nature portfolio

Peer Review File

Gut Phage Biobank: a collection of bacteriophages targeting human commensal bacteria

Corresponding Author: Dr Minfeng Xiao

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
Xing et. al present a resource of phages and phage genomes targeted gut-associated microbes. Using a series of techniques, the group isolates 104 phages targeting gut microbiome associated species, many of which are rare or entirely novel. Following isolation, the group conducts a thorough genomic characterization, as well as some in vitro and in vivo testing. Overall, the manuscript is strong bioinformatically and provides some experimental results that point to the potential utility of the phages.
Line 156 Annotation: There is no mention of using Phold (https://github.com/gbouras13/phold) for genome annotation using structural homology. In our experience with Bacteroides phages (and non-Enterobactericeae phages in general), Phold vastly improves annotations for poorly studied phages.
Throughout: While I appreciate that this is a resource paper, much of the manuscript is a tough read, especially Lines 256-357. There isn't much analysis aside from descriptively stating genomes or rough structural features of the phages. For improved readability, I would suggest supplemental text and perhaps a short main text summary with the highlights.
Lines 431-463: The animal experiments demonstrate that the phages can inhibit the establishment of Dorea in the mouse gut. However they are quite preliminary. It would be more compelling to demonstrate that Dorea can stably colonize post-antibiotic treatment enter a steady state abundance and then test if the phages can deplete the population. Clearly, at Day 1, the Dorea have yet to fully expand so it is difficult to gauge the utility of this phage. Similarly, it would be interesting to know if resistance is observed in the mouse gut. Finally, even though the mouse microbiome is perturbed, the argument was made that the phages would specifically target Dorea. It would be interesting to see if phages altered the mouse microbiome more broadly in the same metagenomic dataset.
Throughout: Bacterial resistance to phage infection is inevitable and a hindrance to phage therapy. There is no mention of the evolution of phage resistance throughout the manuscript. 1) For the phages that are amenable to in vitro experimentation, do you observe genetic resistance to phage (ie a bacterial population that resists further infection post-phage infection)? 2) For all of the plate-reader kill curves, it seems like phages are added at early lag phase. In general, these experiments allow for a subculture of the bacterial strain prior to the phage challenge to assess if the phages can inhibit growth or kill the population. I would recommend repeating the kill curve experiments with early log phase bacteria.

Reviewer #3

(Remarks to the Author)
In this revised manuscript, Xing and colleagues have clarified several of their approaches, streamlined the manuscript to make a more compelling and detailed characterization of the novel gut phages they isolated. The authors have adequately addressed most of my comments and concerns. In responding to the previous reviewers' concerns, new data were added, to the point where I read this revision as a new manuscript.
The discussion is less descriptive and provides interesting contrasts between oxic and anoxic phage-bacteria interactions, something that is particularly relevant for the field. This has also resulted in a more thorough characterization of the isolated phages, which I believe to be valuable for the field.
In general, I think the manuscript could be shortened for improved clarity and some sections of the Discussion could be regrouped or moved for the same purpose.
Minor comments:
- I would still recommend English proof reading, as some terms remain vague or inadequate. I will highlight a few examples,

but I would suggest going over the manuscript more thoroughly. L41: “normalize gut microbiome”; L43: “inhibit”; L67: “accumulate”; L76: “necessary”; L79: “perform”; etc.
- L 40-42: I would also include prebiotics as a way to control gut bacteria.
- L87: a transition or a few words detailing why the focus was on these two bacterial species would be good to include.
- L97-99 and 117-119: the enrichment method should also allow you to detect phages with longer replication times and temperate phages. I believe this is also in accordance with the different types of environment sampled, as animal guts should contain more lysogens.
- Fig. 2A: please use the updated bacterial nomenclature.
- L172-173: remove the last part of the sentence, it doesn’t add any information. On a related note, wouldn’t all genes involved in the lytic cycle also be present in non-temperate phages? I think the absence of integrase and integrase-like genes is a better indicator of replication cycles.
- L175: I don’t understand the term “micro-ecology”. Please rephrase.
- Fig. 2E: I would suggest using different shading and colours between the phage genes linked to a protein type and the bacterial hosts. Or include the protein type in bold letters next to the corresponding phage genes. For example, E. faecalis and Lysis genes are the same colour. Also, consider whether my previous comment about the lytic genes being a specific set of genes not involved in other replication cycles to determine whether this category should remain in this figure panel.
- L188: presumably 17 phage groups?
- L216-218: words missing or two sentences that should part of the same one.
-L228-235: why were only phages from Groups 2 and 4 compared? It seems like Group 4 has the broadest range and could be compared to all other groups.
- L 296: replace “breaking” by “lysing”.
- The TEM photos are beautiful, but I think the scale in the most zoomed-out pictures are off, as it would mean that some phages would be close to measuring 1mm.
- I believe that the N for the mouse experiment is 15, as there were 5 mice per experimental group and there 3 experimental groups. This experiment should be interpreted with caution, as the time frame between Dorea colonization and phage exposure is very short. This is seen by the low levels of engraftment of the bacteria, which take up to 3 days to increase in abundance in the mice. This could be an issue when translating these findings to humans, as typically, the bacterial pathogen would be present in high numbers.
- L470-472: rephrase, I don’t quite understand how these two sentences are connected (or contrast with each other).
- L 488: the quorum-sensing point here is probably correct, but is not intuitive here based on the data presented. I would think that modifications in bacterial metabolism due to culturing are more likely to be involved here.

