Laboratory Animals (Institute for Laboratory Animal Resources, 2011).
389

390 ***RNA extraction and dual RNA sequencing of host and microbiome***

391 RNA libraries were prepared after removing rRNA from the extracted high-quality Total RNA
392 (RNA integrity number (RIN) scores ≥8.6; Supplementary Table S1). Illumina NovaSeq 6000
393 v1.5 sequencing yielded a total of 472 Gnt of paired-end reads (100 nt × 2). This dataset included
394 an average of 314.9 M reads per sample (Supplementary Table S2, Supplementary Note Section
395 2).

396

397 ***Construction of host-microbiome co-expression network and identification of gene modules***

398 **Network analysis**

399 In this study, we utilized gene co-expression network analysis to identify interacting gene
400 modules of host and microbiome. We established three types of networks: host-host, host-
401 microbiome, and microbiome-microbiome, wherein genes were considered as nodes and co-
402 expressing genes were interconnected (Fig. 1D). Networks were defined by an adjacency matrix,
403 where the adjacency a_{ij} between gene i and gene j was defined as follows based on co-
404 expression similarity s_{ij}:

405 $$a_{ij} = \begin{cases} 1\text{, if } s_{ij} > \alpha \\ 0\text{, otherwise} \end{cases}$$

406 The co-expression similarity s_{ij} was computed using Spearman's correlation coefficients for
407 host-host and microbiome-microbiome networks, and absolute values of Spearman's correlation
408 coefficients for host-microbiome networks. The adjacency threshold α was set at 0.80. These
409 three types of networks were then integrated into one.

410 **Identification of gene modules**

411 We developed an algorithm to identify densely interacting gene modules from the host-
412 microbiome co-expression network. This algorithm, based on clique finding, progressively

merges cliques (Fig. 1E). In graph theory, a "clique" refers to a subgraph within the network where all nodes are fully connected. In the context of a gene co-expression network, all genes within a clique indicate a strong co-expression (correlation) amongst them. In this study, we focus only on networks and cliques that include gene nodes from both the host and microbiome, encompassing three types of networks: host-host, host-microbiome, and microbiome-microbiome. Our algorithm identified modules of interacting genes in the gene co-expression network as aggregates of overlapping cliques. It was implemented in three steps: (1) Cliques within the network were found; (2) Beginning with the largest cliques, we examined their overlap with any other cliques. If overlap occurred, we computed the cliqueness scores independently for each of the three types of networks (host-host, host-microbiome, and microbiome-microbiome networks) upon merging the overlapping cliques. The cliqueness scores for the host-host, host-microbiome, and microbiome-microbiome networks were denoted as C_{hh}, C_{hm}, and C_{mm}, respectively; (3) Cliques were merged if they satisfied the condition: $C_{hh} > \tau$, $C_{mm} > \tau$, and $C_{mm} > \tau$. The threshold τ for cliqueness was set at 0.60 based on parameter determination (Supplementary Note Section 7). If there were multiple cliques with minimum cliqueness scores exceeding the threshold, the clique with the highest cliqueness score was selected. Steps (2) and (3) were repeated until no cliques could be merged.

For the host-host networks the cliqueness score C_{hh} between cliques A and B was defined as:

$$C_{hh} = \frac{E}{\frac{1}{2}N(N-1)}$$

Here, N is defined as the number of host gene nodes in the union of cliques A and B, represented as $N = |N_A \cup N_B|$. N_A and N_B denote the number of host gene nodes within clique A and clique B, respectively. E corresponds to the total number of edges between the host nodes included in

435 cliques A and/or B, formulated as $E = E_A + E_B + E_{AB}$. Here, E_A and E_B represent the number of
436 edges between host gene nodes within clique A and clique B, respectively, while E_{AB} denotes the
437 number of edges connecting across host gene nodes within clique A and clique B. The cliqueness
438 score for microbiome-microbiome networks C_{mm} was calculated in the same way.
439 For the host-microbiome network, the cliqueness score C_{hm}, was defined as:

440 $$C_{hm} = \frac{E}{N_{host} N_{microbiome}}$$

441 where N_{host} and $N_{microbiome}$ represent the number of host and microbiome gene nodes
442 respectively in the union of cliques A and B. E corresponds to the total number of edges between
443 the host or microbiome gene nodes included in cliques A and/or B.
444

445 ***Acknowledgements***
446 We thank the scientists of the Central Institute for Experimental Animals for providing support for
447 animal experiments and the members of the Sakakibara Lab at Keio University for helpful
448 discussions.
449 This work was supported by grants from the Japan Agency for Medical Research and Development
450 (AMED PRIME JP19gm6010006). M. Uehara has received funding from JSPS KAKENHI Grant
451 Numbers JP20J21477.
452

453 ***Author contribution***
454 M.U. performed the experiments, conducted the bioinformatics analysis and cowrote the paper;
455 T.I. and E.S. provided common marmoset samples; M.U. and S.H. performed RNA extraction; A.T.
456 performed deep sequencing with high-throughput sequencer; Y.S. designed and supervised the
457 research, analysed the data, and cowrote the paper. All authors have read and approved the

manuscript.

Competing Interests

The authors declare no competing interests.

Data and source code availability

All raw sequence data have been submitted to the DDBJ under project PSUB020951 from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

467 **REFERENCES**

468 1. Yoon HS, Cho CH, Yun MS, Jang SJ, You HJ, Kim JH, et al. Akkermansia muciniphila
469 secretes a glucagon-like peptide-1-inducing protein that improves glucose homeostasis
470 and ameliorates metabolic disease in mice. Nat Microbiol. 2021;6:563–73.
471 2. Vatanen T, Kostic AD, D'Hennezel E, Siljander H, Franzosa EA, Yassour M, et al.
472 Variation in Microbiome LPS Immunogenicity Contributes to Autoimmunity in Humans.
473 Cell. 2016;165:842–53.
474 3. Song X, Sun X, Oh SF, Wu M, Zhang Y, Zheng W, et al. Microbial bile acid metabolites
475 modulate gut RORγ+ regulatory T cell homeostasis. Nature. 2020;577:410–5.
476 4. Priya S, Burns MB, Ward T, Mars RAT, Adamowicz B, Lock EF, et al. Identification of
477 shared and disease-specific host gene–microbiome associations across human diseases
478 using multi-omic integration. Nat Microbiol. 2022;7:780–95.
479 5. Zhang L, Forst CV, Gordon A, Gussin G, Geber AB, Fernandez PJ, et al. Characterization
480 of antibiotic resistance and host-microbiome interactions in the human upper respiratory
481 tract during influenza infection. Microbiome. 2020;8:1–12.
482 6. Fujiogi M, Raita Y, Pérez-Losada M, Freishtat RJ, Celedón JC, Mansbach JM, et al.
483 Integrated relationship of nasopharyngeal airway host response and microbiome
484 associates with bronchiolitis severity. Nat Commun. 2022;13:4970.
485 7. Lloyd-Price J, Arze C, Ananthakrishnan AN, Schirmer M, Avila-Pacheco J, Poon TW, et
486 al. Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. Nature.
487 2019;569:655–62.

