



一株枯草芽孢杆菌原噬菌体 Bsu-yong1 的基因组分析

刘文才¹, 钱敏桦¹, 潘灵婷¹, 林威², 许丽华¹, 任晓清¹, 李登峰^{1*}, 童贻刚²

1 宁波大学海洋学院 海洋生物工程重点实验室, 浙江 宁波 315832

2 北京化工大学生命科学学院, 北京 100029

刘文才, 钱敏桦, 潘灵婷, 林威, 许丽华, 任晓清, 李登峰, 童贻刚. 一株枯草芽孢杆菌原噬菌体 Bsu-yong1 的基因组分析[J]. 微生物学报, 2023, 63(10): 3905-3922.

LIU Wencai, QIAN Minhua, PAN Lingting, LIN Wei, XU Lihua, REN Xiaoqing, LI Dengfeng, TONG Yigang. Genomic analysis of a *Bacillus subtilis* prophage vB_Bsu-yong1[J]. Acta Microbiologica Sinica, 2023, 63(10): 3905-3922.

摘要: 【目的】枯草芽孢杆菌(*Bacillus subtilis*)是在自然界中广泛存在的革兰氏阳性菌, 其抗逆性极强, 能抑制大多数有害菌的繁殖, 是常用的产酶菌, 用其生产的蛋白酶、淀粉酶占全球工业酶产量的 50%。原噬菌体(prophage)整合在宿主基因组中, 可为宿主提供基因和生物学功能, 非常具有研究价值。以往, 有关 *B. subtilis* 原噬菌体的报道主要集中于缺陷型原噬菌体(defective prophage), 本研究对一株非缺陷型活性原噬菌体(active prophage)的基因组进行解析, 以扩充对非缺陷型原噬菌体的认知。【方法】使用丝裂霉素 C 从枯草芽孢杆菌中诱导一株噬菌体, 命名为 *Bacillus* phage Bsu-yong1 (简称 Bsu-yong1)。对 Bsu-yong1 进行负染、透射电镜(transmission electron microscopy, TEM)观察, 用 Illumina MiSeq 测定其基因组序列、综合运用生物信息学工具对其进行基因功能注释和系统进化分析。【结果】Bsu-yong1 与 PBSX 类缺陷型原噬菌体在形态上相似, 但 Bsu-yong1 具有完整的噬菌体基因组, 这与缺陷型原噬菌体不同, 后者在包装过程中不能正确包裹自身的基因组, 而是随机包裹一段宿主染色体。Bsu-yong1 基因组全长为 43 590 bp, G+C 含量为 41%, 含有 62 个开放阅读框(open reading frame, ORF), 呈模块化分布。Bsu-yong1 拥有基因编码 T7SS 效应器 LXG 多态性毒素(T7SS effector LXG polymorphic toxin)、ImmA/IrrE 蛋白和 SM11/KNR4 蛋白。前二者为细菌毒素(toxin), 后者为抗毒素(antitoxin), toxin-antitoxin 是细菌免疫系统重要成员, 参与菌间竞争与环境适应。此前, 尚未有编码 LXG polymorphic toxin 的基因在

资助项目: 国家重点研发计划(2018YFA0903000); 宁波市重点研发计划(2022Z170)

This work was supported by the National Key Research and Development Program of China (2018YFA0903000) and the Key Research and Development Project of Ningbo (2022Z170).

*Corresponding author. E-mail: lidengfeng@nbu.edu.cn

Received: 2023-03-02; Accepted: 2023-05-16; Published online: 2023-05-22

噬菌体中被发现和报道。在基于全基因组比对构建的蛋白谱进化树(proteomic tree)中, Bsu-yong1 与噬菌体 sv105、rho14、vB_BteM-A9Y 聚集形成一个独立的进化支(clade), 基因组比对显示它们基因组的复制与调控模块具有高度保守性, 它们共享 29 个核心基因(core gene), 均具有 PBSX 样形态特征。Bsu-yong1 与其他噬菌体的进化距离较远。将 Bsu-yong1 与所有噬菌体进行比对, 得到的成对序列比较(pairwise sequence comparison, PASC)最大值为 46.72%, 小于属边界值(70%)。

【结论】 vB_Bsu-yong1 在有尾纲中代表一个新的未知的属; 建议构建一个新的科(family), 该科由 Bsu-yong1 与噬菌体 sv105、rho14、vB_BteM-A9Y 组成。vB_Bsu-yong 携带免疫相关基因, 它可能有利于宿主在菌间竞争中获胜和适应环境。本研究丰富了噬菌体基因数据库, 拓展了对芽孢杆菌活性原噬菌体的认知。

关键词: 原噬菌体; 基因组; 枯草芽孢杆菌; 新属; 新科

Genomic analysis of a *Bacillus subtilis* prophage vB_Bsu-yong1

LIU Wencai¹, QIAN Minhua¹, PAN Lingting¹, LIN Wei², XU Lihua¹, REN Xiaoqing¹, LI Dengfeng^{1*}, TONG Yigang²

¹ Key Laboratory of Marine Biotechnology, School of Marine Sciences, Ningbo University, Ningbo 315832, Zhejiang, China

² College of Life Science and Technology, Beijing University of Chemical Technology, Beijing 100029, China

Abstract: [Objective] *Bacillus subtilis*, a Gram-positive bacterium ubiquitous in the nature, has strong stress tolerance and environmental adaptability and can inhibit the propagation of a variety of harmful bacteria. It is a common and important enzyme-producing bacterium, with the yield of proteases and amylases accounting for 50% of the industrial enzyme yield. Prophages integrated in host genomes are capable of providing additional genes and endow hosts with biological properties, demonstrating great research significance. The available studies about *B. subtilis* prophages focus on the defective ones. In this study, we analyzed the genome of an active prophage to expand the knowledge about non-defective prophages. **[Methods]** We induced a phage from *B. subtilis* by using mitomycin C and named the phage as *Bacillus* phage Bsu-yong1. We employed transmission electron microscopy (TEM) to observe the negatively stained Bsu-yong1, Illumina MiSeq to sequence the genome of Bsu-yong1, and bioinformatics tools to annotate the genome. Furthermore, a proteomic tree of Bsu-yong1 was built based on genome-wide sequence alignment. **[Results]** Bsu-yong1 showed the morphology similar to that of PBSX-like defective prophages of *B. subtilis*. Unlike defective prophages which package DNA segments from random portions of the host genome, Bsu-yong1 was capable of packaging its own genome. The full-length genome of Bsu-yong1 was 43 590 bp with the G+C content of 41%. A total of 62 open reading frames (ORFs) were predicted with modular arrangement in the Bsu-yong1 genome. Bsu-yong1 harbored the ORFs encoding T7SS effector LXG polymorphic toxin, IMMA/IrrE and SMI1/KNR4. The former two were bacterial toxins and the later was an antitoxin. Toxins and antitoxins are members of the bacterial

immune system, participating in bacterial competition and environmental adaptation. The gene encoding LXG polymorphic toxin in bacteriophages had never been reported before. A proteomic tree was built based on the whole genome alignment, revealing a long evolution distance between Bsu-yong1 and other phages. In the proteomic tree, Bsu-yong1 clustered with phages sv105, rho14, and vB_BteM-A9Y to form a monophyletic clade. These four phages all had PBSX-like morphology and shared 29 core genes. Pairwise sequence comparison (PASC) illustrated that the highest nucleotide sequence similarity between Bsu-yong1 and known phages was only 46.72%, which was much lower than the threshold (70%) to define a genus. **[Conclusion]** Bsu-yong1 represented a new unknown genus. We suggest the establishment of a novel family composed of Bsu-yong1, sv105, rho14, and vB_BteM-A9Y, which share 29 core genes. Bsu-yong1 carries immunity-related genes, which may benefit the host in bacterial competition and environmental adaptation. The findings enrich the bacteriophage gene database and enrich the knowledge of active *Bacillus* prophages.

Keywords: prophage; genome; *Bacillus subtilis*; novel genus; novel family

噬菌体(phage)是能特异性感染细菌、真菌的病毒。其分布广泛,是自然界中丰度最大的生命体^[1]。噬菌体的生命周期有溶原和裂解 2 种状态^[2]。原噬菌体(prophage)一般只经历溶原周期,此周期中溶原噬菌体将自身的基因组整合到宿主染色体中的特定位点,并可长期随宿主 DNA 复制而同步复制。正常情况下,原噬菌体在溶原周期中并不会大量增殖引起宿主裂解。一些外界因素,如丝裂霉素、高温、紫外线等,可以诱导部分宿主细胞中的原噬菌体大量增殖、释放,并裂解宿主^[3-6]。这些可被诱导的原噬菌体又被称为活性原噬菌体(active prophage)。

噬菌体并不总是细菌的天敌。研究发现噬菌体尤其原噬菌体介导的水平基因转移,能够驱动原核生物基因组的进化^[7],这种基因的水平转移在细菌中是常见的。原噬菌体在细菌演化过程中发挥着关键作用^[8]。这种进化驱动是相互的。细菌可在 DNA 复制过程中获得寄生在其基因组上的原噬菌体的某些基因片段,从而获得新功能或强化某些功能。一些原噬菌体基因可帮助细菌提高种间竞争力和环境适应力。噬菌体亦可以在其 DNA 合成过程中从宿主那里获得某些基因,如

毒力基因、免疫基因等,以获得适应性优势,这些基因不仅能帮助病毒更好地适应宿主,使自己插入和隐藏在宿主的遗传物质内,还能参与调控宿主的某些代谢、免疫途径,增强宿主的某些功能来提升宿主在细菌群落中的竞争能力和对环境的适应能力。

