

RESEARCH ARTICLE

Open Access



Heterogeneity of soil bacterial and bacteriophage communities in three rice agroecosystems and potential impacts of bacteriophage on nutrient cycling

Yajiao Wang^{1,2}, Yu Liu¹, Yuxing Wu², Nan Wu¹, Wenwen Liu¹ and Xifeng Wang^{1*}

Abstract

Background: As genetic entities infecting and replicating only in bacteria, bacteriophages can regulate the community structure and functions of their host bacteria. The ecological roles of bacteriophages in aquatic and forest environments have been widely explored, but those in agroecosystems remains limited. Here, we used metagenomic sequencing to analyze the diversity and interactions of bacteriophages and their host bacteria in soils from three typical rice agroecosystems in China: double cropping in Guangzhou, southern China, rice–wheat rotation cropping in Nanjing, eastern China and early maturing single cropping in Jiamusi, northeastern China. *Enterobacter* phage-NJ was isolated and its functions on soil nitrogen cycling and effect on soil bacterial community structure were verified in pot inoculation experiments and 16S rRNA gene sequencing.

Results: Soil bacterial and viral diversity and predicted functions varied among the three agroecosystems. Genes detected in communities from the three agroecosystems were associated with typical functions: soil bacteria in Jiamusi were significantly enriched in genes related to carbohydrate metabolism, in Nanjing with xenobiotics biodegradation and metabolism, and in Guangzhou with virulence factors and scarce in secondary metabolite biosynthesis, which might lead to a significant occurrence of rice bacterial diseases. The virus community structure varies significantly among the three ecosystems, only 13.39% of the total viral species were shared by the three rice agroecosystems, 59.56% of the viral species were specific to one agroecosystem. Notably, over-represented auxiliary carbohydrate-active enzyme (CAZyme) genes were identified in the viruses, which might assist host bacteria in metabolizing carbon, and 67.43% of these genes were present in Jiamusi. In bacteriophage isolation and inoculation experiments, *Enterobacter* bacteriophage-NJ reduced the nitrogen fixation capacity of soil by lysing N-fixing host bacteria and changed the soil bacterial diversity and community structure.

Conclusion: Our results showed that diversity and function predicted of paddy soil bacteria and viruses varied in the three agroecosystems. Soil bacteriophages can affect nutrient cycling by boosting host metabolism through the carried auxiliary metabolic genes (AMGs) and lysing the host bacteria that are involved in biogeochemical cycles. These findings form a basis for better understanding bacterial and bacteriophage diversity in different rice agroecosystems,

*Correspondence: wangxifeng@caas.cn

¹ State Key Laboratory for Biology of Plant Diseases and Insect Pests, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing 100193, China

Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

laying a solid foundation for further studies of soil microbial communities that support ecofriendly production of healthy rice.

Keywords: Paddy soil, Bacteriophage, Bacteriophage–bacteria interaction, Auxiliary metabolic genes, Nitrogen cycling

Background

Rice, the primary staple food for about 60% of China's population [1, 2] and planted in about 29.5 million ha in China, accounts for 35% of the total area planted in grain crops [3]. Rice-growing areas in China are quite diverse and distributed throughout most of the country and differ in ecological environment, cultivation methods and production levels. The three most important rice agroecosystems are rice double cropping in southern China, which has a subtropical and marginal tropical monsoon climate, rice rotation with other crops in central and eastern China, which has a hot and rainy summer but cold and less raining winter and early-maturing single cropping in northeastern China, which has a cold, long winter and warm, short summer. Different ecological environments and farming methods lead to different soil physicochemical properties and soil microbial diversity [4–6].

The hundreds of millions of soil microbes in soils include bacteria, fungi, archaea, viruses, that compete for nutrients and living space in the soil. Determining the physicochemical properties of soils and their effects on the soil microbial composition and diversity in different regions can help in utilizing local natural resources effectively and improving cultivation techniques to safely produce higher-yielding rice. Soil microbes not only breakdown organic matter and minerals, they also secrete secondary metabolites that can regulate plant growth [7–9]. Microbes in soil can cause disease or help protect against disease [10, 11]. Soil microbial communities and functions have strong spatial variability [12]. Bahram et al. (2018) found contrasting patterns in bacterial diversity in top soils across a latitudinal gradient around the world, with the highest diversity in temperate habitats, and environmental variables caused greater variation in microbial gene diversity than geographic distance did [13]. Soil microbial diversity and multifunctionality (decomposition and nutrient cycling, plant productivity) were demonstrated to be spatially specific and positively correlated in 78 global drylands and 179 locations across Scotland [14]. Elucidating spatial and temporal patterns of soil microbial diversity, community composition, and function will thus contribute to a better understanding of the generation and maintenance of microbial diversity, which will help to harness beneficial microbial

resources for full use of their ecological functions in promoting healthy growth of plants in environmentally safe ecosystems.

Bacteriophages, which can infect and replicate within particular bacterial hosts, are important components of soil microbial community [15]. Bacteriophages can regulate the community structure and function of bacteria and even the whole soil microbes [16], modulate host bacterial populations and diversities by lysing bacterial cells [17], drive host evolution via bacteriophage-mediated horizontal gene transfer [18], and alter host bacterial function by reprogramming their metabolism via the expression of virus-encoded auxiliary metabolic genes (AMGs) [19]. AMGs have a variety of metabolic functions, including carbon, nitrogen and sulfur metabolism [20, 21], and photosynthesis [22]. However, most studies on the function of bacteriophages have focused on marine ecosystems or the human intestinal tract rather than soils in agroecosystems.

Our knowledge about paddy soil microbial diversity and function in different agroecosystems is limited, and even less is known about viral communities and their roles. Studies of viruses in glacier soil [23], mangrove soil [24], mud volcanic soil [25], and Antarctic soil [26] revealed differences in soil viral community structure and largely uncharacterized viral assemblages. In the present study, we used metagenomic sequencing to analyze bacterial and viral community composition, diversity and ecological functions in paddy bulk soils in the three most important rice agroecosystems, double cropping in southern China, rice–wheat rotation cropping in central China, and early-maturing single cropping in northeastern China. We also tested the effects of bacteriophages on soil bacterial community structure and function in amended soils in pot experiments.

Results

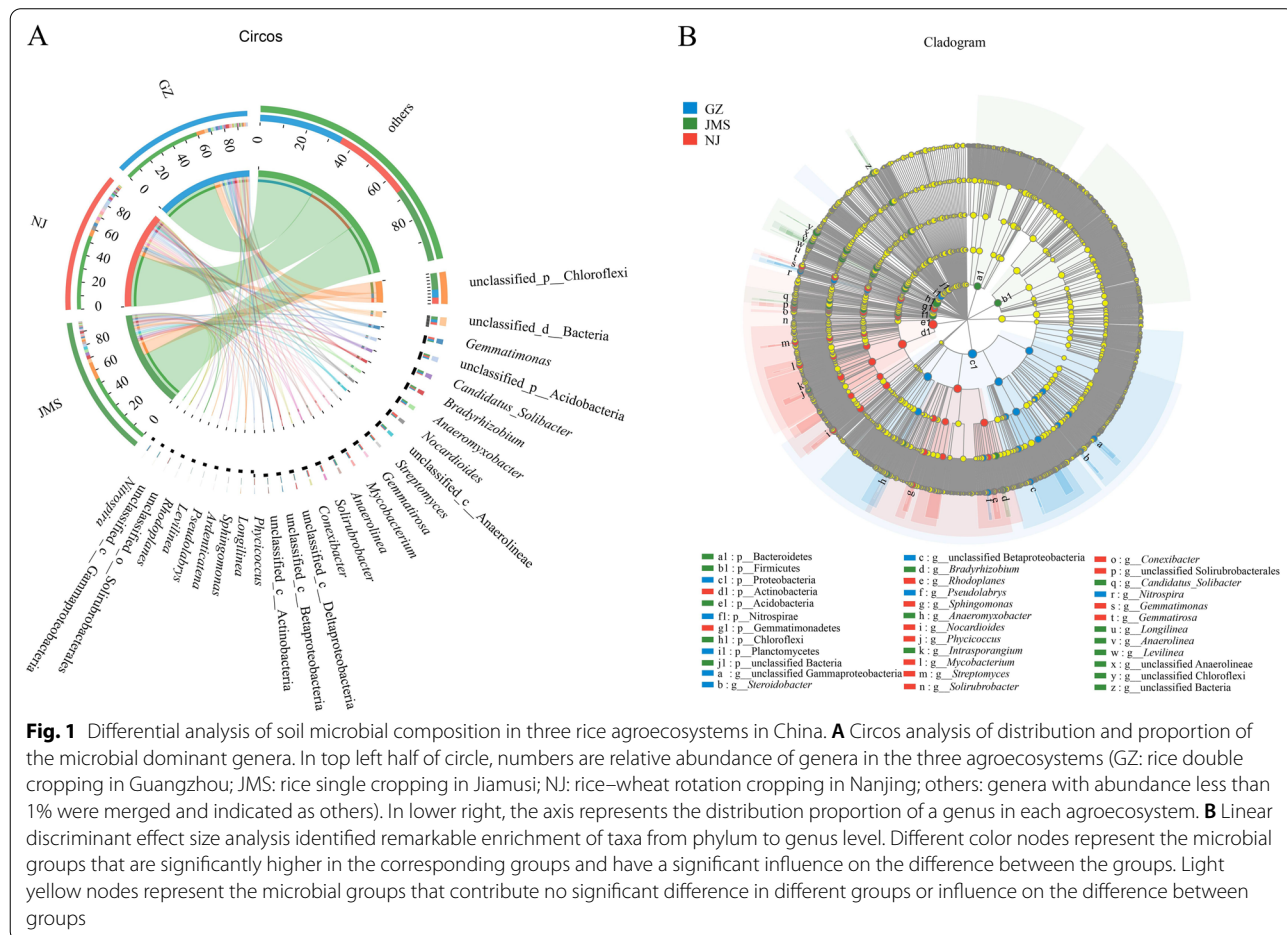
Soil microbial communities and differences in three rice agroecosystems

In the metagenomic sequencing, 1.10 billion raw reads were obtained. After trimming, 1.07 billion clean reads remained. In Guangzhou, Nanjing and Jiamusi, 511,481, 593,018, 848,794 contigs were obtained, respectively, after clean reads were assembled. Based on similarities to entries in the NCBI-NR database, 97.81% of sequences were classified as bacteria, 1.73% as archaea,

0.27% as eukaryota, and 0.026% as viruses. In total, 82 phyla, 1,904 genera and 12,455 microbial species were identified. At the phylum level, Proteobacteria, Actinobacteria, Chloroflexi, Acidobacteria and Bacteroidetes were the top five most-abundant phyla, accounting for more than 78% of all the sequences (Additional file 1: Fig. S1A). According to the LEfSe analysis ($LDA > 3.5$, $p < 0.05$), soils in Guangzhou and Nanjing were enriched in microbes belonging to Proteobacteria and Actinobacteria, respectively, while soils in Jiamusi were enriched in bacteria belonging to Chloroflexi, Acidobacteria and Bacteroidetes (Fig. 1A). At the generic level, there were significant differences in the dominant genera (with abundance greater than 1% and without unclassified genera). There were 10 dominant genera in Guangzhou, 13 in Nanjing and 9 in Jiamusi, only 20% of the total dominant genera were shared by the three rice agroecosystems, and 45% were unique to a single rice agroecosystem (Fig. 1B, Additional file 1: Table S3). The top three most-dominant genera were *Gemmatimonas*, *Anaeromyxobacter* and *Candidatus Solibacter*

in Guangzhou, *Nocardioides*, *Gemmatimonas* and *Solirubrobacter* in Nanjing, *Bradyrhizobium*, *Candidatus Solibacter* and *Anaeromyxobacter* in Jiamusi (Additional file 1: Fig. S1B). It is worth noting that none was shared by the three rice agroecosystems. Our results that the microbial community composition differed among the three rice agroecosystems was confirmed by PCoA analysis. The principal coordinate axis 1 (PCo1) and the principal coordinate axis 2 (PCo2) explained, respectively, 26.66% and 19.22% of the variation in microbial species (Additional file 1: Fig. S2).