488 8. Vertzoni M, Augustijns P, Grimm M, Koziolek M, Lemmens G, Parrott N, et al. Impact of
489 regional differences along the gastrointestinal tract of healthy adults on oral drug
490 absorption: An UNGAP review. Eur J Pharm Sci. 2019;134:153–75.
491 9. Ilhan ZE, Marcus AK, Kang DW, Rittmann BE, Krajmalnik-Brown R. pH-mediated
492 microbial and metabolic interactions in fecal enrichment cultures. mSphere.
493 2017;2:e00047-17.
494 10. Yasuda K, Oh K, Ren B, Tickle TL, Franzosa EA, Wachtman LM, et al. Biogeography of
495 the intestinal mucosal and lumenal microbiome in the rhesus macaque. Cell Host
496 Microbe. 2015;17:385–91.
497 11. Kozik AJ, Nakatsu CH, Chun H, Jones-Hall YL. Comparison of the fecal, cecal, and
498 mucus microbiome in male and female mice after TNBS-induced colitis. PLoS One.
499 2019;14:1–14.
500 12. Zhang L, Wu W, Lee YK, Xie J, Zhang H. Spatial heterogeneity and co-occurrence of
501 mucosal and luminal microbiome across swine intestinal tract. Front Microbiol.
502 2018;9:48
503 13. Uehara M, Inoue T, Kominato M, Hase S, Sasaki E, Toyoda A, et al. Intraintestinal
504 Analysis of the Functional Activity of Microbiomes and Its Application to the Common
505 Marmoset Intestine. mSystems. 2022;7:e00520-22.
506 14. Newman ME. Detecting community structure in networks. Eur Phys J B. 2004;38:321–
507 30.
508 15. Blondel VD, Guillaume JL, Lambiotte R, Lefebvre E. Fast unfolding of communities in
509 large networks. J Stat Mech. 2008;10008.

510 16. Hugenholtz F, de Vos WM. Mouse models for human intestinal microbiota research: a
511 critical evaluation. Cell Mol Life Sci. 2018;75:149–60.
512 17. Kishi N, Sato K, Sasaki E, Okano H. Common marmoset as a new model animal for
513 neuroscience research and genome editing technology. Dev Growth Differ. 2014;56:53–
514 62.
515 18. Inoue T, Okahara N, Okahashi N, Sato K, Ueda M, Atarashi K, et al. 2021. Generation of
516 germ-free common marmosets. Research Square. Preprint.
517 https://assets.researchsquare.com/files/rs-428622/v1_covered.pdf?c=1631871455
518 19. Wang K, Wei G, Liu D. CD19: a biomarker for B cell development, lymphoma diagnosis
519 and therapy. Exp Hematol Oncol. 2012;1:1–7.
520 20. Huse K, Bai B, Hilden VI, Bollum LK, Våtsveen TK, Munthe LA, et al. Mechanism of
521 CD79A and CD79B support for IgM+ B cell fitness through B cell receptor surface
522 expression. J Immunol. 2022;209:2042–53.
523 21. Clark EA, Giltiay N V. CD22: a regulator of innate and adaptive B cell responses and
524 autoimmunity. Front Immunol. 2018;9:2235.
525 22. Pan C, Baumgarth N, Parnes JR. CD72-deficient mice reveal nonredundant roles of
526 CD72 in B cell development and activation. Immunity. 1999;11:495–506.
527 23. M.L. A, K.A. F, C.B. T. T-cell regulation by CD28 and CTLA-4. Nat Rev Immunol.
528 2001;1:220–8.
529 24. Yang Y, Zang Y, Zheng C, Li Z, Gu X, Zhou M, et al. CD3D is associated with immune
530 checkpoints and predicts favorable clinical outcome in colon cancer. Immunotherapy.
531 2020;12:25–35.

532 25. Hubbard N, Hagin D, Sommer K, Song Y, Khan I, Clough C, et al. Targeted gene editing
533 restores regulated CD40L function in X-linked hyper-IgM syndrome. Blood.
534 2016;127:2513–22.
535 26. Sobhani N, Tardiel-Cyril DR, Davtyan A, Generali D, Roudi R, Li Y. CTLA-4 in
536 regulatory T cells for cancer immunotherapy. Cancers. 2021;13:1–18.
537 27. He Y, Hua R, Li B, Gu H, Sun Y, Li Z. Loss of FBP1 promotes proliferation, migration,
538 and invasion by regulating fatty acid metabolism in esophageal squamous cell carcinoma.
539 Aging. 2021;13:4986–8.
540 28. Hirata H, Sugimachi K, Komatsu H, Ueda M, Masuda T, Uchi R, et al. Decreased
541 expression of fructose-1,6-bisphosphatase associates with glucose metabolism and tumor
542 progression in hepatocellular carcinoma. Cancer Res. 2016;76:3265–76.
543 29. Dong Y, Huaying S, Danying W, Chihong Z, Ruibin J, Xiaojiang S, et al. Significance of
544 methylation of FBP1 gene in non-small cell lung cancer. Biomed Res Int. 2018;2018:13–
545 9.
546 30. Li H, Qi Z, Niu Y, Yang Y, Li M, Pang Y, et al. FBP1 regulates proliferation, metastasis,
547 and chemoresistance by participating in C-MYC/STAT3 signaling axis in ovarian cancer.
548 Oncogene. 2021;40:5938–49.
549 31. Hou Y, Wei W, Guan X, Liu Y, Bian G, He D, et al. A diet-microbial metabolism
550 feedforward loop modulates intestinal stem cell renewal in the stressed gut. Nat
551 Commun. 2021;12:271.
552 32. Mardinoglu A, Shoaie S, Bergentall M, Ghaffari P, Zhang C, Larsson E, et al. The gut
553 microbiota modulates host amino acid and glutathione metabolism in mice. Mol Syst
554 Biol. 2015;11:1–15.

555 33. Medina DA, Pinto F, Ovalle A, Thomson P, Garrido D. Prebiotics mediate microbial
556 interactions in a consortium of the infant gut microbiome. Int J Mol Sci. 2017;18:2095.
557 34. Cheng J, Hu J, Geng F, Nie S. Bacteroides utilization for dietary polysaccharides and
558 their beneficial effects on gut health. Food Sci Hum Wellness. 2022;11:1101–10.
559 35. Rosser EC, Piper CJM, Matei DE, Blair PA, Rendeiro AF, Orford M, et al. Microbiota-
560 derived metabolites suppress arthritis by amplifying aryl-hydrocarbon receptor activation
561 in regulatory B cells. Cell Metab. 2020;31:837-851.
562 36. Ding X, Ding X, Ding X, Bin P, Bin P, Bin P, et al. Tryptophan metabolism, regulatory T
563 cells, and inflammatory bowel disease: a mini review. Mediators Inflamm. 2020;2020.
564 37. Ezeji JC, Sarikonda DK, Hopperton A, Erkkila HL, Cohen DE, Martinez SP, et al.
565 Parabacteroides distasonis: intriguing aerotolerant gut anaerobe with emerging
566 antimicrobial resistance and pathogenic and probiotic roles in human health. Gut
567 Microbes. 2021;13:1–27.
568 38. Cekanaviciute E, Yoo BB, Runia TF, Debelius JW, Singh S, Nelson CA, et al. Gut
569 bacteria from multiple sclerosis patients modulate human T cells and exacerbate
570 symptoms in mouse models. Proc Natl Acad Sci USA. 2017;114:10713–8.
571 39. Łaniewski P, Herbst-Kralovetz MM. Bacterial vaginosis and health-associated bacteria
572 modulate the immunometabolic landscape in 3D model of human cervix. NPJ Biofilms
573 Microbiomes. 2021;7:1–17.
574 40. Tomkovich S, Gharaibeh RZ, Dejea CM, Pope JL, Jiang J, Winglee K, et al. Human
575 colon mucosal biofilms and murine host communicate via altered mRNA and microRNA
576 expression during cancer. mSystems. 2020;5:e00451-19.

