

SCIENTIFIC REPORTS



OPEN

Diet-induced reconstruction of mucosal microbiota associated with alterations of epithelium lectin expression and regulation in the maintenance of rumen homeostasis

Hong Shen^{1,2}, Zhongyan Lu³, Zhihui Xu^{1,2} & Zanming Shen³

It is unknown whether lectins of the rumen epithelium contribute to the recognition of mucosal microbes and activation of tolerogenic cytokines in ruminant animals. We applied an integrated method of RNA-seq and 16S rRNA gene sequencing to investigate alterations of epithelial lectin expression and regulation with a diet-induced reconstruction of the mucosal microbiota in the goat rumen. Our results showed that the diversity and richness of the rumen mucosal microbiota were promoted by the dietary concentrate. Meantime, in the rumen epithelium, five lectin genes, namely, sialic acid-binding Ig-like lectin 14 (LOC102180073), C-type lectin domain family 4, member E (CLEC4E), C-type lectin domain family 7, member A (CLEC7A), C-type lectin domain family 16, member A (CLEC16A), and lectin, mannose-binding 2 (LMAN2), were indicated to promote the expression of 8 tolerogenic cytokines, transforming growth factor beta 1 (TGFB1) and 4 enzyme genes involved in retinoic acid biosynthesis via 6 signaling pathways. Analysis of the combined data showed that 9 microbial genera (*Clostridium_IV*, *Desulfobulbus*, *Eubacterium*, *Ochrobactrum*, *Propionibacterium*, *Pseudomonas*, *Slackia*, *Staphylococcus* and *Subdivision5_genera_IS*) were highly related to the expression of functional lectins. These findings provide new insights into the interactions between the rumen epithelium and mucosal microbiota in the maintenance of rumen homeostasis.

The mucosal microbiota is a group of gastrointestinal (GI) microbiota that inhabits the mucus layer of the GI tract. Compared with the group inhabiting the fluid phase of the GI tract, the mucosal microbiota plays a more critical role in the development and performance of the immune functions of the GI epithelium, as well as the whole organism^{1,2}. Throughout evolution, the GI epithelium has developed a complex gene network to regulate the activities of the mucosal microbiota and responses of immune cells to provide a harmonious coexistence between the host's immune system and GI microbiota. In healthy individuals, this network recognizes commensal bacteria in the mucus layer and subsequently activates tolerogenic signals to suppress inappropriate immune responses^{3,4}. Thus, it plays a crucial role in the maintenance of GI homeostasis. However, to date, it is unclear how the GI epithelium recognizes mucosal bacteria and which types of signals are involved in the communication between the GI epithelium and mucosal microbiota.

Lectins are a class of genes whose sequences have several carbohydrate-recognition domains (also referred to as lectin domains)⁵. They are evolutionally conserved pattern-recognition receptors (PRRs) that mediate the agglutination and immobilization of recognized pathogens⁶. Some lectins, such as mannose-binding lectin (MBL) and ficolin, are established as activators of the complement system, which plays important roles in the promotion of the innate and adaptive immunity responses^{7,8}. To date, more than 10 lectin families have been identified in eukaryotes according to sequence differences of the lectin domains⁹. Accumulating evidence has shown

¹College of Life Science, Nanjing Agricultural University, Nanjing, Jiangsu, China. ²Bioinformatics Center, Nanjing Agricultural University, Nanjing, Jiangsu, China. ³Key Lab of Animal Physiology and Biochemistry, College of Veterinary Medicine, Nanjing Agricultural University, Nanjing, Jiangsu, China. Hong Shen and Zhongyan Lu contributed equally to this work. Correspondence and requests for materials should be addressed to Z.L. (email: luzhongyan@njau.edu.cn)

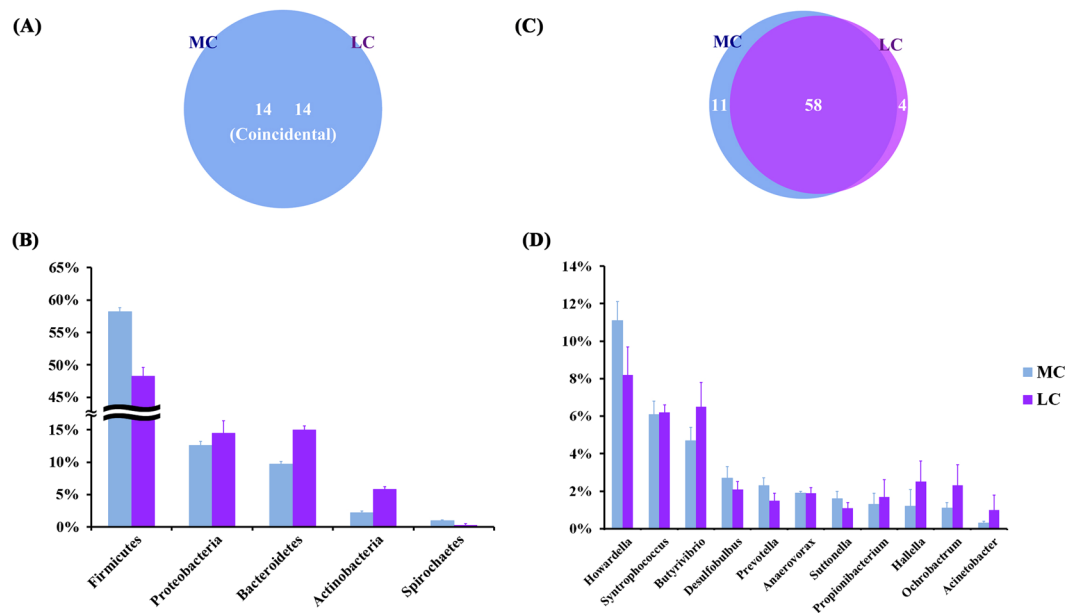


Figure 1. (A) Venn diagram showing the coincidence of phyla between the groups. (B) Phylum-level comparison of bacterial OTUs between the groups. (C) Venn diagram showing the overlap of genera between groups. (D) Genus-level comparison of bacterial OTUs between the groups.

that members from three lectin families, C-type lectins, galectins and siglecs, play crucial roles in suppressing the immune response by recognizing O-glycan-coated cells (including bacteria) in mammals¹⁰. For example, mucin-induced activation of C-type lectins on human dendritic cells (DCs) promotes DC secretion of the tolerogenic cytokines interleukin (IL)-10 and transforming growth factor (TGF)- α ¹¹. Tumor cell-induced activation of siglecs on antigen-presenting cells (APCs) of mice leads to APC secretion of IL-10, TGFs and chemokines¹². Subsequently, tolerogenic cytokines, as well as TGF and retinoic acid (RA), promote the generation of regulatory T cells (Tregs), which play important roles in the suppression of inappropriate immune responses in the GI epithelium¹³. Although there is currently no available report concerning the functions of lectins of the GI epithelium in the regulation of immune tolerance, lectins are present and constantly expressed in the GI epithelium of animals^{14–16}. Accordingly, we infer that these lectins contribute to GI homeostasis by recognizing nonpathogenic bacteria that inhabit the mucus layer and elicit tolerogenic signals in the GI epithelium.