At the species level, we analyzed the relative abundance of four main pathogenic bacteria: *Pseudomonas syringae*, *Acidovorax avenae*, *Xanthomonas oryzae* and *Dickeya zeae*, the results showed that the relative abundance of the four pathogens was the highest in Guangzhou (Additional file 1: Fig. S3A). Similarly, in the field disease evaluation, we found that the disease indices for rice brown spot, brown streak, bacterial blight, bacterial leaf streak and foot rot were the highest in Guangzhou (Additional file 1: Fig. S3B).



Complicated relationship between bacteriophage and bacterial abundance and diversity

Viral and bacterial genes were selected out to analyze their relationship. Metagenomic sequencing only detected viruses that were relatively abundant. In total, 14 families, 29 genera and 368 species of viruses were identified. As in the most-often characterized viromes in various habitats, the annotated viruses in three rice agroecosystems belonged to families Siphoviridae, Podoviridae, Myoviridae, Phycodnaviridae and Herpesviridae. The relative abundance of Phycodnaviridae was highest in Jiamusi, and Siphoviridae, Podoviridae, Myoviridae and Herpesviridae were highest in Nanjing (Fig. 2A). Similar with the trend for the number of bacterial species, the number of virus species was highest in Nanjing (218 virus species, 11,693 bacterial species), followed by Guangzhou (192 virus species, 11,506 bacterial species) and Jiamusi (155 virus species, 11,424 bacterial species). The three rice agroecosystems shared 13.39% of the total virus species, and 27.05% of the total virus species were shared by two agroecosystems; however, 59.56% of the virus species were specific to one agroecosystem.

Bacteriophages that infect the same bacterial genus were grouped together. We analyzed the correlation between the relative abundance of dominant bacteriophage groups (relative abundance greater than 1%) and their host bacterial genera (Table 1). There were 19 dominant bacteriophage groups, the relative abundance of the hosts for three bacteriophage groups was greater than 1%, less than 0.01% for four bacteriophage groups, and between 0.01 and 1% for 12 bacteriophage groups. Thus, the dominant bacteriophages were mainly distributed

in bacteria with moderate abundance. The relationship between the relative abundance of the bacteriophages and that of their hosts is complex. The ratio of viral relative abundance to bacterial relative abundance (VBR) varies widely, between 0 and 27,382. But when VBR was greater than 100 in at least one of the three cropping systems, the abundance of bacteriophage and host bacteria was negatively correlated ($\rho < -0.5$).

Metagenomic analysis of richness (represented by the Chao index) and diversity (represented by the Shannon index) revealed a significant difference in the bacterial and bacteriophage communities of paddy soils in different cropping systems. The Chao index showed that the bacterial richness in Nanjing was significantly higher than in Jiamusi and Guangzhou, while the difference between Jiamusi and Guangzhou was not significant. The viral richness was positively correlated with that of bacteria ($\rho = 0.662$, $p = 0.019$) (Fig. 3A), and the viral richness of Nanjing was significantly higher than that of Jiamusi and Guangzhou. The Shannon index analysis found that the bacterial diversity of Nanjing and Guangzhou did not differ significantly, but both were significantly higher than in Jiamusi. Viral diversity was negatively correlated with that of bacteria ($\rho = -0.876$, $p = 0.001$) (Fig. 3B).

Soil bacterial function and its relationship with environment in three agroecosystems

In the functional annotation of the soil metagenomic reads for the bacterial community using the Kyoto Encyclopedia of Genes and Genomes (KEGG) database, most sequences were associated with metabolism (70.24–72.38%), and far fewer were involved in genetic

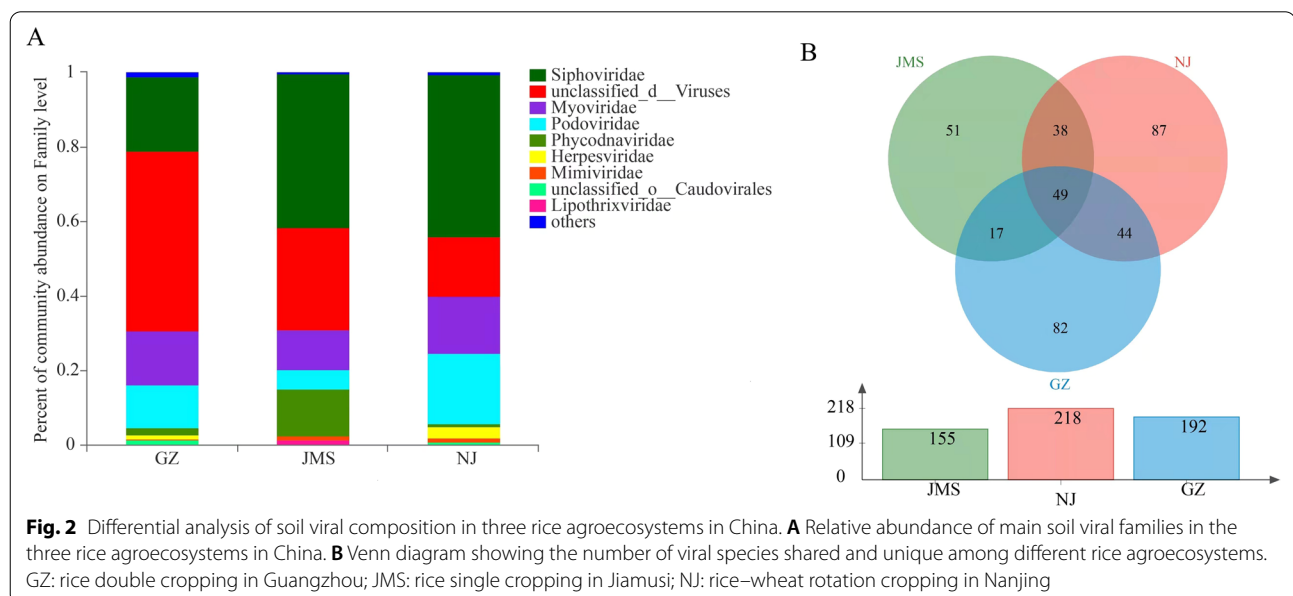
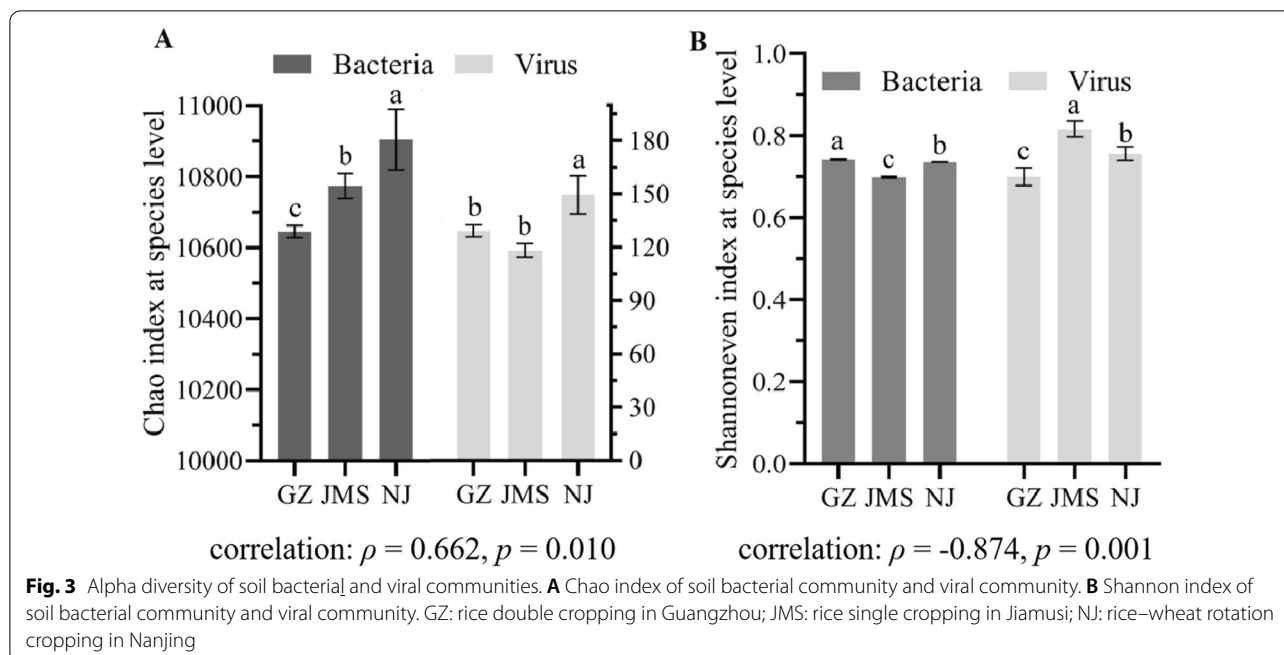


Fig. 2 Differential analysis of soil viral composition in three rice agroecosystems in China. **A** Relative abundance of main soil viral families in the three rice agroecosystems in China. **B** Venn diagram showing the number of viral species shared and unique among different rice agroecosystems. GZ: rice double cropping in Guangzhou; JMS: rice single cropping in Jiamusi; NJ: rice-wheat rotation cropping in Nanjing

Table 1 Correlation between the relative abundance of dominant bacteriophage groups and the host bacterial genera in the three rice agroecosystems studied in China