Version 1:

Reviewer comments:

Reviewer #3

(Remarks to the Author)
In this revised manuscript, Xing and colleagues have streamlined the manuscript to make a more compelling and detailed characterization of the novel gut phages they isolated. The authors have adequately addressed most of my comments and concerns. I am supportive of the work completed and only have a few minor comments left.
Minor comments:
- L32: “while” doesn’t seem to be the correct word.
- L42: replace the term “balance” – what does a “balanced” gut microbiome look like? I would suggest “restore gut microbiome homeostasis” or “restore the diversity of the gut microbiome”. In the same sentence, I would remove “its composition is complex and” to just keep the points that FMT lacks precise regulation and control.
- L53-57: long sentence, consider splitting it.
- L78: “involved”
- L84: this is the first time the acronym GPB is used, detail it.
- L121-122: were the bacterial targets experimentally validated or are they predicted?
- L156: I would replace “phage-bacteria interactions” by “bacterial infection”.
- Fig 2B: what do the dashed black lines refer to?
- Fig 2D: the end of the figure legend doesn’t seem to match the panel. Also, the figure legend could maybe be expanded to highlight the relevance of the figure?
- Fig. 3B: don’t the rows represent phages and the columns the bacteria? Also, remove the text in the legend referring to infection under anoxic infections only, as this condition doesn’t exist for these phages and bacteria.
- L982: “grey lumps” is not adequate. The linkages are shown in grey shading.
- Fig. 3F and 3H: same comment about rows and columns as in panel B.

Before addressing the reviewers' comments in detail, we'd like to recap the summary of the key changes:

1. Improving the Clarity of the Text

We have revised the manuscript to enhance its logical flow and conciseness, ensuring a more streamlined presentation of our findings.

1.1 Restructuring and Refocusing the Results Section

The sections on "Novel crAss-like phages" and "Other novel phages targeting obligate anaerobes" have been consolidated into a single section titled "Novel phages targeting obligate anaerobes", focusing on the key conclusions. Detailed genomic and biological features, previously included in the main text, have been moved to the Supplemental Results for conciseness. This restructuring allows for a more concise presentation while still providing comprehensive data in the supplementary section. Additionally, the section on "Phage-bacteria interactions under oxic and anoxic conditions" has been further refined through systematic data integration and logical reorganization, significantly enhancing the clarity of data presentation.

1.2 Reframing and Refining the Discussion

The Discussion section has been reorganized for improved clarity and logical flow, with specific emphasis on: (1) challenges in isolating gut phages, (2) the complex interactions between GPB phages and their bacterial hosts, and (3) limitations and perspectives of the animal experiments. These changes ensure a more coherent synthesis of our findings and their broader implications.

2. Addressing the Preliminary Nature of the *Dorea*-Phage Experiments

We have carefully addressed the limitations of the *Dorea*-phage experiments, particularly regarding *Dorea* colonization in the Discussion section. Although the relative abundance of *Dorea* in mice was low during phage treatment, significant differences in the abundance of *Dorea* between the phage-treated and control groups were observed three days post-treatment. This suggests that phage therapy can be effective even at low bacterial abundance, providing insights into its potential of early-stage infections or when bacterial pathogens are not yet established. By clearly acknowledging these limitations and adjusting our interpretations accordingly, we provide a more balanced and accurate assessment of our findings.

In accordance with the revisions made to the manuscript, the corresponding figures have also been updated as follows:

- Based on the reviewers' suggestions, we have updated the bacterial nomenclature. Please refer to Figure 2A for these changes. Additionally, we have incorporated Phold for phage functional protein annotation. Figure 2C and Figure 2E now present the newly annotated results obtained by combining Phold with our previous annotation workflow.

- As we have consolidated "Novel crAss-like phages" and "Other novel phages targeting obligate anaerobes" into a single section titled "Novel phages targeting obligate anaerobes," we have revised the figures accordingly. The original Figure 4 has been split into Figure 4 and Supplementary Figures S7-S11. Figure 4 now highlights the key conclusions of the novel

phages targeting obligate anaerobes. Meanwhile, Supplementary Figures S7-S11 provide detailed genomic and biological features of the crAss-like phages, Eggerthella phage, Mediterraneibacter phage, Phocaeicola phage, and Dorea phages, respectively.

Reviewer #1 (Remarks to the Author):

Xing *et. at* present a resource of phages and phage genomes targeted gut-associated microbes. Using a series of techniques, the group isolates 104 phages targeting gut microbiome associated species, many of which are rare or entirely novel. Following isolation, the group conducts a thorough genomic characterization, as well as some in vitro and in vivo testing. Overall, the manuscript is strong bioinformatically and provides some experimental results that point to the potential utility of the phages.

Line 156 Annotation: There is no mention of using Phold (https://github.com/gbouras13/phold) for genome annotation using structural homology. In our experience with Bacteroides phages (and non-Enterobactericeae phages in general), Phold vastly improves annotations for poorly studied phages.

Thank you very much for your valuable suggestion regarding the use of Phold for genome annotation. We have incorporated Phold into our pipeline and applied it on our bacteriophage genomes. As anticipated, Phold enhanced annotation, particularly for structural proteins in poorly studied phages. Integration of Phold with our previous annotation workflow has indeed increased both the annotation coverage and accuracy. We believe this addition strengthens our manuscript. Please refer to the manuscript line 144-145 and Figure 2C.