The rumen is the most important digestive and absorptive organ, as well as the most crucial habitat of symbiotic bacteria, in ruminant animals. Previous studies have shown that the dietary concentrate facilitates rumen fermentation and consequently affects the structure of the mucosal microbiota in the rumen¹⁷. We found that the dietary concentrate promoted immunity of the rumen epithelium by strengthening signaling from toll-like receptors (TLRs)¹⁸. However, it is unknown which members of the lectin families are expressed in the rumen epithelium and which regulatory mechanisms these lectins are used for rumen homeostasis. RNA-seq is a powerful technology for determining the expression and regulatory mechanisms of less-informed gene families. In this study, we shifted the dietary concentrate from 10% to 35% to induce reconstruction of the mucosal microbiota in the goat rumen. After that, we used 16S rRNA gene sequencing to investigate alterations of the mucosal microbiota and used RNA-seq to select the candidate lectins that are involved in the maintenance of rumen homeostasis. Finally, we used statistical tests to determine the regulatory pathways and tolerogenic signals related to the candidate lectins and identify the covariance of the mucosal bacteria and lectin transcripts. By this process, we aimed to understand the regulatory mechanism of lectins in the maintenance of rumen homeostasis.

Results

Histomorphometric Analysis of Rumen Epithelium. The 35% concentrate (MC) diet led to an increase in the ruminal papillae length, width, density, and number of cell layers compared with the 10% concentrate (LC) diet. However, in all samples, the morphology of the rumen epithelium remained complete.

Structure of Epithelium Bacterial Communities. At the phylum level, a total of 14 prokaryotic phyla were identified with 97% similarity, and all of them were common to both groups (Fig. 1A). Firmicutes (23.3–33.2%), Proteobacteria (12.6–14.5%), and Bacteroidetes (9.7–15.0%) were the most abundant among all microbial communities. Compared with the LC group, the relative abundance of Spirochetes was increased by 230%, Verrucomicrobia was increased by 50%, and Firmicutes was increased by 42% in the MC group. By contrast, the relative abundance of the remaining phyla was decreased. The most significant decrease occurred in the phylum Actinobacteria, whose relative abundance was decreased by 62% in the MC group. At the genus level, except for unclassified OTUs, a total of 73 genera were detected in the sequences. Among them, 58 genera were common to both groups (Fig. 1C). The abundances of all of the genera in both groups are shown in Table S2. *Howardella* (11.1–8.2%) was consistently abundant in both groups, except for the unclassified OTUs at the genus level. Four

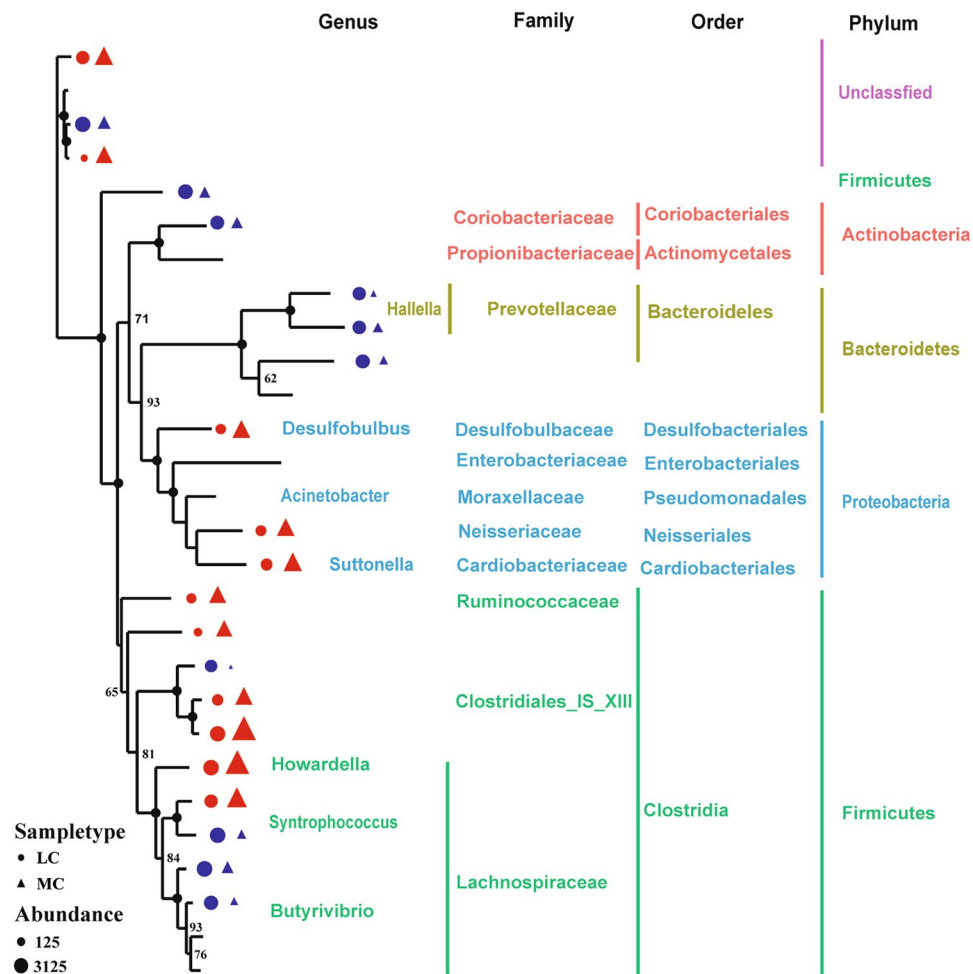


Figure 2. Maximum likelihood tree of 28 detectable OTUs (the relative abundance >1% in the given sample). The complete 16S rRNA gene sequences of the corresponding species in the RDP database were used to construct the tree. The triangle indicates the OTUs in the MC group, and the circle indicates the OTUs in the LC group. Only the OTUs with significantly different ($p < 0.05$) relative abundances are shown behind the branches. The size of the symbol indicates the relative abundance of the OTUs. Red indicates a significant expansion ($p < 0.05$) of the relative abundance of the OTU under the 35% concentrate diet and blue indicates a significant reduction ($p < 0.05$) in the relative abundance of the OTU under a 10% concentrate diet. Only those bootstrap values greater than 60 are shown on the tree. The solid black circles at the nodes stand for a bootstrap value of 100.

genera were only detected in the LC group, and 11 genera were only detected in the MC group. Nonmetric multidimensional scaling (NMDS) plots (Fig. S1) and analysis of similarities (ANOSIM) ($p < 0.05$) revealed the divergence of the community structure in the MC and LC groups.

Diversity and Richness of Microbial Communities. The diversity of the microbial community significantly decreased in the MC group, as indicated by the Shannon index. At the phylum level, the diversity of Bacteroidetes in the MC group was significantly higher than that in the LC group ($p < 0.05$). The diversity of Proteobacteria and Firmicutes in the MC group was significantly lower than that in the LC group (Fig. S2). Maximum likelihood (ML) analysis of 28 detectable OTUs (the relative abundance >1%) showed that the significantly expanded OTUs in the MC group belonged to the families Desulfobulbaceae, Neisseriaceae, Cardiobacteriaceae, Ruminococcaceae, and Clostridiales_IS_XIII (Fig. 2). By contrast, the significantly decreased OTUs belonged to the families Coriobacteriaceae and Prevotellaceae. OTUs belonging to the family Lachnospiraceae showed an inconsistent trend, with 2 OTUs expanding significantly and 3 OTUs shrinking significantly in the MC group.

A rarefaction curve (Fig. S3) was applied to evaluate the richness of the microbiota in each given sample. Compared with the MC group, these curves revealed that the OTU number was significantly increased in the LC group.