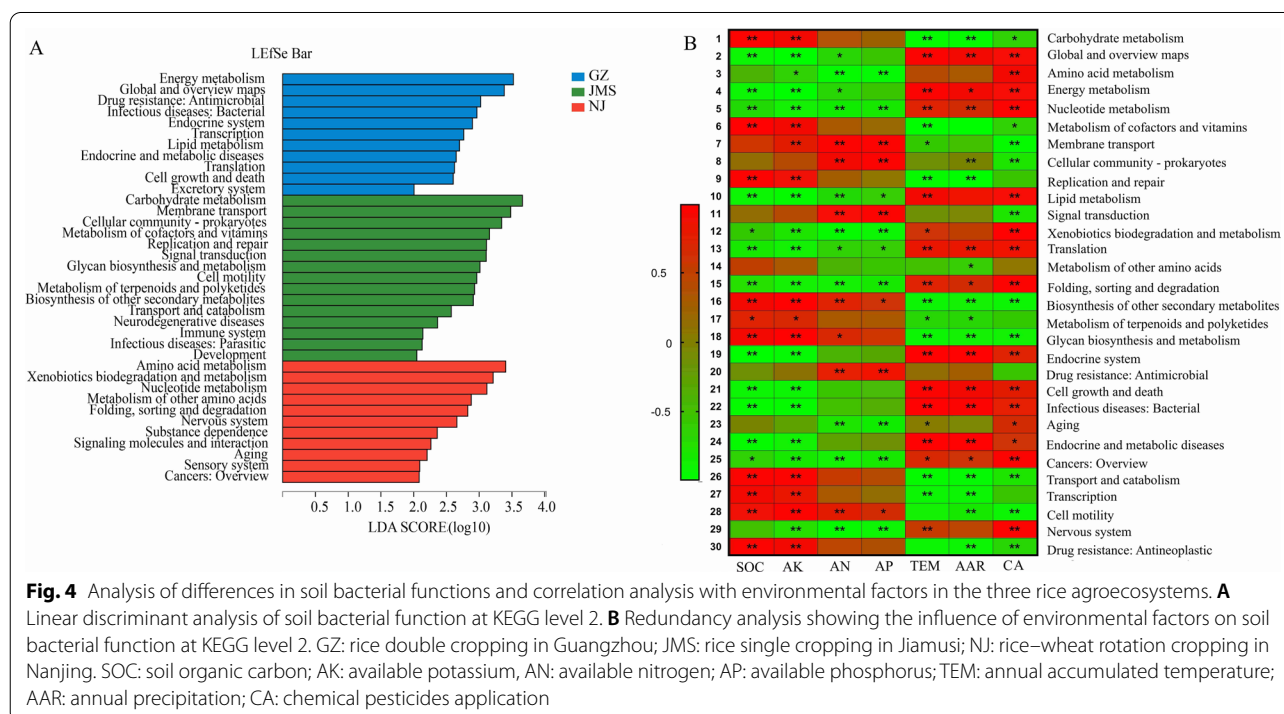
Phages	NJ phage	NJ bacteria	NJ VBR	GZ phage	GZ bacteria	GZ VBR	JMS phage	JMS bacteria	JMS VBR	ρ	P
<i>Achromobacter</i> phage	1.247 ± 0.512	0.845 ± 0.035	1.476	0.252 ± 0.057	1.370 ± 0.039	0.184	0	1.854 ± 0.073	0	-0.855	0.001
<i>Vibrio</i> phage	4.007 ± 1.022	0.024 ± 0.001	164.991	3.067 ± 0.668	0.031 ± 0.0003	98.523	0.254 ± 0.149	0.036 ± 0.001	7.1	-0.856	0.001
<i>Gordonia</i> phage	3.759 ± 1.168	0.080 ± 0.003	47.181	0.507 ± 0.577	0.058 ± 0.002	8.705	20.029 ± 2.19	0.043 ± 0.001	469.223	-0.708	0.01
<i>Streptomyces</i> phage	0.865 ± 0.365	1.977 ± 0.056	0.438	2.955 ± 0.823	1.127 ± 0.059	2.622	2.915 ± 1.688	1.136 ± 0.066	2.566	-0.704	0.011
<i>Enterobacter</i> phage	15.889 ± 0.484	0.005 ± 0.0006	3050.137	0.128 ± 0.101	0.008 ± 0.0007	15.398	0	0.007 ± 0.001	0	-0.684	0.014
<i>Stenotrophomonas</i> phage	7.319 ± 1.331	0.040 ± 0.001	184.654	3.905 ± 1.262	0.045 ± 0.001	87.197	0	0.044 ± 0.0005	0	-0.591	0.043
<i>Salmonella</i> phage	5.390 ± 0.478	0.002 ± 0.0003	2515.275	0.739 ± 0.326	0.003 ± 0.0005	252.934	0	0.003 ± 0.0003	0	-0.572	0.049
<i>Paramoecium bursaria</i> phage	0.401 ± 0.276	0.0009 ± 0.0002	469.215	1.542 ± 0.206	0.0003 ± 0.000007	5020.489	10.980 ± 1.297	0.0004 ± 0.000005	27,382.718	-0.509	0.091
<i>Rhodococcus</i> phage	1.827 ± 1.079	0.393 ± 0.012	4.652	3.983 ± 2.79	0.196 ± 0.009	20.372	3.930 ± 1.682	0.181 ± 0.012	21.693	-0.357	0.254
<i>Burkholderia</i> phage	7.139 ± 1.112	0.226 ± 0.003	31.647	3.873 ± 0.770	0.347 ± 0.005	11.153	0.232 ± 0.186	0.232 ± 0.002	1.002	-0.018	0.955
<i>Caulobacter</i> phage	0.374 ± 0.278	0.087 ± 0.004	4.293	0.972 ± 0.531	0.099 ± 0.002	9.751	0.701 ± 0.201	0.058 ± 0.002	12.128	0.204	0.525
<i>Mycobacterium</i> phage	8.404 ± 7.98	2.405 ± 0.133	3.495	5.269 ± 0.545	1.278 ± 0.087	4.122	5.679 ± 1.066	0.438 ± 0.027	12.961	0.263	0.409
<i>Bacillus</i> phage	2.181 ± 0.428	0.214 ± 0.002	10.170	1.106 ± 0.518	0.240 ± 0.002	4.609	2.670 ± 1.117	0.282 ± 0.005	9.452	0.332	0.291
<i>Synechococcus</i> phage	1.318 ± 0.377	0.059 ± 0.0005	22.252	2.990 ± 0.831	0.082 ± 0.001	36.361	3.436 ± 1.065	0.063 ± 0.001	54.548	0.343	0.275
<i>Pseudomonas</i> phage	6.964 ± 2.194	0.247 ± 0.013	28.207	5.904 ± 1.422	0.334 ± 0.007	17.651	2.486 ± 1.049	0.207 ± 0.002	11.984	0.477	0.116
<i>Ralstonia</i> phage	0.065 ± 0.089	0.063 ± 0.002	1.029	1.046 ± 0.516	0.117 ± 0.0008	8.933	0.501 ± 0.102	0.076 ± 0.001	6.636	0.563	0.046
<i>Rhodiferax</i> phage	0	0.027 ± 0.001	0	1.392 ± 0.386	0.054 ± 0.004	25.709	0	0.044 ± 0.0007	0	0.778	0.003
<i>Actinoplanes</i> phage	1.025 ± 0.482	0.298 ± 0.007	3.446	0.045 ± 0.089	0.147 ± 0.009	0.303	0.523 ± 0.295	0.173 ± 0.011	3.031	0.781	0.003
<i>Arthrobacter</i> phage	4.391 ± 0.892	0.400 ± 0.015	10.969	2.286 ± 0.579	0.118 ± 0.004	19.392	16.613 ± 1.227	0.547 ± 0.079	30.394	0.837	0.001

GZ: rice double cropping system in Guangzhou; JMS: rice single cropping system in Jiamusi; NJ: rice-wheat rotation cropping ecosystem in Nanjing; VBR: the ratio of viral relative abundance to bacterial relative abundance



information processing (8.44–8.93%), environmental information processing (6.79–7.26%), cellular processes (5.49–6.01%), human diseases (3.93–4.57%), and organismal systems (2.71%–2.85%). According to the LEfSe analysis (Fig. 4A), soil bacteria in Jiamusi were significantly enriched in genes associated with carbohydrate

metabolism, replication and repair, biosynthesis of other secondary metabolites which were positively correlated with a high soil organic carbon (SOC) content (Fig. 4B; Additional file 1: Table S2). While soil bacteria in Nanjing were significantly enriched in genes associated with xenobiotics biodegradation and metabolism, nucleotide

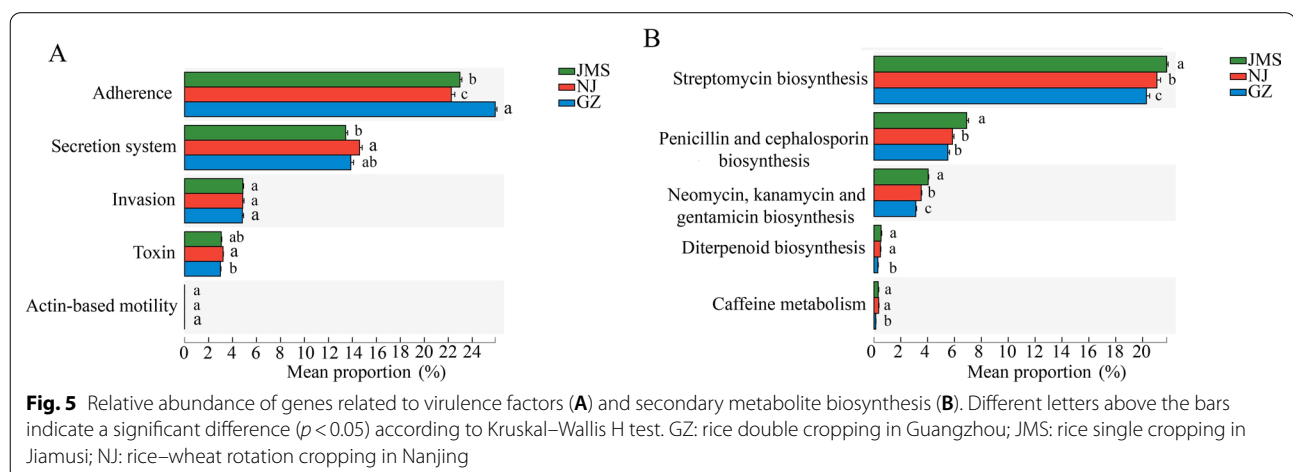


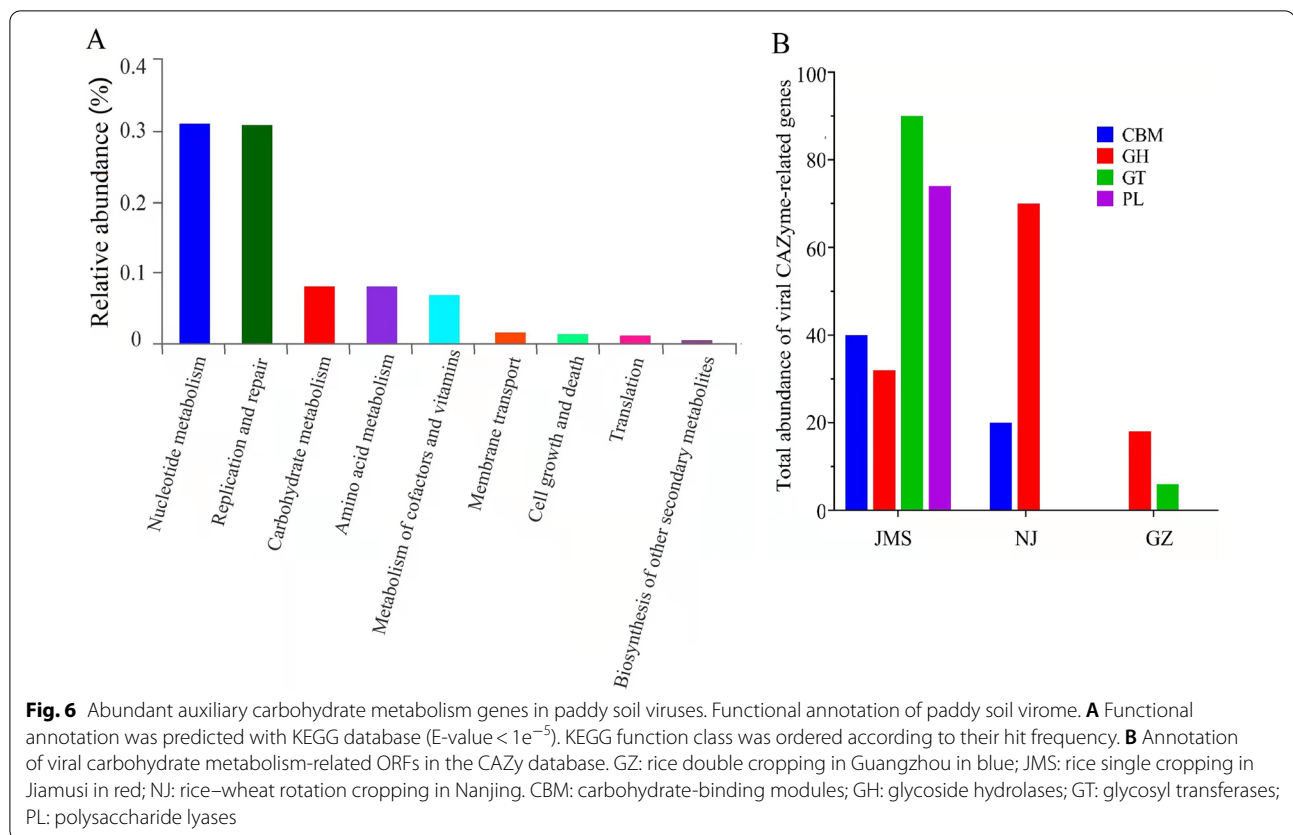
metabolism, folding sorting and degradation, especially for genes involved in biodegradation and metabolism of nitrotoluene, aminobenzoate, naphthalene, steroid and xylene (Additional file 1: Fig. S4). The annual frequency and dosage of chemical pesticide application especially herbicides in Nanjing were higher than those in other two areas (Additional file 1: Table S2). As a result, pesticide residues in the soil were relatively high, and soils were significantly enriched in bacteria with xenobiotics biodegradation genes, presumably to help decompose pesticides in the soil.

