

DOI: 10.5846/stxb202002230312

毕丽, 杜帅, 于丹婷, 张丽梅, 贺纪正, 韩丽丽. 新疆两种土地利用方式下土壤病毒的群落组成与功能特征. 生态学报, 2021, 41(7): 2728-2737.

Bi L, Du S, Yu D T, Zhang L M, He J Z, Han L L. Compositional and functional characteristics of viruses in soil under two types of land-use in Xinjiang, China. Acta Ecologica Sinica, 2021, 41(7): 2728-2737.

新疆两种土地利用方式下土壤病毒的群落组成与功能特征

毕 丽^{1,2}, 杜 帅^{1,2}, 于丹婷³, 张丽梅^{1,2}, 贺纪正³, 韩丽丽^{1,2,*}

1 中国科学院生态环境研究中心, 城市与区域生态国家重点实验室, 北京 100085

2 中国科学院大学, 北京 100049

3 福建师范大学地理科学学院, 福州 350007

摘要: 病毒作为食物网的重要组成部分, 在生态系统中发挥着重要的功能。病毒能够影响宿主的死亡率、群落结构和进化, 以及营养元素循环。但由于技术方法的限制, 对土壤病毒的群落组成和功能特征还知之甚少。为探索不同土地利用方式下土壤病毒组特征, 采集了新疆棉花地和荒漠土壤样品。通过宏病毒组学分析发现, 棉花地土壤和荒漠土壤分别注释到 20 个和 15 个病毒科, 单链 DNA(ssDNA)病毒占优势, 其中微小噬菌体科(*Microviridae*)占比最高。仅在棉花地土壤中检测到花椰菜花叶病毒科(*Caulimoviridae*)、逆转录病毒科(*Retroviridae*)、裸露病毒科(*Nudiviridae*)、多分 DNA 病毒科(*Polydnaviridae*)、杆状病毒科(*Baculoviridae*)和囊泡病毒科(*Ascoviridae*), 其中大部分病毒属于植物病毒和昆虫病毒。本研究推测与土地利用方式相关的人为活动、土壤理化性质以及动植物的差异可能影响土壤病毒的群落组成。通过 Virsorter 共注释到 1824 条病毒 contigs, 主要为微小噬菌体科。进一步利用 SEED 数据库对病毒功能进行注释, 发现两个土壤病毒组注释到的主要功能类似; 在 SEED level 2 水平上, 均以“Phage capsid proteins”和“Phage packaging machinery”占比最高。本研究可为进一步探索土壤病毒生态功能和土壤食物网提供数据支持。

关键词: 土壤; 宏病毒组; 病毒群落组成; 微小噬菌体科

Compositional and functional characteristics of viruses in soil under two types of land-use in Xinjiang, China

BI Li^{1,2}, DU Shuai^{1,2}, YU Danting³, ZHANG Limei^{1,2}, HE Jizheng³, HAN Lili^{1,2,*}

1 State Key Laboratory of Urban and Regional Ecology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing 100085, China

2 University of Chinese Academy of Sciences, Beijing 100049, China

3 Institute of Geography, Fujian Normal University, Fuzhou 350007, China

Abstract: Viruses are an important component of the food web, and have vital functions in the ecosystem. Viruses can affect the mortality, community structure and evolution of their hosts and nutrient cycling. However, we know little about community composition and functional characteristics of viruses in soil, due to technical challenges. To initially explore the characteristics of soil virome under different land-use types, cotton soil and desert soil were collected in Xinjiang. Twenty and fifteen viral families were identified from cotton soil and desert soil, respectively, and ssDNA viruses were predominant in both two samples, with *Microviridae* being the most abundant family. *Caulimoviridae*, *Retroviridae*, *Nudiviridae*,

基金项目: 国家自然科学基金项目(41771289, 41571248)

收稿日期: 2020-02-23; 网络出版日期: 2021-01-27

* 通讯作者 Corresponding author. E-mail: llhan@rcees.ac.cn

Polydnaviridae, *Baculoviridae*, and *Ascoviridae* were only detected in cotton soil, most of which belonged to plant and insect viruses. This study speculates that the difference of human activities, soil physical and chemical properties, plants and animals associated with different land-use types might affect viral communities. A total of 1824 viral contigs were identified by Virsorter and mainly annotated to the family *Microviridae*. Major functions of two viromes were similar by comparison with the SEED database, and at SEED subsystem level 2 annotation, “Phage capsid proteins” and “Phage packaging machinery” accounted for the highest proportion. This study can provide data support for further exploration of soil virus ecological function and soil food web.

Key Words: soil; virome; viral community composition; *Microviridae*

病毒是地球上数量最多、种类最丰富的生物实体,也是生态系统中食物链和食物网的重要组成部分。环境病毒生态学早期主要在海洋生态系统中开展研究,而陆地生态系统病毒学研究在近 10 年才逐渐发展。据报道每克土壤中的病毒颗粒高达 10^7 — 10^9 个^[1],比细菌总数高 1—2 个数量级^[2]。病毒在陆地生态系统中发挥着重要的生态功能^[3-4],如影响宿主的死亡率和群落结构。Li 等通过 T4 型噬菌体标记基因(*g23*)和细菌 16S rRNA 基因测序分析发现 T4 型噬菌体能够影响水稻土壤细菌群落组成^[5]。同时,土壤病毒可能是影响微生物死亡率和周转速率的主要原因之一,它可裂解细胞释放出大量的可溶性和颗粒态有机碳,促进其他微生物的转化和利用,从而影响微生物群落结构和元素的生物地球化学循环^[6-7]。此外,病毒还可通过携带辅助代谢基因(auxiliary metabolic genes, AMGs)影响宿主代谢,从而间接参与元素循环^[4, 8],同时也为宿主提供新的功能基因以拓宽其生态位^[8]。

由于土壤病毒具有庞大的数量体和潜在的生态功能,越来越多的研究人员开始关注土壤病毒。土壤病毒丰度和多样性的研究方法主要包括荧光显微计数法、透射电子显微观察、分离培养、特定标记基因分析、宏基因组学和宏病毒组学等^[9-10]。宏病毒组和宏基因组是目前病毒生态学研究最重要的手段^[3, 10],可以成功避免大量病毒不可培养和缺乏通用标记基因所带来的问题,尤其是宏病毒组学可极大地增加测序序列中病毒基因组的覆盖度。宏病毒组学研究发现在纳米布沙漠土壤中,有尾噬菌体目(*Caudovirales*)丰度最高,其中以长尾噬菌体科(*Siphoviridae*)为主^[11],这与南极土壤中主要病毒类群的分布一致^[12]。而在中国的农田土壤中主要以单链 DNA(ssDNA)病毒为主导,其中以微小噬菌体科(*Microviridae*)和环状病毒科(*Circoviridae*)为主^[13]。近期,Starr 等首次通过宏转录组分析发现,植物根际与非根际土壤中的 RNA 病毒多样性很高,凋落物显著影响病毒和宿主的群落组成,而根系的生长不影响 RNA 病毒的群落组成^[14]。