577 41. Cornick S, Tawiah A, Chadee K. Roles and regulation of the mucus barrier in the gut.
578 Tissue Barriers. 2015;3:1–2.
579 42. Zurkinden L, Sviridov D, Vogt B, Escher G. Downregulation of Cyp7a1 by cholic acid
580 and chenodeoxycholic acid in Cyp27a1/ApoE double knockout mice: differential
581 cardiovascular outcome. Front Endocrinol. 2020;11:1–13.
582 43. Biagioli M, Marchianò S, Carino A, Di Giorgio C, Santucci L, Distrutti E, et al. Bile
583 acids activated receptors in inflammatory bowel disease. Cells. 2021;10:1281.
584 44. Badi SA, Motahhary A, Bahramali G, Masoumi M, Khalili SFS, Ebrahimzadeh N, et al.
585 The regulation of Niemann-Pick C1-Like 1 (NPC1L1) gene expression in opposite
586 direction by Bacteroides spp. and related outer membrane vesicles in Caco-2 cell line. J
587 Diabetes Metab Disord. 2020;19:415–22.
588 45. Wang J, Thingholm LB, Skiecevičie J, Rausch P, Kummen M, Hov JR, et al. Genome-
589 wide association analysis identifies variation in Vitamin D receptor and other host factors
590 influencing the gut microbiota. Nat Genet. 2016;48:1396–406.
591 46. Qin Y, Havulinna AS, Liu Y, Jousilahti P, Ritchie SC, Tokolyi A, et al. Combined effects
592 of host genetics and diet on human gut microbiota and incident disease in a single
593 population cohort. Nat Genet. 2022;54:134–42.
594 47. Zhang Q, Linke V, Overmyer KA, Traeger LL, Kasahara K, Miller IJ, et al. Genetic
595 mapping of microbial and host traits reveals production of immunomodulatory lipids by
596 Akkermansia muciniphila in the murine gut. Nat Microbiol. 2023;8:424–40.
597 48. Zeevi D, Korem T, Godneva A, Bar N, Kurilshikov A, Lotan-Pompan M, et al. Structural
598 variation in the gut microbiome associates with host health. Nature. 2019;568:43–8.

599 49. Wang Y, Baumdicker F, Schweiger P, Kuenzel S, Staubach F. Horizontal gene transfer-
600 mediated bacterial strain variation affects host fitness in Drosophila. BMC Biol. 2021
601 Dec 1;19(1).
602 50. Han B, Sivaramakrishnan P, Lin CCJ, Neve IAA, He J, Tay LWR, et al. Microbial genetic
603 composition tunes host longevity. Cell. 2017;169:1249-1262.
604 51. Zinnen AD, Vichich J, Metzger JM, Gambardella JC, Bondarenko V, Simmons HA, et al.
605 Alpha-synuclein and tau are abundantly expressed in the ENS of the human appendix and
606 monkey cecum. PLoS One. 2022;17:e0269190.
607 52. Fisher RE. The primate appendix: A reassessment. Vol. 261, Anatomical Record. 2000. pp
608 228–36.
609

610 Fig. 1 Overview of the main analytical approach used in this study.
611 This method involves the simultaneous acquisition of the host and microbiome transcriptome
612 (metatranscriptome) from intestinal sites and conducts a gene co-expression network analysis of
613 the host and microbiome community. The procedure is outlined from step A to E. Perform dual
614 RNA-seq of host and microbiome transcriptomes in internal site samples. (A) Separate host and
615 microbiome RNA reads in silico using the host reference genome. (B) Compute the gene
616 expression profiles of the host and microbiome, respectively. (C) Unify gene expression
617 networks between host-host, host-microbiome, and microbiome-microbiome. The host-host and
618 microbiome-microbiome gene expression networks are connected when there is a positive
619 correlation between genes, while the host-microbiome gene expression network is connected
620 when there is either a positive or negative correlation between genes. (D) Identify closely
621 interacting gene modules from the gene expression network using a proposed clique-based

algorithm. (E) This process begins with searching for cliques (E Step1). If overlaps exist between cliques, the cliqueness score is computed in three gene expression networks (host-host, host-microbiome, and microbiome-microbiome) when overlapping cliques are merged (E Step2). If the cliqueness score exceeds the merging threshold τ in all three networks, the cliques are merged into a single module. If the score does not exceed the merging threshold τ in at least one network, the original cliques are retained without merging. This process is repeated until no further modules can be merged (E Step3).

Fig. 2 PCA analysis of gene expression data for host and microbiome

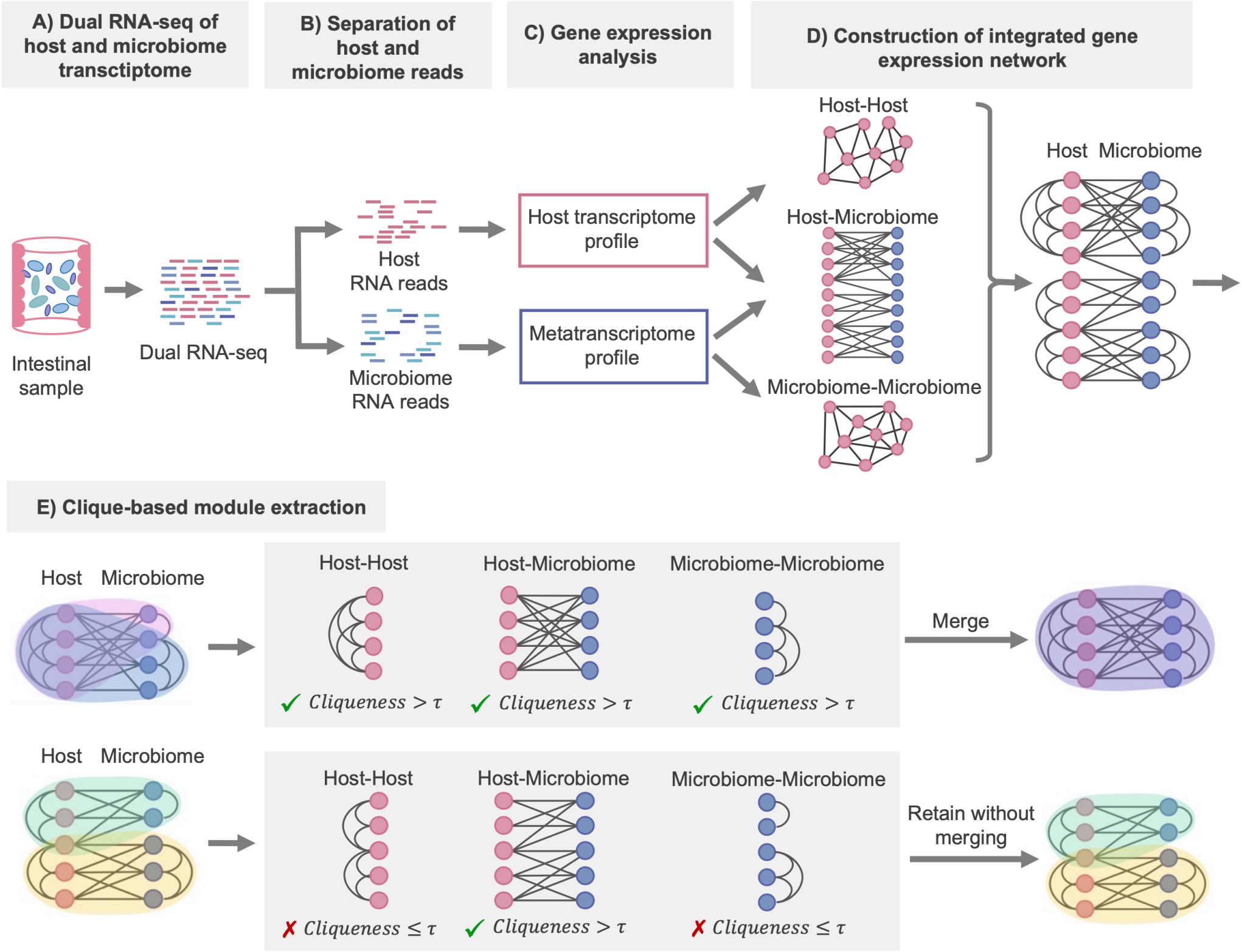
PCA was applied to the host gene expression profiles to visualize groupings according to site (A) and individual (B). PCA was similarly applied to the microbiome gene expression profiles to visualize groupings according to site (C) and individual (D).