Soil bacteria in Guangzhou were significantly enriched in genes associated with bacterial virulence, energy metabolism, translation, cell growth and death, which were positively correlated with annual accumulated temperature and annual precipitation. In addition, the analysis of the virulence factor database VFDB showed that the abundance of genes related to factors (adherence-related genes) in the Guangzhou region was significantly higher than that in Nanjing and Jiamusi, especially for genes for flagella, Trw type IV secretion system, fimbriae, Myf/pH6 antigen, and PI-2, accounting for more than 70% of the total adherence-related genes (Fig. 5C). Functional analysis of biosynthesis of other secondary metabolites showed that the genes related to antimicrobial antibiotic secretion (such as biosynthesis of streptomycin, penicillin and cephalosporin, and neomycin, kanamycin and gentamicin) in Guangzhou were significantly lower compared with that of Nanjing and Jiamusi (Fig. 5D). More serious bacterial diseases in Guangzhou may be related to the higher temperature and humidity and lower secretion of antibiotics, which may be conducive to the reproduction of pathogenic bacteria and to the higher relative abundance of adherence-related genes might aid their infection of the plant host.

Abundant auxiliary carbohydrate metabolic genes in paddy viruses

In the functional annotation of viruses in paddy soils using the KEGG database (Fig. 6A), almost all functions in pathway level 2 were found, but the sequences were notably enriched in a limited number of functions; the top three assigned functions accounted for 70% of the annotated sequences. The top two most-abundant functions were nucleotide metabolism, replication and repair, which are critical for the reproduction and survival of viruses. In addition, viruses have genes related to carbohydrate metabolism and the biosynthesis of secondary metabolites; the relative abundances of genes related to carbohydrate metabolism in viruses were negatively correlated ($\rho = -0.498$, $p = 0.334$) with those in soil bacteria (Additional file 1: Fig. S5A). Similarly, the abundances of genes related to biosynthesis of secondary metabolites in viruses were negatively correlated ($\rho = -0.500$, $p = 0.333$) with those in soil bacteria (Additional file 1: Fig. S5B). These two types of genes carried by bacteriophages might assist host bacteria in carbon metabolism and secondary metabolite synthesis. In the annotation of carbohydrate-active enzymes (CAZymes) using the CAZy database (Fig. 6B), the species and abundance of viral CAZymes differed significantly among the soils of the three cropping systems. Notably, 67.43% of the CAZymes-related genes were present in Jiamusi, which was 9.83 and 2.62 times higher than in Guangzhou and Nanjing. In addition, there were two classes CAZymes (glycoside hydrolases, glycosyl transferases) in Guangzhou, two classes (glycoside hydrolases, carbohydrate-binding modules) in Nanjing, but four classes (glycoside hydrolases, carbohydrate-binding modules, glycosyl transferases, polysaccharide hydrolase) in Jiamusi. Remarkably, polysaccharide lyases (pectate lyase, exopolysaccharide lyase, thiopeptidoglycan lyase) were unique to Jiamusi, implying





potential roles in the decomposition of soil organic carbon.

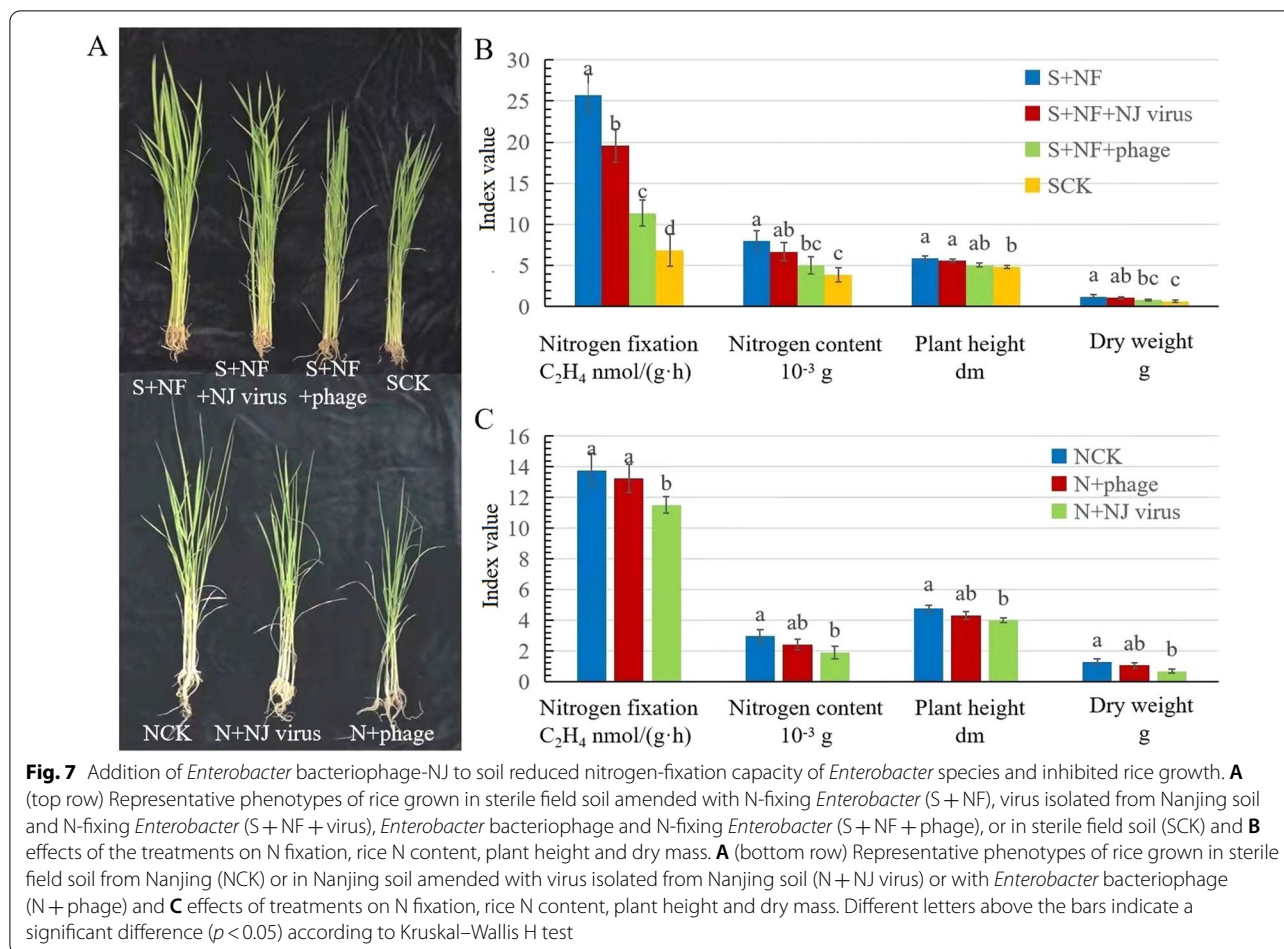
***Enterobacter* bacteriophages can reduce the nitrogen fixation capacity of soil by lysing host bacteria**

In Nanjing, the most abundant bacteriophage was *Enterobacter* phage, with a relative abundance of 14.8%, compared with 0.13% in Guangzhou and 0% in Jiamusi (Additional file 1: Fig. S6A). They were also negatively correlated with the abundance of host bacteria *Enterobacter* ($\rho = -0.805$, $p = 0.404$) (Additional file 1: Fig. S6B). To study its function, the bacteriophage was isolated from the soil in Nanjing and named *Enterobacter* phage-NJ (Additional file 1: Fig. S7A). In the plate lytic experiment, the bacteriophage had lytic activity against nitrogen-fixing *Enterobacter* (Additional file 1: Fig. S7B).

To assess the functional effects of *Enterobacter* phage-NJ on nitrogen fixation and uptake, we added *Enterobacter* phage-NJ to the sterilized soil to which nitrogen-fixing *Enterobacter* (*E. cloacae*, *E. ancerogenus* and *E. ludwigii*) had been added. The results showed that the nitrogen-fixation capacity decreased by 55%, and plant height, fresh mass and nitrogen content of the rice plants were 14.27%, 32.25%, and 37.17%, respectively, lower than in the controls. Similar results were

found when Nanjing soil viruses were added to the sterilized soil, the nitrogen-fixation capacity decreased by 23.72%, and plant height, fresh mass and nitrogen content of the rice plants were 5.45%, 12.90% and 16.27%, respectively (Fig. 7A, B). When *Enterobacter* phage-NJ was added to untreated field soil, the nitrogen-fixation capacity decreased by 16.49%, and plant height, fresh mass, and nitrogen content were 15.88%, 47.66% and 35.47% respectively, lower than in the controls. Similar results were found when Nanjing soil viruses were added to untreated field soil, the nitrogen-fixation capacity decreased by 3.78%, and plant height, fresh mass, and nitrogen content were 9.29%, 17.97%, and 18.58%, respectively (Fig. 7A, C).

Rhizosphere bacterial diversity in untreated field soil after the addition of *Enterobacter* phage-NJ was then studied using 16S rRNA amplicon sequencing. After the addition of *Enterobacter* phage-NJ, the richness and diversity of rice rhizosphere microorganisms was 29.97% and 19.17%, respectively, lower than in the control (Fig. 8A). Moreover, the community structure of rhizosphere microorganisms significantly changed; PCoA demonstrated separate clustering (Fig. 8B). In addition, the metabolic function of rhizosphere microorganisms also significantly decreased, especially carbohydrate



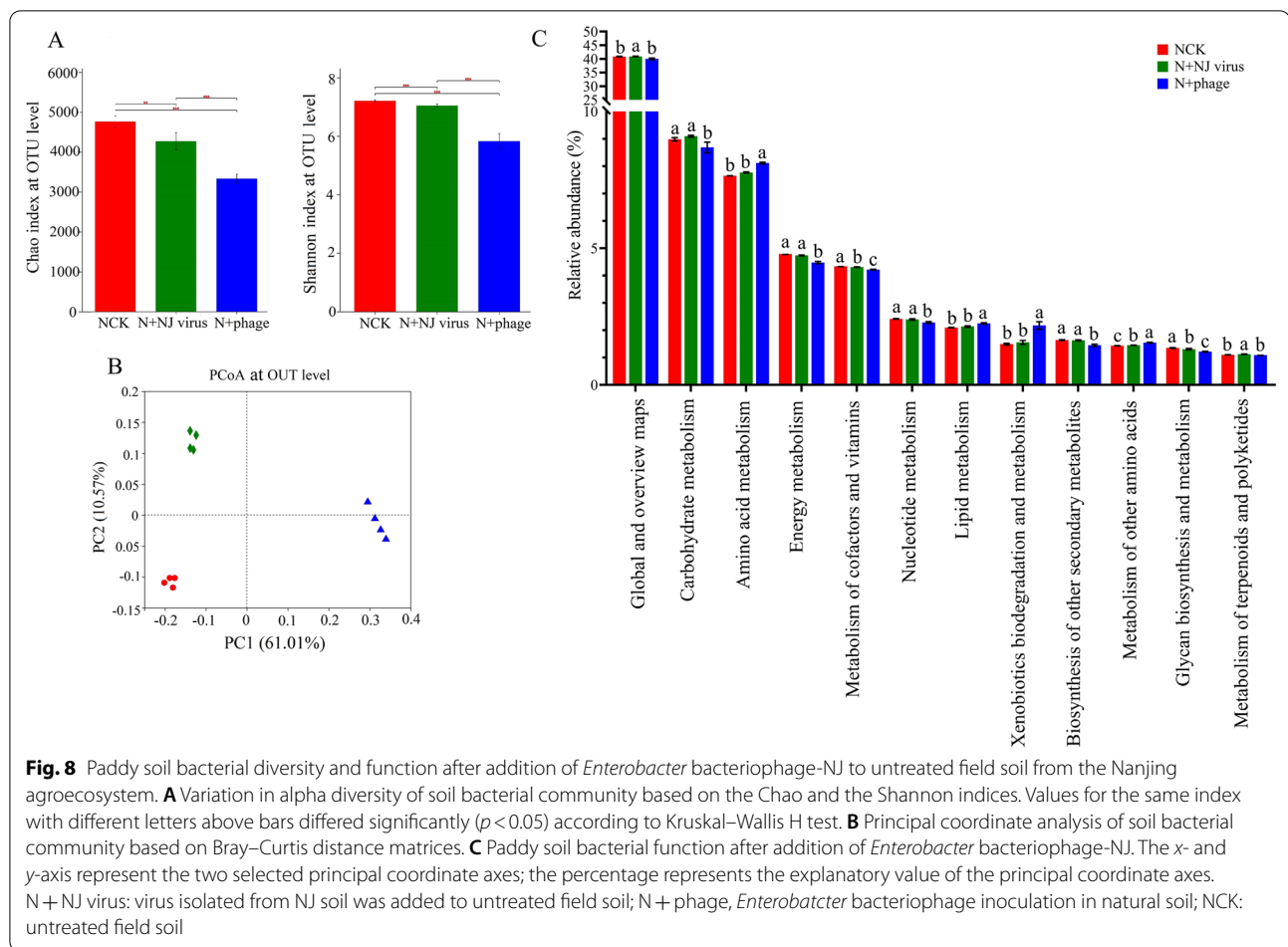
metabolism, energy metabolism and biosynthesis of secondary metabolites (Fig. 8C).