枯草芽孢杆菌(*Bacillus subtilis*)属革兰氏阳性菌,具有典型的芽孢杆菌特征,好氧、能分泌多种酶、抗逆性极强和能抑制大多数有害菌的繁殖^[9-10]。枯草芽孢杆菌是饲料行业常添加的有益菌,也是酶工业领域常用的产酶菌,枯草芽孢杆菌生产的蛋白酶、淀粉酶占整个工业酶行业产量的 50%^[11]。近年来对枯草芽孢杆菌的研究逐渐聚焦到遗传学与分子生物学领域,研究内容包括寻找特定功能的基因,并将其克隆到需要的物种中,或者通过诱变与基因工程手段对枯草芽孢杆菌生产菌进行遗传改造等^[12]。整合在宿主染色体上的原噬菌体与宿主的相互作用机制也被大量研究,如原噬菌体辅助宿主细胞代谢、引起细胞免疫活动、调控细胞毒素生产等^[13-14]。然而,相关研究主要集中在缺陷型原噬菌体(defective prophage)上。缺陷型原噬菌体在枯草芽孢杆菌中

普遍存在,研究表明所有枯草芽孢杆菌菌株的染色体上均携带有 PBSX 或与其类似的缺陷性原噬菌体[被称为 PBSX 样缺陷型原噬菌体 (PBSX-like defective prophage)]^[15-16]。缺陷型原噬菌体在包装过程中不装配自身的基因组 DNA。在 PBSX 病毒粒子中不含有编码自身蛋白的 DNA,而是随机包裹一段约 13 kb 的宿主染色体片段。成熟的 PBSX 噬菌体颗粒为头尾状,其头部直径约为 45 nm,可收缩的尾部长约 200 nm。缺陷型原噬菌体的存在有利于宿主菌在营养贫瘠条件下的正常生长^[17]。

有关非缺陷型 *B. subtilis* 原噬菌体的报道罕见。本课题组用丝裂霉素 C 处理一株自青蟹养殖水体中分离到的枯草芽孢杆菌,成功诱导出 1 株非缺陷型原噬菌体,并命名为 *Bacillus phage Bsu-yong1* (简称 Bsu-yong1)。对 Bsu-yong1 的基因组进行了测序和分析。Bsu-yong1 拥有 1 个从未在噬菌体中被发现和报道的编码毒力蛋白的基因;系统进化分析显示 Bsu-yong1 代表 1 个新的未知的属(genus),与其他 3 个 PBSX 样噬菌体构成 1 个新的科(family)。本研究丰富了噬菌体基因数据库,拓展了对芽孢杆菌活性原噬菌体的认知,为进一步阐明原噬菌体与宿主的相互作用关系奠定了一定的基础。

1 材料与方法

1.1 宿主菌的来源及培养

枯草芽孢杆菌菌株由本课题组分离自青蟹养殖水体,菌株分离与扩大培养均使用 LB 海水培养基(胰蛋白胨 10 g,酵母提取物 5 g,用过滤海水定容至 1 L, pH 调至 7.2,高压灭菌),培养温度为 29 °C;在 BLASTn 比对中,其 16S rRNA 基因序列与枯草芽孢杆菌 KCTC 13429 株 (*B. subtilis* strain KCTC 13429) (NZ_CP029465) 100%同源。

1.2 噬菌体诱导

取新鲜培养的对数期枯草芽孢杆菌菌液 5 mL 于试管中,加入 5 μ L 丝裂霉素(5 mg/mL),使其终浓度为 5 μ g/mL。置于摇床(29 °C、120 r/min)孵育 2 h 后,取裂解液离心 10 min (4 °C、10 000 \times g),取上清依次经 0.45、0.22 μ m 孔径的滤膜过滤,滤过液即噬菌体悬液。

1.3 噬菌体形态观察

取上述噬菌体悬液 1 mL 加入透析袋中,置于 0.01 mol/L 磷酸盐缓冲液(phosphate buffered solution, PBS)中透析 2 h,从透析袋中取出噬菌体悬液,置于 4 °C 备用。制片时,取 10 μ L 噬菌体悬液于铜网上,静置 10 min 后,用 1%磷钨酸染色 30 s,干燥后使用电子投射显微镜 (Hitachi-7650)观察其形态。

1.4 噬菌体活性检验

取新鲜培养的对数期枯草芽孢杆菌菌液 100 μ L,均匀涂布于 LB 海水固体培养基平板上,取 1.2 中获得的噬菌体悬液,滴到涂布有枯草芽孢杆菌的 LB 海水固体培养基平板上,每个平板点 4 个区域,每个区域点 3 μ L。自然晾干后,倒置于培养箱(29 °C)中培养,每隔 1 h 观察 1 次。

1.5 噬菌体 DNA 提取

取噬菌体悬液,添加 RNase A 和 DNase I (终浓度均为 1 μ g/mL),37 °C 过夜消化以去除污染的核酸,80 °C 灭活 15 min 后,高速离心 (35 000 \times g、1 h),弃上清;沉淀用适量 1 \times PBS 充分悬浮后,用高纯度核酸提取试剂盒(罗氏),按照说明书提取噬菌体 DNA。

1.6 全基因组测序与注释

使用 Illumina 的 NEBNext Ultra II DNA 文库制备试剂盒,参照其说明书构建文库。用 Illumina MiSeq 平台进行测序。测序结果的评估用 FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>)进行;再使用 Trimmomaticv0.36 软件删除低质量(Q-value<20)的测序 reads 和接头;

用 SPAdes 3.13.0 软件完成序列拼接。

将噬菌体基因组与 GenBank 中的 nr 数据库中的序列进行 BLASTn (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) 比对, 得到的同源性最高(同源高达 100%)的 4 个序列均为枯草芽孢杆菌基因组。使用在线工具 PHASTER (Phage Search Tool Enhanced Release) (<http://phaster.ca/>)^[18] 预测这 4 株菌中可能的 prophages, 并将这些 prophages 与 Bsu-yong1 比对, 分析 Bsu-yong1 在枯草芽孢杆菌中的整合位置。

使用 RAST 在线工具(<https://rast.nmpdr.org/rast.cgi>)^[19] 进行初步的开放阅读框(open reading frame, ORF) 预测与注释, 再运用 HMMER (<https://www.ebi.ac.uk/Tools/hmmer/search/hmmscan>) 与 HHpred (<https://toolkit.tuebingen.mpg.de/tools/hhpred>)^[20-21] 在线服务器进行比对校准 (E-value $\leq 10^{-5}$, possibility $>96\%$)。分别将各 ORF 在均被上传到 GenBank 中(nr 数据库), 进行 BLASTp (<https://blast.ncbi.nlm.nih.gov/>) 比对。使用在线数据库 (<http://arpcard.mcmaster.ca>) 进行抗生素耐药基因与毒力基因预测。使用在线工具 (http://www.mgc.ac.cn/VFs/search_VFs.htm) 进行毒力基因预测。

1.7 系统进化地位分析

以 2022 年 3 月国际病毒分类委员会 (International Committee on Taxonomy of Viruses, ICTV) 发布的病毒分类系统为标准, 运用在线软件 ViPTree3.3 (<https://www.genome.jp/viptree>) 构建原噬菌体 Bsu-yong1 与 66 个噬菌体的全基因组蛋白谱树。所述的 66 个参考噬菌体的选取标准如下: 首先运用 ViPTree3.3 构建 1 个含有 Bsu-yong1 与数据库中所有 6 202 株病毒的总进化树, 从中选取与 Bsu-yong1 进化距离最近的 2 个噬菌体 *Spizizenvirus* sv105 (NC_048631) 和 *Bacillus* phage rho14 (OM236514); 选取在

BLASTn 比对中与 Bsu-yong1 具有最高同源性的噬菌体 *Bacillus* phage vB_BteM-A9Y (ON528935); 选取在 BLASTn 比对中与 Bsu-yong1 基因组 100% 同源的菌株中预测的 1 个原噬菌体 Jimmer2 (*Brevibacillus* phage Jimmer2) (NC_041976); 选取形态上与 Bsu-yong1 相似的缺陷型原噬菌体 *Bacillus* phage PBP180 (KC847113); 从 ICTV 分类系统中有尾纲(Caudoviricetes) 下的 4 个目 (*Crassvirales*、*Kirjokansivirales*、*Methanobavirales* 和 *Thumleimavirales*) 中各随机选择 2 个代表种; 由 ICTV 分类系统中有尾纲下不属于任何目的独立的 33 个科中, 每科选取 1-2 个代表种。

使用成对序列比较 (pairwise sequence comparison, PASC) 在线工具 (<http://www.ncbi.nlm.nih.gov/sutils/pasc/>)^[22] 计算 Bsu-yong 与数据库中所有病毒间的全基因组核苷酸相似度, 并找到与之相似度最高的病毒(噬菌体)。使用 EasyFig2.2 软件^[23] 比对 Bsu-yong 与同源性最高的噬菌体的基因组, 并输出比对图谱。使用 EzGenome web (<http://www.ezbiocloud.net/ezgenome/ani>)^[24] 分析 Bsu-yong 与同源性最高的噬菌体间的平均核酸一致性值 (average nucleotide identity value, ANI)。使用 GGDC 在线工具 (<http://ggdc.dsmz.de>)^[25] 计算 Bsu-yong 与同源性最高的噬菌体之间的 DNA 分子原位杂交值 (insilico DNA-DNA hybridization value, isDDH)。