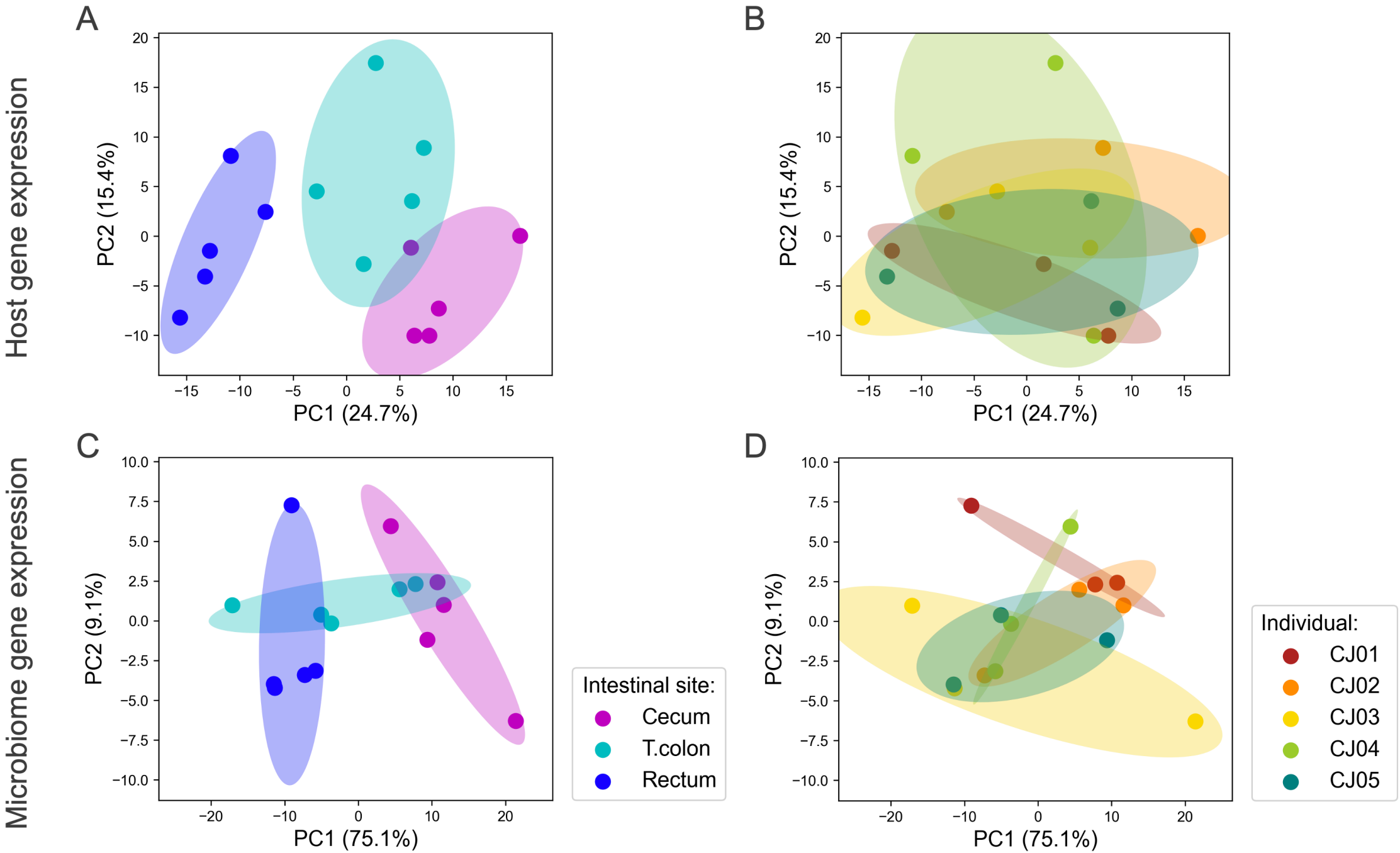
Fig. 3 Variations in host gene expression and functional pathways among intestinal sites

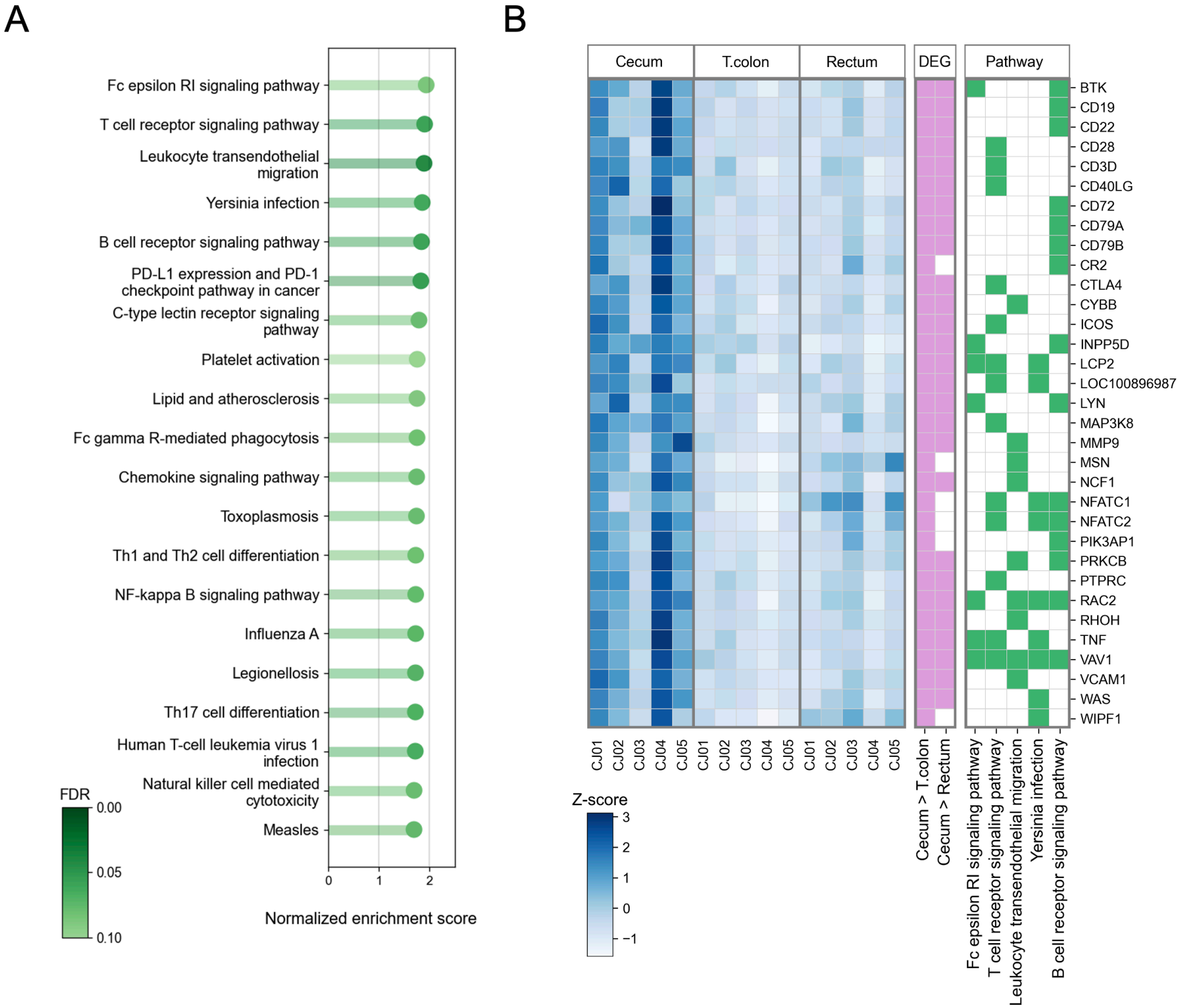
(A) The top 20 functional pathways significantly upregulated in the cecum compared to the transverse colon (FDR$<$0.1), as revealed by gene set enrichment analysis, are presented alongside their Normalized Enrichment Score and FDR. (B) Heatmap of differentially expressed genes belonging to the top 5 enriched functional pathways. The heatmap on the left represents the z-scores of gene expression levels, with genes significantly upregulated in the cecum compared to the transverse colon and rectum (FDR$<$0.01) colored in pink on the middle heatmap. The heatmap on the right indicates the functional pathways to which the genes belong, represented in green.