Discussion

Functional differences of soil bacteria in the three rice agroecosystems

Soil bacteria in three rice agroecosystem performed multiple functions simultaneously and were significantly enriched in typical functions that were specific for an agroecosystem. Soil bacteria in Jiamusi and Nanjing were significantly enriched in genes associated with carbohydrate metabolism, xenobiotics biodegradation and metabolism, respectively, while soil bacteria in Guangzhou were enriched in genes associated with virulence factors, but scarce in genes associated with secondary metabolite biosynthesis, which might lead to the significant occurrence of rice bacterial diseases. These significant differences are likely related to the local environment and cropping system. The effects of environment on soil bacterial function can be divided into direct and indirect effects. Environmental factors especially temperature can directly influence the functions by accelerating

the activity microbes [27, 28] or indirectly affect functions by altering the composition of microbial communities [29]. Soils in Jiamusi were significantly enriched in genes associated with carbohydrate metabolism, which were positively correlated with the high content of soil organic matter. Soil organic matter can provide nutrients and promote multiplication of microbes, ultimately, indirectly promoting the capacity of carbon metabolism [11]. With lower inputs of plant-derived organic carbon, desert communities have lower relative abundances of genes associated with nutrient cycling and the catabolism of plant-derived organic compounds compared with forests, grasslands, and tundra [30]. Our correlation analysis of environment and function showed that soil bacteria in Nanjing were significantly enriched in genes associated with xenobiotics biodegradation and metabolism, which were positively correlated with the large number of pesticide applications. Because of the rice–wheat rotation system in Nanjing, more types and applications of chemicals are used than in the double cropping system in Guangzhou and in the single cropping rice in Jiamusi. Chemical residues are beneficial to the survival



of pesticide-degrading microbes and ultimately result in significant enrichment of the xenobiotics biodegradation and metabolism functions of soil bacteria in Nanjing. The genes associated with bacterial virulence factors in Guangzhou soils were positively correlated with higher annual accumulated temperature and annual precipitation. One of the most important conditions for rice bacterial diseases is high temperature and humidity [31]. In addition, double cropping rice in Guangzhou provides two growing seasons for hosts of bacterial pathogens, so pathogens have more time and opportunity to survive than Nanjing and Jiamusi. These results indicate that functions of soil bacteria vary in different agroecosystems and that environmental factors play important roles through affect microbial community composition.

Differences in soil viruses among the three agroecosystems and their relationship with host bacteria

In our study, there were significant differences in the species and composition of paddy soil viruses among the three agroecosystems. In Nanjing, 40.64% more virus species were detected than in Jiamusi and 13.54% more

than in Guangzhou. A previous study showed that the bacterial host species and abundance may be a key factor controlling viral abundance and distribution [32]. Similarly, we found that soil viral richness was significantly positively correlated with bacterial richness ($\rho = 0.662$, $p = 0.01$). Only 13.39% of the viral species were shared by all three rice agroecosystems, while 59.56% of the viral species were environmentally specific, which is likely associated with the significant differences in temperature, annual precipitation and soil physicochemical properties in the three rice agroecosystems because viral diversity and community structure have been reported to be controlled by factors such as temperature [33], soil moisture [34], and pH [35].

In soils, most viruses are bacteriophages, which can be divided into lysogenic or lytic bacteriophages [15]. The abundance and distribution of bacteriophages and host bacteria influence each other in a complex relationship based on our results. Bacteriophages are obligate intracellular parasites, so their fitness and range limits are influenced by their hosts, and the bacteriophages in turn can affect host abundance and even the community

structure and diversity of the whole bacterial community [36–39]. Lytic bacteriophages can modulate host bacterial populations and diversities by lysing their hosts [40], but lysogenic bacteriophages can increase the host abundance by promoting horizontal gene transfer by lysogenic conversion or transferring packaged bacterial genes to new hosts, which may expand the host metabolic profile and enhance microbial environmental adaptability [41]. In our study, we analyzed the relationships between viruses and their potential hosts. When the ratio of viral abundance to bacterial abundance was greater than 100, there was a significant negative correlation between the abundance of bacteriophage and bacteria, indicating that bacteriophages may be lytic, which negatively regulates the abundance of host bacteria. However, when the ratio was less than 100, the correlation between bacteriophage and bacteria abundance was not significant, perhaps because the bacteriophages may be lysogenic, which have little effect on host bacteria abundance, and/or because bacteriophages overlap in their distribution with multiple susceptible hosts in natural environment, while bacteria can often be infected with several different types of bacteriophages. Furthermore, many factors influence the abundance of bacteriophages and host bacteria, making the relationship between them more complex.

Bacteriophages inhibit soil nitrogen fixation by lysing host bacteria

Lytic bacteriophages have direct and indirect effects on nutrient cycling. They can directly lyse bacteria, which releases organic matter, including nutrients such as carbon, nitrogen and phosphorus. The carbon cycle driven by viruses accounts for 25% of the total carbon cycle in marine agroecosystems [16]. Addition of bacteriophages to soil also increases the NH_4^+ concentration via lysis, resulting in a release of inorganic nitrogen, followed by mineralization [42]. Indirectly, bacteriophage lysis can also slow nutrient transformations by reducing the population of key bacteria involved in nutrient cycling [42]. Higher abundance of T4-like phages increases bacterial death, thereby suppressing soil organic carbon mineralization [43]. In wetlands, bacteriophages can infect sulfate-reducing and methanogenic bacteria, which can potentially repress sulfate reduction (and associated carbon mineralization) and methane production, respectively [44]. However, indirect inhibition of biogeochemical cycling by bacteriophages has mostly been inferred by sequencing analysis without relevant experimental verification. In our study, *Enterobacter* phage-NJ, which was isolated from the soil in Nanjing. It can lyse three species *Enterobacter* bacteria which we could get and were proven to have nitrogen-fixing ability by Agricultural Culture Collection of China. When *Enterobacter*

phages-NJ were added to untreated field soil or sterilized soil with nitrogen-fixing bacteria, the soil nitrogen-fixation capacity decreased, and rice growth and nitrogen content decreased also. This result provides direct evidence that bacteriophages can inhibit biogeochemical cycling through lysis. *Enterobacter* phage was the most abundant in the tested soils from Nanjing, with a relative abundance of 14.8%. When we extracted all viruses from the Nanjing soils and added them into natural field soil, the result was similar to adding bacteriophages alone; that is, soil nitrogen-fixation capacity was reduced. In addition, the soil bacterial community changed, and the bacterial metabolic functions such carbon metabolism and energy metabolism decreased also. These results indicate that bacteriophages can regulate soil bacterial community structure and ecological function and that the differences in soil bacterial function were related to differences in the bacteriophages in the different agroecosystems.

Potential effects of auxiliary metabolic genes of phages on soil microbial function

Bacteriophages can affect nutrient cycling not only by lysing the host but also by metabolically reprogramming their hosts via the horizontal transfer of ecologically important genes and the expression of virus-carried auxiliary metabolic genes (AMGs) [45]. AMGs have been most extensively explored in marine ecosystem and are involved in photosynthesis, carbon metabolism, nitrogen metabolism and phosphate uptake [46, 47]. Those genes are hypothesized to increase viral survival and replication by augmenting key steps in host metabolism. In mangrove soil [24], permafrost soil [48], and maize–barley rotation soil [20], AMGs involved in carbohydrate metabolism are notably over-represented. In the maize–barley rotation soil, three classes of genes—carbohydrate-binding modules, carbohydrate esterases, glycoside hydrolases—were identified. Similarly, in paddy soil, four classes of AMGs of bacteriophages involved in carbon metabolism were identified in our study (glycoside hydrolases, glycosyl transferases, polysaccharide lyases, and carbohydrate-binding modules, with the abundance of glycoside hydrolases the highest), indicating that paddy soil viruses primarily participate in the decomposition of organic carbon, breakdown of complex carbohydrates to increase energy production and boost bacteriophage replication. The species and abundance of AMGs also differed among the three agroecosystems we studied. Compared with the bacteriophages in the other two agroecosystems, the bacteriophages in Jiamusi had the greatest diversity of AMGs genes related to carbon metabolism; there were

four classes of CAZymes in Jiamusi (polysaccharide lyases, glycoside hydrolases, glycosyl transferases, carbohydrate-binding modules), but only two in Guangzhou (glycoside hydrolases, glycosyl transferases) and in Nanjing (glycoside hydrolases, carbohydrate-binding modules). However, soil bacteria in Jiamusi were significantly enriched for carbon metabolism function, likely due to the large diversity of bacteriophage AMGs genes related to carbon metabolism. Remarkably, polysaccharide lyases (pectate lyase, exopolysaccharuronate lyase, thiopeptidoglycan lyase) were unique to Jiamusi. During the infection cycle, bacteriophages need to break down the host bacterial surface barriers such as capsule polysaccharides (CPSs), extracellular polysaccharides (EPSs). Polysaccharide lyases carried by bacteriophages can help to break down the barrier and infect the host [49]. In addition, complex carbohydrates or polysaccharides such as cellulose, xylan, pectin, starch, alginate, mannan, and chitin are major components of plant cell walls, which are difficult to degrade. In Jiamusi, the soil organic matter content is the highest, perhaps because the bacteriophages carry genes for polysaccharide lyases, which might assist host bacteria in degrading organic matter in the soil.

Conclusion

In our analyses to identify and compare the bacterial and viral communities in soils from the three most important rice agroecosystems in China, the communities in Jiamusi were significantly enriched in genes related to carbohydrate metabolism, in Nanjing with xenobiotics biodegradation and metabolism, and in Guangzhou with genes associated with virulence factors, but those involved with secondary metabolite biosynthesis were scarce, which might lead to a significant incidence of rice bacterial diseases. The 368 virus species that were detected from the three agroecosystems likely affect the soil microbial community structure and function. Bacteriophage isolation and inoculation experiments showed that *Enterobacter* phage-NJ not only changed the soil bacterial diversity and community structure, but also reduced the nitrogen-fixation capacity by lysing nitrogen-fixing host bacteria. In addition, auxiliary carbohydrate-active enzyme (CAZyme) genes were notably over-represented in soil viruses, and 67.43% of all CAZyme genes were present in Jiamusi soils, which might assist host bacteria in metabolizing carbon. Our study will contribute to a better understanding of the microbial diversity in the typical rice agroecosystems in China, viral mechanisms that may be involved in this diversity and inform management programs for the safe production of healthy rice.

