Enteric phageome alternations in Type 2 diabetes disease

Qian Chen
Shanghai Jiao Tong University School of Medicine  https://orcid.org/0000-0002-2471-8856

Xiaojing Ma
Shanghai Sixth Peoples Hospital

Chong Li
Fudan University Shanghai Cancer Center

Yun Shen
Shanghai Sixth Peoples Hospital

Wei Zhu
Shanghai Sixth Peoples Hospital

Yan Zhang
Shanghai Jiao Tong University School of Medicine

Xiaokui Guo
Shanghai Jiao Tong University School of Medicine

Jian Zhou
Shanghai Sixth Peoples Hospital

Chang Liu (✉ tiantianlc@sjtu.edu.cn)
Shanghai Jiao Tong University School of Medicine  https://orcid.org/0000-0002-9996-9756

Research

Keywords: Phageome, type 2 diabetes, bacteriome, bacteriophage

Posted Date: November 18th, 2019

DOI: https://doi.org/10.21203/rs.2.17361/v1

License: ☺ ☐ This work is licensed under a Creative Commons Attribution 4.0 International License.  Read Full License
Abstract

Background Type 2 diabetes (T2D) is a complex metabolic disease and has been proved to involve in the alternation of the gut microbiota. The previous studies primarily focused on the changes in bacteriome while ignoring the phage community composition. The extracellular phages could lyse the host bacteria, and thus influence the microbiota through the positive or negative interactions with bacteria. We investigated the change of extracellular phageome and explored its role in T2D pathogenesis.

Results We used a sequencing-based approach to identify the bacteriophage after isolation of VLPs from the fecal samples. We identified 330 phages according to the predicted host bacteria. The phageome characteristics were highly diverse among individuals. In the T2D group, the intestinal phage population is altered and the abundance of 7 identified phages specific to Enterobacteriaceae hosts were found increased markedly. Additionally, the abundance of Enterobacteriaceae bacteria in gut was significantly increased and the systemic LPS elevation was observed in T2D group. Several phage consortia were found to have significant correlations with T2D disease indicators.

Conclusions The alteration of bacteriophages predicted to infect Enterobacteriaceae in the gut was observed in this study, which was expected to be a new source of systemic LPS in T2D patients, and may contribute to the pathogenesis of the disease. The data present in this study revealed the similar variation trend in enteric bacteriome and the correlated bacteriophages, which is likely to shed considerable light on overall understanding the interactions between microbiome and metabolic diseases.

Backgrounds

Type 2 diabetes (T2D) has been one of the leading health issues globally for recent decades [1]. It is characterized by hyperglycemia in the context of insulin resistance and impaired insulin secretion, which is resulting from a complex inheritance-environment interaction along with other risk factors such as sedentary behavior, obesity, and unhealthy dietary habits. T2D and its complications are affecting almost all populations in both developed and developing countries with high rates of diabetes-related morbidity and mortality [2]. In a statistic data in 2015, the number of adult (20~79 years) patients with diabetes mellitus in China came to the highest with 109.6 million, which represent 10.6% of the domestic population, followed by India and USA [3, 4].

As highlighted by plenty of previous studies, gut microbiome was proved to plays a fundamental role in metabolic disorders including T2D [5–7], obesity [8–12], and even cardiovascular diseases [13–15]. Shifts in gut microbiome can result in increased gut permeability, altered metabolism of short-chain fatty acids (SCFA) and vitamins, along with the dysfunction of modulating glucose and lipid homeostasis, regulating satiety, and producing energy [6, 16]. For example, low vitamin D production in gut has been associated with an increased risk of T2D [17–19] and butyrate produced by gut microbiota inversely associated with the degree of insulin resistance [20, 21]. Moreover, acetate has been proved to regulate the functions of islet β-cells via mediating a feed-back loop related to metabolic syndrome in microbiota-dependent way [22]. Evidences from both human studies and animal trials have illuminated that T2D favors lipopolysaccharide (LPS) released from gram-negative bacteria translocating across the attenuated gut barrier, and results in a moderate increase in serum [23–25]. LPS induced chronic low-grade inflammation and is associated with leptin and insulin resistance, which contributes to the establishment of T2D [26].

It is well known that bacteriophages, the dominant constituent of the virome, account for a high abundance in intestinal microbiota with phage-bacterial ratios of ~1:1 [27]. Since phages are bacterial viruses which can lyse the specific bacterial hosts after finishing their progeny replications, correlation between temporal population dynamics of phages and their hosts can be expected. The biological characteristics of phages endow their capability of regulating the abundance of their hosts, thereby affecting the structure of microbiota through the cascade reactions of both positive and negative interactions among the bacteriota. Afterwards, phages could influence intestinal metabolome features through the changed bacteriota [28]. On the other hand, gram-negative phage hosts could result in an increase in LPS. From the above, phage community deserves more attention in research of T2D, both for its contribution in regulating microbiota and in circulating LPS.

