



Article

Potential Functions of the Gastrointestinal Microbiome Inhabiting the Length of the Rat Digest Tract

Dongyao Li ^{1,2}, Haiqin Chen ^{1,2,*}, Jianxin Zhao ^{1,2}, Hao Zhang ^{1,2,3} and Wei Chen ^{1,2,3,4}

¹ State Key Laboratory of Food Science and Technology, Jiangnan University, Wuxi 214122, China; 7130112040@vip.jiangnan.edu.cn (D.L.); jxzha@jiangnan.edu.cn (J.Z.); zhanghao@jiangnan.edu.cn (H.Z.); weichen@jiangnan.edu.cn (W.C.)

² School of Food Science and Technology, Jiangnan University, Wuxi 214122, China

³ National Engineering Research Center for Functional Food, Jiangnan University, Wuxi 214122, China

⁴ Beijing Innovation Center of Food Nutrition and Human Health, Beijing Technology and Business University (BTBU), Beijing 100048, China

* Correspondence: haiqinchen@jiangnan.edu.cn; Tel.: +86-510-8519-7239

Received: 13 February 2019; Accepted: 8 March 2019; Published: 12 March 2019



Abstract: The rat is an important model animal used frequently in biological researches exploring the correlations between gut microbiome and a wide array of diseases. In this study, we used an extended ancestral-state reconstruction algorithm to predict the functional capabilities of the rat gastrointestinal microbiome. Our results indicate an apparent tendency toward metabolic heterogeneity along the longitudinal and transverse axes of the rat gastrointestinal tract (GIT). This heterogeneity was suggested by the enriched small-molecule transport activity and amino acid metabolism in the upper GIT, the aerobic energy metabolism in the stomach and the mucolysis-related metabolism in the lower GIT mucus layer. In contrast to prior results, many functional overlaps were observed when the gastrointestinal microbiomes of different hosts were compared. These overlaps implied that although both the biogeographic location and host genotype were prominent driving forces in shaping the gastrointestinal microbiota, the microbiome functions were similar across hosts when observed under similar physicochemical conditions at identical anatomical sites. Our work effectively complements the rat microbial biogeography dataset we released in 2017 and, thus, contributes to a better understanding and prediction of disease-related alterations in microbial community function.

Keywords: rat; gastrointestinal tract; microbiome function; ecosystem service

1. Introduction

Symbiotic microorganisms located inside and on the surface of the host body provide many biological functions that are not encoded in their host genome [1]. The complementary functions are generally considered as ecosystem services which are crucial to maintaining the host physiological homeostasis [2]. In recent years the gut flora has attracted increasing attention in several fields and is, now, even considered a hidden organ [3], as its collective metabolic activity is somewhat equal to a virtual organ and profoundly influences the host health and disease in both direct [4–6] and indirect manners [7–10].

In the past decade, encouraged by the progress in high-throughput sequencing technology and new developments in bioinformatics, the scientific community has continuously and gradually investigated the gut microbial communities underlying biological functions observed in human- [11,12] and host-associated [13,14] habitats. In 2008, the National Institutes of Health (NIH) initiated the Human Microbiome Project (HMP), which aimed to characterise the microbiota within 5 main human

body habitats (i.e., airway, skin, oral, gut and urogenital) from the aspects of community structure and metabolic potential [15]. In 2010, the MetaHIT project released the first human gut microbial gene catalogue established by metagenomic sequencing [16]. In 2015, the Beijing Genomics Institute (BGI) released the first mouse gut microbial gene catalogue [17]. These catalogues have since served as valuable references for numerous studies. However, most of these studies used functional gene sets derived from faecal samples, although the appropriateness of using faeces as a proxy for the whole gastrointestinal tract (GIT) remains unconfirmed.

In 2017, our research group released the first microbial biogeography of the rat GIT, which systematically characterised the baseline microbial structure, membership, diversity and ecology in this long-standing biomedical research animal model [18]. By comparing different niches along the length of the rat digestive tract, we concluded that faecal samples could not fully represent the microbiota throughout this tract, at least, in terms of taxonomic profiling. In an ecosystem such as the GIT, the collective function is co-determined by the population composition and the activities of each component. The potential metabolic activities of a microbe are encoded by its genome, and many host-associated microorganisms, even those only distantly related, may share functional genes and, thus, thrive in the harsh gastrointestinal (GI) environment [16]. The HMP also indicated that the community function remained relatively stable, despite dramatic variations in the community structure [19].

Therefore, in this study, we took the next step and predicted the functional capabilities of the rat GI microbiome using an extended ancestral-state reconstruction algorithm [20] and projected these functions onto a Kyoto Encyclopaedia of Genes and Genomes (KEGG) orthology and pathway framework [21] to address the following questions: (i) What complementary ecosystem service functions do microbial residents provide along the length of the rat digestive tract? (ii) Does the faecal functional profile fully represent those of other GI segments? (iii) To what extent does heterogeneity in the community composition reflect the differences in the community function across biogeographic locations and hosts?

2. Results and Discussion

The rat genome was published in 2004 and was the third mammalian genome sequenced [22] after the human [23] and mouse genomes [24]. However, the rat intestinal microbiome, as its second genome, was not comprehensively characterised until 2017. Our previously published rat microbial biogeography filled this gap and suggested that the stratification of microbial community and the shift in microbial metabolites might interact, serving as both cause and effect [18].

2.1. Metabolic Heterogeneity in the Community Function

In present study, we aimed to characterise the rat GI microbiome from the perspective of a functional potential. Based on a total of 6009 KEGG orthologue (KO) annotations, we observed an apparent tendency toward a metabolic heterogeneity along the longitudinal and transverse axes of the rat GIT. This tendency was intuitively displayed on the principal component analysis (PCA) ordination plots (Figure 1A–C), where samples were grouped by different biogeographic factors such as the anatomic region (adonis: $R^2 = 0.47$; $p \leq 0.001$), subsite (adonis: $R^2 = 0.60$; $p \leq 0.001$) and niche location (adonis: $R^2 = 0.24$; $p \leq 0.001$). The first principal component (PC1) explained 53.75% of the total heterogeneity and parsed out the niche locations of the samples, while the second principal component (PC2) explained 18.8% and mainly parsed out the anatomic regions.