Methods

Sampling sites and sample collection

Paddy soil samples were collected during heading of rice in 2019 from long-term (more than 20 years) rice farms in Jiamusi (46° 48' N, 130° 22' E, 130 m a.s.l.) in Heilongjiang Province, Nanjing (30° 02' N, 118° 46' E, 30 m a.s.l.) in Jiangsu Province, and Guangzhou (23° 13' N, 113.27° E, 871 m a.s.l.) in Guangdong Province in China. In Jiamusi, Nanjing and Guangzhou, the annual precipitation is 510 mm, 1090 mm and 1720 mm, respectively, with 67%, 73%, and 77% mean humidity, respectively, and the annual accumulated temperature of 2700 °C, 5400 °C and 7500 °C, respectively (<http://www.weather.com.cn>). The three regions different in their rice planting systems (single rice cropping in Jiamusi, rice–wheat rotation in Nanjing, double rice cropping in Guangzhou). Rice was grown at each farm using local practices for cultivation, fertilizer application, and field management. Chemical pesticides were applied as shown in Additional file 1: Table S1. Four plots (5 × 10 m) were randomly selected for collecting 30 cores of bulk soil samples (0–15 cm) at each farm. At the same time, plants were inspected for naturally occurring diseases in four randomly chosen, representative 5-m lengths of rows from each plot. The severity of rice leaf diseases (rice brown spot, brown streak, bacterial blight, bacterial leaf streak and foot rot) was visually scored on a 0 to 5 scale based on the estimated percentage leaf area with lesions: 0, no symptoms; 1, 0.1–6% of total leaf area; 2, 7–12%; 3, 13–25%; 4, 26–50%; 5, 51–100% [50]. A disease severity index was then calculated as \sum (Number of diseased plants with each score × Highest score) / (Total number of plants × Highest score) × 100. Disease incidence of rice bacterial foot rot was investigated as the (Number of plants with blackened, rotting roots / Total number of examined plants) × 100. Soil samples were transported to the laboratory on ice in a cooler. A portion of the soil sample was stored at –80 °C for DNA extraction, and the remainder was stored at 4 °C for analyses of soil chemical properties.

Analyses of soil chemical properties

Soil pH was determined using a Mettler-Toledo FE20-Five Easy Plus™ pH meter (Schwerzenbach, Switzerland) in a 1:5 soil–water (w/v) suspension. Soil organic carbon was measured using the chromic acid titration method [51]. Soil available nitrogen was determined by the alkali diffusion method, soil available phosphorus was determined using the sodium hydrogen carbonate solution-Mo-Sb anti spectrophotometric method, soil available potassium was measured by the ammonium acetate method, followed by flame photometric detection [52].

DNA extraction

Microbial genomic DNA was extracted from 0.5 g soil with the E.Z.N.A. Soil DNA Kit (Omega Bio-TEK, Norcross, GA, USA) according to the manufacturer's instructions. Concentration and purity of extracted DNA were determined with a TBS-380 Mini-Fluorometer (Turner Biosystems, CA, USA) and NanoDrop2000 UV-Vis spectrophotometer (Thermo Scientific, Wilmington, DE, USA), respectively. DNA quality was confirmed using 1.2% agarose gel electrophoresis, with $1\times$ TAE buffer (40 mM Tris-HCl, 40 mM acetate, 1.0 mM EDTA) and ethidium bromide ($0.5\ \mu\text{g mL}^{-1}$) under ultraviolet light [53].

Shotgun metagenomic analysis

The extracted DNA was fragmented to an average size of about 400 bp using Covaris M220 (Gene Company Limited, Hong Kong, China). A paired-end library was constructed using NEXTFLEX[®] Rapid DNA-Seq (Bio Scientific, Austin, TX, USA). Paired-end sequencing was performed using the Illumina HiSeq 4000 platform (Illumina Inc., San Diego, CA, USA) at Majorbio BioPharm Technology Co., Ltd. (Shanghai, China), HiSeq 3000/4000 PE Cluster Kit and HiSeq 3000/4000 SBS Kit according to the manufacturer's instructions (www.illumina.com). Sequence data associated with this project were deposited in the NCBI Short Read Archive database (accession PRJNA732820). The paired-end Illumina reads were trimmed of adaptors, and low-quality reads (length < 50 bp or with a quality value < 20 or having N bases) were removed by fastp version 0.20.0 (<https://github.com/OpenGene/fastp>). Metagenomics data were assembled using MEGAHIT version 1.1.2 (<https://github.com/voutcn/megahit>). Contigs of ≥ 300 bp were selected for further gene prediction and annotation. BLASTP against the NCBI NR database was used for taxonomic annotations of the representative sequences with an e-value cut-off $1e^{-5}$ using Diamond (<http://www.diamondsearch.org/index.php>, version 0.8.35). The KEGG annotation was conducted using Diamond against the Kyoto Encyclopedia of Genes and Genomes database (KEGG) version 92.0 (<http://www.genome.jp/kegg/>) with an e-value cutoff of $1e^{-5}$. Virulence factors were annotated using the Virulence Factors Database (VFDB) database version 2016.03 (<http://www.mgc.ac.cn/VFs/main.htm>) with an e-value cutoff of $1e^{-5}$. Carbohydrate-active enzymes were identified using HMMscan tools and the CarbohydrateActive enZYmes Database (CAZy) version 6.0 (<http://www.cazy.org/>) with an e-value cutoff of $1e^{-5}$. We also selected the bacterial and viral genes and built bacterial and viral sets for taxonomic, KEGG and CAZY annotations. Clustered Regularly Interspaced Short

Palindromic Repeats (CRISPR) were used to predict phage host based on the sequence similarity between the spacers in microbial CRISPR regions and the protospacers in viral genomes [54].

Isolation of soil viruses

Soil sample (1 kg) was suspended in 1 L of SM solution (100 mM NaCl, 8 mM $\text{MgSO}_4\cdot 7\text{H}_2\text{O}$ and 50 mM Tris/HCl, pH 7.5) and shaken for 2 h at 30 °C. The soil particles were removed by centrifugation at $4000\times g$ for 15 min. The supernatant was filtered sequentially through 0.45 and 0.22 μm filters. Then, virus particles in the filtered supernatant were concentrated into 1 mL by centrifugation at $4000\times g$ in 100 kDa centrifugal ultrafiltration tubes.

Isolation of *Enterobacter* bacteriophage from paddy soil

Soil sample (10 g) was suspended in 50 mL Luria Bertani (LB) broth (1% tryptone, 0.5% yeast extract, and 1% NaCl) and dispersed by shaking for 12 h at 30 °C. The soil particles were removed by centrifugation at $4000\times g$ for 15 min, and the supernatant was filter-sterilized through a 0.22 μm filter. One milliliter of filtered supernatant was added to 10 mL of an overnight suspension culture of *Enterobacter cloacae* (proven to be nitrogen-fixing and provided by Agricultural Culture Collection of China) and incubated for 12 h at 30 °C [55]. Then bacterial cells were removed by centrifugation at $4000\times g$ for 15 min, and the supernatant was sterilized through a 0.22 μm filter. Then 0.1 mL of the filtered supernatant was mixed with 1 mL of the *E. cloacae* suspension ($\text{OD}_{600}=1$). This mixture was then added to 2.5 mL of soft agar (0.35% agar prepared in 1% tryptone, 0.5% NaCl, 3 mM MgCl_2 , 3 mM CaCl_2 , and 0.04% [wt / vol] glucose) and poured in a uniform layer onto an LB agar plate. Bacteriophage plaques were identified after an overnight incubation at 30 °C. Bacteriophages were purified via five consecutive transfers of them from individual plaques to new bacterial cell lawns. The purified bacteriophage titer was determined by mixing 100 μL of host cells ($\text{OD}_{600}=1$) and 100 μL of bacteriophage dilution with soft agar and pouring the mixture over LB agar in a plate. Plaques were scored after an overnight incubation at 30 °C. Bacteriophages were kept in buffer (10 mM Tris [pH 7.6], 5 mM $\text{MgSO}_4\cdot 7\text{H}_2\text{O}$, 0.01% gelatin) at 4 °C.

Each isolated *Enterobacter* phage was analyzed for its ability to lyse *E. ancerogenus* and *E. ludwigii* (proven to be nitrogen-fixing and provided by Agricultural Culture Collection of China) using a plate lytic experiment with 0.1 mL bacteriophage mixed with 1 mL of overnight-cultured *Enterobacter* strain and shaken (200 rpm) at room temperature for 20 min, then 3 mL of soft nutrient agar was added, and the mixture was poured on LB agar plates

and incubated overnight at 30 °C. Lytic activity was indicated by transparent plaques [56].

Impact of *Enterobacter* bacteriophages on soil nitrogen-fixation capacity

In a greenhouse experiment, the influences of *Enterobacter* phages on soil nitrogen fixation capacity in untreated field soil and in sterilized field soil were investigated. Rice seeds were surface-sterilized and pregerminated, then seedlings with a primary root length of 1–2 cm were planted in pots (6 seeds per pot) with the untreated field soil. Three treatments with four pots (each pot filled with 1 kg soil) were set up: (1) N+NJ virus, 1 mL virus isolated from 1 kg Nanjing soil was diluted to 50 mL with ddH₂O and added to each pot after planting rice; (2) N+phage, 1 mL *Enterobacter* bacteriophages (10⁷ PFUs) diluted to 50 mL with ddH₂O was added after planting rice; (3) NCK, 50 mL ddH₂O was added as control. After treatment, rice plants were grown in a greenhouse at 26 °C with 14 h of fluorescent light and 10 h of darkness for 7 weeks and watered with 20 mL ddH₂O every 3 days.

Untreated field soil was autoclaved using the liquid cycle for 30 min at 121 °C, then cooled at room temperature, then we repeated this procedure, sterilized soil was stored at 4 °C until use. Mixed *Enterobacter* species (*E. cloacae*-211, *E. ancerogenus*-212, *E. ludwigii*-213, ratio 1:1:1) were cultured in LB broth overnight at 30 °C, followed by centrifugation at 1000×g for 5 min. The pellet was resuspended in ddH₂O and diluted to an optical density OD₆₀₀ of 0.02. Rice seeds were surface-sterilized and pregerminated, then seedlings with primary root length of 1–2 cm were planted in pots (6 seeds per pot) with sterilized soil. Four treatments were set up with four pots per treatment: (1) S+NF, 50 mL *Enterobacter* strains were added after rice planting, then cultured for a week and added with ddH₂O; (2) S+NF+virus, 50 mL *Enterobacter* strains were added after rice planting, then cultured for a week and added with 1 mL virus isolated from 1 kg Nanjing soil, which was diluted to 50 mL with ddH₂O; (3) S+NF+phage, 50 mL *Enterobacter* strains were added after rice planting, then cultured for a week and watered with 1 mL *Enterobacter* bacteriophages (10⁷ PFUs) was diluted to 50 mL with ddH₂O; (4) SCK, sterile water was used as the control. After treatment, rice plants were grown in a greenhouse at 26 °C with 14 h of fluorescent light and 10 h of darkness for 7 weeks and watered with 20 mL ddH₂O every 3 days. After 7 weeks, rhizosphere soil samples were collected [57], rice height and fresh mass were measured. Then rice plants were dried and grounded, nitrogen concentrations were determined with a modified Kjeldahl digestion method [58]. N₂ fixation in soil samples was assessed indirectly using

the acetylene reduction assay measuring ethylene formation from acetylene [59].