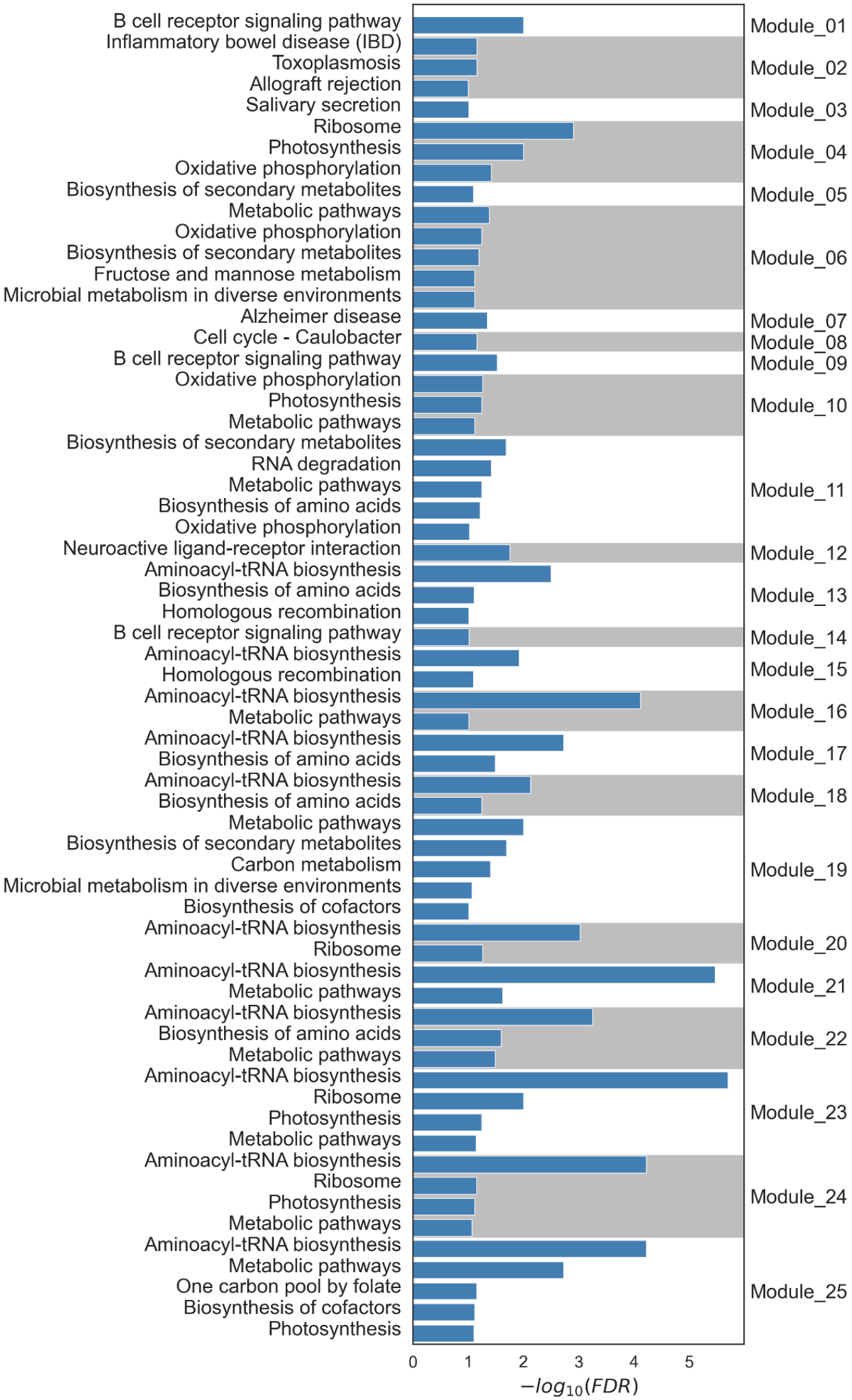
645 Fig. 4 Functional pathways characterized in modules extracted from host-microbiome gene
646 expression networks
647 This figure illustrates the functional characteristics of the modules determined by enrichment
648 analysis of functional pathways (FDR<0.1). The background of the bar graph alternates between
649 grey and white for each module for ease of differentiation.
650
651 Fig. 5 Metabolic pathway-enriched Module_06
652 A module showing a positive correlation between the host gene FBP1, which is highly expressed
653 in the transverse colon, and 40 microbiome genes, including those involved in fructose-mannose
654 metabolism and oxidative phosphorylation, was extracted. These genes were visualized
655 alongside bacterial species nodes that contribute to the expression of the microbiome genes.
656 Edges among genes indicate an absolute Spearman's correlation coefficient greater than 0.8.
657 Edges connecting microbiome genes to bacterial species are shown when the contribution rate of
658 bacterial species to microbiome gene expression exceeds 0.2. The thickness of the edge increases
659 with higher contribution rates.
660
661 Fig. 6 B cell receptor signaling pathway-enriched Module_01
662 Host genes involved in the immune response of B cells, including CD19, CD22, CD79B and
663 PTPN6, which were highly expressed in the cecum, co-expressed with the K01696 (*trpB*). These
664 genes were visualized alongside bacterial species nodes that contribute to the expression of the
665 microbiome genes. K01696 (*trpB*) expression was contributed to by *Parabacteroides distasonis*.
666 Edges among genes indicate an absolute Spearman's correlation coefficient greater than 0.8.
667 Edges connecting microbiome genes to bacterial species are shown when the contribution rate of