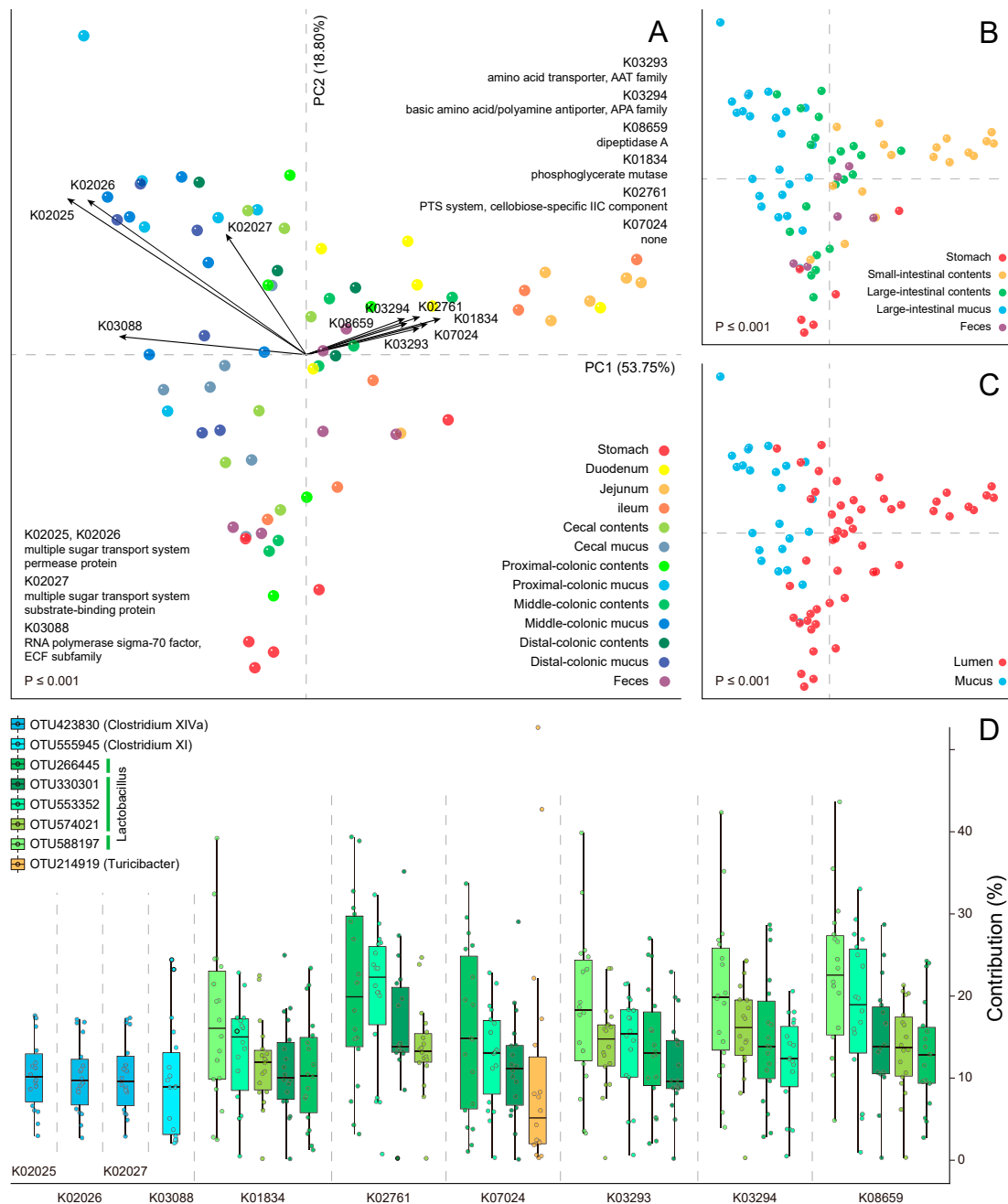


Figure 1. The spatial heterogeneity in the rat microbial community functions: a principal component analysis (PCA) based on the Kyoto Encyclopaedia of Genes and Genomes (KEGG) orthologue (KO) profiles of gut microbiome across the rat gastrointestinal tract (GIT), where potential relationships between metabolic variability and sample metadata are highlighted in panels A to C by different colorations. The permutational multivariate analysis of variance (PERMANOVA) results reveal the grouping of samples from the same (A) sampling site (adonis: $R^2 = 0.60$), (B) anatomic region ($R^2 = 0.47$) and (C) niche location ($R^2 = 0.24$). The p -value of the Monte Carlo permutation test is shown on the lower left. The percentages of variation explained by first principal component (PC1) and second principal component (PC2) are listed along the axes representing them. The top 10 KOs driving the heterogeneity are indicated in panel A. For different anatomic regions, the major contributing operational taxonomic units (OTUs) and their contributions to these KOs are shown in panel (D).

The KOs acting as the strongest drivers of this trend in heterogeneity are plotted as loadings on the PCA plot, with the arrow lengths proportional to their contributions (Figure 1A). In the second

quadrant, 4 KOs point to the colonic mucosal samples, indicating that these KOs play significant roles in the local agglomeration of these samples. Here, 3 KOs, K02025, K02026 and K02027, constituted the elements of a putative multiple sugar transport system, while the fourth, K03088, was an RNA polymerase factor. Similarly, K02025, K02026 and K03088 are also enriched in the mouse gut microbial gene catalogue and drive the separation of the human and mouse gut microbiomes [17]. By partitioning the metagenome functional contributions, we found that the major contributing operational taxonomic unit (OTU) to this ATP-binding cassette (ABC) transporter belonged to the *Clostridium* cluster XIVa, in which many species produce butyrate and specifically colonise the mucins close to the intestinal epithelium [25].