Throughout: While I appreciate that this is a resource paper, much of the manuscript is a tough read, especially Lines 256-357. There isn't much analysis aside from descriptively stating genomes or rough structural features of the phages. For improved readability, I would suggest supplemental text and perhaps a short main text summary with the highlights.

Thank you for your constructive feedback. We have revised the manuscript to enhance its clarity and flow, particularly in the sections between lines 256-357. Following your suggestion, we have focus on the content on the key conclusions by restructuring the sections on "Novel crAss-like phages" and "Other novel phages targeting obligate anaerobes" into a new section titled "Novel phages targeting obligate anaerobes.". Detailed genomic and biological features, previously included in the main text, have been moved to the Supplementary Results, allowing for a more concise presentation while retaining essential details. Please refer to the manuscript line 212-254 and Figure 4 and the Supplementary Results and Figure S7-11.

Additionally, the section on "Phage-bacteria interactions under oxic and anoxic conditions" has been further refined through systematic data integration and logical reorganization, significantly enhancing the clarity of data presentation. Please refer to the manuscript line 172-211.

Lines 431-463: The animal experiments demonstrate that the phages can inhibit the establishment of *Dorea* in the mouse gut. However, they are quite preliminary. It would be more compelling to demonstrate that *Dorea* can stably colonize post-antibiotic treatment enter a steady state abundance and then test if the phages can deplete the population. Clearly, at Day 1, the *Dorea* have yet to fully expand so it is difficult to gauge the utility of this phage. Similarly, it would be interesting to know if resistance is observed in the mouse gut. Finally, even though the mouse microbiome is perturbed, the argument was made that the phages would specifically target *Dorea*. It would be interesting to see if phages altered the mouse microbiome more broadly in the same metagenomic dataset.

1) We sincerely appreciate your insightful comments regarding our animal experiments. We recognize the limitations of the current experimental design, particularly the lack of a stable *Dorea* colonization phase prior to phage treatment. We agree that establishing stable *Dorea* colonization post-antibiotic treatment would provide a more rigorous evaluation of phage-mediated depletion. In the revised manuscript, we have addressed these limitations and discussed the challenges related to *Dorea* colonization. Please refer to the manuscript line 386-402.

Our primary objective was to provide preliminary evidence for phage-mediated inhibition of *Dorea* colonization in the mouse gut. Although the bacterial abundance at Day 1 was low, suggesting incomplete colonization, we observed a significant reduction in *Dorea* levels after three days of phage administration compared to the control group. This indicates that the phages' potential to inhibit *Dorea* even at low abundance, highlighting their inhibitory capacity.

While the results are preliminary, they support the potential of these phages to target *Dorea*. We plan to incorporate your valuable suggestions into future studies to further validate and refine these findings.

2) Regarding phage resistance, we agree that evaluating resistance *in vivo* is an important aspect of phage therapy. However, due to the limited *Dorea* colonization in our current experimental setup, it was challenging to investigate resistance *in vivo*.

In our *in vitro* experiments, 70% of the observed killing curves showed a rise in OD over time, suggesting bacterial regrowth and potential resistance. For example, in Figure S8C-D, phages inhibited *Dorea* growth in the first 8-12 hours, but OD increased afterward, indicating the emergence of resistant bacterial populations. We also observed that different *Dorea* strains exhibited varying resistance patterns. With RTGS0004 developing resistance earlier than RTGS0005, which suggesting strain-specific resistance mechanisms. We have addressed the *in vitro* resistance findings in the revised manuscript. Please refer to the manuscript line 377-385.

These *in vitro* observations provide preliminary insights into *Dorea* phage resistance dynamics. Further studies, particularly *in vivo*, are needed to understand the resistance mechanisms more comprehensively. Future work will focus on optimizing the *Dorea* colonization in the mouse gut and exploring resistance mechanisms in more detail.

3) Thank you for your comment regarding the broader impact of the phages on the mouse microbiome. Investigating whether phages alter the microbiome is indeed an important question

and we have conducted preliminary assessments.

Host-range experiments (as shown in Figure 3) indicated that the phages exhibit high specificity for *Dorea*, minimizing potential impacts on other gut microbiota members. Additionally, since antibiotic treatment is commonly used to reduce background bacterial interference in animal studies, our experimental setup with SPF mice pre-treated with antibiotics allowed us to focus on *Dorea* inhibition. While antibiotic treatment alters gut microbiota composition, we did not investigate the broader effects of phages on the entire microbiome in these animals.

Given the limitations of this study, a detailed analysis of the phages' impact on the overall microbiome would be more appropriate in future research. This could involve metagenomic sequencing and functional analyses to better understand the off-target effects of phage therapy and how microbial interactions contribute to broader shifts in the microbiome.

Throughout: Bacterial resistance to phage infection is inevitable and a hindrance to phage therapy. There is no mention of the evolution of phage resistance throughout the manuscript. 1) For the phages that are amenable to in vitro experimentation, do you observe genetic resistance to phage (i.e. a bacterial population that resists further infection post-phage infection)? 2) For all the plate-reader kill curves, it seems like phages are added at early lag phase. In general, these experiments allow for a subculture of the bacterial strain prior to the phage challenge to assess if the phages can inhibit growth or kill the population. I would recommend repeating the kill curve experiments with early log phase bacteria.