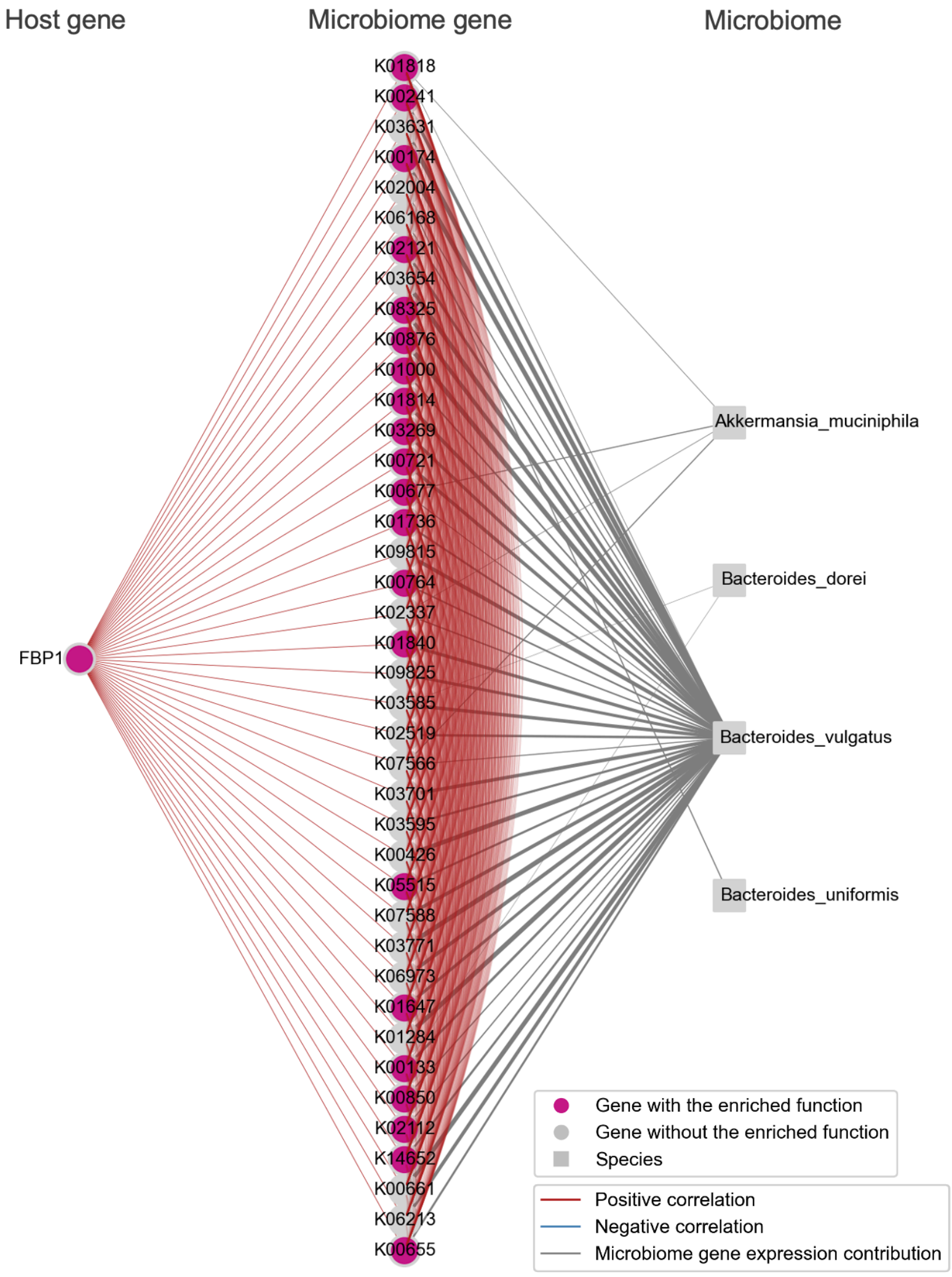
668 bacterial species to microbiome gene expression exceeds 0.2. The thickness of the edge increases
669 with higher contribution rates.
670
671 Fig. 7 Carbon metabolism pathway-enriched Module_19
672 Host genes, including MMP10, which is specifically highly expressed in the cecum, and 40
673 microbiome genes involved in carbon metabolism, were identified. These genes were visualized
674 alongside bacterial species nodes that contribute to the expression of the microbiome genes.
675 Edges among genes indicate an absolute Spearman's correlation coefficient greater than 0.8.
676 Edges connecting microbiome genes to bacterial species are shown when the contribution rate of
677 bacterial species to microbiome gene expression exceeds 0.2. The thickness of the edge increases
678 with higher contribution rates.
679
680 Fig. 8 Module_16 containing host genes and microbiome genes related to the production and
681 breakdown of bile acids. These genes were visualized alongside bacterial species nodes that
682 contribute to the expression of the microbiome genes. Edges among genes indicate an absolute
683 Spearman's correlation coefficient greater than 0.8. Edges connecting microbiome genes to
684 bacterial species are shown when the contribution rate of bacterial species to microbiome gene
685 expression exceeds 0.2. The thickness of the edge increases with higher contribution rates.
686

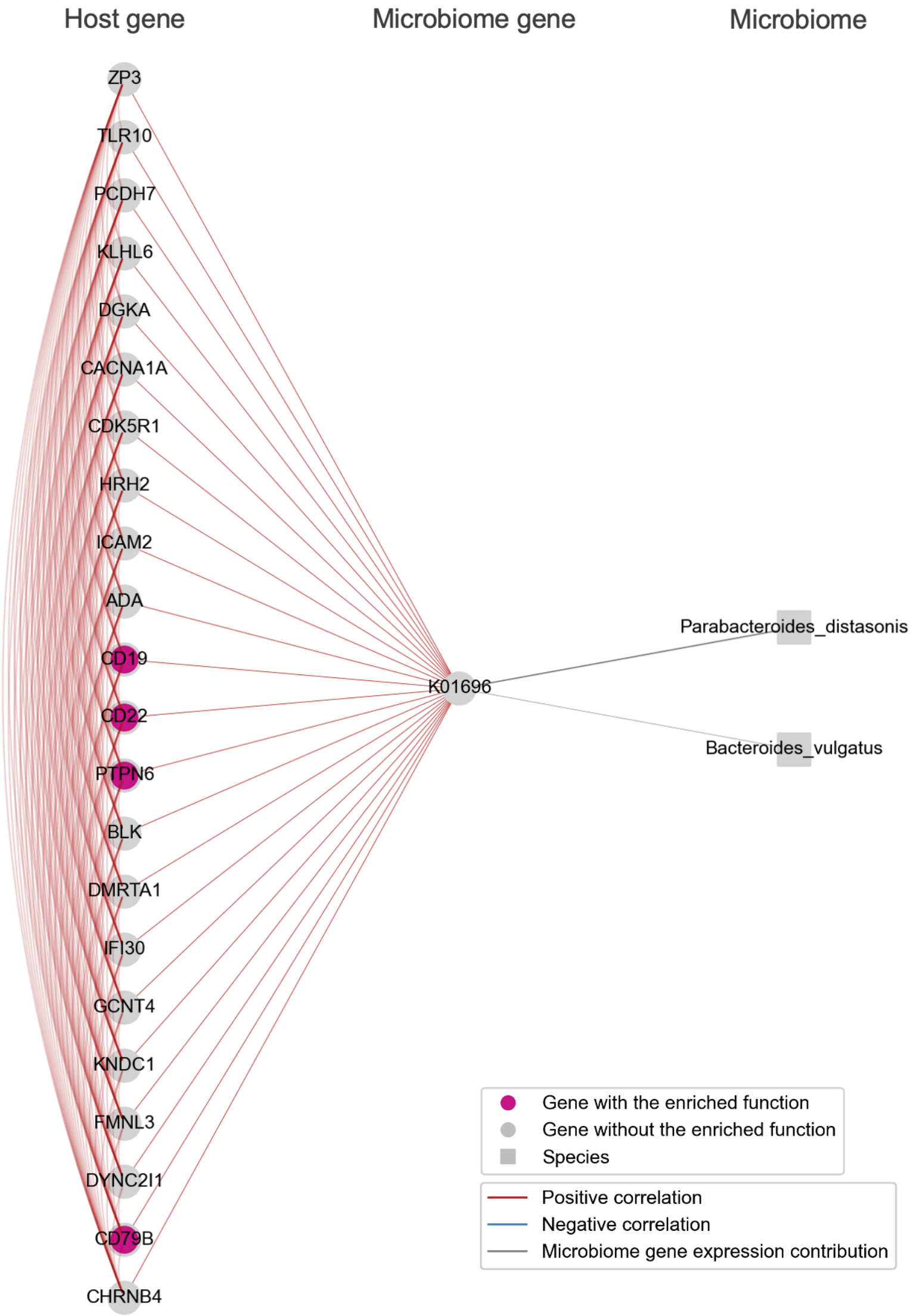

A) Dual RNA-seq of host and microbiome transctiptome
B) Separation of host and microbiome reads
C) Gene expression analysis
D) Construction of integrated gene expression network
Intestinal sample
Dual RNA-seq
Host RNA reads
Microbiome RNA reads
Host transcriptome profile
Metatranscriptome profile
Host-Host
Host-Microbiome
Microbiome-Microbiome
Host Microbiome
E) Clique-based module extraction
Host Microbiome
Host-Host
Host-Microbiome
Microbiome-Microbiome
✓ Cliqueness > τ
✓ Cliqueness > τ
✓ Cliqueness > τ
Merge
Host Microbiome
Host-Host
Host-Microbiome
Microbiome-Microbiome
✗ Cliqueness ≤ τ
✓ Cliqueness > τ
✗ Cliqueness ≤ τ
Retain without merging