Since extracellular phages can infect and lyse bacterial hosts and directly influence the composition of microbiota, we adopt the VLP isolation-dependent method and attempt to explore the changes in the gut extracellular phage community in a cohort with T2D in Shanghai, China. Besides that, we would like to explore whether these changes in the phageome have relationship with the changes in the bacteriome, then discuss the potential relationship between the alteration of phageome and the increased serum LPS. The design diagram is shown in Fig. 1.
Results

1. Fecal phageome composition based on metagenome sequencing

A total of 46 sample including 17 T2D and 29 control were collected. On average, 11.7G (±2.14, SD) bases were sequenced for each sample. From those sequences, we obtained 3,136 phage scaffolds, with an average length of 7,257bp, resulting in a total of 330 species of phage were defined (Additional file 1). 89.5% of the detected genomes of the phages were in form of double-stranded DNA (dsDNA), and less than 0.3% exist as single-stranded DNA (ssDNA). The fecal phage composition is highly variable between individuals both in the T2D group and Ctrl group. There is a higher percentage of ssDNA in D17 (24.3%) than that in remaining samples (Fig. 2a). *Caudovirales* accounts for the most abundant phage community at the order level (89.8% on average). Especially in D03 sample, almost all of the phages detected were defined as *Caudovirales* (>99.9%). While in H18 sample, *Caudovirales* only comprises 64.1% of the phage population (Fig. 2b). Detailed information was shown in Additional file 2.

At the family level, 5 phage families including *Myoviridae*, *Podoviridae*, *Microviridae*, *Siphoviridae*, and *Tectiviridae* were detected. In most samples, the percentage of *Siphoviridae* was the highest with an average of 58.7%, followed by *Myoviridae* with 25.7%. Sample D03 and D15 contain almost only the family of *Podoviridae* and *Siphoviridae*, respectively. Moreover, the abundance of *Microviridae* detected in sample D17 and D08 was higher than that in other samples (Fig. 2c).

Furthermore, these detected phageome were aligned to 27 viral genera. Similar to the above, in the genus level, compositions of the virome in samples were highly diverse. Certain viral genus is dominant in several samples, which result in low diverse indexes. Such as the relative abundance of T1-like virus in D02, T4-like virus in D17, H18, and H24 samples, phiKMV-like in H12, and Lambda-like viruse in D07 and D09 are particularly high (Fig. 2d).

2. VLP identification revealed alterations of phage communities in type 2 diabetes patients

α-diversity was assessed with 4 different diversity analysis methods. Both Ace and Chao index were significantly reduced in the T2D group, while Shannon and Simpson index showed no obvious changes between the two groups (Fig. 3a). For the results of β-diversity analysis, significant different phageome characteristics were observed between T2D patients and nondiabetic population based on principal components analysis (Fig. 3b).

The phages detected were classified according to their putative bacterial hosts. We defined a total of 330 species of phage with 51 putative bacterial hosts, among which 313 in T2D group and 324 in Ctrl group, 307 species were found in both groups (Fig. 3c). The abundances of these phages in the T2D and Ctrl groups were clustered to obtain a heatmap, as shown in Fig. 3d.

3. Phageome characteristics in T2D patients

Of the 330 species of bacteriophage identified sorting by their bacterial hosts, the abundance of 58 species of phage were significantly different between T2D patients and nondiabetic controls. There are 52 species with a cutoff of FDR<0.25, and four bacteriophages involving *Brochothrix* phage_NF5, *Enterococcus* phage_phiFL2A, *Streptococcus* phage_PH10, and *Streptococcus* phage_7201, with FDR<0.05 between the two groups (Additional file 1). The 18 phages with the most significant differences were showed in Fig. 4a.
After being clustered according to the level of the family, shifts of the abundance of *Myoviridae, Podoviridae, Microviridae, Siphoviridae*, and *Tectiviridae* in the T2D group versus Ctrl group were compared, no statistically significance was observed (Fig. 4b).

Host bacterial assignments for curated phage contigs were compared at genus level between the two groups, 7 host bacterial genera were remarkable changed (p ≤ 0.05), including *Brochothrix, Klebsiella, Enterococcus, Bordetella, Shigella, Clostridium*, and *Tetrasphaer* (Fig. 4c). The phages infecting *Brochothrix* and *Klebsiella* had the FDR <0.05 between the T2D group and the Ctrl group.

### 4. Fecal bacteriome features and the alteration in T2D patients

The rarefaction plots reached plateau for bacterial species, indicating the sample size is sufficient to reveal the bacterial population (Fig. 5a). The alpha-diversity presented by Chao1 index and PD whole tree index of the T2D group showed no significant changes (Fig. 5b). Unweighted UniFrac analysis of similarities (ANOSIM) displayed a significant difference in β-diversity between the two groups. Unweighted UniFrac principal coordinate analysis (PCoA) (left) and NMDS (middle) illustrated that the cluster of fecal bacteriome in T2D group is clearly distinct from that of the nondiabetic control. (Fig. 5c)

The bacteria identified in this study was involved in 11 phyla. The detail information was shown in Additional file 3.