Expression Profiles of Lectins in the Rumen Epithelium. The RNA-seq method was used to identify the repertoire of transcribed lectins in the rumen epithelium. It generated a total of 135 M raw reads (average

22.5 M reads per sample, range 20.9–23.5 M) and 127 M clean reads (average 21.2 M reads per sample, range 19.1–22.5 M). On average, 95% of the reads were successfully mapped to the NCBI goat genome annotation release version 101.

One hundred members of the lectin families were found to be encoded on the goat genome. The transcriptome data showed that 35 members were expressed in the rumen epithelium (reads per kilobase million; RPKM >1 in at least one sample) (Table 1). Compared with expression in the LC group, we found that 3 members were significantly upregulated ($\log_2(\text{MC}/\text{LC}) > 1$) and 10 members were significantly downregulated ($\log_2(\text{MC}/\text{LC}) < 1$) in the MC group. By comparing the expression among the members, we found that the expression of galectin-7 (LOC102180339) was highest in all samples (376 RPKM–513 RPKM). Meanwhile, six other members expressed more than 100 RPKM in at least one group.

Related KEGG Pathways of Differentially Expressed Lectins. We used the spearman correlation coefficient (SCC) to quantify gene–gene co-expression for all of the gene-pairs in the transcriptome across all samples. To assess the biological significance of the lectin co-expression network, we checked the related signaling pathways and possible functions of the first neighbors by the Kyoto Encyclopedia of Genes and Genomes (KEGG) annotation. The functions of the significantly different neighbors were classified according to the regulated physiological processes of the rumen epithelium. Accordingly, the regulatory pathways involved in the modulation of immune activities and cytokine productions in this study were assembled and referred to as the immune regulation network (Fig. 3). Other possible types of regulation involved in the cellular community, development, metabolism, and transcription modulation are listed in Table S3.

Five members of the lectin families that had upregulated expression in the LC group were involved in the regulation of immune activities in the rumen epithelium: sialic acid-binding Ig-like lectin 14 (LOC102180073), C-type lectin domain family 16, member A (CLEC16A), C-type lectin domain family 7, member A (CLEC7A), mannose-binding lectin 2 (LMAN2L), and C-type lectin domain family 4, member E (CLEC4E). Except for the negative correlations between the expression of CLEC16A and ataxin 3 (ATXN3) as well as the Fc fragment of immunoglobulin E receptor 1a (FCER1A), there were positive correlations for all gene pairs. Endocytosis, exosome, lysosome, complement and coagulation, inhibition of DNA biosynthesis, immunoglobulin A (IgA) production, platelet activation, and autophagy were directly regulated by neighboring genes. However, leukocyte trans-endothelial migration, leukocyte recruitment, stimulation and amplification of Th2 cell responses, and cytokine production were regulated via activation of specific relevant signaling pathways. The ubiquitin system was involved in lectin modulation of proteolysis.

Differentially Expressed Cytokines. Comparing the expression of the genes encoding ILs, TGFs and chemokines between groups, we found that the expression of 8 cytokine genes and TGFB1 was significantly upregulated in the LC group (Table 2). Among the differentially expressed ILs, IL22 was reported to be secreted by the group 3 innate lymphoid cells in the mucosal tissue and to play important roles in the maintenance of immune homeostasis¹⁹. IL19, belonging to the IL-10 family, was reported to be an anti-inflammatory cytokine²⁰. Other mentioned genes and their functions are listed in Table 2^{21–25}.

Relationships between Lectin Expression and the Relative Abundance of Bacterial Genera. Spearman correlation analysis removed 90% of the OTUs (585/647), which have a SCC with differentially expressed lectins of the correlation coefficient less than 0.6 and *p*-value greater than 0.05. Subsequently, constrained correspondence analysis (CCA) showed positive correlations between functioning lectins (referred to as the lectins that take part in the above-mentioned regulation networks) in 9 expanded microbial genera. The microbial genera that were related to functional lectin transcripts are listed in Fig. 4.

Discussion

The GI tract is covered by a layer of heavily glycosylated proteins, commonly known as mucins. They are an important nutrient source for the microbiota that resides in the mucus layer. Previous studies have shown that the carbohydrate availability in the mucus layer affects the composition of the microbiota found there^{26–28}. In the present study, an LC diet promoted diversification and expansion of the mucosal microbiota, indicating that the amount of energy substrates in the rumen fluid also has an impact on the composition of the mucosal microbiota. To date, the effects of dietary concentrations on the rumen mucosal microbiota are controversial. Mao *et al.*²⁹ reported that a 65% dietary concentrate diet decreased the diversity of the mucosal microbiota in the goat rumen compared with a 0% concentrate diet. However, Wetzels *et al.*³⁰ observed a positive relationship between dietary energy substrates and the diversity and richness of the mucosal microbiota in the goat rumen by using 0%, 30%, and 60% concentrate diets. The damage to the rumen epithelium after weeks of a zero or high concentrate diet might be the reason for these inconsistent results. Based on such considerations, we confirmed the integrity of the ruminal epithelium in all samples by using optical microscopy. Previous studies have shown that microbes with the ability to metabolize mucins have an advantage in colonization and expansion in the mucus layer. Except for unclassified OTUs, the detectable OTUs promoted by the LC diet in this study belong to the families Lachnospiraceae and Prevotellaceae. Members of Lachnospiraceae are reported to have the ability to degrade sialic acids and fucose in the mucus layer^{31,32}. Members of Prevotellaceae can degrade the core structure of the terminal mucin O-glycan³³. Accordingly, our results support the idea that microbes that have the ability to metabolize carbohydrates in the mucosal layer have a competitive advantage in niches under a low energy diet.

Previous studies have indicated that separating microbes from the GI epithelium is a major task of the GI immune system in the maintenance of GI homeostasis³⁴. According to the study by Vaishnavi *et al.*³⁵, the soluble lectins in the mucus layer play important roles in the spatial segregation of the microbiota and GI epithelium. Galectins, which include 15 known mammalian members, are soluble lectins located in the extracellular milieu³⁶.