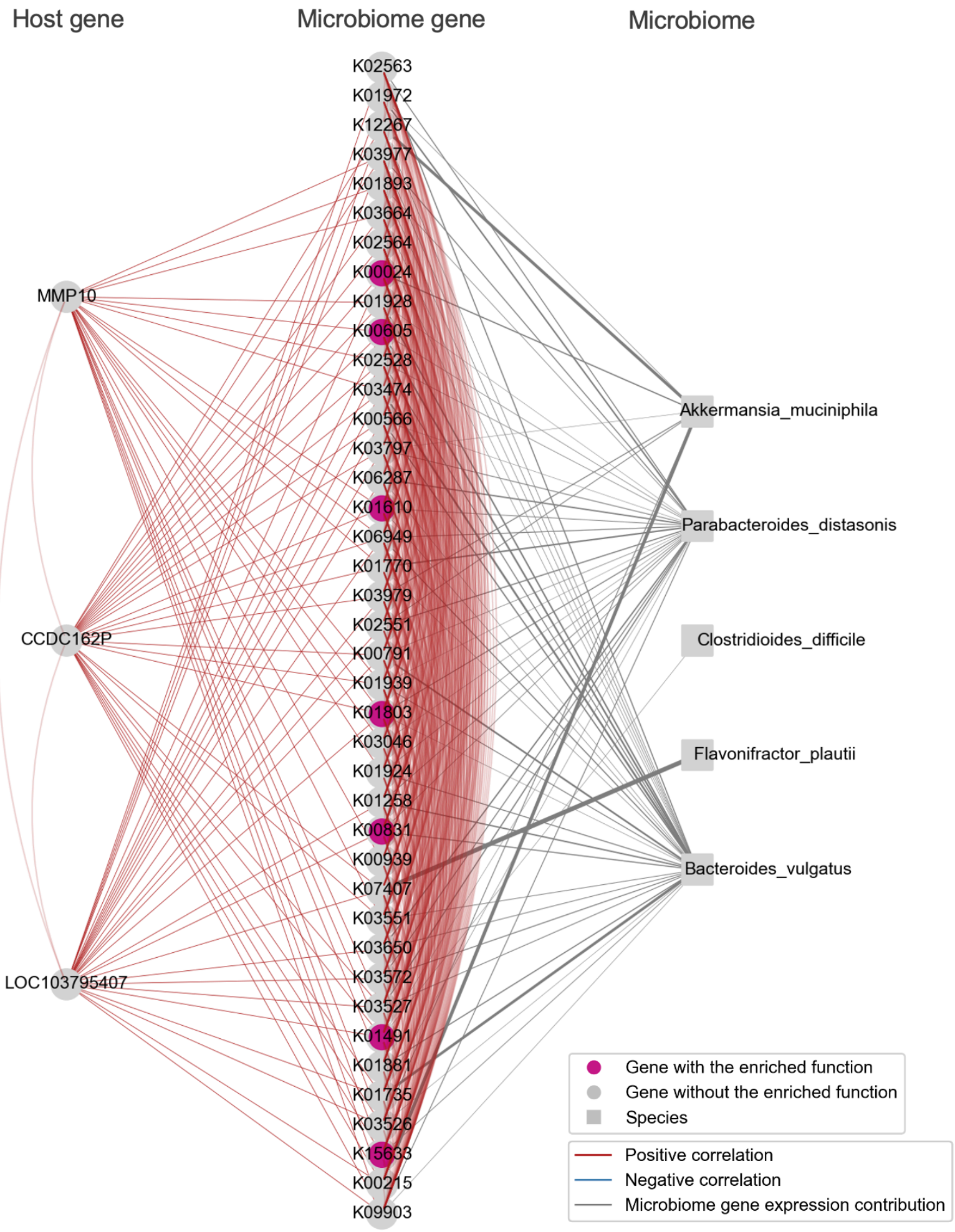

A
B
C
D
Host gene expression
Microbiome gene expression
PC1 (24.7%)
PC2 (15.4%)
PC1 (75.1%)
PC2 (9.1%)
Intestinal site:
Cecum
T.colon
Rectum
Individual:
CJ01
CJ02
CJ03
CJ04
CJ05