2 结果与分析

2.1 原噬菌体的诱导

孵育 2 h 后, 添加了丝裂霉素的 *B. subtilis* 培养液变澄清, 而同期对照组培养液的浊度一直在增高, 二者区别非常明显。

2.2 噬菌体形态观测

电镜观察显示 Bsu-yong1 呈现肌尾噬菌体

样(Myovirus-like)形态特征。具有小的二十面体头部和可收缩的尾(图 1); 头部直径为 (44 ± 1) nm; 尾部长而宽度, 长度为 (222 ± 3) nm, 宽度为 (22 ± 1) nm, 这与之前报道的枯草芽孢杆菌缺陷型原噬菌体 PBPX 和 PBP180 的形态结构相似^[26], 但 Bsu-yong1 的尾部长于 PBPX 和 PBP180 噬菌体的, 后二者的尾部均长约 200 nm^[16]。且 Bsu-yong1 的尾部有明显的尾纤维(如图 1 中的箭头所示), 而在 PBP180 及其他 PBPX 样缺陷型原噬菌体尾部均未发现尾纤维。

2.3 噬菌体活性检验

将噬菌体悬液点在涂布有枯草芽孢杆菌的 LB 海水固体培养基平板上, 培养 6 h 后, 在菌苔上的点样区域出现明显的噬菌斑(图 2)。

2.4 基因组序列测定与分析

使用高通量测序技术获得了完整的 Bsu-yong1 基因组, 测序深度为 656。Bsu-yong1 的基因组全长为 43 590 bp, GC 含量为 41% (图 3), 含有 63 个开放阅读框(ORF)。已将 Bsu-yong1 的完整基因组序列提交至 GenBank, 接收号为 OP918669。

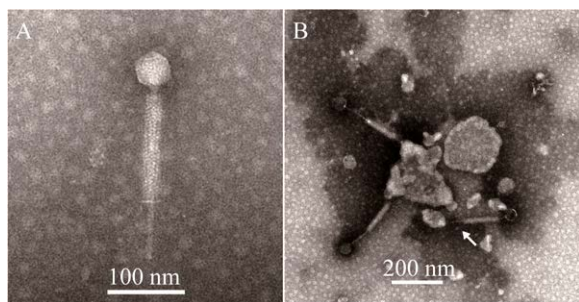


图 1 原噬菌体 Bsu-yong1 的电镜图

Figure 1 Transmission electron micrographs of *Bacillus* prophage Bsu-yong1. A: *Bacillus* phage Bsu-yong1 particles possess an icosahedral head with a diameter of (44 ± 1) nm and a long contractile tail. The length and width of the tail are (222 ± 3) nm and (22 ± 1) nm, respectively. B: Lysed cell debris and several tailed phage-like particles of Bsu-yong1. The tail fibers of Bsu-yong1 were clearly visible (white arrow). Scale bar: 100 nm (A) and 200 nm (B).



图 2 Bsu-yong1 在枯草芽孢杆菌菌苔上形成的噬菌斑(箭头所示)

Figure 2 Plaques developed by phage Bsu-yong1 on *Bacillus subtilis* lawn (arrow).

BLASTn 比对结果显示, 在 GenBank 数据库中与 Bsu-yong1 基因组同源性最高的序列是 4 株芽孢杆菌, 同源覆盖度(query cover)与相似度(identity)均为 100%。这 4 株芽孢杆菌分别是 *B. subtilis* strain BL-01 (CP028812)、*B. inaquosorum* strain DE111 (CP013984)、*B. inaquosorum* strain BSXE-2102 (CP111073)、*B. inaquosorum* strain A65.1 (CP096592), 其中菌株 BSXE-2102 与菌株 A65.1 的基因组序列完全一样, 认为二者实际应为同一菌株。使用 PHASTER 工具对这些菌株的基因组进行原噬菌体预测。结果显示其基因组上均各含有 2 个原噬菌体的基因组序列, 分别为 Jimmer2 (PHASTER 评分为 100 分)和 sv105 (PHASTER 评分为 80–90 分) (图 3)。PHASTER 软件并未成功地在这些菌株的基因组中预测出原噬菌体 Bsu-yong1。PHASTER 软件的分析功能尚有提升空间。通过序列比对, 分析 Bsu-yong1 在这些菌株的基因组中的整合位点(图 4), 发现 Bsu-yong1 与原噬菌体 sv105 在宿主菌基因组上的整合切入位点与切入区域重合度高, Bsu-yong1 与原噬菌体 Jimmer2 的切入位

点与切入区域不重合且相距较远。Bsu-yong1 在宿主菌染色体上的切入位点位于其 DNA 修复蛋白 RadC (DNA repair protein RadC)基因的编码区。

BLASTn 比对中与 Bsu-yong1 相似度最高的病毒是芽孢杆菌噬菌体 vB_BteM-A9Y, 同源覆盖度为 95.82%, 相似度为 28%。

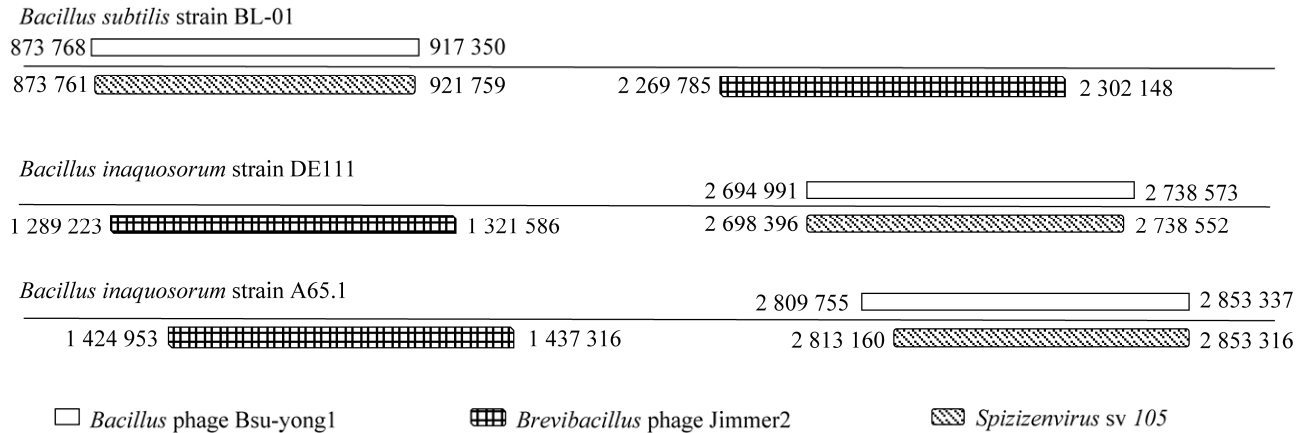


图 3 三株芽孢杆菌的原噬菌体预测结果及其插入位点

Figure 3 Insertion sites of the predicted prophages of the three *Bacillus* bacteria, which were the three top hits in BLASTn alignment of Bsu-yong1 (query cover=100%, identity=100%).

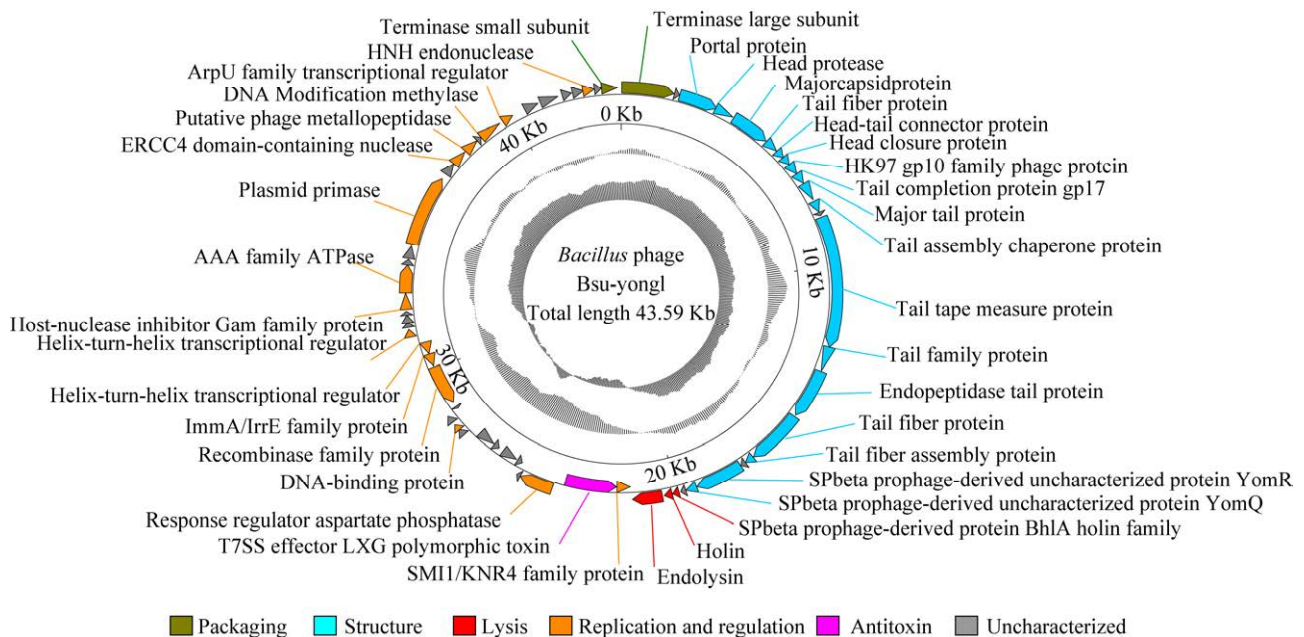


图 4 枯草芽孢杆菌原噬菌体 Bsu-yong1 的基因组图谱

Figure 4 Genome map of prophage Bsu-yong1. The outermost circle represents the 62 ORFs, with different colors represent different functions (clockwise arrow indicates the forward reading frame, counterclockwise arrow indicates the reverse reading frame). The dark circles in the middle represent the GC content (“black” indicates greater than the average GC content compared with the whole genome, and “gray” indicates the opposite). The innermost circle represents the GC skew (G-C/G+C: Outwards indicates>0 and inwards indicates<0).