In the first quadrant, 6 KOs point to the small-intestinal samples. Three of these KOs, K03293, K03294 and K08659, might be involved in the catabolism of free small peptides and amino acids from the upper GIT digestive residue [26]. Interestingly, the corresponding contributing OTUs all belonged to *Lactobacillus*, although the OTU ranks varied between different KOs (Figure 1D). Studies on gut microbiota have rarely reported the proteolytic activity of *Lactobacilli* in GIT. However, *Lactobacillus* strains have occasionally been used as starter cultures to eliminate gluten toxicity [27] and to ferment meat products [28] in the field of food microbiology, suggesting that *Lactobacillus* species may play specific roles in rat small-intestinal habitats. Based on our prior phylogenetic prediction result, the molecular species participating in this process were probably *Lactobacillus intestinalis*, *L. hominis*, *L. johnsonii*, *L. taiwanensis*, *L. gasseri* and *L. saniviri* [18]. Additionally, OTU214919, which was previously annotated as *Turicibacter*, comprised a major contributory phylotype in the orthologous gene family K07024, with a much higher degree of intersubject variation regarding contributions (note: *Turicibacter* is another predominant genus throughout the rat digestive tract).

However, the permutational multivariate analysis of variance (PERMANOVA) result revealed that the individual subject was no longer a significant categorical factor (adonis: $R^2 = 0.09$; $p = 0.18$) with respect to community function, in contrast to the situations regarding structure and membership [18]. To some extent, this weakened stratification was in line with the HMP results, which showed the taxonomic variety versus metabolic stability within a healthy population [19].

2.2. Module-Centric Metabolic Reconstruction of the Rat Gastrointestinal (GI) Microbiome

A KEGG pathway module is a collection of manually defined, tight functional units which allow a higher-level grouping of KO gene families into pathways or functional classifications [29]. To summarise the community function of the rat GIT at a higher level, 6009 KOs were collapsed into 125 metabolic modules via a parsimony approach. The abundance varied significantly in at least one sampling site in more than half (75/125, 60%) of these detected modules, demonstrating the uniqueness of various GI niches (Figure 2A and Table S1). Generally speaking, the structural complexes that processed the environmental information were prevalent in the upper GIT. These complexes were principally transporters of small molecules (saccharides: M00194, M00197, M00201, M00207, M00276 and M00277; polyols: M00200; phosphates: M00222; amino acids: M00230, M00232, M00235 and M00237; peptides: M00333, M00348 and M00349; mineral and organic ions: M00185, M00193, M00299, M00300 and M00319). This phenomenon suggests an active exchange of materials and information between commensal microbiota and the digestive milieu, wherein the macromolecules in foodstuffs are broken down, leading to a relative abundance of free small molecules [30].

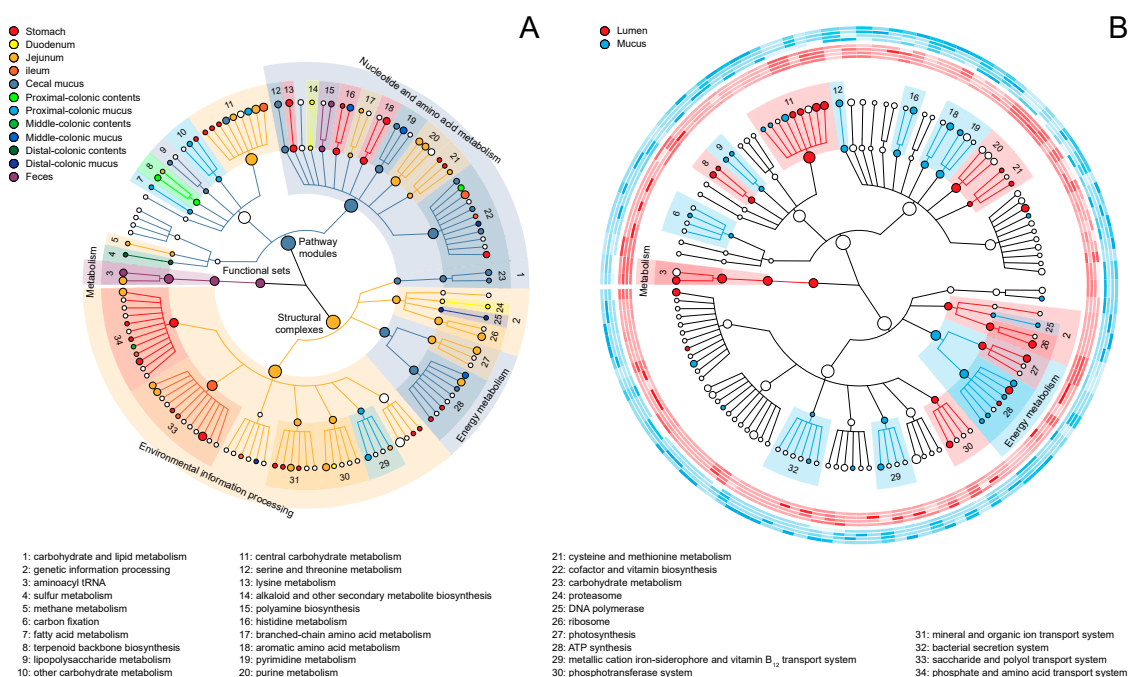


Figure 2. The modular difference in metabolic landscapes across the sampling sites and niche locations: The cladograms illustrate the differentially abundant metabolic modules under the KEGG BRITE hierarchy, determined using the linear discriminant analysis (LDA) effect size (LEfSe) and colored by their most abundant (A) sites and (B) locations in the rat GIT. The dot size is proportional to the global relative abundance of the corresponding module. In panel (A), the whole GI axis was involved in biomarker mining, whereas in panel (B), only the adjacent mucosal and luminal compartments from the large intestine were involved. Around the cladogram are located 8 heatmap rings, where the saturation of each cell represents the average relative abundance of each module in different sampling sites (from within outward: caecal contents; proximal-, middle-, distal-colonic contents; caecal mucus; and proximal-, middle- and distal-colonic mucus). All relative abundances are z-score normalized in the radial direction for visualization. The complete lists of modules are presented in Tables S1 and S2.