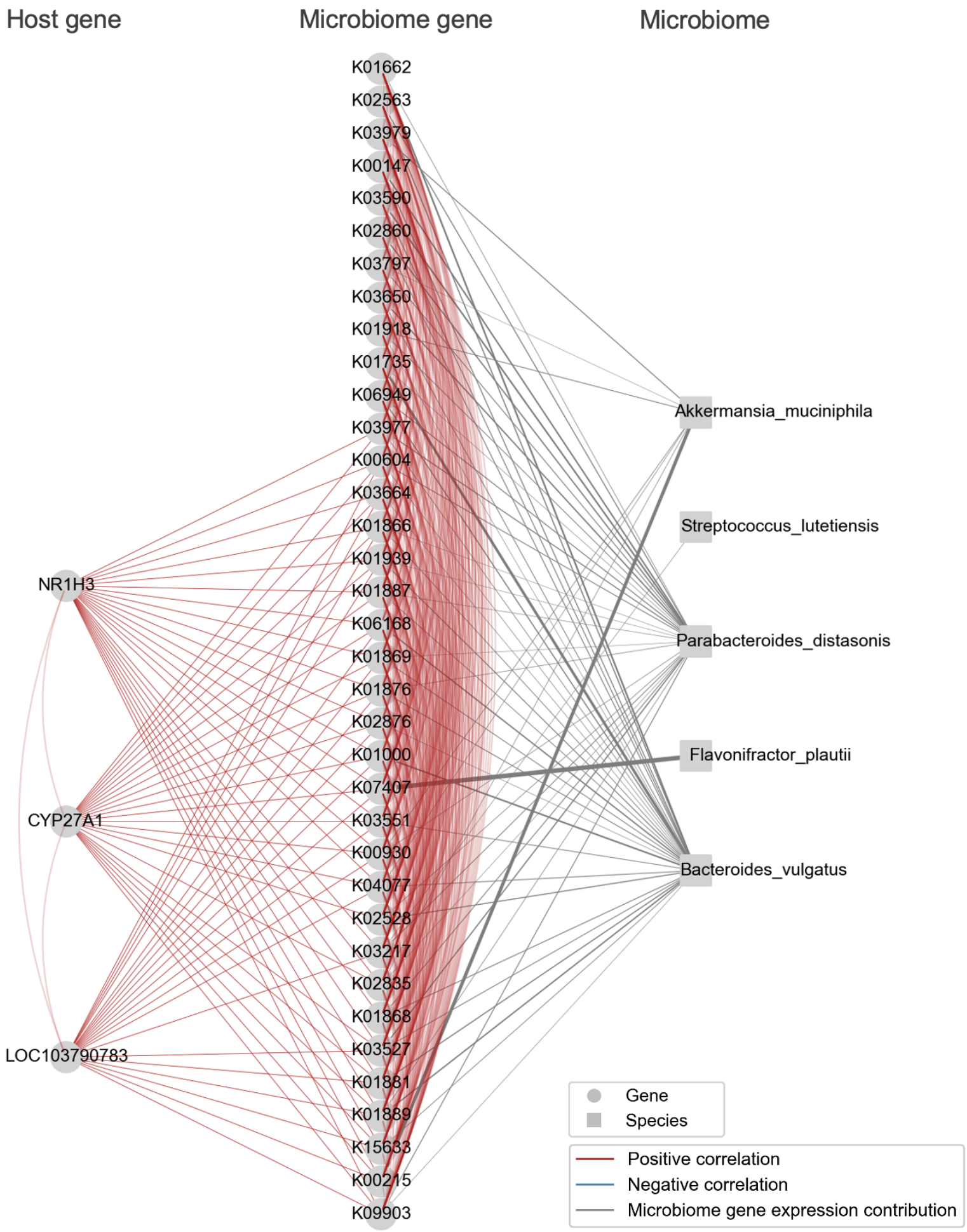

A
Fc epsilon RI signaling pathway
T cell receptor signaling pathway
Leukocyte transendothelial migration
Yersinia infection
B cell receptor signaling pathway
PD-L1 expression and PD-1 checkpoint pathway in cancer
C-type lectin receptor signaling pathway
Platelet activation
Lipid and atherosclerosis
Fc gamma R-mediated phagocytosis
Chemokine signaling pathway
Toxoplasmosis
Th1 and Th2 cell differentiation
NF-kappa B signaling pathway
Influenza A
Legionellosis
Th17 cell differentiation
Human T-cell leukemia virus 1 infection
Natural killer cell mediated cytotoxicity
Measles
FDR
0.00
0.05
0.10
0
1
2
Normalized enrichment score
B
Cecum
T.colon
Rectum
DEG
Pathway
BTK
CD19
CD22
CD28
CD3D
CD40LG
CD72
CD79A
CD79B
CR2
CTLA4
CYBB
ICOS
INPP5D
LCP2
LOC100896987
LYN
MAP3K8
MMP9
MSN
NCF1
NFATC1
NFATC2
PIK3AP1
PRKCB
PTPRC
RAC2
RHOH
TNF
VAV1
VCAM1
WAS
WIPF1
CJ01
CJ02
CJ03
CJ04
CJ05
Cecum > T.colon
Cecum > Rectum
Fc epsilon RI signaling pathway
T cell receptor signaling pathway
Leukocyte transendothelial migration
Yersinia infection
B cell receptor signaling pathway
Z-score
3
2
1
0
−1


B cell receptor signaling pathway
Inflammatory bowel disease (IBD)
Toxoplasmosis
Allograft rejection
Salivary secretion
Ribosome
Photosynthesis
Oxidative phosphorylation
Biosynthesis of secondary metabolites
Metabolic pathways
Oxidative phosphorylation
Biosynthesis of secondary metabolites
Fructose and mannose metabolism
Microbial metabolism in diverse environments
Alzheimer disease
Cell cycle - Caulobacter
B cell receptor signaling pathway
Oxidative phosphorylation
Photosynthesis
Metabolic pathways
Biosynthesis of secondary metabolites
RNA degradation
Metabolic pathways
Biosynthesis of amino acids
Oxidative phosphorylation
Neuroactive ligand-receptor interaction
Aminoacyl-tRNA biosynthesis
Biosynthesis of amino acids
Homologous recombination
B cell receptor signaling pathway
Aminoacyl-tRNA biosynthesis
Homologous recombination
Aminoacyl-tRNA biosynthesis
Metabolic pathways
Aminoacyl-tRNA biosynthesis
Biosynthesis of amino acids
Aminoacyl-tRNA biosynthesis
Biosynthesis of amino acids
Metabolic pathways
Biosynthesis of secondary metabolites
Carbon metabolism
Microbial metabolism in diverse environments
Biosynthesis of cofactors
Aminoacyl-tRNA biosynthesis
Ribosome
Aminoacyl-tRNA biosynthesis
Metabolic pathways
Aminoacyl-tRNA biosynthesis
Biosynthesis of amino acids
Metabolic pathways
Aminoacyl-tRNA biosynthesis
Ribosome
Photosynthesis
Metabolic pathways
Aminoacyl-tRNA biosynthesis
Ribosome
Photosynthesis
Metabolic pathways
Aminoacyl-tRNA biosynthesis
Metabolic pathways
One carbon pool by folate
Biosynthesis of cofactors
Photosynthesis
Module_01
Module_02
Module_03
Module_04
Module_05
Module_06
Module_07
Module_08
Module_09
Module_10
Module_11
Module_12
Module_13
Module_14
Module_15
Module_16
Module_17
Module_18
Module_19
Module_20
Module_21
Module_22
Module_23
Module_24
Module_25
0
1
2
3
4
5
$-log_{10}(FDR)$


Host gene
Microbiome gene
Microbiome
FBP1
K01818
K00241
K03631
K00174
K02004
K06168
K02121
K03654
K08325
K00876
K01000
K01814
K03269
K00721
K00677
K01736
K09815
K00764
K02337
K01840
K09825
K03585
K02519
K07566
K03701
K03595
K00426
K05515
K07588
K03771
K06973
K01647
K01284
K00133
K00850
K02112
K14652
K00661
K06213
K00655
Akkermansia_muciniphila
Bacteroides_dorei
Bacteroides_vulgatus
Bacteroides_uniformis
Gene with the enriched function
Gene without the enriched function
Species
Positive correlation
Negative correlation
Microbiome gene expression contribution


Host gene
Microbiome gene
Microbiome
ZP3
TLR10
PCDH7
KLHL6
DGKA
CACNA1A
CDK5R1
HRH2
ICAM2
ADA
CD19
CD22
PTPN6
BLK
DMRTA1
IFI30
GCNT4
KNDC1
FMNL3
DYNC2I1
CD79B
CHRNB4
K01696
Parabacteroides_distasonis
Bacteroides_vulgatus
Gene with the enriched function
Gene without the enriched function
Species
Positive correlation
Negative correlation
Microbiome gene expression contribution


Host gene
Microbiome gene
Microbiome
MMP10
CCDC162P
LOC103795407
K02563
K01972
K12267
K03977
K01893
K03664
K02564
K00024
K01928
K00605
K02528
K03474
K00566
K03797
K06287
K01610
K06949
K01770
K03979
K02551
K00791
K01939
K01803
K03046
K01924
K01258
K00831
K00939
K07407
K03551
K03650
K03572
K03527
K01491
K01881
K01735
K03526
K15633
K00215
K09903
Akkermansia_muciniphila
Parabacteroides_distasonis
Clostridioides_difficile
Flavonifractor_plautii
Bacteroides_vulgatus
Gene with the enriched function
Gene without the enriched function
Species
Positive correlation
Negative correlation
Microbiome gene expression contribution


Host gene
Microbiome gene
Microbiome
NR1H3
CYP27A1
LOC103790783
K01662
K02563
K03979
K00147
K03590
K02860
K03797
K03650
K01918
K01735
K06949
K03977
K00604
K03664
K01866
K01939
K01887
K06168
K01869
K01876
K02876
K01000
K07407
K03551
K00930
K04077
K02528
K03217
K02885
K01868
K03527
K01881
K01889
K15633
K00215
K09903
Akkermansia_muciniphila
Streptococcus_lutetiensis
Parabacteroides_distasonis
Flavonifractor_plautii
Bacteroides_vulgatus
Gene
Species
Positive correlation
Negative correlation
Microbiome gene expression contribution