目前我国土壤宏病毒组仅有少量的研究,主要关注了农田^[13, 15]、红树林^[16]、冰川^[17]以及海陆交错带土壤^[18]。不同土地利用方式会影响土壤理化性质,例如土壤 pH 和养分^[19],以及微生物群落结构和功能^[20]。但是我们对于不同土地利用方式下,病毒的群落结构和功能特征还缺乏认识。因此本研究从新疆阿克苏沙雅县选取两种典型土地利用方式下的土壤(棉花地和荒漠土壤),通过宏病毒组学技术探究土壤病毒群落结构和功能特征,为进一步了解农业活动对土壤病毒分布的影响提供科学依据。

1 材料与方法

1.1 研究区域与样品采集

采样点位于新疆阿克苏沙雅县(东经 82.54°、北纬 41.32°),该地区属于暖温带沙漠边缘气候区。本研究于 2018 年 6 月采集该地区棉花地土壤和荒漠土壤。棉花地种植时间约 10a,与荒漠土壤之间间隔约 200 m。使用梅花布点法(间隔大于 3 m)布设采样点,每个土壤样品选取 6 个样点,在每个样点采集 0—10 cm 土壤样品。土壤样品通过冰盒运回实验室。所有土壤通过 2 mm 筛去除植物残体后,储存于 4℃ 冰箱用于后期土壤理化性质测定和病毒提取。

1.2 土壤理化性质测定

称取 10 g 土壤样品于 105℃ 烘箱干燥 12 h, 计算土壤含水量 (Soil Water Content, SWC)。称取 5 g 风干土, 用煮沸的蒸馏水以 1:2.5 的土水比混合, 180 rpm 摇床振荡 5 min, 静置 30 min 后, 用 pH 计 (Professional Meter PP-20, Sartorius, Germany) 测定土壤 pH。称取 5 g 风干土, 用煮沸的蒸馏水以 1:5 的土水比混合, 180 rpm 摇床振荡 5 min, 静置 30 min 后, 用电导率仪测定土壤电导率 (Electrical Conductivity, EC)。称取 6 g 鲜土, 用 0.5 M K_2SO_4 溶液以水土比 5:1 浸提, 使用有机碳测定仪 (Multi N/C 3100, Analytikjena, Germany) 测定土壤可溶性有机碳 (Dissolved Organic Carbon, DOC) 和可溶性有机氮 (Dissolved Organic Nitrogen, DON)。称取 6 g 鲜土, 用 1 M KCl 溶液浸提, 并采用连续流动分析仪 (SAN++, Skalar, Holland) 测定铵态氮 NH_4^+-N 和硝态氮 $NO_3^- - N$ 。有效磷 (Available Phosphorus, AP) 用 Olsen 法测定。每种土壤的 6 个样品中两两混合, 设置 3 个重复测定土壤理化性质。

1.3 土壤病毒 DNA 提取及测序

分别将棉花地和荒漠土壤 6 个样点的土壤混合成一个样品, 根据前期的实验方法提取鲜土中的病毒^[13], 每个样品提取 1 次。简述如下: 称取过 2 mm 筛的鲜土 500 g, 与 3 L 的甘氨酸缓冲液 (250 mM, pH=8.5) 混合均匀, 于 150 rpm 振荡 15 min。混合液于 4℃ 1500 rpm 条件下低速离心 2 min。收集上清液, 沉淀继续添加缓冲液, 重悬浮, 混合均匀, 并重复以上步骤两次。利用切向流过滤系统 (Tangential Flow Filter System, QuixStand, GE Healthcare Life Sciences, Pittsburgh, PA, USA) 对上清液进行连续过滤, 依次通过 1 mm, 0.45 μm 和 0.22 μm 的滤柱, 并收集透过液。再将透过液通过 30 kDa 的超滤柱进行浓缩并收集截留液 (体积不大于 100 mL)。浓缩液再经过 0.22 μm 的无菌过滤器过滤, 重复 3 次以去除连续过滤过程中可能带来的杂菌污染。再以 4000 g 的转速通过 30 kDa 超滤离心管 (Merck Millipore Ltd., Tullagreen, Ireland) 进一步浓缩至体积约为 1 mL。接着将病毒浓缩液用 DNase I (Thermo Fisher Scientific, Lithuania, EU) 在 37℃ 孵育 1 h (10 units DNase I/500 μL), 以去除土壤中游离的胞外 DNA。利用细菌 16S rRNA 基因 PCR (引物 27F/1492R)^[21] 检测并确认没有细菌 DNA 污染后进行下一步实验。

使用病毒 DNA 提取试剂盒 (Qiagen AllPrep PowerViral DNA/RNA extraction kit (Qiagen, Hilden, Germany)) 提取浓缩液中的病毒 DNA。随后根据基因组扩增试剂盒 (REPLI-g Mini Kit (Qiagen, Hilden, Germany)) 使用说明对提取的病毒 DNA 进行多重置换扩增, 以得到足量的病毒 DNA 来构建宏病毒组文库。分别利用 Nanodrop 和 Qubit 检测 DNA 的纯度和浓度。再利用超声破碎仪 Covaris M220 (Covaris, Woburn, MA, USA) 将 DNA 随机打断成长度约为 350 bp 的片段, 建立宏病毒组文库后进行 Illumina HiSeq2500 高通量测序。测序工作由北京新科开源基因科技有限公司完成。