Furthermore, the majority of amino acid metabolic modules (M00016, M00017, M00022, M00033, M00034, M00035, M00036 and M00045) were also prevalent in the upper GIT. In addition to the bacterial proteasome module M00342, this prevalence might be associated with the ordination result of PCA (Figure 1A). However, the module M00018, which comprises serine and threonine metabolism, was abundant in the caecal mucus layer, likely because active bacterial mucolysis within this habitat involves the breakdown of serine- and threonine-rich O-linked mucins [31]. The hydroxyl groups of these amino acids are always linked to various uronic acid-rich oligosaccharide side-chains [32], and the corresponding metabolic module M00061 was also detected abundantly in the proximal-colonic mucus layer.

However, in some cases, our attempts to tile a whole pathway suggested the simultaneous existence of alternative modules with distinct internal structures but identical outlines enabling them to fill the same metabolic gaps. In the rat GI microbiome, we identified several such biological situations. For example, either the mevalonate (M00095) or non-mevalonate (M00096) module could perform terpenoid backbone biosynthesis; the former and the latter were enriched in the upper (i.e., jejunal contents) and lower GIT (i.e., proximal-colonic contents), respectively. Similar phenomena were also observed for the modules M00157 (bacterial F-type ATPase; enriched in jejunal contents) and M00159 (prokaryotic V-type ATPase; enriched in middle-colonic contents). These results suggest that when facing various survival challenges, the microbiota inhabiting different GI niches might use distinct strategies to complete the pathway.

We previously found that the rat core microbiota evolved from an aerobic to facultative anaerobic to obligatory anaerobic metabolism along the longitudinal axis [18], in accordance with imaging observations of the oxygen gradient throughout the GIT [33]. This tendency was also evident when biomarkers of community function were mined at different sampling sites; here, functional modules involved in aerobic heterotrophic metabolism (M00008, M00009, M00011, M00012, M00115, M00144 and M00149) were prevalent in the stomach. Besides the longitudinal axis, in the radial direction, an intraluminal oxygen gradient further extends radially from the mucosal tissue interface into the lumen [34]. In the rat lower GIT, both the mucus layer and lumen are densely populated by various microorganisms, which separately form topologically distinct co-occurrence networks [18]. To compare the functions of these closely adjacent compartments, the linear discriminant analysis (LDA) effect size (LEfSe) system was applied only to mucosal and luminal samples from the large intestine, limiting pairwise comparisons performed between the same subsites. The results demonstrated a dramatic abundance variance in approximately a third (39/121, 32.2%) of the modules detected in these samples (Figure 2B and Table S2). Specifically, modules associated with aerobic metabolism (M00009, M00011, M00144, M00149, M00156 and M00159) were over-enriched in the mucus layer, whereas those associated with anaerobic metabolism (M00001 and M00002) were under-enriched.

The GI microbiome is an important supplier of vitamins [35]. Here, we observed elaborate supply patterns in different biogeographic niches. Modules associated with the biosynthesis of vitamins B₂, B₃ and B₁₂ (M00125, M00115 and M00122, respectively) were prevalent in the upper GIT, whereas those associated with the biosynthesis of the vitamins B₁, B₅, B₇ and B₉ (M00127, M00119, M00123 and M00126, respectively) were prevalent in the lower GIT. Particularly, a pairwise comparison of the microbial metagenomes in the mucus layer and lumen revealed a high abundance of the pyridoxal biosynthesis module M00124 (vitamin B₆; cofactor for many enzymes with amino acid substrates [36]) in the former, suggesting ubiquitous proteolytic activities in this niche location.

2.3. Functional Overlaps in the Murine GI Microbiome

Previously, we compared the GIT microbiota of different rodents and concluded that the rat biogeographic map might serve as a new reference for digestive tract-related disease research [18]. In this study, we reanalysed the datasets [37–41] using the identical computational steps applied to rats to further investigate the GIT microbial metabolic similarities and differences among various mammalian hosts. Unsurprisingly, the human samples had the lowest nearest sequenced taxon index (NSTI; mean = 0.094 ± 0.088 (standard deviation)), followed by the rat (0.122 ± 0.039), mouse (0.213 ± 0.065) and woodrat samples (0.218 ± 0.05).

When these samples were mapped on the same PCA plane, a lot of overlaps were observed among different murine hosts (Figure 3A) in contrast to the prior ordination result calculated from the phylotype abundances [18]. This observation was also confirmed when the ecological distances between samples from different hosts were calculated. Here, the Euclidean distances calculated from the OTU abundances were significantly higher than those calculated from the KO profiles (nonparametric paired Wilcoxon-signed-rank test, $p < 2.2 \times 10^{-16}$). Furthermore, regarding the functional overlap, human GI microbiome was more similar to the rat microbiome than the mouse microbiome within the large-intestinal mucus layer (5405 overlapping KOs between humans and rats versus 4583 overlapping KOs between humans and mice). The pattern of overlap resembled the pattern observed when the human and mouse gut catalogues were compared at the KO level (4969 overlapping KOs between humans and mice) [17].