LEfSe is used for identifying the most varied abundant bacterial taxa at different levels, and for data analyzing and visualizing key species which are identified as differentiating factors between diabetes patients and health controls (Fig. 5d). With a log LDA score above 2.0, we found an enriched abundance of OTUs contributed by *Phascolarctobacterium, Paraprevotella, Odoribacter, Clostridium XIVb, Butyricimonas, Shinella, Anaerofilum, Methanobrevibacter*, and *Streptophyta* among health controls, while the T2D patients had increased abundance of *Erysipelotrichaceae_incetae_sedis, Allisnonella, Lactobacillus, Dialister*, and *Megasphaera*. Moreover, *Enterobacteriaceae* bacteria is enriched in the T2D group (Fig. 5e).

129 bacterial genera were identified by 16S rDNA sequencing. Due to the technological limitations of 16S rDNA sequencing, among the 129 bacterial genera, 21 phylogenetic types bacteria failed to be identified to the exact genera. Among the 105 genera, the abundance of 14 genera was significantly different between the two groups. Excluding *Methanobrevibacter, Anaerofilum*, and *Shinella*, of which the abundance was too low to be visualized in figures, we calculated the differences in abundance of the other 11 bacterial genera between the T2D group and Ctrl group. And the results were presented in the boxplot in Fig. 5f.

### 5. Correlations between Type 2 Diabetes-Associated Changes in the Phageome and Bacterial Microbiome

In this study, the relationship between the phageome and the bacteriome was also investigated. Using 16S rDNA sequencing, 105 bacterial genera were identified. At the same time, the phageome we identified corresponded to 51 bacterial genera hosts.

Spearman correlation analysis was used to assess the correlations between the most abundant bacterial genera and bacteriophages in both T2D and control group. Moderately increased correlations of bacteria and phages were found in T2D group. (Fig.6a)
The bacteria detected and the hosts of bacteriophages could not completely overlap, only 16 putative host genera were identified in the detected bacterial group. These bacteria genera contained *Klebsiella, Enterococcus, Lactococcus, Aggregatibacter, Pseudomonas, Pseudoalteromonas, Rhodococcus, Bacteroides, Bacillus, Corynebacterium, Staphylococcus, Streptococcus, Lactobacillus, Acinetobacter, Actinomyces*, and *Haemophilus*.

Pearson correlation analysis was used to evaluate possible linear relation between bacteriophages and their host bacteria. The results indicated neither significant positive nor negative linear correlation between them (Fig. 6b).

6. **Circulating LPS elevated in T2D patients**

The LPS concentration in serum samples from T2D patients (n=13) and nondiabetic subjects (n=13) were assessed using tachypleus amebocyte lysate (TAL) based method. The results showed a significant increase in the LPS concentration of the T2D patients (Fig. 7c), as demonstrated in many studies [30, 31]. We hypothesized that the elevated level of serum LPS in T2D patients may result from the intensified lysis of the intestinal gram-negative bacteria under the action of phages. Based on this hypothesis, we analyzed the difference in the abundance of phageome mapped to gram-negative bacterial phages and gram-positive bacterial phages (Additional file 1). The results indicated that the relative abundance of gram-negative bacterial phages in the T2D group slightly higher than that in Ctrl group (Fig. 7a). While the gram-positive bacteriophages changed in the contrary direction. In addition, the most abundant 6 bacterial phyla *Actinobacteria, Bacteroidetes, Firmicutes, Fusobacteria, Proteobacteria*, and *Verrucomicrobia* were analyzed to explore the changes of gram-positive bacteria and gram-negative bacteria abundances in the T2D group and the Ctrl group, respectively. However, no significant differences of the mentioned bacteria were observed in T2D group or Ctrl group (Fig. 7b). *Enterobacteriaceae* bacteria are typical gram-negative bacteria, and the abundance of them increased in the T2D group. In view of that, the alteration of the relative abundance of *Enterobacteriaceae* specific phages were further evaluated and showed an increase in the T2D group with statistically significant (Fig. 7d).

7. **Increasing correlations between fecal phage community and T2D clinical indexes.**

Correlations between the phageome and T2D disease indicators were evaluated. The results that comprehensively presented in table 1 indicated that there were some significant correlations between several specific phage consortia and 6 T2D indexes, which referred to fasting blood glucose (blood glucose0), fasting insulin (insulin0), 0.5hrs insulin after meal (insulin30), 2hrs insulin after meal (insulin120), high sensitive-C reactive protein, and free thyroxine.
Table 1 Phage communities with significant correlation with T2D disease indicators