1.4 宏病毒组数据分析

测序得到的原始数据通过 fastp 软件, 选择默认参数^[22] 过滤接头并去除低质量序列, 最后得到 clean reads。利用 SortMeRNA v2.1 去除 clean reads 中的核糖体序列^[23], 利用 bbmap 比对 NCBI UniVec 数据库 (ftp://ftp.ncbi.nlm.nih.gov/pub/UniVec, March 2017) 去除可移动遗传元件。然后利用 metaSPAdes^[24] 对 clean reads 进行组装拼接得到原始 contigs。Clean reads 与病毒数据库进行 BLASTx 比对和物种注释 (E-Value < $1e^{-5}$, score 值 > 50)^[25], 病毒数据库综合了 NCBI non-redundant (NR) 数据库, Refseq virus 数据库和 PHAST 网站上的噬菌体数据库 (截至 2018 年 8 月)^[26-28]。比对得到的病毒序列使用 Prodigal 以默认参数进行开放阅读框 (ORFs) 预测^[29], 并与 NR 和 RefSeq 病毒蛋白数据库进行比对 (Diamond, BLASTp, E-Value < $1e^{-5}$, identity > 60%)^[25], 最后在 MEGAN 6 中利用 SEED 数据库对病毒进行功能分类^[30]。通过 Virsorter 预测病毒 contigs^[31], contigs 仅选取于具有更高可信度的 categories 1, 2, 5。病毒基因组示意图通过 Easyfig 进行可视化^[32]。

1.5 序列提交

本研究中的两个宏病毒组原始序列已经提交至 NCBI Sequence Read Archive (SRA), 获取号为 PRJNA599555。

2 结果与分析

2.1 土壤理化性质

如表 1 所示,新疆棉花地土壤含水量、DOC、DON 和 NO_3^- -N 极显著高于荒漠土壤 ($P < 0.01$), NH_4^+ -N 显著高于荒漠土壤 ($P < 0.05$)。其中棉花地土壤含水量 (10.92%—11.12%) 和 NO_3^- -N (4.74—5.84 mg/kg) 约是荒漠土壤相应理化性质数值的 5 倍。而棉花地土壤电导率、pH 和 AP 极显著低于荒漠土壤 ($P < 0.01$)。所有检测的理化性质在两个土壤样品中都存在显著的差异。

表 1 土壤理化性质
Table 1 Soil physicochemical properties

样品 Samples	含水量 SWC/%	pH	电导率 EC/ ($\mu\text{S}/\text{cm}$)	可溶性有机碳 DOC/ (mg/kg)	可溶性有机氮 DON/ (mg/kg)	铵态氮 NH_4^+ -N/ (mg/kg)	硝态氮 NO_3^- -N/ (mg/kg)	有效磷 AP/ (mg/kg)
A-1	11.12	8.5	723.2	99.95	35.68	3.12	4.97	11.59
A-2	11.04	8.4	766.9	118.23	48.74	3.69	4.74	11.75
A-3	10.92	8.5	675.6	103.91	37.31	3.10	5.84	13.22
B-1	2.78	9.0	2741.0	50.05	9.07	2.52	0.87	18.47
B-2	2.78	9.0	2927.0	64.72	10.28	2.00	1.15	20.02
B-3	2.73	9.1	2818.0	61.95	10.26	1.69	1.07	20.34

A:新疆棉花地土壤, Cotton Soil in Xinjiang; B:新疆荒漠土壤, Desert Soil in Xinjiang; 含水量: Soil Water Content (SWC); 电导率: Electrical Conductivity (EC); 可溶性有机碳: Dissolved Organic Carbon (DOC); 可溶性有机氮: Dissolved Organic Nitrogen (DON); 有效磷: Available Phosphorus (AP)

2.2 土壤病毒组信息

通过宏病毒组测序,从新疆棉花地和荒漠土壤样品中分别获得 78072330 和 80219056 条 clean reads, GC 含量分别为 53.56% 和 49.45%。在棉花地和荒漠土壤宏病毒组中分别有 12.34% 和 11.57% 的 reads 比对上病毒数据库。对 clean reads 进行拼接,在棉花地和荒漠土壤宏病毒组中分别获得 3050 和 2771 条 contigs (>1000 bp), 最长的两条 contigs 分别为 13586 bp 和 17062 bp (表 2)。

表 2 病毒组信息
Table 2 Overview of sequences of two viromes

样品 Samples	总的 clean reads 条数 Number of clean reads	clean reads 测序深度 Total size of clean reads/Gbp	GC 百分含量 GC content/%	比对率 Hit/%	Contig 总条数 Number of contigs/ (>1000 bp)	最长 contig 长度 Length of the longest contig/bp
A	78072330	11.70	53.56	12.34	3050	13586
B	80219056	12.03	49.45	11.57	2771	17062

A:新疆棉花地土壤 Cotton Soil in Xinjiang; B:新疆荒漠土壤 Desert Soil in Xinjiang; Clean reads: 质控后数据; GC 含量: 鸟嘌呤和胞嘧啶所占的比率

2.3 病毒群落结构

通过病毒数据库注释,新疆棉花地和荒漠土壤宏病毒组分别注释到 20 个和 15 个病毒科和一些未分类的病毒 (图 1)。在两种土壤中,均是 ssDNA 病毒占绝对优势,其中微小噬菌体科占比最高,分别为 69% 和 78%。微小噬菌体科主要包括蘑菇状噬菌体亚科 (*Gokushovirinae*), 分别占据棉花地和荒漠土壤宏病毒组的 35% 和 39%。环状病毒科在棉花地和荒漠土壤中所占比例分别为 2.0% 和 0.7%, 矮缩病毒科 (*Nanoviridae*) 在两种土壤中所占比例分别为 0.4% 和 0.2%。两种土壤中均检测到 *Genomoviridae*、双生病毒科 (*Geminiviridae*)、短尾噬菌体科 (*Podoviridae*)、肌尾噬菌体科 (*Myoviridae*)、长尾噬菌体科、*Alphasatellitidae*、*Smacoviridae* 和脱氧核糖核酸病毒科 (*Phycodnaviridae*) 等。此外在棉花地土壤中特有的病毒有花椰菜花叶病毒科 (*Caulimoviridae*)、逆转

录病毒科 (*Retroviridae*)、裸露病毒科 (*Nudiviridae*)、多分 DNA 病毒科 (*Polydnaviridae*)、杆状病毒科 (*Baculoviridae*) 和囊泡病毒科 (*Ascoviridae*)。而乳头瘤病毒科 (*Papillomaviridae*) 仅在荒漠土壤病毒组中检测到。