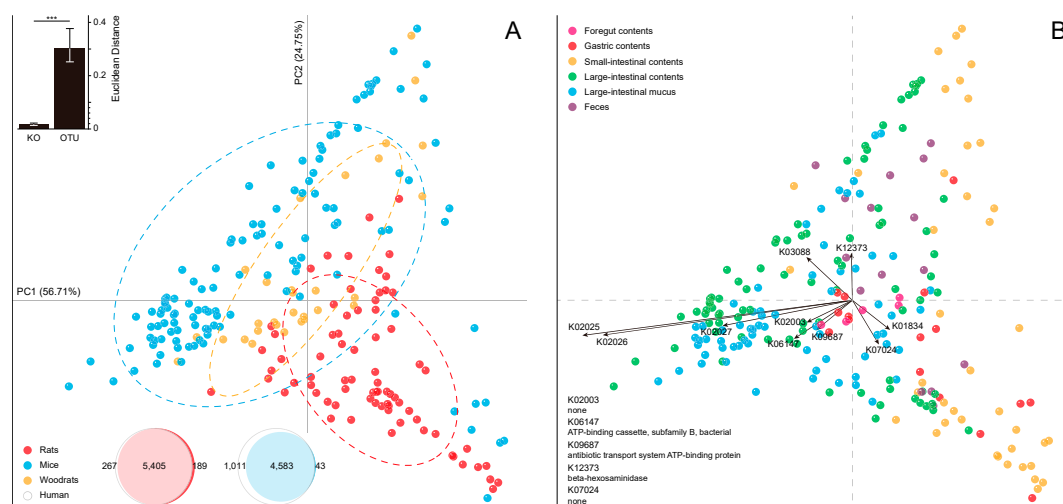


Figure 3. The similarities and differences in metabolic potentials among the GI microbiomes from different hosts: the PCA based on the KO profiles of GI microbiome used 227 samples from mice, rats and woodrats, grouped by (A) host genotype (adonis: $R^2 = 0.27$; $p \leq 0.001$) and (B) anatomic region ($R^2 = 0.09$; $p \leq 0.001$) in different colorations. The percentages of variation explained by PC1 and PC2 are listed along the axes representing them. The top 10 KOs driving the difference are indicated in panel (B). In panel (A), the normal confidence ellipses cover 67% of the samples belonging to the corresponding hosts. The insert graph on the upper left summarizes the Euclidean distances between samples from the same region but different hosts, calculated from KO profiles and OTU abundances respectively. The bar height represents the median inter-sample distance and the error bars range from the 25th to 75th quantile. The insert graph on the lower left shows the Venn diagrams demonstrating the functional overlaps between different mammalian hosts in the large-intestinal mucosal microbiome.

Besides the host genotype (adonis: $R^2 = 0.27$; $p \leq 0.001$), the anatomic region was also identified as a relatively weak categorical factor ($R^2 = 0.09$; $p \leq 0.001$) and was mainly parsed out during PC1 (56.71% variation explained). Coincidentally, some of the loading KOs (K02025, K02026, K02027, K03088 and K01834) in rats also appeared on the PCA plot when 2 other rodents were considered (Figure 3B). In the third quadrant, K02025, K02026, K02027, K06147 and K09687, which comprise an ABC transporter, again pointed to the large-intestinal samples, while K01834 still pointed toward the small-intestinal samples in the fourth quadrant. Moreover, K12373, which corresponds to a beta-hexosaminidase involved in keratan sulfate degradation, fell into the second quadrant, suggesting that this orthologous gene family plays a major role in the bacterial mucolysis specific to the mouse large-intestinal microbiome.

3. Materials and Methods

The raw reads of 75 samples sequenced during our prior project, PRJNA324666 [18], were downloaded from the NCBI Short Read Archive (SRA). The quality control and operational taxonomic unit (OTU) picking were performed as previously described [18], whereas *de novo* OTUs were removed by keeping only those with matching Greengenes IDs [42]. Chimeras (identified using Broad Institute Chimera Slayer [43]) and singletons were also removed from the OTU table. During OTU picking, 84.23% of reads were mapped to references at a similarity level of 97%. Given the distances between the sequencing reads and reference seeds, however, the cluster representative names were permuted 10 times (with replacements) within the OTU table to avoid introducing biases when estimating the gene families present in microorganisms from the Phylogenetic Investigation of Communities by Reconstruction of Unobserved States' (PICRUSt) pre-calculated files [20]. Next, the PICRUSt pipeline (version 1.1.3) was used to predict the sample-wide KEGG orthologue (KO) profiles based on the 10 permuted tables and to estimate the contributions of each OTU to the given

KOs. The average correlation between the permutations within a sample exceeded 0.99 (Spearman's rho). Hence, we calculated the arithmetically averaged abundances of KO groups per sample and used these values in all the downstream analyses. We further retrieved and processed raw reads from 5 other projects (mice: PRJNA178786 (35 samples retrieved) [37], PRJEB2233 (2 samples) [38] and figshare1499145 (85 samples) [39]; woodrats: PRJNA197212 (30 samples) [40]; and human: mgp1982 (17 samples) [41]) as described previously and above and combined these data with the KO profiles of rat samples.

The HUMAnN pipeline (HMP Unified Metabolic Analysis Network; version 0.99) [44] was selected to reconstruct the metabolic modules and thus collapse the predicted functions into a higher level. The parameter settings and processing modules within this pipeline were optimised for the best overall performance as previously described [44]; here, a parsimony approach [45] was used to conservatively assign gene sets to hierarchical metabolic modules. To normalise the sequencing depth, we used total-sum scalings to the KO profile and module abundance matrices and thus formed the substrate for the following statistical analyses.

Functional similarities between the samples were measured using Euclidean distances based on the KO profile matrix. PERMANOVA was used to statistically describe the relationships between the metabolic variability and sample metadata, and the significance of these relationships was tested using 999 Monte Carlo permutations. PCA was performed to map the KO profiles in a multidimensional space to a 2-dimensional plane. The calculation of the Euclidean distance, PERMANOVA and PCA were all performed using the R-package VEGAN (version 2.5-2) [46]. The PCA results were visualised using EMPEROR [47], and both sample scores and KO loadings were plotted in the same coordinates. The LefSe biomarker discovery tool (LDA Effect Size; version 1.0) [48] was applied to the module abundances from HUMAnN to determine the over- or under-enriched metabolic modules in at least one sampling site and in different niche locations; here, the alpha value of 0.01 and logarithmic LDA score threshold of 2.0 were applied. The metabolic modules were hierarchically organised in a dendrogram according to the KEGG BRITE database and visualised using GraphAn (Graphical Phylogenetic Analysis; version 1.1.3) [49]. Venn diagrams describing the function overlaps between hosts were constructed using the R-package VennDiagram (version 1.6.20 <https://CRAN.R-project.org/package=VennDiagram>).