<table>
<thead>
<tr>
<th>Bacteriophages</th>
<th>Bacteriophages</th>
<th>Bacteriophages</th>
</tr>
</thead>
<tbody>
<tr>
<td>InS0 Bacillus_phage_phi29</td>
<td>InS Bacillus_phage_Fah</td>
<td>InS Bacillus_phage_Fah</td>
</tr>
<tr>
<td>Bacillus_phage_Fah</td>
<td>Bacillus_phage_phi29</td>
<td>Bacillus_phage_B103</td>
</tr>
<tr>
<td>Salmonella_phage_ST64T</td>
<td>Salmonella_phage_ST64T</td>
<td>Salmonella_phage_ST64T</td>
</tr>
<tr>
<td>Bacillus_phage_B103</td>
<td>Lactococcus_phage_KSY1#</td>
<td>Cyanophages_9515-10a</td>
</tr>
<tr>
<td>Enterobacteria_phage_933W_sensu_lato</td>
<td>Enterobacteria_phage_CC31</td>
<td>Enterobacteria_phage_EF62phi</td>
</tr>
<tr>
<td>Lactococcus_phage_936_sensu_lato</td>
<td>Enterococcus_phage_EF62phi</td>
<td>Klebsiella_phage_K11#</td>
</tr>
<tr>
<td>Bordetella_phage_BPP-1</td>
<td>Staphylococcus_phage_K11#</td>
<td>Rhodococcus_phage_ReqiPepy6</td>
</tr>
<tr>
<td>Burkholderia_phage_phi1026b</td>
<td>Enterobacteria_phage_phiEcoM-GJ1#</td>
<td>Yersinia_phage_phiR1-37#</td>
</tr>
<tr>
<td>Clostridium_phage_phiCP13O</td>
<td>Staphylococcus_phage_187#</td>
<td>Staphylococcus_phage_phiPVL-CN125</td>
</tr>
<tr>
<td>Enterobacteria_phage_phiEcoM-GJ1#</td>
<td>Staphylococcus_phage_phiHN3#</td>
<td>Listeria_phage_B054</td>
</tr>
<tr>
<td>Escherichia_phage_vB_EcoP_G7C</td>
<td>Myxococcus_phage_Mx8#</td>
<td>Staphylococcus_phage_phiPVL-CN125</td>
</tr>
<tr>
<td>Silicibacter_phage_DSS3phi2</td>
<td>Staphylococcus_phage_phiPVL-CN125</td>
<td>Enterococcus_phage_EFAP-1</td>
</tr>
<tr>
<td>Staphylococcus_phage_phieTA3#</td>
<td>Enterococcus_phage_EFAP-1</td>
<td>Streptococcus_pyogenes_phage_315.1#</td>
</tr>
<tr>
<td>Enterobacteria_phage_SV</td>
<td>Staphylococcus_pyogenes_phage_315.1#</td>
<td>Klebsiella_phage_KP15</td>
</tr>
<tr>
<td>Mycobacterium_phage_Trixx#</td>
<td>Enterobacteria_phage_RTP</td>
<td>Thermus_phage_PhiYS40</td>
</tr>
<tr>
<td>Mycobacterium_phage_244#</td>
<td>Shigella_phage_Shf1</td>
<td></td>
</tr>
<tr>
<td>Lactobacillus_prophage_lj928</td>
<td>Mycobacterium_phage_Yoshi</td>
<td></td>
</tr>
<tr>
<td>Streptococcus_phage_SMP</td>
<td>Enterobacteria_phage_JK06</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Klebsiella_phage_phiK02</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Enterobacteria_phage_TLS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Enterobacteria_phage_T1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Microbacterium_phage_Min1</td>
<td></td>
</tr>
</tbody>
</table>

InS0, fasting insulin

InS30, 0.5hrs insulin after meal

InS120, 2hrs insulin after meal

<table>
<thead>
<tr>
<th>InS CRP</th>
<th>Bacteriophages</th>
<th>Free thyroid</th>
<th>Bacteriophages</th>
<th>InS CRP</th>
<th>Bacteriophages</th>
<th>Free thyroid</th>
<th>Bacteriophages</th>
</tr>
</thead>
<tbody>
<tr>
<td>hs</td>
<td>Salmonella_phage_PhiSH19#</td>
<td>Enterobacteria_phage_RB49</td>
<td>GLU0 Lactococcus_phage_KSY1#</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Listeria_phage_A500#</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Prochlorococcus_phage_P-SSM4#</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pseudomonas_phage_M6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pseudomonas_phage_PAju2#</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Salmonella_phage_ST160#</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Staphylococcus_phage_TEM126</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Synechococcus_phage_S-CBS2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Salmonella_phage_g341c</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Thermus_phage_PhiYS40#</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Enterobacteria_phage_phiP27#</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pantoea_phage_LIMElight#</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Burkholderia_phage_Bcep1</td>
<td>Enterobacteria_phage_Ef24C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Erwinia_phage_phiEt88#</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Salmonella_phage_ST64B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Streptococcus_phage_SMP#</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bacillus_phage_phi29</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Enterococcus_phage_phiEF24C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Staphylococcus_phage_53_sensu_lato#</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Enterobacteria_phage_P22#</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bacillus_phage_B103</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lactobacillus_prophage_lj965</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

GLU0, fasting blood glucose
Among the disease indicators in Additional file 4, only those with statistically significant correlations (FDR<0.05) were presented in this figure.