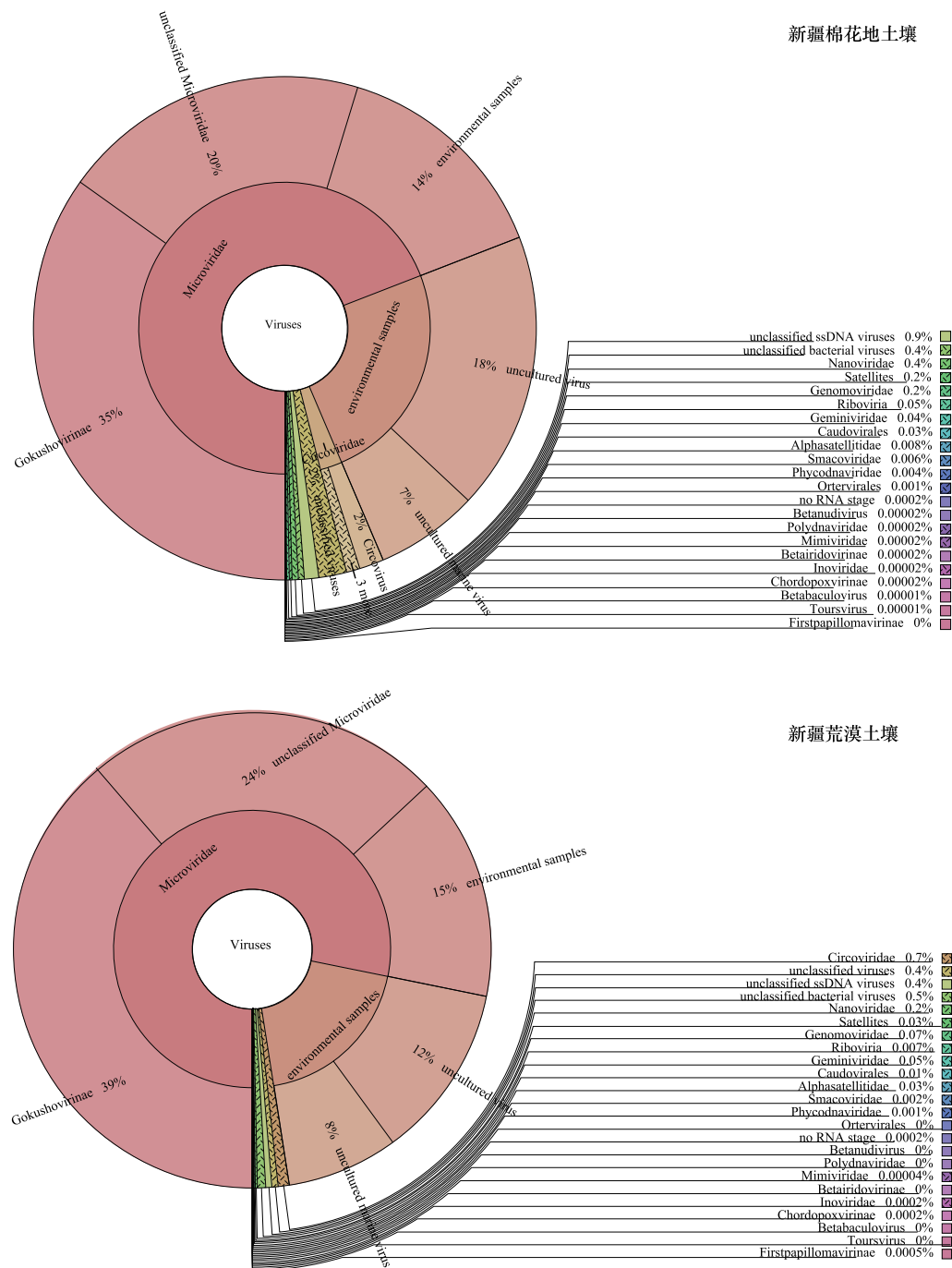


图 1 两个土壤病毒组物种注释结果(基于 reads 注释)

Fig.1 Taxonomic composition of viromes in Xinjiang cotton soil and desert soil based on reads annotation

viruses; 病毒; environmental samples; 环境样品; unclassified viruses; 未分类的病毒; uncultured marine virus; 未培养的海洋病毒; uncultured virus; 未培养的病毒; Microviridae; 微小噬菌体科; unclassified Microviridae; 未分类的微小噬菌体科; Gokushovirinae; 蘑菇状噬菌体亚科; Circoviridae; 环状病毒科; Circovirus; 环状病毒; unclassified ssDNA viruses; 未分类的单链 DNA 病毒; unclassified bacterial viruses; 未分类的细菌病毒; Nanoviridae; 矮缩病毒科; Satellites; 卫星病毒; Geminiviridae; 双生病毒科; Caudovirales; 有尾噬菌体目; Alphasatellitidae; 阿尔法卫星病毒科; Smacoviridae; 小环状 DNA 病毒科; Phycodnaviridae; 脱氧核糖核酸病毒科; Ortervirales; 逆转录病毒目; Polydnaviridae; 多分 DNA 病毒科; Mimiviridae; 拟菌病毒科; Inoviridae; 丝状病毒科; Chordopoxvirinae; 脊椎动物痘病毒亚科; Betabaculovirus; 乙型肝炎病毒属

2.4 病毒功能特征

利用 MEGAN 的 SEED 数据库对两种土壤的病毒功能进行分类, Level 1 水平的主要病毒功能占比类似, 均是“Phages, Prophages, Transposable elements”和“Phages, Prophages, Transposable elements, Plasmids”占比最高(图 2)。在棉花地土壤中注释到更多的病毒功能分类单元, 除以上两种功能外还注释到“Amino Acids and Derivatives”、“Carbohydrates”、“Cofactors, Vitamins, Prosthetic Groups, Pigments”和“Photosynthesis”。在 level 2 水平上, 两个土壤一共注释到 30 个病毒功能分类单元, 其中“Phage capsid proteins”所占比例最高, 其次是“Phage packaging machinery”和“Phage head and packaging”(图 2)。