Bsu-yong1 的全部 62 个 ORF 中, 有 38 个预测为编码已知功能蛋白, 占比为 61.3% (表 1)。Bsu-yong1 的基因分布呈现典型的模块化, 功能相关的基因相邻(图 4)。Bsu-yong1 基因组含有 4 大基因功能模块: 溶菌(lysis)、DNA 包装(DNA packaging)、病毒结构(structure)、复制与调节(replication and regulation) (图 4)。溶菌模块的 3 个 ORF (ORF26、ORF27、ORF28)分别编码穿孔素(holin)和细菌溶解酶(lysin 又称溶素) (图 4, 表

1); DNA 包装模块的 2 个 ORF (ORF1、ORF 62) 分别编码末端酶大亚基和末端酶小亚基(图 4, 表 1); 病毒结构模块中有 17 个 ORFs (ORF3–ORF21)被注释成编码噬菌体结构蛋白(图 4, 表 1); 复制与调节模块中有 16 个 ORF (ORFs: 26、27、28、35、38、39、40、41、45、46、49、51、52、54、55、60)被注释为编码参与 DNA 复制、代谢与基因表达调控相关的蛋白(图 4, 表 1)。

表 1 Bsu-yong1 各 ORF 功能预测和 BLASTp 比对结果

Table 1 Function prediction and top BLASTp hits of Bsu-yong1 ORFs

ORF	Size (aa)	Prediction function	Top BLAST hit ^a	Identity ^b (aa)	E value ^c
1	569	Terminase large subunit	ref WP_060399185.1 MULTISPECIES: terminase large subunit [<i>Bacillus subtilis</i>]	100% (569/569)	0.0
2	56	Hypothetical protein	ref WP_178938629.1 MULTISPECIES: hypothetical protein [<i>Bacillus</i>]	100% (56/56)	3E-28
3	426	Portal protein	ref WP_060399184.1 MULTISPECIES: phage portal protein [<i>Bacillus subtilis</i>]	100% (426/426)	0.0
4	208	Head protease	ref WP_060399183.1 MULTISPECIES: HK97 family phage prohead protease [<i>Bacillus subtilis</i>]	100% (208/208)	1E-151
5	431	Major capsid protein	ref WP_060399182.1 MULTISPECIES: phage major capsid protein [<i>Bacillus subtilis</i>]	100% (431/431)	0.0
6	150	Tail fiber protein	ref WP_060399181.1 MULTISPECIES: collagen-like protein [<i>Bacillus subtilis</i>]	100% (150/150)	2E-100
7	100	Head-tail connector protein	ref WP_060399180.1 MULTISPECIES: head-tail connector protein [<i>Bacillus subtilis</i>]	100% (100/100)	5E-67
8	104	Head closure protein	ref WP_060399179.1 MULTISPECIES: phage head closure protein [<i>Bacillus subtilis</i>]	100% 104/104)	4E-71
9	132	Mu-like protein gpG	ref WP_060399178.1 MULTISPECIES: HK97 gp10 family phage protein [<i>Bacillus</i>]	100% (132/132)	1E-90
10	130	Tail completion protein gp17	ref WP_060399177.1 MULTISPECIES: hypothetical protein [<i>Bacillus subtilis</i>]	100% (130/130)	7E-90
11	204	Major tail protein	ref WP_060399176.1 MULTISPECIES: phage tail protein [<i>Bacillus subtilis</i>]	100% (204/204)	5E-146
12	125	Tail assembly chaperone protein	ref WP_060399175.1 MULTISPECIES: hypothetical protein [<i>Bacillus subtilis</i>]	100% (125/125)	1E-82
13	40	Hypothetical protein	ref QHM08352.1 hypothetical protein C7M28_00059 [<i>Bacillus subtilis</i>]	100% (33/33)	7E-15
14	1495	Tail tape measure protein	ref WP_060399174.1 MULTISPECIES: phage tail tape measure protein [<i>Bacillus subtilis</i>]	100% (1495/1495)	0.0
15	279	Tail family protein	ref WP_060399173.1 MULTISPECIES: phage tail family protein [<i>Bacillus subtilis</i>]	100% (279/279)	0.0

(待续)

(续表 1)

ORF	Size (aa)	Prediction function	Top BLAST hit ^a	Identity ^b (aa)	E value ^c
16	567	Endopeptidase tail protein	ref WP_060399172.1 MULTISPECIES: phage tail protein [<i>Bacillus subtilis</i>]	100% (567/567)	0.0
17	632	Tail fiber protein	ref WP_060399171.1 MULTISPECIES: hypothetical protein [<i>Bacillus subtilis</i>]	100% (632/632)	0.0
18	102	Tail fiber assembly protein	ref WP_060399170.1 MULTISPECIES: hypothetical protein [<i>Bacillus subtilis</i>]	100% (102/102)	3E-66
19	57	Hypothetical protein	ref WP_164906782.1 MULTISPECIES: hypothetical protein [<i>Bacillus</i>]	100% (57/57)	4E-34
20	529	SPbeta prophage-derived uncharacterized protein YomR	ref WP_060399169.1 MULTISPECIES: BppU family phage baseplate upper protein [<i>Bacillus subtilis</i>]	100% (529/529)	0.0
21	123	SPbeta prophage-derived uncharacterized protein YomQ	ref WP_060399168.1 MULTISPECIES: hypothetical protein [<i>Bacillus subtilis</i>]	100% (123/123)	5E-85
22	57	Hypothetical protein	ref WP_080429041.1 MULTISPECIES: XkdX family protein [<i>Bacillus subtilis</i>]	100% (57/57)	4E-34
23	70	SPbeta prophage-derived protein BhlAholin family	ref WP_031600555.1 MULTISPECIES: hypothetical protein [<i>Bacillus</i>]	100% (70/70)	5E-43
24	87	Holin	ref WP_060399167.1 MULTISPECIES: phage holin [<i>Bacillus subtilis</i>]	100% (87/87)	2E-55
25	325	Endolysin	ref WP_060399166.1 MULTISPECIES: peptidoglycan-binding protein [<i>Bacillus subtilis</i>]	100% (325/325)	0.0
26	152	SMI1/KNR4 family protein	ref WP_060399165.1 MULTISPECIES: SMI1/KNR4 family protein [<i>Bacillus subtilis</i>]	100% (152/152)	5E-105
27	589	T7SS effector LXG polymorphic toxin	ref WP_060399164.1 MULTISPECIES: T7SS effector LXG polymorphic toxin [<i>Bacillus subtilis</i>]	100% (589/589)	0.0
28	370	Response regulator aspartate phosphatase	ref WP_060399163.1 MULTISPECIES: tetratricopeptide repeat protein [<i>Bacillus subtilis</i>]	100% (370/370)	0.0
29	44	Hypothetical protein	ref CAF1840838.1 hypothetical protein NRS6167_02201 [<i>Bacillus subtilis</i>]	100% (44/44)	4E-21
30	62	Hypothetical protein	ref WP_060399162.1 hypothetical protein [<i>Bacillus subtilis</i>]	100% (62/62)	1E-36
31	181	Hypothetical protein	ref WP_060399161.1 MULTISPECIES: hypothetical protein [<i>Bacillus subtilis</i>]	100% (181/181)	2E-126
32	71	Hypothetical protein	ref WP_060399160.1 MULTISPECIES: hypothetical protein [<i>Bacillus subtilis</i>]	100% (71/71)	7E-45
33	203	Hypothetical protein	ref WP_060399159.1 MULTISPECIES: hypothetical protein [<i>Bacillus subtilis</i>]	100% (203/203)	3E-145
34	80	Hypothetical protein	ref WP_060399158.1 MULTISPECIES: hypothetical protein [<i>Bacillus subtilis</i>]	100% (80/80)	8E-51
35	76	DNA-binding protein	ref WP_060399157.1 MULTISPECIES: DNA-binding protein [<i>Bacillus subtilis</i>]	100% (76/76)	4E-45
36	84	Hypothetical protein	ref POD84434.1 hypothetical protein S101384_03525 [<i>Bacillus subtilis</i> subsp. <i>subtilis</i>]	100% (84/84)	2E-54
37	29	Hypothetical protein	ref CAF1840716.1 hypothetical protein NRS6167_02193 [<i>Bacillus subtilis</i>]	100% (29/29)	3E-11
38	475	Recombinase family protein	ref PPA37087.1 recombinase family protein [<i>Bacillus subtilis</i>]	100% (475/475)	0.0

(待续)

(续表 1)