Illumina MiSeq sequencing

Region V3-V4 of the bacterial 16S rRNA gene was amplified from each sample using primers 338F (5'-ACTCCTACGGGAGGCAGCAG-3') and 806R (5'-GGACTACHVGGGTWTCTAAT-3') in a thermocycler PCR system (GeneAmp 9700, ABI, Foster, CA, United States); barcode was added before the primer to build different libraries in the same sequencing pool. The PCR mixture contained 4 µL 5× TransStart FastPfu buffer, 2 µL 2.5 mM dNTPs, 0.8 µL 5 µM forward primer, 0.8 µL 5 µM reverse primer, 0.4 µL TransStart FastPfu DNA Polymerase, 10 ng template DNA, with ddH₂O added to reach 20 µL. PCR cycling conditions were initial denaturation at 95 °C for 3 min; 35 cycles of denaturing at 95 °C for 30 s, annealing at 55 °C for 30 s and extension at 72 °C for 45 s; and single extension at 72 °C for 10 min, and end at 4 °C. The PCR product was extracted from 2% agarose gel and purified using the AxyPrep DNA Gel Extraction Kit (Axygen Biosciences, Union City, CA, USA), then quantified using a Quantus Fluorometer (Promega, Madison, WI, USA). Purified amplicons were paired-end sequenced (2×300) on an Illumina MiSeq platform (Illumina, San Diego, CA, USA) according to standard protocols by Majorbio Bio-Pharm Technology Co. Ltd. (Shanghai, China). The raw reads were deposited in the NCBI Sequence Read Archive (SRA) database (accession No. PRJNA780198).

The raw sequencing reads were demultiplexed, quality-filtered by Trimmomatic and merged by FLASH [60]. Operational taxonomic units (OTUs) with 97% similarity cutoff were clustered using UPARSE version 7.1 (<http://drive5.com/uparse/>), and chimeric sequences were identified and removed. The abundance of OTUs was rarefied to the lowest number of shared OTUs to remove the effect of sequencing depth across samples. The taxonomy of each 16S gene sequence was analyzed using the UNITE version 7.2 (<http://unite.ut.ee>) database with a confidence threshold of 70% [61].

Statistical analyses

Shannon and Chao indices were calculated by Mothur version v.1.30 (<http://www.mothur.org/>) with the vegan package to compare the alpha diversity of soil microbial communities. Principal coordinate analysis (PCoA) was performed in R version 3.0.2 (<https://www.r-project.org/>), R Core package using Bray–Curtis distances to compare soil microbial communities. Circos figures were obtained from Circos-0.67–7 (<http://circos.ca/>) and reflected the proportion of prominent genera in each sample [62]. Redundancy analysis (RDA) was

executed in R version 3.0.2 to analyze the relationship between dominant (abundance > 1%) genera or function at KEGG level 2 and environmental factors. Viral genes of virus were picked out to create a viral gene set to analyze the relative abundance of viruses and its functions in soil virus. Spearman's correlation between virus and bacteria or microbial function and environmental factors were calculated using the `rcorr` function in the R package "Hmisc" (v. 4.0-3). Data normally distributed were tested for significant differences with Tukey test after a one-way analysis of variance (ANOVA); data not normally distributed were tested with nonparametric Kruskal–Wallis test using SPSS version 16.0 (SPSS, Chicago, IL, USA). Figures were drawn in R version 3.0.2.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40793-022-00410-8>.

Additional file 1. Table S1. Chemical applications (apps.) on farms using three rice cropping systems. GZ: Rice double cropping ecosystems in Guangzhou; JMS: Rice single cropping ecosystems in Jiamusi; NJ: Rice-wheat rotation cropping ecosystems in Nanjing. **Table S2.** Means \pm standard deviation for environmental factors variables at the three agroecosystem sites in China. **Table S3.** Percentage distribution of the dominant genera in soils of three rice cropping agroecosystems. GZ: rice double cropping in Guangzhou; JMS: rice single cropping in Jiamusi; NJ: rice-wheat rotation in Nanjing. **Fig. S1.** Relative abundance of main microbial phyla (A) and genera (B) in soils of three rice cropping agroecosystems. GZ: rice double cropping in Guangzhou; JMS: rice single cropping in Jiamusi; NJ: rice-wheat rotation in Nanjing. **Fig. S2.** Principal coordinate analysis (PCoA) of soil microbial community based on Bray–Curtis distance matrices. The percentage represents the explanatory value of the PCoA axes to the difference in sample composition; distances between symbols on the ordination plot reflect relative dissimilarities. GZ: rice double cropping in Guangzhou; JMS: rice single cropping in Jiamusi; NJ: rice-wheat rotation cropping in Nanjing. **Fig. S3.** Relative abundance of potentially pathogenic bacteria (A) and disease index for four rice diseases in the three agroecosystems in China. Different letters above the bars indicate a significant difference ($p < 0.05$) according to Kruskal–Wallis H test. GZ: rice double cropping in Guangzhou; JMS: rice single cropping in Jiamusi; NJ: rice-wheat rotation cropping in Nanjing. **Fig. S4.** Relative abundance of genes related to xenobiotics biodegradation and metabolism in viruses and bacteria in soils of three rice agroecosystems. GZ: rice double cropping in Guangzhou; JMS: rice single cropping in Jiamusi; NJ: rice-wheat rotation in Nanjing. Different letters above the bars indicate a significant difference ($p < 0.05$) according to Kruskal–Wallis H test. **Fig. S5.** Relative abundance of genes related to carbohydrate (CHO) metabolism (A) and biosynthesis of secondary (Sec) metabolites (B) in viruses and bacteria in soils of three rice agroecosystems. GZ: rice double cropping in Guangzhou; JMS: rice single cropping in Jiamusi; NJ: rice-wheat rotation in Nanjing. Different letters above the bars indicate a significant difference ($p < 0.05$) according to Kruskal–Wallis H test. **Fig. S6.** Relative abundance of *Enterobacter* phage among soil viruses (A) and relative abundance of *Enterobacter* among soil bacteria (B) in the three rice agroecosystems. GZ: rice double cropping in Guangzhou; JMS: rice single cropping in Jiamusi; NJ: rice-wheat rotation in Nanjing. Different letters above the bars indicate a significant difference ($p < 0.05$) according to Kruskal–Wallis H test. **Fig. S7.** (A) Transmission electron micrograph of *Enterobacter* phage-NJ. (B) Plate lytic experiment of *Enterobacter* phage-NJ against nitrogen-fixing *Enterobacter cloacae*.

Acknowledgements

Authors thank Dr. B. E. Hazen (Willows End scientific editing and writing, United States) for critically reading and revising the manuscript.

Authors' contributions

Y.W. Conceptualization, Methodology, Investigation, Writing – original draft. Y.L. Software, Formal analysis, Investigation. Y.W. Software and statistics. N.W. Investigation, bioinformatic and statistics. W. L. Sampling campaign, Investigation. X.W. Conceptualization, Formal analysis, Validation, Writing – review & editing, Supervision. All authors read and approved the final manuscript.

Funding

This work was funded by the National Key R&D Program of China: Inter-Governmental S&T Cooperation Proposal (Grant No. 2019YFE0108500) and the National Natural Science Foundation of China Project U1704234.

Availability of data and materials

All the raw sequence data for metagenomic sequencing were submitted to the NCBI Sequence Read Archive (<http://www.ncbi.nlm.gov>) database under accession number PRJNA732820 and PRJNA780198.

Declarations

Ethics approval and consent to participate

The study does not contain any studies with human participants or animals that require specific ethical approval.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹State Key Laboratory for Biology of Plant Diseases and Insect Pests, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing 100193, China. ²Institute of Plant Protection, Hebei Academy of Agricultural and Forestry Sciences, Baoding 071000, China.

Received: 17 December 2021 Accepted: 19 March 2022

Published online: 06 April 2022

References

- Yuan L. Development of hybrid rice to ensure food security. *Rice Sci.* 2014;21:1–2.
- Fang H, Zhang Z, Li D, Liu K, Zhang K, Zhang W, et al. Temporal dynamics of paddy soil structure as affected by different fertilization strategies investigated with soil shrinkage curve. *Soil Tillage Res.* 2019;187:102–9. <https://doi.org/10.1016/j.still.2018.12.006>.
- Liu W, Zhang X, Wu N, Ren Y, Wang X. High diversity and functional complementation of alimentary canal microbiota ensure small brown planthopper to adapt different biogeographic environments. *Front Microbiol.* 2020;10:2953. <https://doi.org/10.3389/fmicb.2019.02953/full>.
- Wang Z, Yang Y, Xia Y, Wu T, Zhu J, Yang J, et al. Time-course relationship between environmental factors and microbial diversity in tobacco soil. *Sci Rep.* 2019;9:1–11. <https://doi.org/10.1038/s41598-019-55859-4>.
- De Quadros PD, Zhalnina K, Davis-Richardson A, Fagen JR, Drew J, Bayer C, et al. The effect of tillage system and crop rotation on soil microbial diversity and composition in a subtropical acrisol. *Diversity.* 2012;4:375–95.
- Wu N, Shi W, Liu W, Gao Z, Han L, Wang X. Differential impact of Bt-transgenic rice plantings on bacterial community in three niches over consecutive years. *Ecotoxicol Environ Saf.* 2021;223: 112569. <https://doi.org/10.1016/j.ecoenv.2021.112569>.
- Talbot JM, Bruns TD, Smith DP, Branco S, Glassman SI, Erlandson S, et al. Independent roles of ectomycorrhizal and saprotrophic communities in soil organic matter decomposition. *Soil Biol Biochem.* 2013;57:282–91. <https://doi.org/10.1016/j.soilbio.2012.10.004>.
- Liu D, Wang H, An S, Bhople P, Davlatbekov F. Geographic distance and soil microbial biomass carbon drive biogeographical distribution of fungal communities in Chinese Loess Plateau soils. *Sci Total Environ.* 2019;660:1058–69. <https://doi.org/10.1016/j.scitotenv.2019.01.097>.