Item	Symbol	Description	MC (RPKM)	LC (RPKM)	Log ₂ (MC/LC)	Lectin Family
1	LOC102180073	sialic acid-binding Ig-like lectin 14	0.03	1.33	-5.47	siglec
2	FCN1	ficolin 1	0.93	4.98	-2.42	galectin
3	CLEC7A	C-type lectin domain family 7, member A	0.65	1.84	-1.50	CLR
4	LOC102181776	sialic acid-binding Ig-like lectin 14	0.78	1.89	-1.28	siglec
5	LOC102184901	C-type lectin domain family 2 member D11-like	65.59	147.84	-1.17	CLR
6	CLEC4E	C-type lectin domain family 4, member E	1.27	2.82	-1.15	CLR
7	LGALS9	lectin, galactoside-binding, soluble, 9	5.40	11.44	-1.08	galectin
8	CLEC16A	C-type lectin domain family 16, member A	2.11	4.35	-1.04	CLR
9	LGALS3BP	lectin, galactoside-binding, soluble, 3 binding protein	3.24	6.65	-1.04	galectin
10	LMAN2	lectin, mannose-binding 2	30.76	62.86	-1.03	CLR
11	LGALS15	lectin, galactoside-binding, soluble, 15	1.50	0.64	1.23	galectin
12	LOC106503212	C-type lectin domain family 10 member A-like	3.99	1.96	1.03	CLR
13	LOC102189655	C-type lectin domain family 2 member D11-like	2.48	1.24	1.00	CLR
14	LOC102180339	galectin-7	393.68	486.38	-0.31	galectin
15	LOC102180072	galectin-7	280.25	327.80	-0.23	galectin
16	CLEC3B	C-type lectin domain family 3, member B	263.21	205.23	0.36	CLR
17	LGALS3	lectin, galactoside-binding, soluble, 3	99.80	108.23	-0.12	galectin
18	LGALS	lectin, galactoside-binding-like	114.98	99.52	0.21	galectin
19	LGALS1	lectin, galactoside-binding, soluble, 1	148.30	86.79	0.77	galectin
20	OS9	OS9, endoplasmic reticulum lectin	42.99	49.58	-0.21	CLR
21	ERLEC1	endoplasmic reticulum lectin 1	25.02	21.84	0.20	CLR
22	LOC106504001	C-type lectin domain family 1 member A-like	20.86	20.86	0.00	CLR
23	LMAN1	lectin, mannose-binding, 1	12.35	14.85	0.00	CLR
24	LGALS8	lectin, galactoside-binding, soluble, 8	14.36	14.59	0.27	galectin
25	LGALS4	lectin, galactoside-binding, soluble, 4	10.33	7.80	0.02	galectin
26	LOC102180574	C-type lectin domain family 6 member A	7.15	6.63	-0.41	CLR
27	COLEC12	collectin sub-family member 12	6.82	5.96	-0.11	CLR
28	LOC102180291	C-type lectin domain family 4 member A	4.92	3.52	-0.19	CLR
29	CLEC5A	C-type lectin domain family 5, member A	1.63	2.01	-0.48	CLR
30	CLEC12A	C-type lectin domain family 12, member A	2.17	1.82	0.30	CLR
31	CLEC2A	C-type lectin domain family 2, member A	2.89	1.69	-0.25	CLR
32	CLEC1A	C-type lectin domain family 1, member A	2.47	1.51	-0.77	CLR
33	LMAN2L	lectin, mannose-binding 2-like	0.89	1.50	-0.71	CLR
34	CLEC14A	C-type lectin domain family 14, member A	0.88	1.36	0.75	CLR
35	LOC102189932	C-type lectin domain family 2 member H-like	1.66	1.02	0.63	CLR

Table 1. Members of the lectin families expressed and observed in the present study.

In the present study, expression of galectin-7 was consistently highest in all samples, thereby indicating the constitutive roles of this lectin in the maintenance of rumen homeostasis. However, its expression was not synchronized with the expansion of the mucosal microbiota, suggesting a loose relationship between galectin-7 and the activities of the microbes.

To date, most C-type lectins (CLRs) and all known sialic acid-binding immunoglobulin-like lectins (siglecs) have been shown to be membrane-bound proteins that are found on the cell surface and that contain one or more carbohydrate-recognition domains (CRDs) that are responsible for glycan binding¹⁰. In the present study, one siglec gene (LOC102180073) and four CLR genes (CLEC4E, CLEC7A, CLEC16A, and LMAN2) were shown to be involved in 13 types of immune activities. Moreover, their expression was significantly increased with the expansion of the mucosal microbiota under the LC diet. Previous studies have shown that upon recognition of specific pathogens, innate immune cell CLRs contribute to DC differentiation and immune response programs by affecting cytokine production^{37–41}. A well-described example is the signaling function of the activated CLR dectin-1 in the recognition of fungal α -glucan structures. Activated dectin-1 of DCs triggers expression of TNF- α , IL-10, and IL-12 by activating the p38-Erk-JNK kinase cascade and transcription factor NF- κ B⁴⁰. The CLR DC-specific intercellular adhesion molecule-3-grabbing nonintegrin (SIGN) of DCs has also been shown to signal upon recognition of pathogen glycan components, which trigger expression of IL-10 via TLR-dependent NF- κ B signaling pathway³⁷. In the present study, we observed that activation of CLRs had an impact on cytokine production by affecting 6 signaling pathways. In addition to the NF- κ B and TLR signaling pathways, the chemokine signaling pathway, TNF signaling pathway, cytosolic DNA-sensing pathway, and RIG-I-like receptor signaling pathway were also involved in cytokine activation under the LC diet. Moreover, the co-expression network constructed in this study predicted that 4 additional mechanisms were involved in the programming functions of CLRs on the innate and adaptive immune responses, namely, leukocyte recruitment, leukocyte trans-endothelial migration,

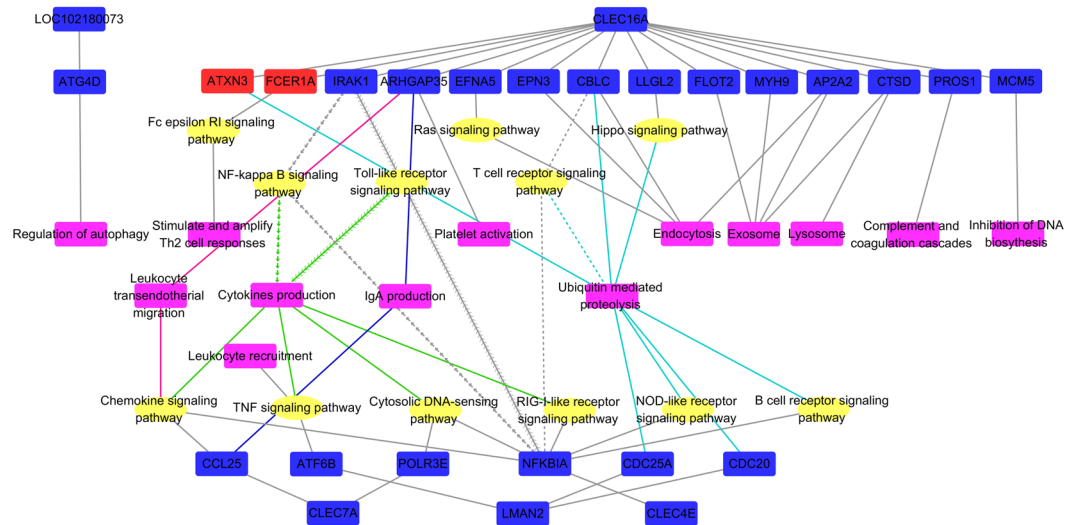


Figure 3. Regulation network of the lectin genes related to immune activities. The functions of the first neighbors were predicted by KEGG pathway analysis. Genes that show positive correlation with the 35% concentrate feeding are in red, and the genes that show negative correlation with the 35% concentrate diet are in blue. Signaling pathways are in yellow, and regulatory functions are in violet. The genes regulating the same signaling pathways are given the same line shapes and the genes and signaling pathways regulating the same functions are given the same color.

stimulation and amplification of Th2 cell responses, and platelet activation. Except for the inhibitory effects on the stimulation and amplification of Th2 cell responses, other activities appear to be enhanced by the LC diet. These activities, which are associated with an expanded microbiota, seem to create an alert environment for the immune system, with proper suppression of the inflammatory responses at the barrier site.