1) Thank you for your insightful comment regarding the evolution of bacterial resistance to phage infection. As mentioned in our response, we observed preliminary resistance phenotypes in our *in vitro* experiments. However, we did not isolate resistant strains post-co-culture or conduct whole-genome sequencing on the bacterial populations to identify specific genetic mechanisms of resistance in this study.

We acknowledge that bacterial resistance is a key challenge in phage therapy. Several strategies have been proposed to mitigate the development of resistance, including: (1) repeated selection of phages against resistant bacteria to develop phage cocktails [1]; (2) screening for phages with broad-spectrum antibacterial activity targeting multiple bacterial receptors [2]; and (3) combining phages with small molecule drugs or phage derivatives to enhance efficacy [3,4].

2) Thank you for your valuable feedback. To clarify, in the kill curve experiments, bacterial glycerol stock was sub-cultured and inoculated to grow to the early log phase (OD600 ≈ 0.3). Phages were then added at MOI 1 and 10. To facilitate comparison of killing curves across different samples, the phage-bacteria mixtures were diluted to an OD600 of approximately 0.1 before starting measurement. This experimental setup ensures accurate capture of bacterial killing dynamics during the early log phase. We have updated the manuscript to include this clarification for greater transparency. Please refer to lines 625-629 for the revised description.

Reviewer #3 (remarks to the author):

In this revised manuscript, Xing and colleagues have clarified several of their approaches, streamlined the manuscript to make a more compelling and detailed characterization of the novel gut phages they isolated. The authors have adequately addressed most of my comments and concerns. In responding to the previous reviewers' concerns, new data were added to the point where I read this revision as a new manuscript.

The discussion is less descriptive and provides interesting contrasts **between oxic and anoxic phage-bacteria interactions**, something that is particularly relevant for the field. This has also resulted in a more thorough characterization of the isolated phages, which I believe to be valuable for the field.
In general, I think the manuscript could be **shortened** for improved clarity and some sections of the **Discussion could be regrouped or moved** for the same purpose.

We sincerely appreciate your suggestion to improve the clarity and flow of the manuscript. In response, we have revised the manuscript from two key dimensions to address your concerns:

1) Structural Reorganization for Improved Clarity:

   Based on your suggestion, we have revised the manuscript to improve its clarity and flow. The sections on "Novel crAss-like phages" and "Other novel phages targeting obligate anaerobes" have been consolidated into a single section titled "Novel phages targeting obligate anaerobes", focusing on the key conclusions. Detailed genomic and biological features, previously included in the main text, have been moved to the Supplemental Results for conciseness. This restructuring allows for a more concise presentation while still providing comprehensive data in the supplementary section. Additionally, the section on "Phage-bacteria interactions under oxic and anoxic conditions" has been further refined through systematic data integration and logical reorganization, significantly enhancing the clarity of data presentation.

2) Reframing and Refining the Discussion:

   The Discussion section has been reorganized for improved clarity and logical flow, with specific emphasis on: (1) challenges in isolating gut phages, (2) the complex interactions between GPB phages and their bacterial hosts, and (3) limitations and perspectives of the animal experiments. These changes ensure a more coherent synthesis of our findings and their broader implications.

Minor comments:

- I would still recommend English proof reading, as some terms remain vague or inadequate. I will highlight a few examples, but I would suggest going over the manuscript more thoroughly. L41: "normalize gut microbiome"; L43: "inhibit"; L67: "accumulate"; L76: "necessary"; L79: "perform"; etc.

Thank you for your valuable feedback and suggestions. We have carefully reviewed your comments and have made comprehensive revisions to improve the manuscript's language and

clarity. Below are our responses to the specific examples you highlighted:

- Line 42 (old Line 41): We have revised "normalize gut microbiome" to "restore the balance of the gut microbiome" to convey the effect of FMT more accurately. The new sentence is " While FMT may restore the balance of the gut microbiome, its composition is complex and lacks precise regulation and control currently."
- Line 43: We have replaced "inhibit" with "target and infect" to more precisely describe the action of bacteriophages on bacteria. The new sentence is "Bacteriophages, which specifically target and infect bacteria, have been successfully used in the treatment of multidrug-resistant infections."
- Line 66 (old Line 67): We have changed "accumulate" to "isolate and characterize" to better emphasize the need for active research and analysis, specifically conveying the process of obtaining and studying phages.
- Line 75 (old Line 76): We have rewritten the entire sentence to " Despite these approaches, developing systematic method for isolating novel or previously undescribed phages remains a critical challenge."
- Line 78 (old Line 79): We have substituted "perform" with "conduct" to describe our research process more professionally. The new sentence is " We conduct infection matrix analysis to assess the host specificity and infectivity of the phages under both oxic and anoxic conditions, …"

In addition to these specific revisions, we have also conducted a thorough proofreading of the entire manuscript. We have carefully reviewed each term and sentence to ensure that the language is clear, precise, and appropriate for the scientific context. We have made several adjustments to improve readability and coherence throughout the text.

- L 40-42: I would also include prebiotics as a way to control gut bacteria.

Thank you for your suggestion to include prebiotics. We have revised the manuscript to incorporate this important point. The updated sentence now reads:

Line 39-42: "Interventions such as probiotics, prebiotics, and fecal microbiota transplantation (FMT) have been explored to address chronic diseases associated with gut bacteria. Probiotics offer therapeutic benefits through beneficial microorganisms, while prebiotics selectively stimulate the growth of beneficial bacteria in the gut."

- L87: a transition or a few words detailing why the focus was on these two bacterial species would be good to include.