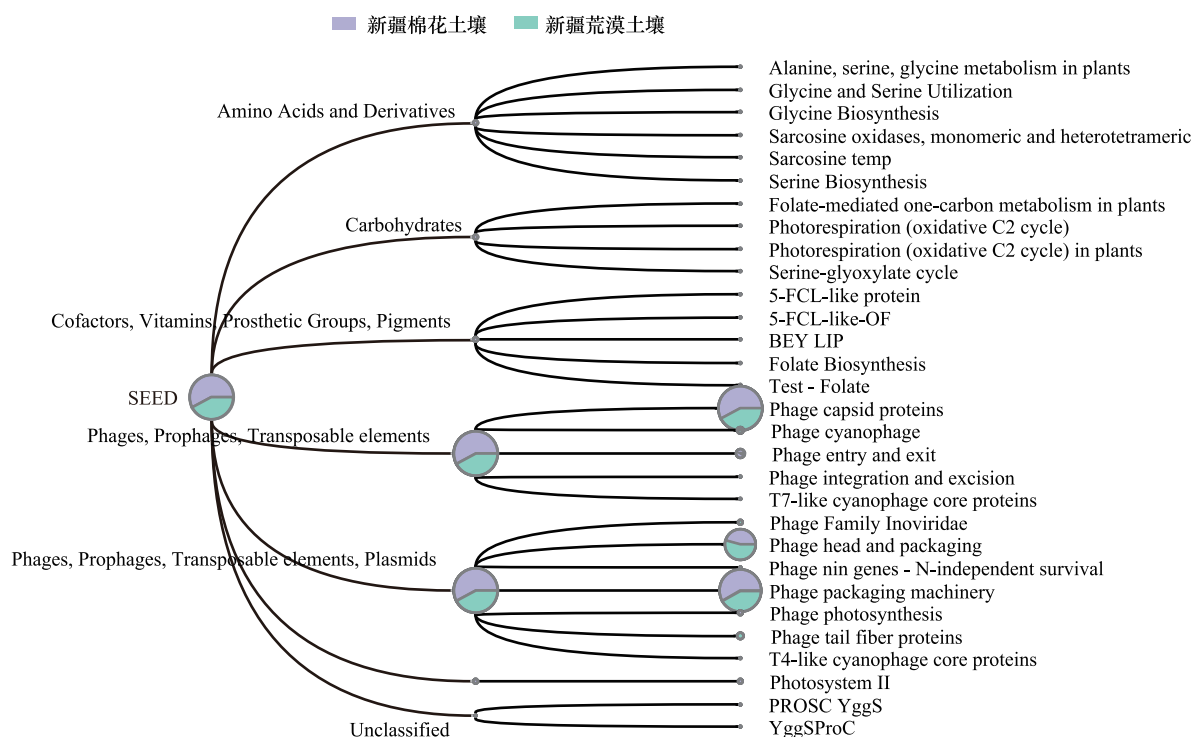


图 2 病毒功能注释

Fig.2 Functional profile of two viromes

Amino Acids and Derivatives:氨基酸及其衍生物;Carbohydrates:碳水化合物;Cofactors, Vitamins, Prosthetic Groups, Pigments:辅助因子,维生素,辅基和色素;Phages, Prophages, Transposable elements:噬菌体、原噬菌体、转座因子;Phages, Prophages, Transposable elements, Plasmids:噬菌体、原噬菌体、转座因子、质粒;Photosynthesis:光合作用;Alanine, serine, glycine metabolism in plants:植物丙氨酸、丝氨酸、甘氨酸代谢;Glycine and Serine Utilization:甘氨酸和丝氨酸利用;Glycine Biosynthesis:甘氨酸生物合成;Sarcosine oxidases, monomeric and heterotetrameric:肌氨酸氧化酶、单体和异四聚体;Sarcosine temp;Serine Biosynthesis:丝氨酸生物合成;Folate-mediated one-carbon metabolism in plants:叶酸介导的植物一碳代谢;Photorespiration:光呼吸;Serine-glyoxylate cycle:丝氨酸-乙醛酸循环;5-FCL-like protein:类 5-FCL 蛋白;5-FCL-like-OF;BEY LIP;Folate Biosynthesis:叶酸生物合成;Test - Folate;Phage capsid proteins:噬菌体衣壳蛋白;Phage cyanophage:噬藻体;Phage entry and exit;Phage integration and excision:噬菌体整合与删除;T7-like cyanophage core proteins:T7 型噬藻体核心蛋白;Phage Family Inoviridae:丝状病毒科噬菌体;Phage head and packaging:噬菌体头部和组装;Phage nin genes-N-independent survival;Phage packaging machinery;Phage photosynthesis:噬菌体光合作用;Phage tail fiber proteins:噬菌体尾纤维蛋白;T4-like cyanophage core proteins:T4 型噬藻体核心蛋白;Photosystem II:光系统 II;PROSC YggS;YggSProC

2.5 病毒基因组分析

通过 Virsorter 预测病毒序列和功能注释, 在棉花田土壤病毒组中注释到 711 条病毒 contigs (> 1 kb), 在荒漠土壤病毒组中注释到 1113 条病毒 contigs (> 1 kb)。两个土壤中的病毒 contigs 均主要注释到微小噬菌体科。大于 5 kb 的 contigs 一共有 147 条, 选择其中最长的两条 contigs 分析其功能基因组组成(图 3)。Contig 1 携带 14 个基因, 其中包括 8 个病毒相关功能基因 (protein affiliated to virus), 5 个未知功能基因 (unknown)

function) 和 1 个假设蛋白(hypothetical protein)。Contig 2 携带 18 个基因,其中包括 9 个病毒相关功能基因,1 个未知功能基因和 8 个假设蛋白(图 3)。

3 讨论

3.1 新疆两种土地利用方式下土壤病毒群落组成

土壤微生物在自然和人工生态系统中都扮演着十分重要的角色。标记基因和宏基因组学等研究方法极大地推动了我们土壤微生物的认知和应用,例如合理利用微生物来提高土壤肥力、增加作物产量,增强我们对陆地生态系统响应环境变化的理解^[33]。病毒也在陆地生态系统中发挥着重要的作用,例如通过侵染微生物来调控宿主代谢、影响宿主死亡率等^[34]。加州大学戴维斯分校 Joanne B. Emerson 副教授提出土壤病毒学是土壤微生物学研究的前沿热点,并指出未来可以探究土壤病毒对有机质降解、碳氮元素循环、温室气体排放以及农业生产的影响^[35],从而补充土壤微生物对生态过程的贡献。