Alteration of microbial metabolic products in the mucus layer is a crucial reason for the change in lectin expression in the ruminal epithelium. However, since all known bacteria are coated with glycans, microbial composition alterations might also be a reason for the change in lectin expression. Over 20 pathogens are known to be recognized by lectins in animals. The ability to synthesize or capture glycans from their hosts and incorporate them into their own glycoconjugates is considered to be an important strategy to enable pathogens to escape an immune attack from the host's immune system¹⁰. For example, an interaction between the high-mannose structures on the pathogen *Mycobacterium tuberculosis* and the CLR DC-SIGN on DCs leads to the production of anti-inflammatory cytokines, such as IL-10⁴². The glycan on the pathogen *Helicobacter pylori* activates the DC-SIGN on DCs, thereby inducing the conversion of Th1 into Th2⁴³. In the present study, we predicted microbes that correlated with the expression of functioning lectins by statistical methods. Among them, members of *Desulfobulbus* are important mucosa-associated sulfate-reducing bacteria (SRB) in the human colon. They are thought to play roles in the pathogenesis of chronic inflammatory disorders of the colon⁴⁴. Bacteria belonging to the lineage of the subdivision5_genera_IS have important functions in the suppression of immune responses and maintenance of GI homeostasis^{45–47}. *Staphylococcus*, *Propionibacterium* and *Ochrobactrum* are potential pathogens to animals^{48–50}. Members of *Eubacterium* and *Clostridium_IV* have been revealed to be major butyrate-producing bacteria in the human colon⁵¹. Members of *Slackia* have been shown to exhibit a high adhesion capacity to human Caco-2 cells⁵² and to also be involved in the pathogenesis of prostate cancer⁵³. However, little information is presently available on the cell-wall components and metabolites of the gut bacteria that belong to these clades. Other than correlation analysis, we have not been able to make further inferences regarding the interactions between the microbes and functioned lectins studied here. Further studies are required to identify their functions and significance in the maintenance of GI homeostasis.

In summary, we found that the low concentrate diet promoted colonization of opportunistic pathogens in the mucus layer of the rumen. Such alterations in the compositions of the mucosal microbiota induced upregulated expression of five membrane-bound lectins on the rumen epithelium. Subsequently, these five functional lectins contributed to the downregulated expression of 9 tolerogenic cytokines, TGFβ1 and 4 enzyme genes involved in retinoic acid biosynthesis by downregulating the expression of genes related to six signaling pathways. Besides, the decrease of 9 microbial clades in the mucus layer was indicated to be correlated to upregulated expression of these functional lectins in the rumen epithelium. In sum, our study indicates that the lectins of the rumen epithelium play important roles in the transduction of microbial signals from the mucus layer to the immune system. They recognize specific bacteria in the mucus layer and, subsequently, induce tolerogenic signals in the rumen epithelium to suppress the attack of immune cells on nonpathogenic bacteria in the mucus layer. This study expands our fundamental knowledge concerning the roles of evolutionarily conserved lectins in the maintenance of rumen homeostasis and provides new insight into the interactions between the mucosal microbiota and rumen epithelium.

Symbol	Description	MC (RPKM)	LC (RPKM)	Log2 (MC/LC)	Reported Function in GI epithelium
IL22	interleukin 22	2.11	11.27	-2.42	Group 3 ILC derived homeostatic cytokine ¹⁶
IL19	interleukin 19	84.30	306.20	-1.86	IL-10 family anti-inflammatory cytokine ¹⁷
IL1B	interleukin 1 beta	6.95	14.43	-1.05	Pro-inflammatory cytokine
IL26	interleukin 26	7.69	3.34	1.20	Regulation of intercellular adhesion ²¹
IL33	interleukin 33	24.37	11.53	1.08	Regulation of cell cycle ²⁰
LOC102170310	C-C motif chemokine 15	12.80	29.18	-1.19	Recruitment of monocytes, T cells, eosinophils ¹⁹
LOC102181154	C-C motif chemokine 3	1.21	3.35	-1.47	Recruitment of monocytes, T cells, NK cells, basophils, eosinophils ¹⁹
CXCL8	C-X-C motif chemokine ligand 8	41.09	153.22	-1.90	Recruitment of neutrophils, T cells, basophils, endothelial cells ¹⁹
CX3CL1	C-X3-C motif chemokine ligand 1	0.80	2.35	-1.56	Recruitment of effector T cells ¹⁹
CXCL11	C-X-C motif chemokine ligand 11	2.06	5.08	-1.30	Recruitment of T cells ¹⁹
LOC102174969	C-C motif chemokine 8	11.57	5.16	1.17	Recruitment of monocytes, T cells, eosinophils, basophils, NK cells ¹⁹
CXCL14	C-X-C motif chemokine ligand 14	194.03	86.38	1.17	Recruitment of neutrophils, NK cells ¹⁹
TGFB1	transforming growth factor beta 1	1.39	3.54	-1.35	Regulation of DCs' conversion ⁵
RDH13	retinol dehydrogenase 13	2.86	6.31	-1.14	Metabolize retinol to retinaldehyde ¹⁸
ALDH1A3	aldehyde dehydrogenase 1 family member A3	11.92	25.31	-1.09	Metabolize retinaldehyde to retinoic acid ¹⁸
ALDH16A1	aldehyde dehydrogenase 16 family member A1	2.33	4.73	-1.02	Metabolize retinaldehyde to retinoic acid ¹⁸
ALDH4A1	aldehyde dehydrogenase 4 family member A1	2.90	8.32	-1.52	Metabolize retinaldehyde to retinoic acid ¹⁸
ALDH1A1	aldehyde dehydrogenase 1 family member A1	1420.73	558.50	1.35	Metabolize retinaldehyde to retinoic acid ¹⁸

Table 2. Differentially expressed genes encoding interleukins, transforming growth factors and chemokines.

Methods

Ethics approval. This study was approved by the Animal Care and Use Committee of Nanjing Agricultural University, in compliance with the Regulations for the Administration of Affairs Concerning Experimental Animals (The State Science and Technology Commission of P. R. China, 1988) and the Care and Use of Animals (Nanjing Agricultural University, 1999).

Animals. Six male goats (Boer × Yangtze River Delta White, aged 4 months) were randomly allocated into two groups and received either a diet of 65% hay plus 35% concentrate (MC group, $n = 3$) or a diet of 90% hay plus 10% concentrate (LC group, $n = 3$) (Table S1). All goats were fed with two equal portions of the designated diet at 0800 and 1700 daily for 28 days. Water was freely available to all goats during the experimental period. On day 28, the goats were killed at a local slaughterhouse.

Sample Collection. On day 28, all goats were slaughtered at 8 h after the morning feeding. Rumen tissue from the ventral blind sac was quickly excised and gently washed using ice-cold phosphate-buffered saline (PBS; Ph 7.4). The epithelium was subsequently separated from the muscle layers and cut into 1–2 cm² pieces. One piece was placed on ice and promptly used for the extraction of microbial DNA. Five pieces were immediately fixed in 4% paraformaldehyde (PFA) (Sigma, St. Louis, MO) for histomorphometric microscopy analysis. The remaining pieces were stored at -80 °C for later extraction of epithelial RNA.

Morphological Study. Ten papillae per animal were prepared for optical microscopy analyzed according to the description of Odongo *et al.*⁵⁴. In brief, single PFA-fixed papilla were embedded in paraffin and sectioned to a thickness of 6 µm. Each section was stained with hematoxylin and eosin and then mounted on a slide for microscopic analysis. Image Pro Plus software (Media Cybernetics, Silver Spring, MD) was used to observe the integrity of the epithelium in all samples.