9. Khan MS, Zaidi A, Ahemad M, Oves M, Wani PA. Plant growth promotion by phosphate solubilizing fungi—current perspective. *Arch Agron Soil Sci.* 2010;56:73–98. <https://doi.org/10.1080/03650340902806469>.
10. Xu W, Wang K, Wang H, Liu Z, Shi Y, Gao Z, et al. Evaluation of the biocontrol potential of *Bacillus* sp. WB against *Fusarium oxysporum* f. sp. *niveum*. *Biol Control.* 2020;147:10428. <https://doi.org/10.1016/j.biocontrol.2020.104288>.
11. Wang Y, Ji L, Li Q, Wu Y, Li C, Kong L, et al. Effects of long-term bare fallow during the winter-wheat growth season on the soil chemical properties, fungal community composition, and the occurrence of maize fungal diseases in North China. *Plant Dis.* 2021. <https://doi.org/10.1094/PDIS-11-20-2492-RE>.
12. Shi Y, Li Y, Xiang X, Sun R, Yang T, He D, et al. Spatial scale affects the relative role of stochasticity versus determinism in soil bacterial communities in wheat fields across the North China Plain. *Microbiome.* 2018;6:27. <https://doi.org/10.1186/s40168-018-0409-4>.
13. Bahram M, Hildebrand F, Forslund SK, Anderson JL, Soudzilovskaia NA, Bodegom PM, et al. Structure and function of the global topsoil microbiome. *Nature.* 2018;560:233–7.
14. Maestre FT, Delgado-Baquerizo M, Jeffries TC, Eldridge DJ, Ochoa V, Gozalo B, et al. Increasing aridity reduces soil microbial diversity and abundance in global drylands. *Proc Natl Acad Sci.* 2015;112:15684–9. <https://doi.org/10.1073/pnas.1516684112>.
15. Emerson JB. Soil viruses: a new hope. *mSystems.* 2019;4:1–4. <https://doi.org/10.1128/mSystems.00120-19>.
16. Suttle CA. Marine viruses—major players in the global ecosystem. *Nat Rev Microbiol.* 2007;5:801–12.
17. Brum JR, Sullivan MB. Rising to the challenge: accelerated pace of discovery transforms marine virology. *Nat Rev Microbiol.* 2015;13:147–59. <https://doi.org/10.1038/nrmicro3404>.
18. Mann NH, Cook A, Millard A, Bailey S, Clokie M. Bacterial photosynthesis genes in a virus. *Nature.* 2003;424:741–741.
19. Roux S, Brum JR, Dutilh BE, Sunagawa S, Duhaime MB, Loy A, et al. Ecogenomics and potential biogeochemical impacts of globally abundant ocean viruses. *Nature.* 2016;537:689–93. <https://doi.org/10.1038/nature19366>.
20. Bi L, Yu D, Du S, Zhang L, Zhang L, Wu C, et al. Diversity and potential biogeochemical impacts of viruses in bulk and rhizosphere soils. *Environ Microbiol.* 2021;23:588–99. <https://doi.org/10.1111/1462-2920.15010>.
21. Middelboe M, Brussaard C. Marine viruses: key players in marine ecosystems. *Viruses.* 2017;9:302.
22. Sullivan MB, Lindell D, Lee JA, Thompson LR, Bielawski JP, Chisholm SW. Prevalence and evolution of core photosystem II genes in marine cyanobacterial viruses and their hosts. *Moran NA, editor. PLoS Biol.* 2006;4:234. <https://doi.org/10.1371/journal.pbio.0040234>.
23. Han LL, Yu DT, Zhang LM, Wang JT, He JZ. Unique community structure of viruses in a glacier soil of the Tianshan Mountains, China. *J Soils Sediments.* 2017;17:852–60. <https://doi.org/10.1007/s11368-016-1583-2>.
24. Jin M, Guo X, Zhang R, Qu W, Gao B, Zeng R. Diversities and potential biogeochemical impacts of mangrove soil viruses. *Microbiome.* 2019;7:1–15. <https://doi.org/10.1186/s40168-019-0675-9>.
25. Yu D-T, He J-Z, Zhang L-M, Han L-L. Viral metagenomics analysis and eight novel viral genomes identified from the Dushanzi mud volcanic soil in Xinjiang, China. *J Soils Sediments.* 2019;19:81–90. <https://doi.org/10.1007/s11368-018-2045-9>.
26. Adriaenssens EM, Kramer R, Van Goethem MW, Makhalyane TP, Hogg I, Cowan DA. Environmental drivers of viral community composition in Antarctic soils identified by viromics. *Microbiome.* 2017;5:83. <https://doi.org/10.1186/s40168-017-0301-7>.
27. Wolters V, Silver WL, Bignell DE, Coleman DC, Lavelle P, Van Der Putten WH, et al. Effects of global changes on above- and belowground biodiversity in terrestrial ecosystems: implications for ecosystem functioning. *Bioscience.* 2000;50:1089–98. [https://doi.org/10.1641/0006-3568\(2000\)050\[1089:EOGCOA\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2000)050[1089:EOGCOA]2.0.CO;2).
28. Wardle DA, Bardgett RD, Klironomos JN, Setälä H, van der Putten WH, Wall DH. Ecological linkages between aboveground and belowground biota. *Science (80-).* 2004;304:1629–33. <https://doi.org/10.1126/science.1094875>.
29. Jing X, Sanders NJ, Shi Y, Chu H, Classen AT, Zhao K, et al. The links between ecosystem multifunctionality and above- and belowground biodiversity are mediated by climate. *Nat Commun.* 2015;6:8159.
30. Fierer N, Leff JW, Adams BJ, Nielsen UN, Bates ST, Lauber CL, et al. Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. *Proc Natl Acad Sci.* 2012;109:21390–5. <https://doi.org/10.1073/pnas.1215210110>.
31. Melotto M, Underwood W, He SY. Role of stomata in plant innate immunity and foliar bacterial diseases. *Annu Rev Phytopathol.* 2008;46:101–22. <https://doi.org/10.1146/annurev.phyto.121107.104959>.
32. Williamson KE, Radosevich M, Smith DW, Wommack KE. Incidence of lysogeny within temperate and extreme soil environments. *Environ Microbiol.* 2007;9:2563–74. <https://doi.org/10.1111/j.1462-2920.2007.01374.x>.
33. Wen K, Ortmann AC, Suttle CA. Accurate estimation of viral abundance by epifluorescence microscopy. *Kevin Wen.* 2004;70:3862–7.
34. Zhuang J, Jin Y. Virus retention and transport through Al-oxide coated sand columns: effects of ionic strength and composition. *J Contam Hydrol.* 2003;60:193–209.
35. Loveland J, Ryan J, Amy G, Harvey R. The reversibility of virus attachment to mineral surfaces. *Colloids Surf A Physicochem Eng Asp.* 1996;107:205–21.
36. Zünd M, Ruscheweyh H-J, Field CM, Meyer N, Cuenca M, Hoces D, et al. High throughput sequencing provides exact genomic locations of inducible prophages and accurate phage-to-host ratios in gut microbial strains. *Microbiome.* 2021;9:77. <https://doi.org/10.1186/s40168-021-01033-w>.
37. Williamson KE, Radosevich M, Wommack KE. Abundance and diversity of viruses in six Delaware soils. *Appl Environ Microbiol.* 2005;71:3119–25. <https://doi.org/10.1128/AEM.71.6.3119-3125.2005>.
38. Gómez P, Buckling A. Bacteria-phage antagonistic coevolution in soil. *Science (80-).* 2011;332:106–9. <https://doi.org/10.1126/science.1198767>.
39. Wigington CH, Sonderegger D, Brussaard CPD, Buchan A, Finke JF, Fuhrman JA, et al. Re-examination of the relationship between marine virus and microbial cell abundances. *Nat Microbiol.* 2016;1:15024.
40. Zhang J, Gao Q, Zhang Q, Wang T, Yue H, Wu L, et al. Bacteriophage-prokaryote dynamics and interaction within anaerobic digestion processes across time and space. *Microbiome.* 2017;5:57. <https://doi.org/10.1186/s40168-017-0272-8>.
41. Howard-Varona C, Hargreaves KR, Abedon ST, Sullivan MB. Lysogeny in nature: mechanisms, impact and ecology of temperate phages. *ISME J.* 2017;11:1511–20. <https://doi.org/10.1038/ismej.2017.16>.
42. Braga LPP, Spor A, Kot W, Breuil M-C, Hansen LH, Setubal JC, et al. Impact of phages on soil bacterial communities and nitrogen availability under different assembly scenarios. *Microbiome.* 2020;8:52. <https://doi.org/10.1186/s40168-020-00822-z>.
43. Wei X, Ge T, Wu C, Wang S, Mason-Jones K, Li Y, et al. T4-like phages reveal the potential role of viruses in soil organic matter mineralization. *Environ Sci Technol.* 2021;55:6440–8. <https://doi.org/10.1021/acs.est.0c06014>.
44. Dalcin Martins P, Danczak RE, Roux S, Frank J, Borton MA, Wolfe RA, et al. Viral and metabolic controls on high rates of microbial sulfur and carbon cycling in wetland ecosystems. *Microbiome.* 2018;6:1–17. <https://doi.org/10.1186/s40168-018-0522-4>.
45. Kimura M, Jia Z-J, Nakayama N, Asakawa S. Ecology of viruses in soils: past, present and future perspectives. *Soil Sci Plant Nutr.* 2008;54:1–32. <https://doi.org/10.1111/j.1747-0765.2007.00197.x>.
46. Hurwitz BL, U'Ren JM. Viral metabolic reprogramming in marine ecosystems. *Curr Opin Microbiol.* 2016;31:161–8. <https://doi.org/10.1016/j.mib.2016.04.002>.
47. Emmerling M, Dauner M, Ponti A, Fiaux J, Hochuli M, Szyperski T, et al. Metabolic flux responses to pyruvate kinase knockout in *Escherichia coli*. *J Bacteriol.* 2002;184:152–64. <https://doi.org/10.1128/JB.184.1.152-164.2002>.
48. Trubl G, Bin JH, Roux S, Emerson JB, Solonenko N, Vik DR, et al. Soil viruses are underexplored players in ecosystem carbon processing. *bioRxiv.* 2018;3:1–21. <https://doi.org/10.1128/msystems.00076-18>.
49. Latka A, Maciejewska B, Majkowska-Skropek G, Briers Y, Drulis-Kawa Z. Bacteriophage-encoded virion-associated enzymes to overcome the carbohydrate barriers during the infection process. *Appl Microbiol Biotechnol.* 2017;101:3103–19. <https://doi.org/10.1007/s00253-017-8224-6>.
50. Rajarajeswari NVL, Muralidharan K. Assessments of farm yield and district production loss from bacterial leaf blight epidemics in rice. *Crop Prot.* 2006;25:244–52.
51. Walkley A, Black IA. An examination of the degtjareff method for determining soil organic matter, and a proposed modification of the chromic acid titration method. *Soil Sci.* 1934;37:29–38.
52. Hanway JJ. Growth stages of Corn (*Zea mays*, L.) 1. *Agron J.* 1963;55:487–92. <https://doi.org/10.2134/agronj1963.00021962005500050024x>.

53. Liu J, Yu Z, Wang X, Jin J, Liu X, Wang G. The distribution characteristics of the major capsid gene (g23) of T4-type phages in paddy floodwater in Northeast China. *Soil Sci Plant Nutr.* 2016;62:133–9. <https://doi.org/10.1080/00380768.2016.1163507>.
54. Huang D, Yu P, Ye M, Schwarz C, Alvarez P. Enhanced mutualistic symbiosis between soil phages and bacteria with elevated chromium-induced environmental stress. *Microbiome.* 2021;9:1–19. <https://doi.org/10.1186/s40168-021-01074-1>.
55. Cross T, Schoff C, Chudoff D, Graves L, Broomell H, Terry K, et al. An optimized enrichment technique for the isolation of *Arthrobacter* bacteriophage species from soil sample isolates. *J Vis Exp.* 2015;2015:1–9.
56. Summer EJ, Liu M, Gill JJ, Grant M, Chan-Cortes TN, Ferguson L, et al. Genomic and functional analyses of *Rhodococcus equi* phages ReqiPepy6, ReqiPoco6, ReqiPine5, and ReqiDocB7. *Appl Environ Microbiol.* 2011;77:669–83. <https://doi.org/10.1128/AEM.01952-10>.
57. Berlanas C, Berbegal M, Elena G, Laidani M, Cibriain JF, Sagües A, et al. The fungal and bacterial rhizosphere microbiome associated with grapevine rootstock genotypes in mature and young vineyards. *Front Microbiol.* 2019;10:1–16. <https://doi.org/10.3389/fmicb.2019.01142>.
58. Nelson DW, Sommers LE. Determination of total nitrogen in plant material. *Agron J.* 1973;65:109–12. <https://doi.org/10.2134/agronj1973.00021962006500010033x>.
59. Lopez-Lozano NE, Carcaño-Montiel MG, Bashan Y. Using native trees and cacti to improve soil potential nitrogen fixation during long-term restoration of arid lands. *Plant Soil.* 2016;403:317–29. <https://doi.org/10.1007/s11104-016-2807-3>.
60. Magoc T, Salzberg SL. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics.* 2011;27:2957–63. <https://doi.org/10.1093/bioinformatics/btr507>.
61. Kõljalg U, Nilsson RH, Abarenkov K, Tedersoo L, Taylor AFS, Bahram M, et al. Towards a unified paradigm for sequence-based identification of fungi. *Mol Ecol.* 2013;22:5271–7. <https://doi.org/10.1111/mec.12481>.
62. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, et al. Circo: an information aesthetic for comparative genomics. *Genome Res.* 2009;19:1639–45. <https://doi.org/10.1101/gr.092759.109>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