**Discussion**

As a typical disease of metabolic syndrome, T2D is known to involve complex cellular and molecular mechanisms, leading to dysregulated glucose homeostasis in the body, while microbiota has been proved to be one of the fundamental factors influencing T2D development. Phage community are an important component of microbiota, containing both bacteriophages and the genetic elements of phages integrated in bacterial genome, representing the phage particles and prophages respectively, with the ability to modify the microbiota or participate in the horizontal transfer of host genes. When it comes to the technical routes for investigating phageome, a metagenome-dependent method and a VLP isolation-dependent method are in use [32–35]. A previous study using the metagenomic data analysis has investigated the comprehensive DNA phageome containing both the genomes of phage particles and the genetic elements of prophages in patients with T2D for the first time [36]. In present study, we used a method based on virus-like particles isolation and focused on the extracellular phages, which could directly regulate the microbiota, thus to affect the physiology of the human host. Clinical studies have shown that phage filtrate can achieve good clinical results in patients with severe diarrhea caused by *Clostridoides difficile* [37]. It also confirmed the regulation of extracellular phage on micro-ecological systems.

Viruses are divided into 152 families based on the latest (2019) report by the International Committee for the Taxonomy of Viruses (ICTV, Virus Metadata Repository: version September 9, 2019; MSL34). The ICTV report divided the RNA phages into only two families; Cystoviridae with genus Cystovirus, and Leviviridae, with 2 genera Levivirus and Allolevivirus [38]. Despite the genome of phages could be DNA or RNA, we focused on the abundant DNA phageome in this study. The study above based on metagenomic analysis revealed an increase in the number of gut phages and an elevation of the relative pOTU number in the T2D group [36]. Using the VLP isolation-based method in our study, no changes in abundance of the 5 phage families were found between the T2D group and Ctrl group (Fig. 4b). Regardless of the differences in the structures of phageome between the two study cohorts, such divergence illustrated that the study results depend on the chosen research methods. Nevertheless, alternations of phage community were discovered when the phages were clustered according to the bacterial hosts (Fig. 4a, 4c).

Remarkably, phages host in *Klebsiella* bacteria and *Shigella* bacteria that are the most concerned *Enterobacteriaceae* bacteria, increased in the T2D group. Then the changes in the relative abundance of all the detected *Enterobacteriaceae* specific phages between the two groups were assessed, and showed a statistically significant increase in the T2D group (Fig. 7d). Meanwhile, the bacteria belonging to family *Enterobacteriaceae* significantly increased in T2D patients compared with nondiabetic controls (Fig. 5e). Family *Enterobacteriaceae* is a typical cluster of gram-negative bacteria and is known as a bacterial family enriched in opportunistic pathogens including *Escherichia*, *Shigella*, *Klebsiella* and *Salmonella*. It’s well known that the LPS could be released from gram-negative bacteria, leading to the systematic subclinical inflammatory and, affecting the insulin sensitivity [39]. Given a fact of that, we speculated that the alteration of gram-negative bacteria and their corresponding phages might lead to a circulating LPS elevation.

To confirm our hypothesis above, the relative abundance of phages whose putative host are gram-negative bacteria in the T2D group and Ctrl group were compared, and an increasing trend were found in the T2D group (Fig. 7a). At the same time, the relative abundance of gram-negative bacteria in the T2D group were also evaluated, and showed no obvious changes (Fig. 7b). Based on the experimental evidence above, we postulated a novel pathway via which the phage component in microbiota influence the pathogenesis of T2D (Fig. 8). That is, an enhanced level of *Enterobacteriaceae* bacteria, and their specific phages provided basis for the further intensified lysis of *Enterobacteriaceae* bacteria under pathological status of T2D, leading to an increase in serum LPS and then the development or aggravation of T2D. The structures of LPS in different bacteria differ a lot [40], so their pathological consequences also be various. Since *Enterobacteriaceae* received more attention in the traditional infection disease, the possible pathogenic mechanisms linking *Enterobacteriaceae* bacteria-phages-LPS in T2D are supposed to be deeply explored.

Studies have revealed a depletion in universal butyrate-producing bacteria as well as an increase in some opportunistic pathogens in T2D patients [41–44]. In a metagenomic study carried by a Chinese group, the depleted butyrate-producing bacteria involve *Clostridiales* sp. SS3/4, *Eubacterium rectale*, *Faecalibacterium prausnitzii*, *Roseburia intestinalis* and *Roseburia inulinivorans*. While the opportunistic pathogens include *Bacteroides caccae*, *Clostridium hathewayi*, *Eggerthella lenta* and *Escherichia coli* [41]. From the data of 16S rDNA sequencing processed in this study, similarly, butyrate-producing bacteria *Phascolarctobacterium* and *Megasphaera* decreased in T2D
patients. And *Clostridium* XIVb, which could contribute in the maintenance of gut homeostasis [45] was also depleted in the T2D group (Fig. 5f).

The correlation analysis of the most abundant phages and bacterial genera showed a slightly strengthened correlation in disease status, and most of the correlations were positive correlations (Fig. 6a), which is inverse to a previous study about the phages in gut mucosal of ulcerative colitis [29]. But the mechanism is not clear. The relative abundance of phages and their putative host bacteria in the fecal samples of neither T2D patients nor nondiabetic individuals showed no linear correlation (Fig. 6b). That could be explained from the perspective of the dynamic relationship of bacteria-phage. Because phages not only live in bacteria, but also lyse their host bacteria, which have been described as ‘predator-prey’ dynamic model. Additionally, there are also ‘kill-the-winner’ and ‘arms-race’ dynamics at the same time [46].