Microbial DNA Extraction and Sequencing. To detach the tightly attached bacteria, the ruminal epithelium was placed in a 1.5 ml tubes with 0.7 ml of PBS and several plastic beads and moderately shaken on a vortex for 30 second. The ruminal epithelium was moved to a new tube and processed with the detaching step once again. Subsequently, the metagenomic DNA was extracted from the PBS mixture by using a Bacterial DNA Kit (Omega, Shanghai, China). The DNA concentration was determined using a Nanodrop 1000 (Thermo Fisher Scientific, Wilmington, DE, USA) and stored at -20 °C until further processing. The 16S rRNA gene library preparation was performed by using polymerase chain reaction (PCR) amplification of the V3–V4 region. The universal primers 338F (5'-ACTCCTACGGGAGGCAGCAG-3') and 806R (5'-GGACTACHVGGGTWTCTAAT-3')⁵⁵, which included TruSeq adapter sequences and indices, were used in the PCRs. All libraries were sequenced using an Illumina MiSeq platform (Illumina, San Diego, California, USA) at Biomarker Technologies, Beijing, China.

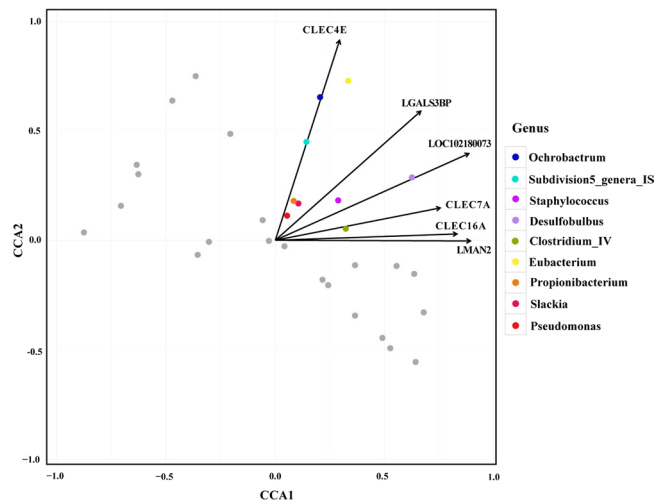


Figure 4. Constrained correspondence analysis revealing the correlations between the relative abundance of the reduced microbial clades and the expression of the functional lectin genes.

Ruminal Microbiota Analysis. Paired reads were filtered for quality (Q30) and joined by FLASH v1.2.11⁵⁶. Sequences that contained read lengths shorter than 400 bp and greater than 550 bp were removed by PRINSEQ v0.20.4⁵⁷. The remaining sequences were then classified into operational taxonomic units (OTUs) by the RDP classifier release 9.0⁵⁸ at the default setting. OTUs, whose counts were more than 3 in at least one of the samples, were hierarchically summed at all taxonomic levels, and the counts were normalized to the relative abundance for each sample. The diversity of the microbial communities was estimated using the R program phyloseq package⁵⁹. For a deeper analysis of the diversity of the major evolutionary clades in the mucosal microbiota, OTUs were filtered to a relative abundance of at least 1% in at least one sample. Then, MUSCLE v3.8.31⁶⁰ was used to align the complete 16S rRNA sequences of the corresponding species in the RDP database. RaxML v8⁶¹ and the GTR model were used to construct the maximum likelihood (ML) trees. The R program ape package⁶² was used to plot the tree.

Epithelial RNA Extraction and Sequencing. Total RNA was extracted from the ruminal epithelium using the RNeasy Mini Kit (Qiagen, Shanghai, China) according to the manufacturer's instructions. RNA was quantified using a NanoDrop 1000 spectrophotometer and its integrity was evaluated by using the RNA 6000 Assay Kit of the Agilent Bioanalyzer 2100 system (Agilent Technologies, CA, USA). High-quality RNA (RNA Integrity number >9.0) was processed using the NEBNext Ultra RNA Library Prep Kit (New England Biolabs, Beijing, China) following the manufacturer's instruction. All libraries were sequenced via paired-end chemistry (PE125) on an Illumina HiSeq2500 platform (Illumina, San Diego, California, USA) at Biomarker Technologies, Beijing, China.

Transcriptome Assembling and Differentially Expressed Gene Identification. Low-quality reads (including more than 50% low-quality bases (<Q30) and more than 10% ambiguous bases (N)) were first removed by using PRINSEQ v0.20.4. The NCBI goat genome annotation release version 101 was used to construct the reference genome by using Bowtie release 1.2.0⁶³. High-quality reads were mapped to the reference genome using TopHat v2.1.0⁶⁴ with standard parameters. Each SAM output file from the TopHat alignment was used in the Cuffdiff program of Cufflinks version 2.2.1⁶⁵ as input files to test for differential gene expression. In the Cuffdiff program, only mapped reads were used to estimate the gene expression level of each gene transcript and the gene expression values were subsequently normalized to the RPKM. In this study, only genes with more than 1 RPKM in at least one of the samples were considered expressed.

Gene Co-expression Network. The co-expressed genes were identified by computing the SCC between pairs of genes across six samples by using R program. Only expressed genes (RPKM > 1 in at least one sample) were used in the correlation analysis. A threshold for the SCC larger than 0.8 and *p*-value less than 0.05 was used to identify significantly co-expressed genes. The expression levels of the neighbors nearest to the differentially expressed lectins on the co-expression network were compared between the groups. The differentially expressed neighbors were analyzed from the pathways in the KEGG on the KABAS version 3.0 web server⁶⁶. Finally, cytoscape version 3.4.0⁶⁷ was applied to visualize the gene co-regulation network.

Dimensionality Reduction for Microbial Features and Multi Data Integration. To improve the power to associate microbial composition with the transcriptional activity of the host genes, we reduced the dimensionality of the microbial features by calculating their SCC with the differentially expressed lectins. The procedure was the same as that used in the construction of the gene co-expression network. The OTUs, whose SCC were larger than 0.6 and *p*-value was less than 0.05 with at least one type of investigated lectins, were used in the following analysis. Next, the relationships between the relative abundance of the selected OTUs and the

expressed RPKM of the differentially expressed lectins were explored using the CCA in the R program vegan package⁶⁸. The expressed RPKM of the lectins was used as the environmental factors in the CCA analysis. The R program ggplot2 package⁶⁹ was used to generate the visual interpretation of the gene-microbiota relationships. The location of each microbial genus was calculated as the centroid of the corresponding OTUs within the given genus.

Availability of data and material. The meta-genome and transcriptome data are available in the NCBI under BioProject PRJNA339481.