Thank you for this suggestion. We have revised the sentence to address this point. The updated version now reads:

" We analyze the prevalence and relative abundance of *Mediterraneibacter* and *Dorea* bacterial strains, as previous studies have highlighted their association with obesity and T2D in European cohorts, along with the distribution of their corresponding phages, to verify whether similar association exist in Asian disease and health cohorts." Please refer to the manuscript line 85-88.

- L97-99 and 117-119: the enrichment method should also allow you to detect phages with longer replication times and temperate phages. I believe this is also in accordance with the different types of environments sampled, as animal guts should contain more lysogens.

Thank you for your insightful comment regarding the potential of our enrichment method to detect phages with longer replication times and temperate phages. While our study didn't extensively focus on temperate phages, we acknowledge the importance of your point. We'd like to address three aspects of our approach:

1) We have carefully reviewed our results regarding the enrichment method and its potential to detect temperate phages, particularly in gut-associated environments. In our dataset, we identified 17 phages (20.99%) carrying lysogeny-related genes among the 81 phages isolated using traditional methods. In contrast, among the 32 phages obtained through enrichment-based methods (including metagenomics- guided method), 28.13% harbored lysogeny-associated genes. This suggests that enrichment method may better capture non-traditional lytic phages or uncover diverse phage lifecycles present in fecal samples. Additionally, we note that all phages obtained via the enrichment method originated from animal-associated environments (mixed fecal samples and wastewater), aligning with your opinion that animal guts should contain more lysogens.

2) We encountered challenges in inducing sufficient prophage yields from gut bacteria for comprehensive biological and genomic analyses. As noted in our discussion, to improve the capture of temperate or carrier-state phages, which are challenging to form plaques, optimizing plaque-free culturing approach and incorporating sequencing-guided detection will be essential.

3) Our study primarily focused on gut bacterial hosts associated with diseases. Given our future aim to potentially use phages to inhibit these bacteria and alleviate disease, lytic phages presented a more immediate application scenario. However, when optimizing methods for prophage recovery in future studies, we plan to emphasize iterative enrichment strategies and the analysis of more complex samples to improve the detection of temperate and carrier-state phages.

- Fig. 2A: please use the updated bacterial nomenclature.

Thank you for your valuable suggestion. We have updated the bacterial taxonomy in the manuscript according to the NCBI taxonomy (as of March 17, 2025). After the update, the phages we isolated target a total of 4 bacterial phyla, 6 class, 9 order, 10 families, 14 genera, and 17 species. These include 51 phages targeting Bacillota, 30 phages targeting Pseudomonadota, 16 phages targeting Bacteroidota, and 7 phages targeting Actinomycetota.

Additionally, we have revised the entire manuscript and updated the corresponding figures to reflect these changes. Please refer to the manuscript line 119-120 and Figure 2A.

- L172-173: remove the last part of the sentence, it doesn’t add any information. On a related note, wouldn’t all genes involved in the lytic cycle also be present in non-temperate phages? I

think the absence of integrase and integrase-like genes is a better indicator of replication cycles.

Thank you for your valuable comments, and we apologize for any confusion. Our intention was to highlight that the 37 phages in question possess genes for all key stages of the lytic cycle, which may confer a higher degree of autonomy compared to phages lacking genes for certain stages. Please refer to the manuscript line 157-161. Furthermore, as per your recommendation, we have removed the last part of the sentence ("potentially exhibiting a higher degree of autonomy during their reproductive process") to better clarify our intended meaning.

We also agree with your opinion that the genes involved in the lytic cycle would be present in both temperate and non-temperate phages. Temperate phages can indeed carry genes encoding lytic proteins, which allow them to switch to a lytic cycle, thereby transforming into virulent phages. This allows them to escape from the compromised host bacteria and seek new hosts.

Regarding the integrase genes, we agree that the absence of integrase and integrase-like genes is generally a useful indicator for distinguishing replication cycles. However, our study presents an interesting finding: we identified lysogeny-related genes in the genomes of 24 phages that stably form plaques. This suggests that while these phages may integrate their genomes into bacterial chromosomes, they also retain the ability to lyse bacteria.

- L175: I don't understand the term "micro-ecology". Please rephrase.

Thank you for your feedback regarding the term "micro-ecology." We have revised the sentence for clarity:

" Previous studies have shown that some phages carry auxiliary metabolic genes (AMGs), which can alter host metabolism and influence the entire microbial ecosystem" Please refer to the manuscript line 162-163.

Additionally, we have reviewed the full manuscript for similar instances and revised ensure consistency and clarity throughout:

L407: "…, our understanding of phage-mediated impacts on microbial diversity, ecosystem dynamics, evolution, and human health will significantly improve".

- Fig. 2E: I would suggest using different shading and colours between the phage genes linked to a protein type and the bacterial hosts. Or include the protein type in bold letters next to the corresponding phage genes. For example, E. faecalis and Lysis genes are the same colour. Also, consider whether my previous comment about the lytic genes being a specific set of genes not involved in other replication cycles to determine whether this category should remain in this figure panel.

1) Thank you for your suggestion regarding Figure 2E. We have included the protein type in bold letters next to the corresponding phage genes to enhance the readability of the figure.

2) Thank you for your thoughtful comment on the classification of lytic genes in the figure panel. Among the phages we isolated, some contained lysogeny-related genes while also harboring lytic-associated genes and exhibiting plaque formation. This observation suggests

that gut-associated phages may follow a unique pattern of life cycle transitions. Therefore, we believe it is valuable to retain this data—not to emphasize their specific life cycle classification, but rather to highlight the presence of these genes within the phage genomes. We hope this clarification addresses your concern.