4. Conclusions

In summary, we comprehensively characterised the putative biological functions of the rat GI microbiome, which complemented our previously released microbial biogeography project [18]. The rat digestive system composes many different niches with distinct physicochemical conditions and immune responses. We observed a number of specific functional clades ubiquitous within and characteristic to each of these niches. This spatially functional heterogeneity was consonant with our prior phylotype-based analysis, which revealed the tendency of GIT microbiota to form a stratified community structure and membership [18]. From the perspective of community function, we thus further confirmed that the faecal gene set cannot represent the full ecosystem service repertoire provided by the entire rat GIT microbiome. Furthermore, a comparison of the GI microbial metagenomes from different hosts revealed many functional overlaps. This phenomenon implies that although the GI microbiota is shaped significantly by the host genotypic background, many microbiome functions are shared across hosts, consistent with the similar physicochemical conditions at identical anatomical sites.

Supplementary Materials: Supplementary materials can be found at <http://www.mdpi.com/1422-0067/20/5/1232/s1>. Table S1. Modules differentially abundant in at least one sampling site as determined by LefSe. Table S2. Modules differentially abundant in different niche locations as determined by LefSe.

Author Contributions: Conceptualization, D.L. and H.C.; methodology, D.L.; software, D.L.; formal analysis, D.L.; resources, H.C., J.Z., H.Z. and W.C.; writing—original draft preparation, D.L.; writing—review and editing, H.C.; visualization, D.L.; funding acquisition, H.C., J.Z., H.Z. and W.C.

Funding: This study was supported by the National Natural Science Foundation of China (31820103010, 31722041), the Fundamental Research Funds for the Central Universities (JUSRP51702A), the National First-Class Discipline Program of Food Science and Technology (JUFSTR20180102) and the Program of Collaborative Innovation Center of Food Safety and Quality Control in Jiangsu Province.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Gill, S.R.; Pop, M.; DeBoy, R.T.; Eckburg, P.B.; Turnbaugh, P.J.; Samuel, B.S.; Gordon, J.I.; Relman, D.A.; Fraser-Liggett, C.M.; Nelson, K.E. Metagenomic analysis of the human distal gut microbiome. *Science* **2006**, *312*, 1355–1359. [[CrossRef](#)]
2. Costello, E.K.; Stagaman, K.; Dethlefsen, L.; Bohannan, B.J.M.; Relman, D.A. The application of ecological theory toward an understanding of the human microbiome. *Science* **2012**, *336*, 1255–1262. [[CrossRef](#)]
3. O'Hara, A.M.; Shanahan, F. The gut flora as a forgotten organ. *EMBO Rep.* **2006**, *7*, 688–693. [[CrossRef](#)]
4. Manichanh, C.; Rigottier-Gois, L.; Bonnaud, E.; Gloux, K.; Pelletier, E.; Frangeul, L.; Nalin, R.; Jarrin, C.; Chardon, P.; Marteau, P.; et al. Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. *Gut* **2006**, *55*, 205–211. [[CrossRef](#)]
5. Xavier, R.J.; Podolsky, D.K. Unravelling the pathogenesis of inflammatory bowel disease. *Nature* **2007**, *448*, 427–434. [[CrossRef](#)]
6. Wong, S.H.; Zhao, L.; Zhang, X.; Nakatsu, G.; Han, J.; Xu, W.; Xiao, X.; Kwong, T.N.Y.; Tsoi, H.; Wu, W.K.K.; et al. Gavage of Fecal Samples from Patients with Colorectal Cancer Promotes Intestinal Carcinogenesis in Germ-Free and Conventional Mice. *Gastroenterology* **2017**, *153*, 1621–1633. [[CrossRef](#)]
7. Qin, J.; Li, Y.; Cai, Z.; Li, S.; Zhu, J.; Zhang, F.; Liang, S.; Zhang, W.; Guan, Y.; Shen, D.; et al. A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* **2012**, *490*, 55–60. [[CrossRef](#)]
8. Turnbaugh, P.J.; Ley, R.E.; Mahowald, M.A.; Magrini, V.; Mardis, E.R.; Gordon, J.I. An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* **2006**, *444*, 1027–1031. [[CrossRef](#)]
9. Qin, N.; Yang, F.; Li, A.; Prifti, E.; Chen, Y.; Shao, L.; Guo, J.; Le Chatelier, E.; Yao, J.; Wu, L.; et al. Alterations of the human gut microbiome in liver cirrhosis. *Nature* **2014**, *513*, 59–64. [[CrossRef](#)]
10. Foster, J.A.; Neufeld, K.-A.M. Gut–brain axis: How the microbiome influences anxiety and depression. *Trends Neurosci.* **2013**, *36*, 305–312. [[CrossRef](#)]
11. Turnbaugh, P.J.; Hamady, M.; Yatsunencko, T.; Cantarel, B.L.; Duncan, A.; Ley, R.E.; Sogin, M.L.; Jones, W.J.; Roe, B.A.; Affourtit, J.P.; et al. A core gut microbiome in obese and lean twins. *Nature* **2009**, *457*, 480–484. [[CrossRef](#)]
12. Arumugam, M.; Raes, J.; Pelletier, E.; Le Paslier, D.; Yamada, T.; Mende, D.R.; Fernandes, G.R.; Tap, J.; Bruls, T.; Batto, J.M.; et al. Enterotypes of the human gut microbiome. *Nature* **2011**, *473*, 174–180. [[CrossRef](#)]
13. Muegge, B.D.; Kuczynski, J.; Knights, D.; Clemente, J.C.; González, A.; Fontana, L.; Henrissat, B.; Knight, R.; Gordon, J.I. Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. *Science* **2011**, *332*, 970–974. [[CrossRef](#)]
14. Warnecke, F.; Luginbühl, P.; Ivanova, N.; Ghassemian, M.; Richardson, T.H.; Stege, J.T.; Cayouette, M.; McHardy, A.C.; Djordjevic, G.; Aboushadi, N.; et al. Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. *Nature* **2007**, *450*, 560–565. [[CrossRef](#)]
15. Methé, B.A.; Nelson, K.E.; Pop, M.; Creasy, H.H.; Giglio, M.G.; Huttenhower, C.; Gevers, D.; Petrosino, J.F.; Abubucker, S.; Badger, J.H.; et al. A framework for human microbiome research. *Nature* **2012**, *486*, 215–221.
16. Qin, J.; Li, R.; Raes, J.; Arumugam, M.; Burgdorf, K.S.; Manichanh, C.; Nielsen, T.; Pons, N.; Levenez, F.; Yamada, T.; et al. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **2010**, *464*, 59–65. [[CrossRef](#)]
17. Xiao, L.; Feng, Q.; Liang, S.; Sonne, S.B.; Xia, Z.; Qiu, X.; Li, X.; Long, H.; Zhang, J.; Zhang, D.; et al. A catalog of the mouse gut metagenome. *Nat. Biotechnol.* **2015**, *33*, 1103–1108. [[CrossRef](#)]
18. Li, D.; Chen, H.; Mao, B.; Yang, Q.; Zhao, J.; Gu, Z.; Zhang, H.; Chen, Y.Q.; Chen, W. Microbial Biogeography and Core Microbiota of the Rat Digestive Tract. *Sci. Rep.* **2017**, *7*, 45840. [[CrossRef](#)]
19. Huttenhower, C.; Gevers, D.; Knight, R.; Abubucker, S.; Badger, J.H.; Chinwalla, A.T.; Creasy, H.H.; Earl, A.M.; Fitzgerald, M.G.; Fulton, R.S.; et al. Structure, function and diversity of the healthy human microbiome. *Nature* **2012**, *486*, 207–214.