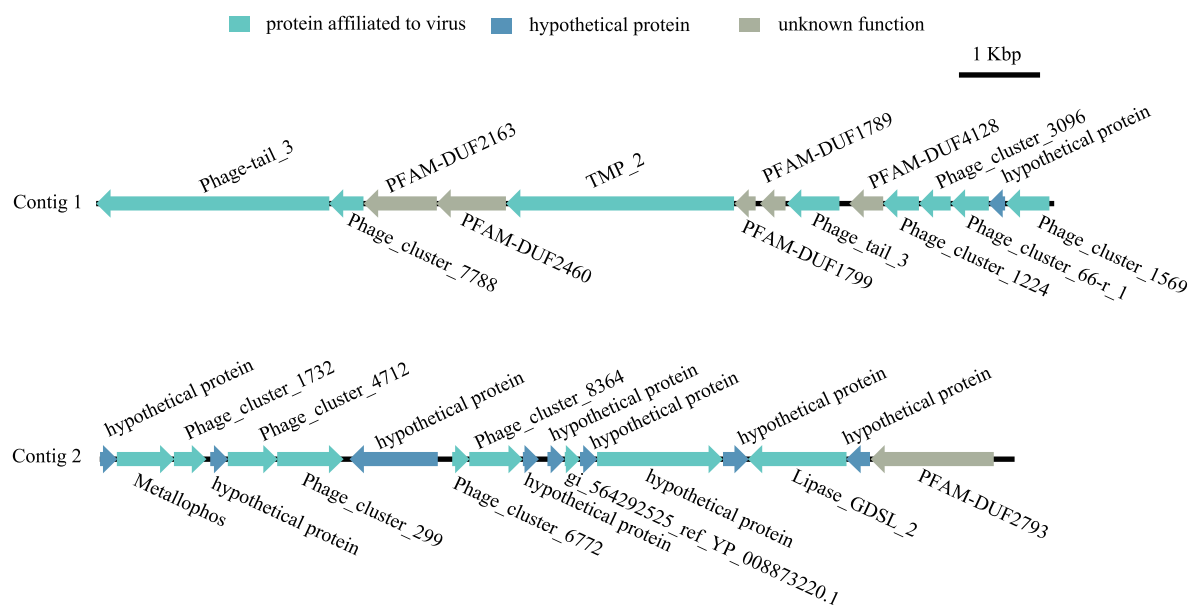


图 3 基因组示意图

Fig.3 Map analysis of viral contigs

protein affiliated to virus:病毒相关蛋白; hypothetical protein:假设蛋白; unknown function:未知功能

本研究通过宏病毒组学分析发现,新疆两种不同土地利用方式下的土壤均富集到大量的 ssDNA 病毒,其中以微小噬菌体科为主。ssDNA 病毒一般包括丝状病毒科(*Inoviridae*)和微小噬菌体科两个噬菌体病毒科和五个真核病毒科^[36]。微小噬菌体科是具有环状单链 DNA 基因组的二十面体的噬菌体家族,主要侵染肠杆菌、细胞内寄生菌和螺原体^[37]。微小噬菌体科在感染周期中的潜伏期短、子代释放量大^[38],从而广泛分布于环境中,如海洋、污水、沉积物和人体肠道等均有分布。2015 年 Brian Reavy 等人通过宏病毒组学研究发现,苏格兰滨海土壤和农田土壤中 ssDNA 病毒数量占比最大,以微小噬菌体科为主(分别占 84.6%和 50.32%)^[39]。而在红树林土壤中以动物病毒环状病毒科占主导,其次是微小噬菌体科^[16]。有研究指出多重置换扩增可能会偏好性地放大 ssDNA 病毒^[40],从而导致结果中 ssDNA 病毒比例偏高。在棉花地和荒漠土壤中,dsDNA 病毒有尾噬菌体目(包括长尾噬菌体科、肌尾噬菌体科和短尾噬菌体科)所占比例比较低。有尾噬菌体目在部分土壤样品中也占据主导地位,例如纳米布沙漠土壤^[11]、美国某些农田土壤^[41]以及我国新疆冰川土壤^[17]。

相比荒漠土壤,新疆棉花地土壤宏病毒组注释到更多的病毒科。这暗示着人为活动可能增加土壤病毒组成的复杂程度。王银亚等人通过磷脂脂肪酸(PLFA)方法分析发现,相比荒漠土壤,开垦土壤(农田)微生物

总 PLFA 含量和土壤微生物多样性指数显著增加。这可能是因为在开垦荒漠土壤的过程中,土壤水分、全氮和有机碳含量增加、盐分降低,从而影响微生物群落^[42]。在本研究中,棉花地土壤与荒漠土壤的理化性质差异显著,相比荒漠土壤,棉花地土壤具有更高的含水率、DOC、DON、 $\text{NH}_4^+\text{-N}$ 和 $\text{NO}_3^-\text{-N}$,以及更低的 pH。宿主群落组成和多样性以及土壤理化性质的差异可能导致棉花地土壤和荒漠土壤病毒组成的差异。先前的研究发现土壤病毒的群落组成受环境因子的影响,例如在消融的永久冻土中,宿主组成、土壤 pH、水分和土壤深度是影响病毒群落组成的主要因素^[2],而在南极土壤中,病毒群落组成主要受到土壤 pH、钙离子浓度和海拔的影响^[12]。但对于环境因子如何影响病毒群落结构还需要更多的实验和数据进行解释。另外,农业活动,如灌溉、有机肥施用也可能引入外源病毒。仅在棉花地土壤中检测到有花椰菜花叶病毒科、逆转录病毒科、裸露病毒科、多分 DNA 病毒科、杆状病毒科和囊泡病毒科,其中花椰菜花叶病毒科是侵染植物的 dsDNA 病毒,裸露病毒科、多分 DNA 病毒科、杆状病毒科和囊泡病毒科是昆虫相关的病毒。此外,植物病毒矮缩病毒科和双生病毒科在棉花地中检测出的比例均高于荒漠土壤。植物病毒主要通过昆虫进行传播^[43],相对于荒漠而言,昆虫较多的棉花地土壤更可能广泛传播植物病毒。本研究推测不同土地利用方式可能导致了宿主生物组成(微生物和动植物)的差异,从而影响土壤病毒的分布。