- L188: presumably 17 phage groups?

We apologize for this confusion. The term "17" does not refer to 17 phage groups, but rather to 17 phage-bacteria interaction profiles. Each interaction profile represents phages that specifically infect a particular bacterial species and are unable to infect other species. We studied 17 different host bacterial species, which is why our results show 17 distinct groups in the figure. These profiles reflects the high host specificity observed in our study, where phages that target one bacterial species cannot infect bacteria of other species.

We have revised the sentence for clarity: "The results showed 17 distinct phage-bacteria interaction profiles, with each profile representing phages specific to particular bacterial species." Please refer to the manuscript line 173-175.

- L216-218: words missing or two sentences that should part of the same one.

Thank you for pointing out the issue in lines 216-218. The original sentence was intended to convey that for *C. sakazakii* phages, all 22 isolates could infect the three host strains under oxic conditions. However, under anoxic conditions, only 5 isolates infected all three host strains, and 4 other isolates infected one of the host strains. That is to say, the host range of *C. sakazakii* phages varied under different oxygen conditions, with a narrower host range observed under anoxic conditions.

We have observed that oxygen conditions can influence the infectivity of gut phages targeting multiple bacterial species. To streamline the manuscript and condense this information, we have integrated these observations into a single sentence. The revised sentence now reads: "Oxygen conditions influenced infectivity: 25 of 26 E. faecalis phages formed plaques under both conditions, *C. sakazakii* phages infected fewer hosts under oxic conditions, while *Lacticaseibacillus paracasei* phage, isolated under anoxic conditions, only formed plaques under such conditions." Please refer to the manuscript line 183-185.

-L228-235: why were only phages from Groups 2 and 4 compared? It seems like Group 4 has the broadest range and could be compared to all other groups.

Thank you for your insightful question regarding the comparison of phages from Groups 2 and 4. We appreciate the opportunity to clarify our approach.

We grouped the phages based on whole-genome similarity (≥97%), resulting in four groups. Groups 1 and 3 each contain only one phage, while groups 2 and 4 contain multiple phage isolates. To better explore the relationship between genetic features and host range phenotypes, we primarily focused on comparing the differences within each group rather than between groups. This approach was chosen because, despite the high genomic similarity within groups

2 and 4, phages from these groups exhibited different host ranges, making within-group comparisons both meaningful and insightful.

We specifically chose to focus on groups 2 and 4 because these groups contained multiple isolates, which allowed us to observe intra-group variability in host range. While it is true that phages from group 4 demonstrated the broadest host range, comparing group 4 with the other groups was not the primary focus of this study. Our goal was to investigate the differences in host interactions among phylogenetically related phages within the same group, rather than comparing the host ranges of phages across different groups. Moreover, the relatively high genomic similarity within groups 2 and 4 made them particularly suitable for this analysis.

Interestingly, despite the high genomic similarity of the phages in both groups, we found that specific tail-related proteins (tail protein in Group 2 and a tail tape measure protein in Group 4) exhibited amino acid differences in each group. This finding suggests that these specific tail-related proteins may be crucial factors influencing the host range of these phages.

We have revised the paragraph to better understanding: "For *Mediterraneibacter gnavus* (formerly *Ruminococcus gnavus*) phages, seven were classified into 4 genomic groups, with Group 2 and 4 showing variable host infectivity within groups (Figure 3F). In Group 2, comparative genomic analysis revealed a 7-amino acids deletion in the tail protein of phage CPB0986 and a 162-amino acids deletion in CPB1098 relative to CPB1099 (Figure 3G, S5). In Group 4, the tail tape measure protein of CPB1094 showed a 21-amino acids deletion compared to CPB1088 (Figure 3G and Figure S6), suggesting these structural variations may influence host range." Please refer to the manuscript line 193-199.

- L 296: replace "breaking" by "lysing".

Thank you for your valuable suggestion regarding the terminology used in our manuscript. We have carefully considered your feedback and have revised the sentence to enhance its precision and professionalism. Rather than simply replacing "breaking" with "lysing," we have rephrased the sentence to incorporate the more specific and scientific term "non-lytic replication." This change not only addresses your concern but also provides a clearer and more accurate description of the phenomenon we are discussing. The updated text now reads:

Line 231-233 "In contrast, crAss-like phages from other families may possess distinct lifecycles or host interactions (e.g., non-lytic replication), limiting plaque assays detection."

We believe this modification better reflects the nuances of the phage-host interactions and aligns with the standards of scientific communication.

- The TEM photos are beautiful, but I think the scale in the most zoomed-out pictures are off, as it would mean that some phages would be close to measuring 1mm.

We apologize for this confusion. We would like to clarify the scale bars in the TEM images. The scale bar in the TEM images of panels Figure S7 (old Figure 4A), Figure S9A (old Figure 4D), and Figure S10B (old Figure 4M) corresponds to 100 nm, while the scale bars in Figure

S11A (old Figure 4G) and Figure S8A (old Figure 4J) correspond to 50 nm. Meanwhile, we have included images of the phage plaques with a scale of 10 mm to provide further context for the sizes. We have implemented revisions in the figures to explicitly differentiate TEM micrographs from phage plaque images, please refer to Figure S7-S11.