References

- Kabat, A. M., Srinivasan, N. & Maloy, K. J. Modulation of immune development and function by intestinal microbiota. *Trends in immunology* **35**, 507–517, doi:10.1016/j.it.2014.07.010 (2014).
- Shanahan, F. The host-microbe interface within the gut. *Best practice & research. Clinical gastroenterology* **16**, 915–931 (2002).
- Kelly, D. & Mulder, I. E. Microbiome and immunological interactions. *Nutrition reviews* **70**(Suppl 1), S18–30, doi:10.1111/j.1753-4887.2012.00498.x (2012).
- Swiatczak, B. & Cohen, I. R. Gut feelings of safety: tolerance to the microbiota mediated by innate immune receptors. *Microbiology and immunology* **59**, 573–585, doi:10.1111/1348-0421.12318 (2015).
- Sharon, N. & Lis, H. History of lectins: from hemagglutinins to biological recognition molecules. *Glycobiology* **14**, 53R–62R, doi:10.1093/glycob/cwh122 (2004).
- Vasta, G. R. *et al.* Structural and functional diversity of the lectin repertoire in teleost fish: relevance to innate and adaptive immunity. *Developmental and comparative immunology* **35**, 1388–1399, doi:10.1016/j.dci.2011.08.011 (2011).
- Endo, Y., Matsushita, M. & Fujita, T. Role of ficolin in innate immunity and its molecular basis. *Immunobiology* **212**, 371–379, doi:10.1016/j.imbio.2006.11.014 (2007).
- Farrokh, M. *et al.* Mannose-binding Lectin Mediated Complement Pathway in Autoimmune Neurological Disorders. *Iranian journal of allergy, asthma, and immunology* **15**, 251–256 (2016).
- Dang, L. & Van Damme, E. J. Genome-wide identification and domain organization of lectin domains in cucumber. *Plant physiology and biochemistry: PPB/Societe francaise de physiologie vegetale* **108**, 165–176, doi:10.1016/j.plaphy.2016.07.009 (2016).
- van Kooyk, Y. & Rabinovich, G. A. Protein-glycan interactions in the control of innate and adaptive immune responses. *Nature immunology* **9**, 593–601, doi:10.1038/ni.f.203 (2008).
- Shan, M. *et al.* Mucus enhances gut homeostasis and oral tolerance by delivering immunoregulatory signals. *Science* **342**, 447–453, doi:10.1126/science.1237910 (2013).
- Cornelissen, L. A. & Van Vliet, S. J. A Bitter Sweet Symphony: Immune Responses to Altered O-glycan Epitopes in Cancer. *Biomolecules* **6**, doi:10.3390/biom6020026 (2016).
- Coombes, J. L. *et al.* A functionally specialized population of mucosal CD103+ DCs induces Foxp3+ regulatory T cells via a TGF-beta and retinoic acid-dependent mechanism. *The Journal of experimental medicine* **204**, 1757–1764, doi:10.1084/jem.20070590 (2007).
- Argayosa, A. M. & Lee, Y. C. Identification of (L)-fucose-binding proteins from the Nile tilapia (*Oreochromis niloticus* L.) serum. *Fish & shellfish immunology* **27**, 478–485, doi:10.1016/j.fsi.2009.06.012 (2009).
- Pan, S., Tang, J. & Gu, X. Isolation and characterization of a novel fucose-binding lectin from the gill of bighead carp (*Aristichthys nobilis*). *Veterinary immunology and immunopathology* **133**, 154–164, doi:10.1016/j.vetimm.2009.07.015 (2010).
- Salerno, G. *et al.* F-type lectin from the sea bass (*Dicentrarchus labrax*): purification, cDNA cloning, tissue expression and localization, and opsonic activity. *Fish & shellfish immunology* **27**, 143–153, doi:10.1016/j.fsi.2009.01.004 (2009).
- Petri, R. M. *et al.* Changes in the Rumen Epimural Bacterial Diversity of Beef Cattle as Affected by Diet and Induced Ruminal Acidosis. *Applied and environmental microbiology* **79**, 3744–3755, doi:10.1128/AEM.03983-12 (2013).
- Shen, H., Lu, Z., Chen, Z., Wu, Y. & Shen, Z. Rapid Fermentable Substance Modulates Interactions between Ruminal Commensals and Toll-Like Receptors in Promotion of Immune Tolerance of Goat Rumen. *Frontiers in microbiology* **7**, 10.3389/fmicb.2016.01812, doi:10.3389/fmicb.2016.01812 (2016).
- Moretta, L. & Locatelli, F. Innate lymphoid cells in normal and disease: An introductory overview. *Immunology letters*. doi:10.1016/j.imlet.2016.07.008 (2016).
- Horiuchi, H. *et al.* Interleukin-19 acts as a negative autocrine regulator of activated microglia. *PLoS one* **10**, e0118640, doi:10.1371/journal.pone.0118640 (2015).
- Duester, G. Retinoic acid synthesis and signaling during early organogenesis. *Cell* **134**, 921–931, doi:10.1016/j.cell.2008.09.002 (2008).
- Iliev, I. D., Mileti, E., Matteoli, G., Chieppa, M. & Rescigno, M. Intestinal epithelial cells promote colitis-protective regulatory T-cell differentiation through dendritic cell conditioning. *Mucosal immunology* **2**, 340–350, doi:10.1038/mi.2009.13 (2009).
- Le, Y., Zhou, Y., Iribarren, P. & Wang, J. Chemokines and chemokine receptors: their manifold roles in homeostasis and disease. *Cellular & molecular immunology* **1**, 95–104 (2004).
- Shelite, T. R. *et al.* IL-33-Dependent Endothelial Activation Contributes to Apoptosis and Renal Injury in Orientia tsutsugamushi-Infected Mice. *PLoS neglected tropical diseases* **10**, e0004467, doi:10.1371/journal.pntd.0004467 (2016).
- Traks, T. *et al.* Association analysis of class II cytokine and receptor genes in vitiligo patients. *Human immunology* **77**, 375–381, doi:10.1016/j.humimm.2015.09.050 (2016).
- Martens, E. C., Chiang, H. C. & Gordon, J. I. Mucosal glycan foraging enhances fitness and transmission of a saccharolytic human gut bacterial symbiont. *Cell host & microbe* **4**, 447–457, doi:10.1016/j.chom.2008.09.007 (2008).
- Ng, K. M. *et al.* Microbiota-liberated host sugars facilitate post-antibiotic expansion of enteric pathogens. *Nature* **502**, 96–99, doi:10.1038/nature12503 (2013).
- Tong, M., Jacobs, J. P., McHardy, I. H. & Braun, J. Sampling of intestinal microbiota and targeted amplification of bacterial 16S rRNA genes for microbial ecologic analysis. *Current protocols in immunology/edited by John E. Coligan ... [et al.]* **107**(7), 41 41–11, doi:10.1002/0471142735.im0741s107 (2014).
- Mao, S., Zhang, M., Liu, J. & Zhu, W. Characterising the bacterial microbiota across the gastrointestinal tracts of dairy cattle: membership and potential function. *Scientific reports* **5**, 16116, doi:10.1038/srep16116 (2015).
- Wetzels, S. U. *et al.* Pyrosequencing reveals shifts in the bacterial epimural community relative to dietary concentrate amount in goats. *Journal of dairy science* **98**, 5572–5587, doi:10.3168/jds.2014-9166 (2015).
- Almagro-Moreno, S. & Boyd, E. F. Insights into the evolution of sialic acid catabolism among bacteria. *BMC evolutionary biology* **9**, 118, doi:10.1186/1471-2148-9-118 (2009).
- Croft, E. H. *et al.* Utilisation of mucin glycans by the human gut symbiont *Ruminococcus gnavus* is strain-dependent. *PLoS one* **8**, e76341, doi:10.1371/journal.pone.0076341 (2013).
- Rho, J. H. *et al.* A novel mechanism for desulfation of mucin: identification and cloning of a mucin-desulfating glycosidase (sulfoglycosidase) from *Prevotella* strain RS2. *Journal of bacteriology* **187**, 1543–1551, doi:10.1128/JB.187.5.1543-1551.2005 (2005).
- Hakansson, A. & Molin, G. Gut microbiota and inflammation. *Nutrients* **3**, 637–682, doi:10.3390/nu3060637 (2011).