In the last part of our study, we presented plentiful phages correlated with T2D disease indicators including fasting blood glucose, fasting insulin, 0.5hrs insulin after meal, 2hrs insulin after meal, hs-CRP, and free thyroxine (Table 1). Meanwhile, only a few bacterial genera have significant correlation with each T2D indicator (Additional file 5). Changes in disease-associated phageome might be more sensitive than those in the bacteriome.

It is urgent requirement that microbiome related researches should be closely integrated with clinical study, the indicator of bacteriome changes in disease status have been confirmed to be valuable auxiliary measures in clinical diagnosis [47], and expected to serve the precision medicine. It is undeniable that with the continuous improvement of the phageome researches and someday the completing of the standard operating procedure, phagesome variation may occupy a place in clinical diagnosis.

**Conclusions**

In this study, we investigated the alternations of intestinal extracellular phaegeome in patients with T2D. The data present in this study revealed the similar variation trend in enteric bacteriome and the correlated bacteriophages, the increase of bacteriophages predicted to infect *Enterobacteriaceae* in the gut was speculated as a novel origin of systemic LPS, which may contribute to the pathogenesis of the disease. Furthermore, we found some phage consortium related to disease indicators of T2D, which is likely to shed considerable light on overall understanding the interactions between microbiome and metabolic diseases.

**Methods**

1. **Human subjects**

   Patients with type 2 diabetes involved in this study were recruited from Shanghai Jiao Tong University Affiliated Sixth People's Hospital. Fecal samples and serum samples were collected, divided into two aliquots respectively, and stored in -80 °C refrigerator. Healthy volunteers aged from 20 to 50 years old in control group were recruited from Shanghai Jiao Tong University Affiliated Sixth People's Hospital. All subjects in the nondiabetic control group were confirmed without diabetes by oral glucose tolerance test (OGTT), and excluded other metabolic diseases. Clinical characteristics of the T2D patients were presented in table 2. The detailed test results of nondiabetic controls are not shown in this study, while the relevant clinical features and demographic information recorded for each patient are listed in Additional file 4.
Table 2 Clinical characteristics of the T2D patients involved in this study

<table>
<thead>
<tr>
<th>Variables</th>
<th>Values in D (mean ±SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blood pressure (mmHg)</td>
<td></td>
</tr>
<tr>
<td>Systolic</td>
<td>132±21</td>
</tr>
<tr>
<td>Diastolic</td>
<td>77±6</td>
</tr>
<tr>
<td>Plasma glucose (mmol/L)</td>
<td></td>
</tr>
<tr>
<td>Fasting</td>
<td>7.24±2.34</td>
</tr>
<tr>
<td>0.5hrs post meal</td>
<td>10.23±1.90</td>
</tr>
<tr>
<td>2hrs post meal</td>
<td>12.41±4.22</td>
</tr>
<tr>
<td>Insulin (μU/mL)</td>
<td></td>
</tr>
<tr>
<td>Fasting</td>
<td>28.98±45.03</td>
</tr>
<tr>
<td>0.5hrs post meal</td>
<td>78.01±80.18</td>
</tr>
<tr>
<td>2hrs Post meal</td>
<td>98.02±88.61</td>
</tr>
<tr>
<td>Glycated hemoglobin (%)</td>
<td>8±2</td>
</tr>
<tr>
<td>Serum triglycerides (mmol/L)</td>
<td>1.45±0.75</td>
</tr>
<tr>
<td>Serum cholesterol (mmol/L)</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>4.32±1.09</td>
</tr>
<tr>
<td>LDL-c</td>
<td>2.65±0.97</td>
</tr>
<tr>
<td>HDL -c</td>
<td>1.07±0.22</td>
</tr>
<tr>
<td>Free thyroxine (pmol/L)</td>
<td>16.87±2.61</td>
</tr>
<tr>
<td>Hypersensitive C-reaction protein (mg/L)</td>
<td>0.90±1.20</td>
</tr>
</tbody>
</table>

2. Diagnosis of diabetes

We used 2010 American Diabetes Association (ADA)’s criteria [48] for the diagnosis of diabetes. Diabetes was defined as fasting glucose ≥ 7.0 mmol/L, and/or 2-h glucose ≥ 11.1 mmol/L and/or HbA1c ≥ 6.5%. Those using antidiabetic drugs in the examination were also included into the type 2 diabetes cases.