ORF	Size (aa)	Prediction function	Top BLAST hit ^a	Identity ^b (aa)	E value ^c
39	148	ImmA/IrrE family metallo-endopeptidase	ref WP_060399203.1 MULTISPECIES: ImmA/IrrE family metallo-endopeptidase [<i>Bacillus</i>]	100% (148/148)	6E-106
40	144	Helix-turn-helix transcriptional regulator	ref WP_060399202.1 MULTISPECIES: helix-turn-helix transcriptional regulator [<i>Bacillus subtilis</i>]	100% (144/144)	9E-99
41	90	Helix-turn-helix transcriptional regulator	ref WP_060399201.1 MULTISPECIES: helix-turn-helix transcriptional regulator [<i>Bacillus</i>]	100% (90/90)	4E-58
42	51	Hypothetical protein	ref WP_166638362.1 MULTISPECIES: hypothetical protein [<i>Bacillus</i>]	100% (51/51)	3E-27
43	91	Hypothetical protein	ref WP_041055414.1 MULTISPECIES: hypothetical protein [<i>Bacillus</i>]	100% (91/91)	3E-59
44	45	Hypothetical protein	ref WP_164906783.1 MULTISPECIES: hypothetical protein [<i>Bacillus</i>]	100% (45/45)	6E-24
45	184	Host-nuclease inhibitor Gam family protein	ref WP_060399200.1 MULTISPECIES: host-nuclease inhibitor Gam family protein [<i>Bacillus subtilis</i>]	100% (184/184)	2E-130
46	312	AAA family ATPase	ref WP_060399199.1 MULTISPECIES: AAA family ATPase [<i>Bacillus subtilis</i>]	100% (312/312)	0.0
47	65	Hypothetical protein	ref WP_019260095.1 MULTISPECIES: hypothetical protein [<i>Bacillus</i>]	100% (65/65)	5E-38
48	146	DUF669 domain-containing protein	ref WP_060399198.1 MULTISPECIES: DUF669 domain-containing protein [<i>Bacillus subtilis</i>]	100% (146/146)	1E-104
49	805	Phage/plasmid primase	ref WP_060399197.1 MULTISPECIES: phage/plasmid primase, P4 family [<i>Bacillus subtilis</i>]	100% (805/805)	0.0
50	145	Hypothetical protein	ref WP_060399196.1 MULTISPECIES: hypothetical protein [<i>Bacillus subtilis</i>]	100% (145/145)	6E-102
51	179	ERCC4-type nuclease	ref WP_060399195.1 MULTISPECIES: ERCC4 domain-containing protein [<i>Bacillus subtilis</i>]	100% (179/179)	4E-128
52	171	Putative phage metallopeptidase	ref WP_060399194.1 MULTISPECIES: putative metallopeptidase [<i>Bacillus subtilis</i>]	100% (171/171)	1E-123
53	54	Hypothetical protein	ref CAF1841383.1 hypothetical protein NRS6167_02237 [<i>Bacillus subtilis</i>]	100% (54/54)	2E-31
54	251	DNA cytosine methyltransferase	ref WP_060399193.1 MULTISPECIES: DNA cytosine methyltransferase [<i>Bacillus subtilis</i>]	100% (251/251)	0.0
55	125	ArpU family transcriptional regulator	ref WP_019260088.1 MULTISPECIES: ArpU family transcriptional regulator [<i>Bacillus subtilis</i>]	100% (125/125)	7E-84
56	157	Hypothetical protein	ref WP_060399191.1 MULTISPECIES: hypothetical protein [<i>Bacillus subtilis</i>]	100% (157/157)	3E-112
57	200	Hypothetical protein	ref WP_060399190.1 MULTISPECIES: hypothetical protein [<i>Bacillus subtilis</i>]	100% (200/200)	2E-143
58	104	Hypothetical protein	ref WP_060399189.1 MULTISPECIES: hypothetical protein [<i>Bacillus subtilis</i>]	100% (104/104)	3E-69
59	122	Hypothetical protein	ref WP_060399188.1 MULTISPECIES: hypothetical protein [<i>Bacillus subtilis</i>]	100% (114/114)	4E-79
60	127	HNH endonuclease	ref WP_060399187.1 MULTISPECIES: HNH endonuclease [<i>Bacillus subtilis</i>]	100% (121/121)	6E-86
61	73	Hypothetical protein	ref CAF1841293.1 hypothetical protein NRS6167_02230 [<i>Bacillus subtilis</i>]	100% (73/73)	5E-44
62	171	Terminase small subunit	ref WP_060399186.1 MULTISPECIES: phage terminase small subunit P27 family [<i>Bacillus subtilis</i> group]	100% (171/171)	2E-123

^a: The most closely related protein and its organism. "No hits" indicates no significant hits detected for a particular amino acid sequence. ^b: Percent Per identity for top hits in BLASTP searches. Numbers in parentheses provide length of each alignment. ^c: The probability of obtaining a match by chance as determined by BLASTP analysis.

Bsu-yong1 的 ORF24 被注释为 *holin* 基因, *Holin* 是噬菌体编码的小跨膜蛋白, 在噬菌体感染后特定时间内在宿主细胞膜上形成非特异性穿孔, 从而控制噬菌体感染周期的时长, 并在最佳时间点裂解宿主细胞。由于 *Holin* 可以调节噬菌体感染周期, 它也被称为控制噬菌体感染的时钟^[27]。ORF25 被预测为 *endolysin* 基因, 噬菌体的 *Endolysin* 是肽聚糖水解酶, 被称为内容素, 又称为细菌溶解酶, 具有高特异性, 在感染周期结束时参与细菌裂解^[28]。噬菌体穿孔素在细胞膜上形成孔洞, 使得内容素可以抵达细菌细胞壁上的肽聚糖靶点, 并对细菌细胞壁进行切割水解, 最终导致细菌裂解、死亡。

Bsu-yong1 的 ORF45 被预测为编码宿主核酸酶抑制蛋白(host-nuclease inhibitor Gam family protein), ORF60 被预测为编码 HNH 核酸内切酶(HNH endonuclease), 二者均参与介导噬菌体基因组切入整合到宿主基因组中。核酸内切酶可以切开宿主菌的特异性位点, 随后在宿主核酸酶抑制蛋白的作用下, 噬菌体的单链 DNA 能和宿主基因组 DNA 之间进行同源重组^[29-30]。

毒素与抗毒素(toxin-antitoxin)在细菌免疫以适应环境和获得竞争优势中起着重要的作用。Bsu-yong1 的 ORF26 被预测为编码 SMI1/KNR4 家族蛋白。SMI1/KNR4 family protein 广泛存在于细菌、真核生物和各种 DNA 病毒中^[31]。其最初在酿酒酵母胞中被发现, 它不仅在几丁质沉积和细胞壁装配中起调节作用, 更作为抗毒素(antitoxin)赋予酿酒酵母抗性, 以拮抗其他酵母分泌的毒素^[32]。编码 SMI1/KNR4 family protein 的基因也广泛存在于乳酸杆菌属(*Lactobacillus*)、链球菌属(*Streptococcus*)、

Geobacillus、土芽孢杆菌属(*Listeria*)、芽孢杆菌属(*Bacillus*)等细菌中^[33]。Zhang 等^[34]进行的生物信息学研究表明, 细菌中的 SMI1/KNR4 家族蛋白是潜在的免疫蛋白。宏基因组研究表明编码 SMI1/KNR4 家族蛋白的基因广泛存在于 DNA 病毒中^[34]。噬菌体携带抗毒素基因可能有利于宿主菌在竞争和环境适应中获得生存优势, 以利于噬菌体自身, 具体机制有待深入研究。Bsu-yong1 的 ORF27 被注释为编码 T7SS 效应器 LXG 多态性毒素(T7SS effector LXG polymorphic toxin)。编码 LXG polymorphic toxin 的基因以往只被报道存在于革兰氏阳性菌中, 其中在链球菌中被报道最多^[35-36]。LXG polymorphic toxin 最初是通过生物信息学鉴定的, 属于 VII 型分泌系统(type VII secretory system, T7SS), 被革兰氏阳性细菌用来感染宿主或使细菌性竞争对手中毒。它的毒力作用主要体现在识别和清除周围环境中的其他种类的细菌或同种但非近源的菌株(closely related strains)。Chatterjee 等^[36]的研究发现, 链球菌表达的 LXG 毒素被分泌到胞外环境中, 周边同一环境中的革兰氏阳性菌和 LXG 毒素接触后, 除少数革兰氏阳性菌会产生拮抗作用外, 可抑制其他细菌的繁殖。产生拮抗作用的革兰氏阳性菌中的 *essB* 免疫基因的表达能抑制 T7SS 系统表达的 LXG 毒素的毒性。少数革兰氏阳性菌在分类地位上与链球菌有紧密的亲缘关系; 而不产生拮抗作用的革兰氏阳性菌与该链球菌亲缘关系不紧密, 会被 T7SS 系统表达的 LXG 毒素杀死。这一机制使携带 LXG 毒素基因的革兰氏阳性宿主菌在与其他菌株的种内和种间竞争中占据优势^[36]。此前, 编码 LXG polymorphic toxin 的基因未在噬菌体中被发现、报道, 本文属首次。长期以来, 病毒

与其宿主之间的相互作用导致了二者的基因组进化,一方面宿主通过从噬菌体中获取外源核苷酸序列进化出多种抗性机制;另一方面,病毒可以在其 DNA 合成过程中从宿主那里获得某些基因,以获得适应性优势。*Bsu-yong1* 这个毒力基因可能帮助宿主在复杂环境的竞争中占据优势。*Bsu-yong1* 的 ORF39 被注释为编码 ImmA/IrrE 家族蛋白(ImmA/IrrE family protein), ImmA/IrrE 家族蛋白又称 ImmA/IrrE 家族金属内肽酶。ImmA/IrrE 家族蛋白是毒素-抗毒素(toxin-antitoxin)系统 ToxN/AbiQsystem 的成员。其中 *immA* 基因被发现在枯草芽孢杆菌与其原噬菌体的可移动保守遗传元件中广泛存在,能调控转录。在电离、紫外线辐射或丝裂霉素 C 诱导突变等条件下,细菌需要进行移动遗传元件(例如 ICEBs1)的水平转移、DNA 修复和存活所需基因的转录,ImmA/IrrE 金属蛋白酶能破坏阻遏蛋白(ImmR/Ddro),从而使含有 *immA/irrE* 基因菌细胞获得竞争优势^[37]。此外,噬菌体可通过 ImmA 直接参与和控制宿主细胞的免疫防御,这能帮助宿主抵抗其他病毒的入侵^[35-36]。*Bsu-yong1* 携带的毒素与抗毒素基因可能有助于枯草芽孢杆菌的环境适应和竞争。