- I believe that the N for the mouse experiment is 15, as there were 5 mice per experimental group and there 3 experimental groups. This experiment should be interpreted with caution, as the time frame between *Dorea* colonization and phage exposure is very short. This is seen by the low levels of engraftment of the bacteria, which take up to 3 days to increase in abundance in the mice. This could be an issue when translating these findings to humans, as typically, the bacterial pathogen would be present in high numbers.

1) Thank you for your attention to detail regarding the number of mice in our experiment. We apologize for any confusion caused by our description. To clarify, the correct number of mice used in the study was 15, as you indicated. These 15 mice were initially treated with antibiotics, and then randomly divided into three groups (n=5 per group). One group (n=5) was treated with saline suspension for the entire experiment. The remaining two groups (n=10) were subjected to bacterial colonization starting from Day 0. From Day 1 onward, these two colonization groups were further divided. One group (n=5) received phage treatment, while the other group (n=5) was treated with a vehicle (saline) as a control.

We have revised the relevant section of the manuscript to clarify this. The updated text now reads:

Line 321-331: "…, we treated specific-pathogen-free (SPF) mice (n=15) with a one-week regimen of four antibiotics, followed by a two-day washout period. The mice were then randomly allocated into three groups (n=5 per group). The control group (n=5) received a saline suspension throughout the experiment. The remaining two groups (n=10), designated as the phage-treated and phage-free groups, were colonized with 107 CFU of a specific bacterial suspension one day prior to the initial phage administration (Day 0). Starting on Day 1, the phage-treated group (n=5) was administered 108 PFU (MOI 10) of corresponding phages via oral gavage daily from Day 1 to Day 4, while the phage-free group (n=5) received a saline vehicle at equivalent time points. We collected fecal samples 12 hours post-gavage to quantify residual host bacteria via CFU enumeration and to assess the relative abundance of the host bacteria and associated phages through metagenomic sequencing (see Methods)"

2) Thank you for your insightful comment regarding the interpretation of our experiment. We appreciate your perspective and agree that the short time frame between *Dorea* colonization and phage exposure warrants careful consideration. We acknowledge that the relative abundance of *Dorea* in the mice was relatively low at the time of phage treatment. However, despite these relatively low bacterial colonization levels, we were still able to observe significant differences in the abundance of *Dorea* between the phage-treated and control groups three days after phage administration. These findings suggest that phage treatment can have an effect even when bacterial populations are not yet abundant, which may provide valuable insights into phage therapy's potential effectiveness in early stages of infection or when

bacterial pathogens are not in high numbers. We appreciate your feedback and have systematically discussed the limitations of the animal experiment and the issues related to *Dorea* colonization. Please refer to the manuscript line 390-395.

- L470-472: rephrase, I don't quite understand how these two sentences are connected (or contrast with each other).

We aimed to highlight the contrast between findings from bioinformatics studies based on metagenomic data and those from experimental studies using isolated gut phages. While genomic analyses of public datasets (GPD) suggest 35.49% of phages exhibit cross-species infectivity (including 0.13% with phylum-spanning potential), our systematic experimental testing of 17 bacterial species using GPB isolates revealed stringent strain/species-level specificity. Crucially, this narrow host range aligns with prior experimental studies of cultured phages, reinforcing the necessity to synthesize computational and experimental approaches. To clarify this distinction, we revised the text as follows:

Line 352-356: "Bioinformatic analyses of GPD revealed that 35.49% of phages infect multiple species, with some exhibiting extensive host range, even spanning across phyla (0.13%)18. In contrast, GPB phages showed high host species specificity, particularly those targeting Bacteroides and Parabacteroides, consistent with prior study on isolated gut phages."

- L 488: the quorum-sensing point here is probably correct but is not intuitive here based on the data presented. I would think that modifications in bacterial metabolism due to culturing are more likely to be involved here.

Thank you for your insightful comment regarding the explanation of the discrepancy between predicted and isolated phages. We have revised this section to include your suggestion about the potential role of bacterial metabolism modifications due to culturing. Revised text:

L 374-377: "This discrepancy may be attributed to several factors, including bacterial metabolic changes during culturing, adaptive immune mechanisms, and quorum-sensing mechanisms, which may influence phage-host dynamics and our ability to isolate phages."

1 Federici, S. *et al.* Targeted suppression of human IBD-associated gut microbiota commensals by phage consortia for treatment of intestinal inflammation. *Cell* **185**, 2879-2898 e2824, doi:10.1016/j.cell.2022.07.003 (2022).

2 Rotman, E. *et al.* Rapid design of bacteriophage cocktails to suppress the burden and virulence of gut-resident carbapenem-resistant Klebsiella pneumoniae. *Cell Host Microbe* **32**, 1988-2003 e1988, doi:10.1016/j.chom.2024.09.004 (2024).

3 Baker, Z. R. *et al.* Sustained in situ protein production and release in the mammalian gut by an engineered bacteriophage. *Nat Biotechnol*, doi:10.1038/s41587-025-02570-7 (2025).

4 Dong, X. *et al.* Bioinorganic hybrid bacteriophage for modulation of intestinal microbiota to remodel tumor-immune microenvironment against colorectal cancer. *Sci Adv* **6**, eaba1590, doi:10.1126/sciadv.aba1590 (2020).

Our responses to the comments are in blue.

We would like to thank you for your thorough reading and the detailed feedbacks.