3.2 土壤病毒功能特性

棉花地和荒漠土壤宏病毒组的主要病毒功能注释结果相似,均主要注释到与病毒结构相关的功能,其中“Phage capsid proteins”和“Phage packaging machinery”占比最高(图 2)。两条全长病毒基因组序列的分析也表明可注释的编码基因均以病毒结构相关功能为主,未知功能蛋白仍占较高比例(图 3)。“phage capsid proteins”是病毒相关的衣壳蛋白,可作为某些特定病毒类群的标记基因,例如 T4 噬菌体的 *g23* 基因^[44-45]和微小噬菌体科的 VP1 蛋白^[46]。在纳米布沙漠土壤病毒功能中,“phages, prophages, transposable elements, plasmids”占比最高(49%)^[11]。这也与一些美国农田土壤病毒的主要功能类似^[41],而部分病毒功能分类单元在农田表层和深层土壤中的相对丰度差异较大,在深层土壤中检测到更多的噬菌体结构相关功能(如噬菌体衣壳、尾部和颈部纤维)。但在中国南方红壤中,仅有少部分病毒功能在玉米根际与非根际土壤之间存在显著差异^[15]。此外,相对于其他土壤,棉花地和荒漠土壤病毒组功能类群较为单一^[2, 41]。这也可能受到土样样品数量的限制,对于不同土地利用方式下土壤病毒功能特征与差异还需要进一步扩大样品进行探索。此外,由于目前我们对病毒功能基因的分析主要依赖现有数据库,而相关数据库还不完善,功能注释仍存在局限性,因此土壤病毒功能特征还有待进一步地探索和挖掘。

4 结论

本研究通过宏病毒学方法分析了新疆两种土地利用方式下土壤病毒的群落组成和功能特征。棉花地和荒漠土壤的理化性质存在显著差异,两种土壤分别注释到 20 个和 15 个病毒科,均以 ssDNA 病毒为优势类群,其中以微小噬菌体科占绝对优势。在棉花地土壤中检测到更高比例的植物病毒(矮缩病毒科和双生病毒科)。此外,仅在棉花地土壤中检测到与植物相关的花椰菜花叶病毒科,与昆虫相关的裸露病毒科、多分 DNA 病毒科、杆状病毒科和囊泡病毒科。这些结果表明土壤病毒群落组成可能受到与土地利用方式相关的人为活动、土壤理化性质和动植物的影响。在两种土壤中共获得 1824 条病毒 contigs,主要匹配到微小噬菌体科。用 SEED 数据库对病毒功能进行注释,棉花地和荒漠土壤病毒组的功能注释均以病毒结构相关功能基因占比最高。本研究探索了新疆农田土壤和荒漠土壤病毒组的特征,未来研究人员还需要在更大时空尺度和基于更多样品的土地利用方式上探究土壤病毒生态学,并进一步阐明土壤病毒对陆地生态过程的贡献。

参考文献(References):

- [1] Williamson K E, Fuhrmann J J, Wommack K E, Radosevich M. Viruses in soil ecosystems: an unknown quantity within an unexplored territory. *Annual Review of Virology*, 2017, 4: 201-219.
- [2] Emerson J B, Roux S, Brum J R, Bolduc B, Woodcroft B J, Jang H B, Singleton C M, Solden L M, Naas A E, Boyd J A, Hodgkins S B, Wilson

- R M, Trubl G, Li C S, Frolking S, Pope P B, Wrighton K C, Crill P M, Chanton J P, Saleska S R, Tyson G W, Rich V I, Sullivan M B. Host-linked soil viral ecology along a permafrost thaw gradient. *Nature Microbiology*, 2018, 3(8): 870-880.
- [3] Pratama A A, van Elsas J D. The 'neglected' soil virome-potential role and impact. *Trends in Microbiology*, 2018, 26(8): 649-662.
- [4] Trubl G, Jang H B, Roux S, Emerson J B, Solonenko N, Vik D R, Solden L, Ellenbogen J, Runyon A T, Bolduc B, Woodcroft B J, Saleska S R, Tyson G W, Wrighton K C, Sullivan M B, Rich V I. Soil viruses are underexplored players in ecosystem carbon processing. *mSystems*, 2018, 3(5): e00076-18.
- [5] Li Y, Liu H Y, Pan H, Zhu X Y, Liu C, Zhang Q C, Luo Y, Di H J, Xu J M. T4-type viruses: important impacts on shaping bacterial community along a chronosequence of 2000-year old paddy soils. *Soil Biology and Biochemistry*, 2019, 128: 89-99.
- [6] Kuzyakov Y, Mason-Jones K. Viruses in soil: nano-scale undead drivers of microbial life, biogeochemical turnover and ecosystem functions. *Soil Biology and Biochemistry*, 2018, 127: 305-317.
- [7] Weinbauer M G. Ecology of prokaryotic viruses. *FEMS Microbiology Reviews*, 2004, 28(2): 127-181.
- [8] Hurwitz B L, U'Ren J M. Viral metabolic reprogramming in marine ecosystems. *Current Opinion in Microbiology*, 2016, 31: 161-168.
- [9] 韩丽丽, 于丹婷, 贺纪正. 土壤病毒生态学研究方法. *生态学报*, 2017, 37(6): 1749-1756.
- [10] Trubl G, Hyman P, Roux S, Abedon S T. Coming-of-age characterization of soil viruses: a user's guide to virus isolation, detection within metagenomes, and viromics. *Soil Systems*, 2020, 4(2): 23.
- [11] Adriaenssens E M, van Zyl L, de Maayer P, Rubagotti E, Rybicki E, Tuffin M, Cowan D A. Metagenomic analysis of the viral community in Namib Desert hypoliths. *Environmental Microbiology*, 2015, 17(2): 480-495.
- [12] Adriaenssens E M, Kramer R, van Goethem M W, Makhalyane T P, Hogg I, Cowan D A. Environmental drivers of viral community composition in Antarctic soils identified by viromics. *Microbiome*, 2017, 5: 83.
- [13] Han L L, Yu D T, Zhang L M, Shen J P, He J Z. Genetic and functional diversity of ubiquitous DNA viruses in selected Chinese agricultural soils. *Scientific Reports*, 2017, 7: 45142.
- [14] Starr E P, Nuccio E E, Pett-Ridge J, Banfield J F, Firestone M K. Metatranscriptomic reconstruction reveals RNA viruses with the potential to shape carbon cycling in soil. *Proceedings of the National Academy of Sciences of the United States of America*, 2019, 116(51): 25900-25908.
- [15] Bi L, Yu D T, Du S, Zhang L M, Zhang L Y, Wu C F, Xiong C, Han L L, He J Z. Diversity and potential biogeochemical impacts of viruses in bulk and rhizosphere soils. *Environmental Microbiology*, 2020, doi: 10.1111/1462-2920.15010.
- [16] Jin M, Guo X, Zhang R, Qu W, Gao B L, Zeng R Y. Diversities and potential biogeochemical impacts of mangrove soil viruses. *Microbiome*, 2019, 7: 58.
- [17] Han L L, Yu D T, Zhang L M, Wang J T, He J Z. Unique community structure of viruses in a glacier soil of the Tianshan Mountains, China. *Journal of Soils and Sediments*, 2017, 17(3): 852-860.
- [18] Yu D T, Han L L, Zhang L M, He J Z. Diversity and distribution characteristics of viruses in soils of a marine-terrestrial ecotone in east China. *Microbial Ecology*, 2018, 75(2): 375-386.
- [19] McGrath D A, Smith C K, Gholz H L, de Assis Oliveira F. Effects of land-use change on soil nutrient dynamics in Amazonia. *Ecosystems*, 2001, 4(7): 625-645.
- [20] de Vries F T, Thébaud E, Liiri M, Birkhofer K, Tsiafouli M A, Bjørnlund L, Jørgensen H B, Brady M V, Christensen S, de Ruiter P C, d'Hertefeldt T, Frouz J, Hedlund K, Hemerik L, Hol W H G, Hotes S, Mortimer S R, Setälä H, Sgardelis S P, Uteseny K, van der Putten W H, Wolters V, Bardgett R D. Soil food web properties explain ecosystem services across European land use systems. *Proceedings of the National Academy of Sciences of the United States of America*, 2013, 110(35): 14296-14301.
- [21] Suzuki M T, Taylor L T, DeLong E F. Quantitative analysis of small-subunit rRNA genes in mixed microbial populations via 5'-nuclease assays. *Applied and Environmental Microbiology*, 2000, 66(11): 4605-4614.
- [22] Chen S F, Zhou Y Q, Chen Y R, Gu J. Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, 2018, 34(17): i884-i890.
- [23] Kopylova E, Noé L, Touzet H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics*, 2012, 28(24): 3211-3217.
- [24] Nurk S, Meleshko D, Korobeynikov A, Pevzner P A. metaSPAdes: a new versatile metagenomic assembler. *Genome Research*, 2017, 27(5): 824-834.
- [25] Buchfink B, Xie C, Huson D H. Fast and sensitive protein alignment using DIAMOND. *Nature Methods*, 2015, 12(1): 59-60.
- [26] NCBI Resource Coordinators. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Research*, 2016, 44(D1): D7-D19.
- [27] Pruitt K D, Tatusova T, Maglott D R. NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Research*, 2007, 35(S1): D61-D65.