20. Langille, M.G.I.; Zaneveld, J.; Caporaso, J.G.; McDonald, D.; Knights, D.; Reyes, J.A.; Clemente, J.C.; Burkepile, D.E.; Thurber, R.L.V.; Knight, R.; et al. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat. Biotechnol.* **2013**, *31*, 814–821. [[CrossRef](#)]
21. Kanehisa, M.; Furumichi, M.; Tanabe, M.; Sato, Y.; Morishima, K. KEGG: New perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* **2017**, *45*, D353–D361. [[CrossRef](#)]
22. Gibbs, R.A.; Weinstock, G.M.; Metzker, M.L.; Muzny, D.M.; Sodergren, E.J.; Scherer, S.; Scott, G.; Steffen, D.; Worley, K.C.; Burch, P.E.; et al. Genome sequence of the Brown Norway rat yields insights into mammalian evolution. *Nature* **2004**, *428*, 493–521.
23. Venter, J.C.; Adams, M.D.; Myers, E.W.; Li, P.W.; Mural, R.J.; Sutton, G.G.; Smith, H.O.; Yandell, M.; Evans, C.A.; Holt, R.A.; et al. The Sequence of the Human Genome. *Science* **2001**, *291*, 1304–1351. [[CrossRef](#)]
24. Waterston, R.H.; Lindblad-Toh, K.; Birney, E.; Rogers, J.; Abril, J.F.; Agarwal, P.; Agarwala, R.; Ainscough, R.; Alexandersson, M.; An, P.; et al. Initial sequencing and comparative analysis of the mouse genome. *Nature* **2002**, *420*, 520–562. [[PubMed](#)]
25. Van den Abbeele, P.; Belzer, C.; Goossens, M.; Kleerebezem, M.; De Vos, W.M.; Thas, O.; De Weirldt, R.; Kerckhof, F.M.; Van De Wiele, T. Butyrate-producing Clostridium cluster XIVa species specifically colonize mucins in an *in vitro* gut model. *ISME J.* **2013**, *7*, 949–961. [[CrossRef](#)]
26. Stern, M.O.; Santos, K.A.; Satter, L.D. Protein Degradation in Rumen and Amino Acid Absorption in Small Intestine of Lactating Dairy Cattle Fed Heat-Treated Whole Soybeans. *J. Dairy Sci.* **1985**, *68*, 45–56. [[CrossRef](#)]
27. Rizzello, C.G.; De Angelis, M.; Di Cagno, R.; Camarca, A.; Silano, M.; Losito, I.; De Vincenzi, M.; De Bari, M.D.; Palmisano, F.; Maurano, F.; et al. Highly efficient gluten degradation by lactobacilli and fungal proteases during food processing: New perspectives for celiac disease. *Appl. Environ. Microbiol.* **2007**, *73*, 4499–4507. [[CrossRef](#)] [[PubMed](#)]
28. Fadda, S.; Oliver, G.; Vignolo, G. Protein degradation by *Lactobacillus plantarum* and *Lactobacillus casei* in a sausage model system. *J. Food Sci.* **2002**, *67*, 1179–1183. [[CrossRef](#)]
29. Muto, A.; Kotera, M.; Tokimatsu, T.; Nakagawa, Z.; Goto, S.; Kanehisa, M. Modular architecture of metabolic pathways revealed by conserved sequences of reactions. *J. Chem. Inf. Model.* **2013**, *53*, 613–622. [[CrossRef](#)]
30. Borgström, B.; Dahlqvist, A.; Lundh, G.; Sjövall, J. Studies of intestinal digestion and absorption in the human. *J. Clin. Investig.* **1957**, *36*, 1521–1536. [[CrossRef](#)]
31. Gendler, S.J.; Lancaster, C.A.; Taylor-Papadimitriou, J.; Duhig, T.; Peat, N.; Burchell, J.; Pemberton, L.; Lalani, E.N.; Wilson, D. Molecular cloning and expression of human tumor-associated polymorphic epithelial mucin. *J. Biol. Chem.* **1990**, *265*, 15286–15293. [[PubMed](#)]
32. Salyers, A.A.; Vercellotti, J.R.; West, S.E.H.; Wilkins, T.D. Fermentation of mucin and plant polysaccharides by strains of Bacteroides from the human colon. *Appl. Environ. Microbiol.* **1977**, *33*, 319–322. [[PubMed](#)]
33. He, G.; Shankar, R.A.; Chzhan, M.; Samouilov, A.; Kuppusamy, P.; Zweier, J.L. Noninvasive measurement of anatomic structure and intraluminal oxygenation in the gastrointestinal tract of living mice with spatial and spectral EPR imaging. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 4586–4591. [[CrossRef](#)] [[PubMed](#)]
34. Albenberg, L.; Esipova, T.V.; Judge, C.P.; Bittinger, K.; Chen, J.; Laughlin, A.; Grunberg, S.; Baldassano, R.N.; Lewis, J.D.; Li, H.; et al. Correlation between intraluminal oxygen gradient and radial partitioning of intestinal microbiota. *Gastroenterology* **2014**, *147*, 1055–1063. [[CrossRef](#)]
35. LeBlanc, J.G.; Milani, C.; de Giori, G.S.; Sesma, F.; van Sinderen, D.; Ventura, M. Bacteria as vitamin suppliers to their host: A gut microbiota perspective. *Curr. Opin. Biotechnol.* **2013**, *24*, 160–168. [[CrossRef](#)]
36. Christen, P.; Mehta, P.K. From cofactor to enzymes. The molecular evolution of pyridoxal-5'-phosphate-dependent enzymes. *The Chem. Rec.* **2001**, *1*, 436–447. [[CrossRef](#)]
37. Gu, S.; Chen, D.; Zhang, J.N.; Lv, X.; Wang, K.; Duan, L.P.; Nie, Y.; Wu, X.L. Bacterial community mapping of the mouse gastrointestinal tract. *PLoS ONE* **2013**, *8*, e74957. [[CrossRef](#)]
38. Nava, G.M.; Friedrichsen, H.J.; Stappenbeck, T.S. Spatial organization of intestinal microbiota in the mouse ascending colon. *ISME J.* **2011**, *5*, 627–638. [[CrossRef](#)]
39. Li, H.; Limenitakis, J.P.; Fuhrer, T.; Geuking, M.B.; Lawson, M.A.; Wyss, M.; Brugiroux, S.; Keller, I.; Macpherson, J.A.; Rupp, S.; et al. The outer mucus layer hosts a distinct intestinal microbial niche. *Nat. Commun.* **2015**, *6*, 8292. [[CrossRef](#)]
40. Kohl, K.D.; Miller, A.W.; Marvin, J.E.; Mackie, R.; Dearing, M.D. Herbivorous rodents (*Neotoma spp.*) harbour abundant and active foregut microbiota. *Environ. Microbiol.* **2014**, *16*, 2869–2878. [[CrossRef](#)]