Bsu-yong1 的 ORF1 和 ORF62 分别被预测为编码末端酶大亚基和小亚基,有尾噬菌体在组装过程中,使用一个 DNA 包装引擎来组装和封闭它们的基因组。由末端酶小亚基充当分子媒介物的角色,识别噬菌体 DNA,并将其转移给末端酶大亚基,由后者进行 DNA 的切割和包装。但是,末端酶大亚基和末端酶小亚基如何与脱氧核糖核酸结合的分子机制尚不清楚^[38-39]。

2.5 进化地位分析

在基于全基因组构建的 proteomic tree 中,*Bsu-yong1* 与芽孢杆菌噬菌体 sv105、rho14 和

vB_BteM-A9Y 聚在一起,形成 1 个独立的进化支(clade),但 *Bsu-yong1* 与这些噬菌体的进化距离较远(图 5)。将 *Bsu-yong1* 与已知数据库中所有噬菌体进行比对,得到的 PASC (全基因组核苷酸序列相似度)的最大值为 46.72% (表 2),与 *Bsu-yong* 具有最大 PASC 值的是芽孢杆菌噬菌体 vB_BteM-A9Y,这与 BLASTn 的结果一致(如前所述,在 BLASTn 比对中,与 *Bsu-yong* 基因组具有最大相似度的噬菌体是 vB_BteM-A9Y)。PASC 在线分析工具是国际病毒分类委员会 ICTV 推荐的用于分析核苷酸序列相似度的工具,ICTV 下的细菌与古菌病毒分会(Bacterial and Archaeal Viruses Subcommittee, BAVS)关于噬菌体属的界定标准是 PASC<70%为 1 个新的属^[40]。噬菌体 *Bsu-yong1* 与所有已知噬菌体间的 PASC 值均小于该边界值,说明 *Bsu-yong1* 在有尾纲(Caudoviricetes)下代表 1 个新的属。

从系统进化树(图 5)可见,*Bsu-yong1*、sv105、rho14 和 vB_BteM-A9Y 构成的 clade 与其他噬菌体间具有很远的进化距离,这一距离大于有尾纲下多数科间的进化距离。结果显示,它们可能构成 1 个新的科(family)。在该科中,只有 sv105 在国际病毒分类委员会 ICTV 的分类系统中被定义了分类地位,它属于有尾纲下的 1 个独立的属 *Spizizenvirus*^[41]。在 ICTV 当前的分类系统中,*Spizizenvirus* sv105 是 *Spizizenvirus* 属的唯一成员。系统进化树(图 5)显示,*Spizizenvirus* sv105 与 rho14 间的进化距离小。二者的 ANI 值为 93.49% (表 2),小于定义一个种的边界值(95%),因而二者分别为独立的种。二者间的 PASC 值为 84.85%,大于属边界值(70%),因而,建议将 rho14 作为 *Spizizenvirus* 属的第 2 个成员。vB_BteM-A9Y 与 *Bsu-yong1*、sv105、rho14 间的 PASC 值均低于 70%,因而,它代表 1 个新的属。

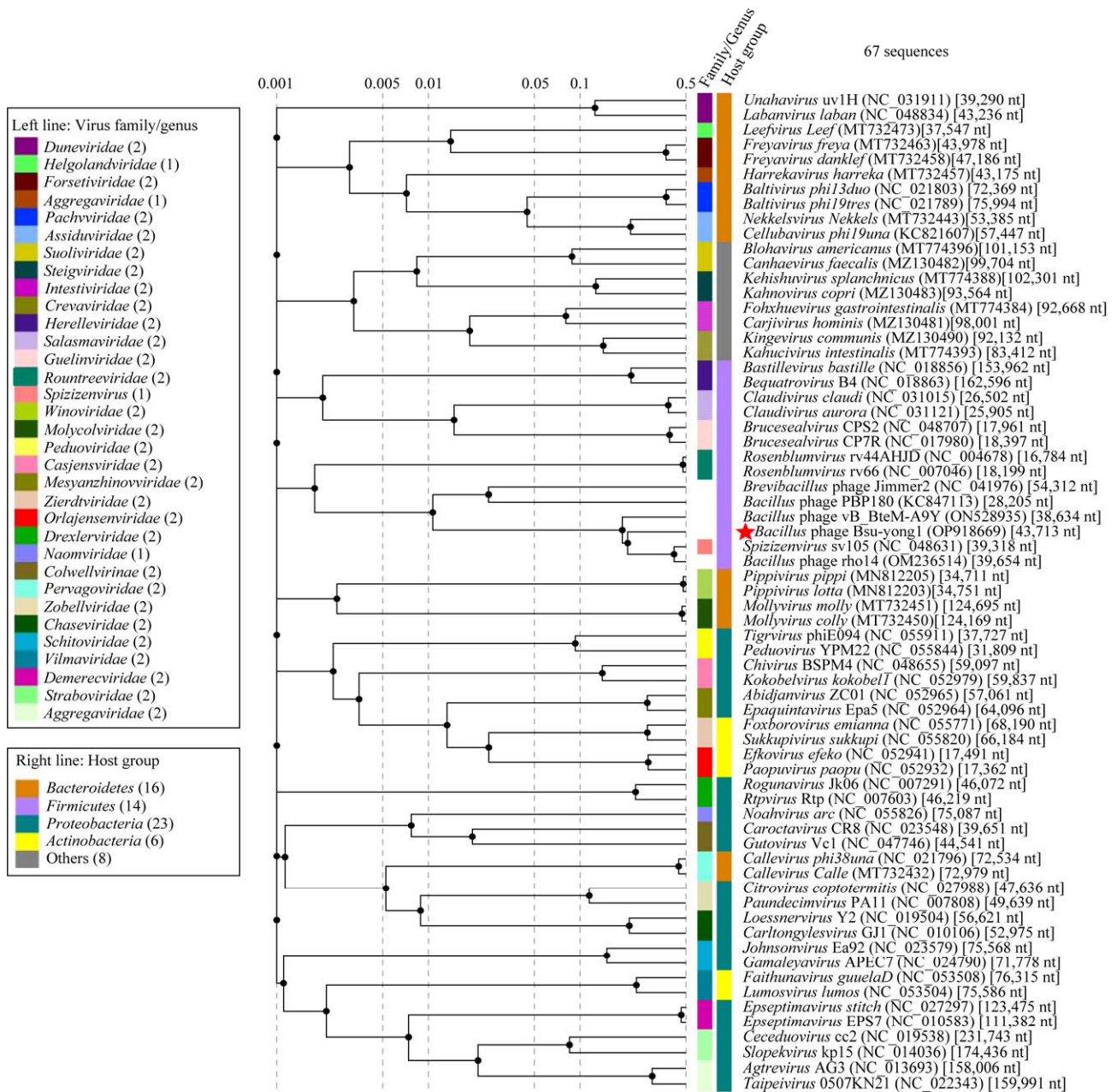


图 5 基于全基因组的 proteomic tree

Figure 5 Proteomic tree of *Bsu-yong1*, *Spizizenvirus sv105*, *Bacillus phage rho14*, *Bacillus phage vB_BteM-A9Y*, *Brevibacillus phage Jimmer2*, *Bacillus phage PBP180*, and 61 classified phages of the 33 families of *Caudoviricetes*.

全基因组比较图(图 6)显示, 所建的新科中的噬菌体间的复制与调控模块具有高度保守性, 该模块中, 除个别基因外, 各基因间的相似度高 于 64%, 该同源模块长度分别占各噬菌体基因组的总长的 46%左右。该科内 4 个噬菌体共有的核心基因(core gene)有 29 个, 分别编码:

SMI1/KNR4 家族蛋白 (SMI1/KNR4 family protein)、T7SS 效应器 LXG 多态性毒素(T7SS effector LXG polymorphic toxin)、应答调控子天冬氨酸磷酸酶 (response regulator aspartate phosphatase)、DNA 结合蛋白 (DNA-binding protein)、重组酶家族蛋白 (recombinase family

protein)、ImmA/IrE 家族金属内肽酶(ImmA/IrE family metallo-endopeptidase)、H-T-H 转录调控因子(H-T-H transcriptional regulator)、宿主核酸酶抑制剂 Gam 家族蛋白(host-nuclease inhibitor Gam family protein)、AAA 家族 ATP 酶(AAA family ATPase)、含 DUF669 结构域的蛋白(DUF669 domain-containing protein)、质粒原酶(plasmid primase)、ERCC4 型核酸酶(ERCC4-type nuclease)、噬菌体金属肽酶(putative phage metalloproteinase)、DNA 胞嘧啶甲基转移酶(DNA cytosine methyltransferase)和 ArpU 家族转录调控因子(ArpU family transcriptional regulator)。Bsu-yong1 与噬菌体 sv105 共有 31 个 ORF 有高的相似度(64%–100%)，占 Bsu-yong1 全部 ORF

的 50%，它们分别是 12 个复制与调控相关的 ORF、4 个与结构相关的 ORF、2 个与裂解相关的 ORF、1 个与 DNA 包装相关的 ORF 和 12 个未知功能的 ORF。Bsu-yong1 与噬菌体 vB_BteM-A9Y 共有 31 个 ORF 有高的相似度(64%–100%)，占 Bsu-yong1 全部 ORF 的 50%，它们分别是 12 个与复制与调控相关的 ORF、4 个与结构相关的 ORF、2 个与裂解相关的 ORF 和 13 个未知功能的 ORF。Bsu-yong1 与噬菌体 rho14 共有 26 个 ORF 有高的相似度(64%–100%)，占 Bsu-yong1 全部 ORFs 的 41.93%，它们包括 10 个复制与调控相关的 ORF、2 个与结构相关的 ORF、1 个与裂解相关的 ORF 和 13 个未知功能的 ORF。