3. Samples treatment for viral-like particles (VLP) isolation, and the DNA extraction

One aliquot of the stool samples was weighed to about 1.0 g and placed in sterile SM buffer (100 mM NaCl, 8 mM MgSO$_4$·7H$_2$O, 50 mM Tris-Cl (pH7.5), 0.01% gelatin (w/v)). Then the fecal samples were vortexed until they were thoroughly mixed with SM buffer. The sample treatment process is based on previous studies [32, 49, 50]. In brief, centrifuged the sample homogenates (3,500×g for 30min at 4 ºC). Followed by another centrifugation (10,000×g for 20min at 4 ºC) of the supernatants, the resulting supernatants was collected and filtered through a 0.45 μm and 0.22 μm Millipore filter sequentially. After collecting the filtrate, the samples were centrifuged (33,000rpm for 4h at 4ºC) in a SW41Ti swinging bucket rotor (Beckman). The 1.5 g ml$^{-1}$ layer was recovered for the enriched bacteriophages [49]. This procedure was illustrated in Fig 1.
The 1.5 g ml\(^{-1}\) layer was collected from the step gradient and used for the source of the phageome. With the method described previously [32], the DNA was extracted from each sample, followed by the amplifying of bacterial 16S rRNA and eukaryotic 18S rRNA in one aliquot of the DNA to assess the absence of detectable non-viral DNA.

Before the DNA sequencing, the total DNA was amplified using GenomiPhi V2 kit (GE Healthcare) to generate enough genes for library construction.

4. 16S rDNA gene sequencing

Total DNA of one aliquot of the fecal samples were extracted using QIAamp Fast DNA Stool Mini Kit. The 16S rDNA high-throughput sequencing was performed using the Illumina HiSeq PE250. Variable regions V3–V4 on 16S rDNA genes of bacterial genome were amplified with forward primer F341 5’-ACTCCTACGGGRSGCAGCAG-3’ and reverse primer R806 5’-GGACTACVGGGTATCTAATC-3’. The raw paired end reads were assembled by pandaseq with overlap nucleotides. Then, the reads were quality-filtered. The raw data were then subjected to a quality control procedure using UPARSE. The qualified reads were clustered to generate operational taxonomic units (OTUs) at the 97% similarity level using USEARCH. Principal components analysis (PCA), heatmap analysis, Bray-Curtis similarity cluster, and species abundance analysis were performed using R program.

5. Metagenomic sequencing of DNA from Phages derived from VLP particles

Subject phage DNA was first sheared into ~400bp-long fragments on Covaris S2 (Covaris, US). The resulting DNA fragments were used to construct sequencing library according to manufacturer’s instruction (NEXTflex\(^\text{TM}\) DNA Sequencing Kit compatible with the Biomek\(^\text{®}\) FXp (Bio Scientific, US). DNA library were sequenced on Illumina\(^\text{®}\) X-ten platform with a read length of 150 bp.

Raw reads for each sample were preprocessed using the Trimmomatic[51]. This included trimming adapter and removing reads with low-quality and insufficient length via the following parameters: SLIDINGWINDOW:4:5, MINLEN:50. The obtained clean reads were first mapped to human (hg19) reference genomes by bowtie2 [52]. The resulting unmapped reads were further assembled into scaffolds by velvet software (version, 1.2.10, https://www.ebi.ac.uk/~zerbino/velvet/). The VLP scaffolds were performed using NCBI blast with the POGs database (1e-10).

6. Circulating endotoxin levels measurement

Serum endotoxin was assayed using a chromogenic limulus amebocyte lysate (LAL) test, which is a quantitative test for gram-negative bacterial endotoxin (BIOENDO, Xiamen). Gram-negative bacterial endotoxin catalyzes the activation of a proenzyme in the LAL. The activated enzyme catalyzes the splitting of p-nitroaniline (pNA) from the colorless substrate Ac-Ile-Glu-Ala-Arg-pNA. The pNA released was measured photometrically at 405~410 nm following termination of the reaction. The correlation between the absorbance and the endotoxin concentration is linear in the 0.1~1.0 EU/ml range. All samples were run in duplicate within the same plate.

7. Statistics analysis
Statistical analyses were performed using the R package and GraphPad Prism. T-tests were performed to assess differences in the α-diversity of the T2D group and Ctrl group. For the comparison of continuous variables, the Mann-Whitney U-test was conducted for two groups. Multiple test correction was performed using the Benjamini and Hochberg method [53] and false discovery rate (FDR) was obtained. Spearman’s rank test was performed for correlation analysis. Pearson correlation analysis was used for evaluating the correlation between bacteriophages and the corresponding bacterial hosts.

**Declarations**

**Ethics approval and consent to participate**

The study was approved by the Ethics Committees of Shanghai Jiao Tong University Affiliated Sixth People’s Hospital. This study was registered at www.chictr.org.cn with clinical trial registration number ChiCTR-IPR-17011324. All subjects provided written informed consent before enrollment.

**Consent for publication**

Not applicable

**Availability of data and materials**

Raw sequencing data of 16S rDNA gene were deposited in the Sequence Read Archive (SRA) at NCBI under the accession ID PRJNA587207. Raw data of the phageome is being submitted to the SRA.

**Competing interests**

The authors clear no conflict of interest.

**Authors’ contributions**

Chang Liu, Jian Zhou, and Xiaokui Guo designed the study; Qian Chen performed the "wet experiments" and wrote the manuscript; Chong Li and Qian Chen conducted the analyses; Chang Liu, Jian Zhou and Chong Li edited the manuscript; Xiaojing Ma, Yun Shen, Wei Zhu, and Yan Zhang recruited the study subjects. Xiaojing Ma and Jian Zhou provided the clinical information about the patients with T2D.