35. Vaishnava, S. *et al.* The antibacterial lectin RegIII γ promotes the spatial segregation of microbiota and host in the intestine. *Science* **334**, 255–258, doi:10.1126/science.1209791 (2011).
36. Rabinovich, G. A., Toscano, M. A., Jackson, S. S. & Vasta, G. R. Functions of cell surface galectin-glycoprotein lattices. *Current opinion in structural biology* **17**, 513–520, doi:10.1016/j.sbi.2007.09.002 (2007).
37. Gringhuis, S. I. *et al.* C-type lectin DC-SIGN modulates Toll-like receptor signaling via Raf-1 kinase-dependent acetylation of transcription factor NF- κ B. *Immunity* **26**, 605–616, doi:10.1016/j.immuni.2007.03.012 (2007).
38. Hodges, A. *et al.* Activation of the lectin DC-SIGN induces an immature dendritic cell phenotype triggering Rho-GTPase activity required for HIV-1 replication. *Nature immunology* **8**, 569–577, doi:10.1038/ni1470 (2007).
39. Robinson, M. J., Sancho, D., Slack, E. C., LeibundGut-Landmann, S. & Reis e Sousa, C. Myeloid C-type lectins in innate immunity. *Nature immunology* **7**, 1258–1265, doi:10.1038/ni1417 (2006).
40. Rogers, N. C. *et al.* Syk-dependent cytokine induction by Dectin-1 reveals a novel pattern recognition pathway for C type lectins. *Immunity* **22**, 507–517, doi:10.1016/j.immuni.2005.03.004 (2005).
41. van Vliet, S. J., Saeland, E. & van Kooyk, Y. Sweet preferences of MGL: carbohydrate specificity and function. *Trends in immunology* **29**, 83–90, doi:10.1016/j.it.2007.10.010 (2008).
42. van Kooyk, Y. & Geijtenbeek, T. B. DC-SIGN: escape mechanism for pathogens. *Nature reviews. Immunology* **3**, 697–709, doi:10.1038/nri1182 (2003).
43. Bergman, M., Del Prete, G., van Kooyk, Y. & Appelmelk, B. Helicobacter pylori phase variation, immune modulation and gastric autoimmunity. *Nature reviews. Microbiology* **4**, 151–159, doi:10.1038/nrmicro1344 (2006).
44. Nava, G. M. *et al.* Hydrogenotrophic microbiota distinguish native Africans from African and European Americans. *Environmental microbiology reports* **4**, 307–315, doi:10.1111/j.1758-2229.2012.00334.x (2012).
45. Derrien, M. *et al.* Modulation of Mucosal Immune Response, Tolerance, and Proliferation in Mice Colonized by the Mucin-Degrader Akkermansia muciniphila. *Frontiers in microbiology* **2**, 166, doi:10.3389/fmicb.2011.00166 (2011).
46. Png, C. W. *et al.* Mucolytic bacteria with increased prevalence in IBD mucosa augment in vitro utilization of mucin by other bacteria. *The American journal of gastroenterology* **105**, 2420–2428, doi:10.1038/ajg.2010.281 (2010).
47. Swidsinski, A. *et al.* Acute appendicitis is characterised by local invasion with Fusobacterium nucleatum/necrophorum. *Gut* **60**, 34–40, doi:10.1136/gut.2009.191320 (2011).
48. Haviari, S. *et al.* Outbreak of Achromobacter xylosoxidans and Ochrobactrum anthropi Infections after Prostate Biopsies, France, 2014. *Emerging infectious diseases* **22**, 1412–1419, doi:10.3201/eid2208.151423 (2016).
49. Kwon, H. H. & Suh, D. H. Recent progress in the research about Propionibacterium acnes strain diversity and acne: pathogen or bystander? *International journal of dermatology*. doi:10.1111/ijd.13282 (2016).
50. Pai, M. P. Anti-Infective Dosing for Obese Adult Patients: A Focus on Newer Drugs to Treat Methicillin-Resistant Staphylococcus aureus Acute Bacterial Skin and Skin Structure Infections. *Clinical therapeutics*. doi:10.1016/j.clinthera.2016.07.094 (2016).
51. Riviere, A., Selak, M., Lantin, D., Leroy, F. & De Vuyst, L. Bifidobacteria and Butyrate-Producing Colon Bacteria: Importance and Strategies for Their Stimulation in the Human Gut. *Frontiers in microbiology* **7**, 979, doi:10.3389/fmicb.2016.00979 (2016).
52. Cho, G. S. *et al.* Quantification of Slackia and Eggerthella spp. in Human Feces and Adhesion of Representatives Strains to Caco-2 Cells. *Frontiers in microbiology* **7**, 658, doi:10.3389/fmicb.2016.00658 (2016).
53. Nagata, Y. *et al.* Relationship of serum levels and dietary intake of isoflavone, and the novel bacterium Slackia sp. strain NATTS with the risk of prostate cancer: a case-control study among Japanese men. *International urology and nephrology* **48**, 1453–1460, doi:10.1007/s11255-016-1335-7 (2016).
54. Odongo, N. E. *et al.* Effects of mild heat stress and grain challenge on acid-base balance and rumen tissue histology in lambs. *Journal of animal science* **84**, 447–455 (2006).
55. Mori, H. *et al.* Design and experimental application of a novel non-degenerate universal primer set that amplifies prokaryotic 16S rRNA genes with a low possibility to amplify eukaryotic rRNA genes. *DNA research: an international journal for rapid publication of reports on genes and genomes* **21**, 217–227, doi:10.1093/dnares/dst052 (2014).
56. Magoc, T. & Salzberg, S. L. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **27**, 2957–2963, doi:10.1093/bioinformatics/btr507 (2011).
57. Schmieder, R. & Edwards, R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* **27**, 863–864, doi:10.1093/bioinformatics/btr026 (2011).
58. Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and environmental microbiology* **73**, 5261–5267, doi:10.1128/AEM.00062-07 (2007).
59. McMurdie, P. J. & Holmes, S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS one* **8**, e61217, doi:10.1371/journal.pone.0061217 (2013).
60. Edgar, R. C. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC bioinformatics* **5**, 113, doi:10.1186/1471-2105-5-113 (2004).
61. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313, doi:10.1093/bioinformatics/btu033 (2014).
62. Paradis, E., Claude, J. & Strimmer, K. APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics* **20**, 289–290 (2004).
63. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome biology* **10**, R25, doi:10.1186/gb-2009-10-3-r25 (2009).
64. Kim, D. *et al.* TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome biology* **14**, R36, doi:10.1186/gb-2013-14-4-r36 (2013).
65. Trapnell, C. *et al.* Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature biotechnology* **28**, 511–515, doi:10.1038/nbt.1621 (2010).
66. Xie, C. *et al.* KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. *Nucleic acids research* **39**, W316–322, doi:10.1093/nar/gkr483 (2011).
67. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research* **13**, 2498–2504, doi:10.1101/gr.1239303 (2003).
68. vegan: Community Ecology Package v. 2.4-0 (2016).
69. Wickham, H. ggplot2: *Elegant Graphics for Data Analysis* (Springer-Verlag New York, 2009).

Acknowledgements

This work was supported by the Project Grant Natural Science Foundation of Jiangsu Province (BK20150654), the Independent Innovation Project of Nanjing Agriculture University (KYZZ01628), and the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD).

Author Contributions

Z.Y.L. designed the research. H.S. wrote the paper. H.S. and Z.H.X. analyzed data. Z.Y.L. and Z.H.X. performed the experiments. Z.M.S. edited the version of the paper to be published.

Additional Information

Supplementary information accompanies this paper at doi:[10.1038/s41598-017-03478-2](https://doi.org/10.1038/s41598-017-03478-2)

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017