- [28] Zhou Y, Liang Y J, Lynch K H, Dennis J J, Wishart D S. PHAST: a fast phage search tool. *Nucleic Acids Research*, 2011, 39 (S1): W347-W352.
- [29] Hyatt D, Chen G L, LoCascio P F, Land M L, Larimer F W, Hauser L J. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics*, 2010, 11: 119.
- [30] Huson D H, Auch A F, Qi J, Schuster S C. MEGAN analysis of metagenomic data. *Genome Research*, 2007, 17(3): 377-386.
- [31] Roux S, Enault F, Hurwitz B L, Sullivan M B. VirSorter: mining viral signal from microbial genomic data. *PeerJ*, 2015, 3(348): e985
- [32] Sullivan M J, Petty N K, Beatson S A. Easyfig: a genome comparison visualizer. *Bioinformatics*, 2011, 27(7): 1009-1010.
- [33] Fierer N. Embracing the unknown: disentangling the complexities of the soil microbiome. *Nature Reviews Microbiology*, 2017, 15 (10): 579-590.
- [34] Rohwer F, Thurber R V. Viruses manipulate the marine environment. *Nature*, 2009, 459(7244): 207-212.
- [35] Emerson J B. Soil viruses: a new hope. *mSystems*, 2019, 4(3): e00120-19.
- [36] King A M Q, Adams M J, Carstens E B, Lefkowitz E J. *Virus Taxonomy: Ninth Report of the International Committee on Taxonomy of Viruses*. Amsterdam; Elsevier, 2011.
- [37] Brentlinger K L, Hafenstein S, Novak C R, Fane B A, Borgon R, McKenna R, Agbandje-McKenna M. *Microviridae*, a family divided: Isolation, characterization, and genome sequence of ϕ MH2K, a bacteriophage of the obligate intracellular parasitic bacterium *Bdellovibrio bacteriovorus*. *Journal of Bacteriology*, 2002, 184(4): 1089-1094.
- [38] Labonté J M, Suttle C A. Metagenomic and whole-genome analysis reveals new lineages of gokushoviruses and biogeographic separation in the sea. *Frontiers in Microbiology*, 2013, 4: 404.
- [39] Reavy B, Swanson M M, Cock P J A, Dawson L, Freitag T E, Singh B K, Torrance L, Mushegian A R, Taliany M. Distinct circular single-stranded DNA viruses exist in different soil types. *Applied and Environmental Microbiology*, 2015, 81(12): 3934-3945.
- [40] Kim K H, Bae J W. Amplification methods bias metagenomic libraries of uncultured single-stranded and double-stranded DNA viruses. *Applied and Environmental Microbiology*, 2011, 77(21): 7663-7668.
- [41] Liang X L, Wagner R E, Zhuang J, DeBruyn J M, Wilhelm S W, Liu F, Yang L, Staton M E, Sherfy A C, Radosevich M. Viral abundance and diversity vary with depth in a southeastern United States agricultural ultisol. *Soil Biology and Biochemistry*, 2019, 137: 107546.
- [42] 王银亚, 李晨华, 马健. 开垦对荒漠土壤微生物群落结构特征的影响. *中国沙漠*, 2017, 37(3): 514-522.
- [43] Whitfield A E, Falk B W, Rotenberg D. Insect vector-mediated transmission of plant viruses. *Virology*, 2015, 479-480: 278-289.
- [44] Filée J, Tétart F, Suttle C A, Krusch H M. Marine T4-type bacteriophages, a ubiquitous component of the dark matter of the biosphere. *Proceedings of the National Academy of Sciences of the United States of America*, 2005, 102(35): 12471-12476.
- [45] Liu J J, Wang G H, Wang Q, Liu J D, Jin J, Liu X B. Phylogenetic diversity and assemblage of major capsid genes (*g23*) of T4-type bacteriophages in paddy field soils during rice growth season in Northeast China. *Soil Science and Plant Nutrition*, 2012, 58(4): 435-444.
- [46] Quaiser A, Dufresne A, Ballaud F, Roux S, Zivanovic Y, Colombet J, Sime-Ngando T, Francez A J. Diversity and comparative genomics of *Microviridae* in *Sphagnum*-dominated peatlands. *Frontiers in Microbiology*, 2015, 6: 375.