Reviewer #3:

In this revised manuscript, Xing and colleagues have streamlined the manuscript to make a more compelling and detailed characterization of the novel gut phages they isolated. The authors have adequately addressed most of my comments and concerns. I am supportive of the work completed and only have a few minor comments left.

- L32: “while” doesn’t seem to be the correct word.

Thank you for raising this question. We have revised the original text accordingly.

Revision (Manuscript, Introduction, L32): “Cohort analysis revealed a higher prevalence of *Mediterraneibacter* and *Dorea* **and** a lower prevalence of *Mediterraneibacter* phages in Asian disease population.”

- L42: replace the term “balance” – what does a “balanced” gut microbiome look like? I would suggest “restore gut microbiome homeostasis” or “restore the diversity of the gut microbiome”. In the same sentence, I would remove “its composition is complex and” to just keep the points that FMT lacks precise regulation and control.

Thank you for your suggestions. We have directly modified the original text.

Revision (Manuscript, Introduction, L42-43): “While FMT may **restore gut microbiome homeostasis**, **it currently lacks precise regulation and control.**”

- L53-57: long sentence, consider splitting it.

Thank you for your suggestions. We have split the long sentence into two to enhance readability, making the sentence following "for example" stand alone.

- L78: “involved”

Thank you for raising this question. We have revised the original text accordingly.

Revision (Manuscript, Introduction, L77): “We analyze their genomic characteristics and the genes **involved** in phage-bacteria interactions.”

- L84: this is the first time the acronym GPB is used, detail it.

Thanks for the careful read. GPB stands for "Gut Phage Biobank". We have added the full term for the acronym GPB at the beginning of this paragraph.

Revision (Manuscript, Introduction, L77): "In this study, we construct **a gut phage biobank (GPB)** through comprehensive **isolation of** phages targeting abundant or disease-associated gut commensal bacteria."

- L121-122: were the bacterial targets experimentally validated or are they predicted?

Thank you for your question. All bacterial targets were experimentally obtained, meaning the phages were isolated using specific bacteria as hosts. To clarify this process, we have changed the term "95% (99/104) target" to "95% (99/104) of the isolated phages target", which better reflects our action of isolating the phages.

Revision (Manuscript, Results, L131): "Notably, 95% (99/104) **of the isolated phages target** disease-associated bacteria, …"

- L156: I would replace "phage-bacteria interactions" by "bacterial infection".

Thank you for your suggestions. We have directly modified the original text.

Revision (Manuscript, Results, L167): "…, which are essential for **bacterial infection**."

- Fig 2B: what do the dashed black lines refer to?

Thank you for your question. The two dashed black lines correspond to phage genome sizes of 20 kb and 200 kb, respectively, highlighting the broad genome size range of the GPB phages. These sizes align with the reported genome ranges for small-genome phages (*Salasmaviridae* phage, 12-20 kb[1]) and large-genome phages (jumbo phage, >200 kb[2]; Lak megaphages, 203-735 kb[3]).

To help people better understand the figure, we have added the following content to the legend of Figure 2B:

Revision (Figure2B, legend): "…**The dashed black lines at 20 kb and 200 kb highlight the broad genome size range of the GPB phages**."

- Fig 2D: the end of the figure legend doesn't seem to match the panel. Also, the figure legend could maybe be expanded to highlight the relevance of the figure?

Thank you for your suggestions. We expanded the legend of Figure 2D to match its relevance to the figure better.

Revision (Figure 2D, legend): "**The genome lengths of each GPB phage and the sizes of all genes within each genome. The pink lines show the genome lengths of each GPB phage, arranged from largest to smallest from left to right. The box colors indicate genome lengths of corresponding phages, with** orange represent**ing** phages with genome lengths over 80 Kb, blue represent**ing those between** 20 **and** 80 Kb, and yellow represent**ing those** less than 20 Kb. **Box lengths represent the interquartile range (IQR) of the data, and whiskers extend to the lowest and highest values within 1.5 times the IQR from the first and third quartiles, respectively. Scatter plots** represent **genes whose sizes exceed 1.5 times the range.** …"

- Fig. 3B: don't the rows represent phages and the columns the bacteria? Also, remove the text in the legend referring to infection under anoxic infections only, as this condition doesn't exist for these phages and bacteria.

Thanks for the careful read. We have corrected the errors in the wording of columns and rows, and the same corrections have been made in Figure 3F and H. Additionally, we have removed the text about "anoxic infections."

Revision (Figure 3B, F and H, legends): "The **columns** represent bacteria strains …. The **rows** represent phages…"

"The **columns** represent bacteria strain, and the **rows** represent phages…"

- L982: "grey lumps" is not adequate. The linkages are shown in grey shading.

Thank you for your suggestions. We have directly modified the original text.

Revision (Figure3C, legend): "… were linked with grey **shading**"

- Fig. 3F and 3H: same comment about rows and columns as in panel B.

We have revised the original text accordingly as mentioned above.

Reference

1 Shen, J. *et al.* Large-scale phage cultivation for commensal human gut bacteria. *Cell Host Microbe* **31**, 665-677.e667 (2023). https://doi.org:10.1016/j.chom.2023.03.013

2 Al-Shayeb, B. *et al.* Clades of huge phages from across Earth's ecosystems. *Nature* **578**,

425-431 (2020). https://doi.org:10.1038/s41586-020-2007-4

3 Cook, R. *et al.* Decoding huge phage diversity: a taxonomic classification of Lak megaphages. *J Gen Virol* **105** (2024). https://doi.org:10.1099/jgv.0.001997