**Acknowledgements**

We thank Jiahui Peng, Dongjun Dai, Xingxing He, and Lingwen Ying sincerely for their assistance in sampling. And we thank Lin Lyu for his help in the data processing.

**Funding**

The authors gratefully acknowledge the financial supported by the Shanghai Municipal Education Commission-Gaofeng Clinical Medicine Grant Support (20161430).

**References**


**Supplementary Information**

**Additional file 1:**

1. The number of the pOTUs of the detected 330 species of bacteriophages in each sample.
2. The relative abundance of bacteriophages in each sample. The differences of each bacteriophages between the T2D group and Ctrl group were assessed, and p value as well as FDR were presented.
3. The relative abundance of bacteriophages classified according to different host bacteria, excluding several bacteria genera including *Prochlorococcus*, *Synechococcus*, and *cyanobacteria*.

**Additional file 2:**

Proportion of various phages at different classification levels corresponding to Fig. 2. Data are shown in percentages.

**Additional file 3:**

The OTU table and taxa percent table of the detected bacteria according to 16S rDNA sequencing.

**Additional file 4:**

Clinical features and demographic information recorded for each patient with T2D.

**Additional file 5:**

Bacteria genera with significant correlation with T2D disease indicators, with a cutoff of FDR<0.05.

**Figures**
Figure 1

Schematic diagram of this study. The light-blue part is a sketch map of the preliminary exploratory herein, including exploring the relationship between phageome and putative host bacteriome, correlation between phageome and T2D disease indicators, discussion of the potential connection between the increase in serum LPS and the changes of bacteriophages.
Figure 2

Phageome features were highly diverse among different individuals. a Virus classification derived from sequence alignment. Phage community composition at the order level (b), family level (c), and genus level (d). The samples starting with D and H presented those from T2D patients and nondiabetic individuals respectively.
Figure 3

Illustration of the identified phageome characteristics. a α-diversity indexes of phageome in the T2D (D) group and Ctrl group (Ctrl). b Principal component analysis (PCA) of the phageome in the T2D group and Ctrl group. PCA plot of phageome samples from diabetes patients (blue dots) and controls (orange triangles). The first component (dim1) explained 21.8% of the total variance, the dim2 is 10.8%. c Venn diagram for the number distribution of phages classified by bacterial hosts in two groups. d The heatmap of the identified phages in T2D patients and nondiabetic controls according to their corresponding bacterial hosts. **p< 0.01.
Figure 4

Differential abundance of phage infected bacteria taxonomy between T2D patients and controls. a 18 bacteriophages with the most significant difference in the T2D group. b Differentially detected viruses at the level of families. c Phages clustered by 7 bacterial genera hosts changed significantly in the T2D group. *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001. pentagram: FDR < 0.05.
Figure 5

Alterations of the fecal bacteriome in the study. a Rarefaction curve for the fecal samples. b Box-plots for α-diversity metrics with the index of Chao1 (left) and PD (right) whole tree. c Unweighted unifrac PCoA, NMDS, and Anosim of bacteriome in the T2D group and Ctrl group. Taxonomic cladogram (d) and LDA (e) were obtained using LEfSe analysis of the 16S sequences in the T2D group (in blue) and Ctrl group (in orange). f The significantly altered bacterial genera in the T2D group. *p< 0.05, **p< 0.01, ***p< 0.001.
Figure 6

Bacteria-phage correlations in fecal samples in T2D patients and nondiabetic controls. a Results of spearman correlation analysis between the most abundant 20 bacteria genera and the most abundant 30 species of phage [29]. Blue circles indicate positive correlations and red circles indicate negative correlations. The size and shading indicate the magnitude of the correlation where darker shades denote more intensive correlations than light ones. b Correlations of the relative abundance of phages and their bacterial hosts in the T2D group and Ctrl group. The relative abundance was represented by the normalized OTU number. Optimal regression line for 95% confidence interval were plotted.
Figure 7

Changes in the abundance of the Enterobacteriaceae specific phages, gram-negative bacteria and their bacteriophages, and the related events in the serum. a Normalized OTU number of gram-positive bacterial phages and gram-negative bacterial phages in the T2D group and Ctrl group. b Changes in the relative abundance of gram-positive bacteria and gram-negative bacteria in both T2D group and Ctrl group. c The concentration of the concentration of serum LPS in the T2D group and Ctrl group. In each group, n=13. d The relative abundance of Enterobacteriaceae specific phages in two groups. *p< 0.05.
Figure 8

Schema graph of the speculated mechanism of how phagome alternations influence the development of T2D. Elevated levels of both Enterobacteriaceae specific phages and Enterobacteriaceae bacteria were found in the enteric microbiome of T2D patients, which might lead to the increase in circulating LPS and the systematic inflammatory subsequently, and result in the insulin resistance. Arrows present the conformed mechanisms by previous studies.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile3.xlsx
- Additionalfile2.xlsx
- Additionalfile5.xlsx
- Additionalfile1.xlsx
- Additionalfile4.docx