表 2 芽孢杆菌噬菌体 Bsu-yong1、vB_BteM-A9Y、sv105 和 rho14 的基因组相似性分析值
Table 2 The genome similarity analysis of prophage Bsu-yong1, vB_BteM-A9Y, sv105, and rho14

Phage	BLSATn (%)	ANI (%)	isDDH (%)	PASC (%)
Bsu-yong1				
<i>Bacillus</i> phage rho14	25	83.12	18.80	46.44
<i>Bacillus</i> phage vB_BteM-A9Y	28	93.14	20.20	46.72
<i>Spizizenvirus</i> sv105	24	79.72	18.80	44.86
rho14				
<i>Bacillus</i> phage vB_BteM-A9Y	26	85.02	19.40	43.38
<i>Spizizenvirus</i> sv105	88	93.49	86.60	84.85
vB_BteM-A9Y				
<i>Spizizenvirus</i> sv105	34	83.16	20.60	44.70

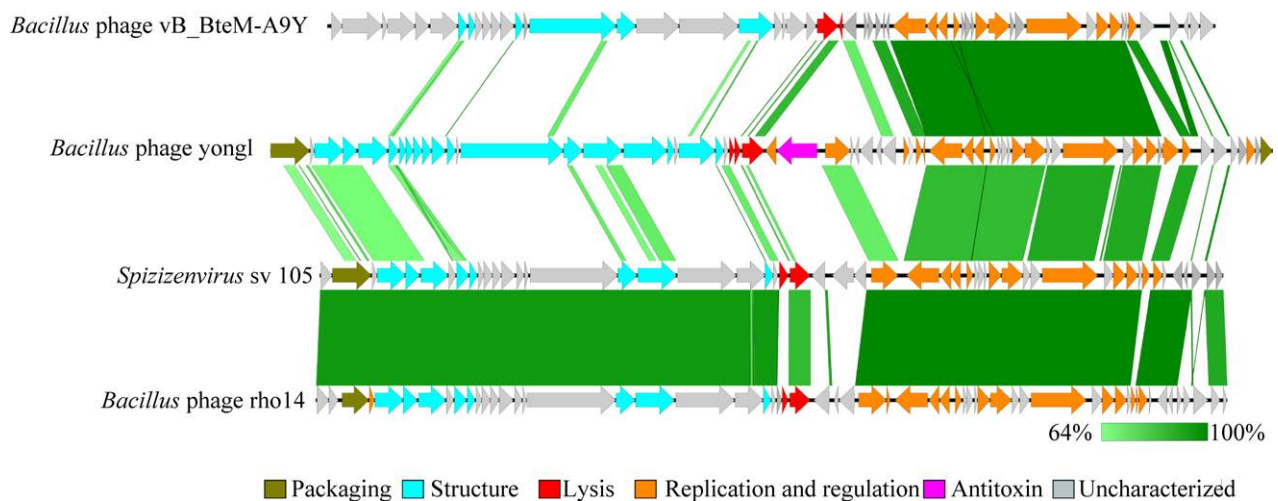


图 6 芽孢杆菌噬菌体 Bsu-yong1、vB_BteM-A9Y、sv105 和 rho14 的基因组比较图

Figure 6 Genome comparison of the *Bacillus* prophage Bsu-yong1, vB_BteM-A9Y, sv105, and rho14.

3 讨论与结论

如前言所述,有关枯草芽孢杆菌噬菌体的研究,主要集中在 PBSX 类缺陷型原噬菌体 (PBSX-like defective prophages) 上。PBSX 类缺陷型原噬菌体在芽孢杆菌中广泛存在。PBSX 基因组位于宿主 (*B. subtilis* 168) 染色体的 *metA-metC* 区^[42]。PBSX 的基因组大小被估计为约 54 kb^[17,43]。PBSX 粒子及 PBSX 样噬菌体粒子都有 1 个较小的头部 (直径 41–45 nm) 和 1 个相对于其他大多数肌尾样噬菌体更长的、可收缩的尾部 (长 190–240 nm, 宽约 18 nm)。比较特殊的是, PBSX 缺陷型噬菌体粒子中不包裹自身基因组 DNA, 而是随机包含约 13 kb 的宿主 DNA 片段^[44,45]。Okamoto 等提出, PBSX 缺陷型噬菌体可能是从非缺陷型噬菌体进化而来的^[44]。以往报道的 PBSX 类原噬菌体绝大多数是缺陷型的, 其病毒粒子中只是随机包裹宿主 DNA 的片段, 有关 PBSX 及 PBSX 样噬菌体的基因组究竟有多大, 其进化来源噬菌体究竟是什么样的, 并不清楚。本研究成功诱导了 1 株非缺陷型 PBSX 样噬菌体, 并分析了其基因组, 具有重要的学术意义。

缺陷型 PBSX 样原噬菌体与宿主的相互作用已被大量研究。由于缺陷型 PBSX 样原噬菌广泛存在于芽孢杆菌中, 这促进了基因在芽孢杆菌间的水平转移, PBSX 在芽孢杆菌进化中发挥了重要作用^[44]。不止于此, 由分离得到的芽孢杆菌属细菌的基因组几乎都含有缺陷型 PBSX 样原噬菌体基因, 这表明 PBSX 对于芽孢杆菌属细菌来说似乎是不可缺少的一部分, 可以帮助宿主细胞在复杂的环境中存活^[44-45]。Buxton^[46]的研究发现, 芽孢杆菌基因组上缺失缺陷型 PBSX 样原噬菌体区域的突变体在培养基中生长不良, 这表明 PBSX 可能有利于宿主在营养不良的环境中

存活, Piggot 等也证明了 PBSX 的缺失会影响细胞正常的代谢活动^[17]。本研究发现非缺陷型噬菌体 Bsu-yong1 基因组拥有编码毒素或抗毒素 (toxin and antitoxin), 它们分别是 LXG polymorphic toxin、ImmA/IrrE family protein 和 SMI1/KNR4 family protein。其中, 编码 LXG polymorphic toxin 的基因是首次在噬菌体基因组中被发现和报道。这些基因的发现暗示 Bsu-yong1 对其宿主的帮助可能不止于基因的水平转移以促进细菌的进化, Bsu-yong1 还能帮助宿主在复杂的环境中获得竞争优势和更好的环境适应能力。此外, Bsu-yong1 有 24 个 ORF 被预测为编码未知功能蛋白, 占全部 62 个 ORF 的 38.7%。后续对 Bsu-yong1 基因进行进一步的深入研究, 很有可能会为深入认识原噬菌体与宿主的关系, 为研究枯草芽孢杆菌细胞代谢、免疫调控提供新的思路与方向。

Bsu-yong1 与所有已知噬菌体基因组间的 PASC 值 ($\leq 46.73\%$) 均小于 ICTV 界定的属边界值 (70%), 它代表 1 个新的属。从系统进化树中 (图 5), 噬菌体 Bsu-yong1、sv105、rho14 和 vB_BteM-A9Y 在基于全基因组序列相似性构建的系统进化树中, 聚成 1 个独立的 clade, 该 clade 与其他噬菌体间的进化距离大于有尾纲下多数科间的进化距离; 这 4 个芽孢杆菌噬菌体的复制与调控基因模块高度同源, 具有 29 个核心蛋白 (core protein)。结果提示它们构成 1 个新的独立的科。本研究丰富了噬菌体基因数据库, 拓展了对芽孢杆菌活性原噬菌体的认知, 拓展了对噬菌体分类多样性的认知。

参考文献

- [1] CLOKIE MRJ, MILLARD AD, LETAROV AV, HEAPHY S. Phages in nature[J]. Bacteriophage, 2011, 1(1): 31-45.