41. Stearns, J.C.; Lynch, M.D.; Senadheera, D.B.; Tenenbaum, H.C.; Goldberg, M.B.; Cvitkovitch, D.G.; Croitoru, K.; Moreno-Hagelsieb, G.; Neufeld, J.D. Bacterial biogeography of the human digestive tract. *Sci. Rep.* **2011**, *1*, 170. [[CrossRef](#)]
42. DeSantis, T.Z.; Hugenholtz, P.; Larsen, N.; Rojas, M.; Brodie, E.L.; Keller, K.; Huber, T.; Dalevi, D.; Hu, P.; Andersen, G.L. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* **2006**, *72*, 5069–5072. [[CrossRef](#)]
43. Haas, B.J.; Gevers, D.; Earl, A.M.; Feldgarden, M.; Ward, D.V.; Giannoukos, G.; Ciulla, D.; Tabbaa, D.; Highlander, S.K.; Sodergren, E.; et al. Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. *Genome Res.* **2011**, *21*, 494–504. [[CrossRef](#)]
44. Abubucker, S.; Segata, N.; Goll, J.; Schubert, A.M.; Izard, J.; Cantarel, B.L.; Rodriguez-Mueller, B.; Zucker, J.; Thiagarajan, M.; Henrissat, B.; et al. Metabolic Reconstruction for Metagenomic Data and Its Application to the Human Microbiome. *PLoS Comput. Biol.* **2012**, *8*, e1002358. [[CrossRef](#)]
45. Ye, Y.; Doak, T.G. A Parsimony Approach to Biological Pathway Reconstruction/Inference for Metagenomes. *PLoS Comput. Biol.* **2011**, *5*, e1000465.
46. Dixon, P. VEGAN, a package of R functions for community ecology. *J. Veg. Sci.* **2003**, *14*, 927–930. [[CrossRef](#)]
47. Vázquez-Baeza, Y.; Pirrung, M.; Gonzalez, A.; Knight, R. EMPERor: A tool for visualizing high-throughput microbial community data. *Gigascience* **2013**, *2*, 16. [[CrossRef](#)]
48. Segata, N.; Izard, J.; Waldron, L.; Gevers, D.; Miropolsky, L.; Garrett, W.S.; Huttenhower, C. Metagenomic biomarker discovery and explanation. *Genome Biol.* **2011**, *12*, R60. [[CrossRef](#)]
49. Asnicar, F.; Weingart, G.; Tickle, T.L.; Huttenhower, C.; Segata, N. Compact graphical representation of phylogenetic data and metadata with GraPhlAn. *PeerJ* **2015**, *3*, e1029. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).