- [2] FORTIER LC, SEKULOVIC O. Importance of prophages to evolution and virulence of bacterial pathogens[J]. *Virulence*, 2013, 4(5): 354-365.
- [3] ŁOŚM, CZYZ A, SELL E, WĘGRZYN A, NEUBAUER P, WĘGRZYN G. Bacteriophage contamination: is there a simple method to reduce its deleterious effects in laboratory cultures and biotechnological factories?[J]. *Journal of Applied Genetics*, 2004, 45(1): 111-120.
- [4] WEINBAUER MG. Ecology of prokaryotic viruses[J]. *FEMS Microbiology Reviews*, 2004, 28(2): 127-181.
- [5] 沈萍, 陈向东. 微生物学[M]. 8 版. 北京: 高等教育出版社, 2016.
- SHEN P, CHEN XD. *Microbiology*[M]. 8th ed. Beijing: Higher Education Press, 2016 (in Chinese).
- [6] SONG WC, SUN HX, ZHANG C, CHENG L, PENG Y, DENG ZQ, WANG D, WANG Y, HU M, LIU WE, YANG HM, SHEN Y, LI JH, YOU LC, XIAO MF. Prophage hunter: an integrative hunting tool for active prophages[J]. *Nucleic Acids Research*, 2019, 47(W1): W74-W80.
- [7] TOUCHON M, SOUSA JAM, ROCHA EPC. Embracing the enemy: the diversification of microbial gene repertoires by phage-mediated horizontal gene transfer[J]. *Current Opinion in Microbiology*, 2017, 38: 66-73.
- [8] REIS-CUNHA JL, BARTHOLOMEU DC, MANSON AL, EARL AM, CERQUEIRA GC. ProphET, prophage estimation tool: a stand-alone prophage sequence prediction tool with self-updating reference database[J]. *PLoS One*, 2019, 14(10): e0223364.
- [9] REN JJ, HE WH, LI CY, HE S, NIU DZ. Purification and identification of a novel antifungal protein from *Bacillus subtilis* XB-1[J]. *World Journal of Microbiology and Biotechnology*, 2019, 35(10): 1-7.
- [10] 朱瑾, 朱红军, 曲湘勇. 畜禽臭气危害及枯草芽孢杆菌在畜禽生产中的应用[J]. *畜牧与兽医*, 2019, 51(10): 127-130.
- ZHU J, ZHU HJ, QU XY. Odor of livestock and poultry and application of *Bacillus subtilis* to them[J]. *Animal Husbandry & Veterinary Medicine*, 2019, 51(10): 127-130 (in Chinese).
- [11] 刘兴, 姚晶, 夏元军. 枯草芽孢杆菌在固态发酵饲料中的应用研究进展[J]. *饲料研究*, 2020, 43(2): 113-116.
- LIU X, YAO J, XIA YJ. Research progress on the application of *Bacillus subtilis* in solid-state fermented feed[J]. *Feed Research*, 2020, 43(2): 113-116 (in Chinese).
- [12] 王红革, 李文清, 徐柏年, 罗进贤. 地衣芽孢杆菌 α -淀粉酶基因在枯草芽孢杆菌中的诱导表达[J]. *微生物学报*, 1997, 37(2): 101-106.
- WANG HG, LI WQ, XU BN, LUO JX. Inducible expression of α -amylase gene of *Bacillus licheniformis* in *Bacillus subtilis*[J]. *Acta Microbiologica Sinica*, 1997, 37(2): 101-106 (in Chinese).
- [13] FRIEDMAN DI, OLSON ER, GEORGOPOULOS C, TILLY K, HERSKOWITZ I, BANUETT F. Interactions of bacteriophage and host macromolecules in the growth of bacteriophage lambda[J]. *Microbiological Reviews*, 1984, 48(4): 299-325.
- [14] MIRZAEI MK, MAURICE CF. Ménage à trois in the human gut: interactions between host, bacteria and phages[J]. *Nature Reviews Microbiology*, 2017, 15(7): 397-408.
- [15] ANDERSON LM, BOTT KF. DNA packaging by the *Bacillus subtilis* defective bacteriophage PBSX[J]. *Journal of Virology*, 1985, 54(3): 773-780.
- [16] MAUËL C, KARAMATA D. Characterization of proteins induced by mitomycin C treatment of *Bacillus subtilis*[J]. *Journal of Virology*, 1984, 49(3): 806-812.
- [17] PIGGOT PJ, BUXTON RS. Bacteriophage PBSX-induced deletion mutants of *Bacillus subtilis* 168 constitutive for alkaline phosphatase[J]. *Microbiology*, 1982, 128(4): 663-669.
- [18] ARNDT D, MARCU A, LIANG YJ, WISHART DS. PHAST, PHASTER and PHASTEST: tools for finding prophage in bacterial genomes[J]. *Briefings in Bioinformatics*, 2019, 20(4): 1560-1567.
- [19] AZIZ RK, BARTELS D, BEST AA, DEJONGH M, DISZ T, EDWARDS RA, FORMSMA K, GERDES S, GLASS EM, KUBAL M, MEYER F, OLSEN GJ, OLSON R, OSTERMAN AL, OVERBEEK RA, MCNEIL LK, PAARMANN D, PACZIAN T, PARRELLO B, PUSCH GD, et al. The RAST Server: rapid annotations using subsystems technology[J]. *BMC Genomics*, 2008, 9: 75.
- [20] POTTER SC, LUCIANI A, EDDY SR, PARK Y, LOPEZ R, FINN RD. HMMER web server: 2018 update[J]. *Nucleic Acids Research*, 2018, 46(W1): W200-W204.

- [21] ZIMMERMANN L, STEPHENS A, NAM SZ, RAU D, KÜBLER J, LOZAJIC M, GABLER F, SÖDING J, LUPAS AN, ALVA V. A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core[J]. *Journal of Molecular Biology*, 2018, 430(15): 2237-2243.
- [22] BAO YM, CHETVERNIN V, TATUSOVA T. Improvements to pairwise sequence comparison (PASC): a genome-based web tool for virus classification[J]. *Archives of Virology*, 2014, 159(12): 3293-3304.
- [23] SULLIVAN MJ, PETTY NK, BEATSON SA. Easyfig: a genome comparison visualizer[J]. *Bioinformatics*, 2011, 27(7): 1009-1010.
- [24] FIGUERAS MJ, BEAZ-HIDALGO R, HOSSAIN MJ, LILES MR. Taxonomic affiliation of new genomes should be verified using average nucleotide identity and multilocus phylogenetic analysis[J]. *Genome Announcements*, 2014, 2(6): e00927-14.
- [25] MEIER-KOLTHOFF JP, AUCH AF, KLENK HP, GÖKER M. Genome sequence-based species delimitation with confidence intervals and improved distance functions[J]. *BMC Bioinformatics*, 2013, 14(1): 1-14.
- [26] JIN TT, ZHANG XM, ZHANG Y, HU ZS, FU ZW, FAN JP, WU M, WANG Y, SHEN P, CHEN XD. Biological and genomic analysis of a PBSX-like defective phage induced from *Bacillus pumilus* AB94180[J]. *Archives of Virology*, 2014, 159(4): 739-752.
- [27] WANG IN, SMITH DL, YOUNG R. Holins: the protein clocks of bacteriophage infections[J]. *Annual Review of Microbiology*, 2000, 54: 799-825.
- [28] RODRÍGUEZ-RUBIO L, MARTÍNEZ B, DONOVAN DM, RODRÍGUEZ A, GARCÍA P. Bacteriophage virion-associated peptidoglycan hydrolases: potential new enzybiotics[J]. *Critical Reviews in Microbiology*, 2013, 39(4): 427-434.
- [29] KIM YG, CHA J, CHANDRASEGARAN S. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 1996, 93(3): 1156-1160.
- [30] WILLIAMS RJ. Restriction endonuclease[J]. *Molecular Biotechnology*, 2003, 23(3): 225-243.
- [31] LESAGE G, SDICU AM, MÉNARD P, SHAPIRO J, HUSSEIN S, BUSSEY H. Analysis of β -1,3-glucan assembly in *Saccharomyces cerevisiae* using a synthetic interaction network and altered sensitivity to caspofungin[J]. *Genetics*, 2004, 167(1): 35-49.
- [32] TREMBLAY O, THOW Z, GEDDES-MCALISTER J, MERRILL A. Several new putative bacterial ADP-ribosyltransferase toxins are revealed from *in silico* data mining, including the novel toxin vorin, encoded by the fire blight pathogen *Erwinia amylovora*[J]. *Toxins*, 2020, 12(12): 792.
- [33] SURACHAT K, SANGKET U, DEACHAMAG P, CHOTIGEAT W. *In silico* analysis of protein toxin and bacteriocins from *Lactobacillus paracasei* SD1 genome and available online databases[J]. *PLoS One*, 2017, 12(8): e0183548.
- [34] ZHANG DP, IYER LM, ARAVIND L. A novel immunity system for bacterial nucleic acid degrading toxins and its recruitment in various eukaryotic and DNA viral systems[J]. *Nucleic Acids Research*, 2011, 39(11): 4532-4552.
- [35] TRAN HK R, GREBENC DW, KLEIN TA, WHITNEY JC. Bacterial type VII secretion: an important player in host-microbe and microbe-microbe interactions[J]. *Molecular Microbiology*, 2021, 115(3): 478-489.
- [36] CHATTERJEE A, WILLET JLE, DUNNY GM, DUERKOP BA. Phage infection and sub-lethal antibiotic exposure mediate *Enterococcus faecalis* type VII secretion system dependent inhibition of bystander bacteria[J]. *PLoS Genetics*, 2021, 17(1): e1009204.
- [37] GÓMEZ LA, ALVAREZ FI, MOLINA RE, SOTO-SHARA R, DAZA-CASTRO C, FLORES MR, LEÓN Y, OÑATE AA. A zinc-dependent metalloproteinase of *Brucella abortus* is required in the intracellular adaptation of macrophages[J]. *Frontiers in Microbiology*, 2020, 11: 1586.
- [38] LIN H, SIMON MN, BLACK LW. Purification and characterization of the small subunit of phage T4 terminase, gp16, required for DNA packaging[J]. *Journal of Biological Chemistry*, 1997, 272(6): 3495-3501.
- [39] BLACK LW. DNA packaging and cutting by phage terminases: control in phage T4 by a synaptic mechanism[J]. *BioEssays*, 1995, 17(12): 1025-1030.

- [40] ADRIAENSSENS EM, SULLIVAN MB, KNEZEVIC P, ZYL LJ, SARKAR BL, DUTILH BE, ALFENAS-ZERBINI P, ŁOBOCKA M, TONG YG, BRISTER JR, MORENO SWITT AI, KLUMPP J, AZIZ RK, BARYLSKI J, UCHIYAMA J, EDWARDS RA, KROPINSKI AM, PETTY NK, CLOKIE MRJ, KUSHKINA AI, et al. Taxonomy of prokaryotic viruses: 2018-2019 update from the ICTV bacterial and archaeal viruses subcommittee[J]. *Archives of Virology*, 2020, 165(5): 1253-1260.
- [41] TURNER D, KROPINSKI AM, ADRIAENSSENS EM. A roadmap for genome-based phage taxonomy[J]. *Viruses*, 2021, 13(3): 506.
- [42] WOOD HE, DAWSON MT, DEVINE KM, MCCONNELL DJ. Characterization of PBSX, a defective prophage of *Bacillus subtilis*[J]. *Journal of Bacteriology*, 1990, 172(5): 2667-2674.
- [43] THURM P, GARRO AJ. Bacteriophage-specific protein synthesis during induction of the defective *Bacillus subtilis* bacteriophage PBSX[J]. *Journal of Virology*, 1975, 16(1): 179-183.
- [44] OKAMOTO K, MUDD JA, MARMUR J. Conversion of *Bacillus subtilis* DNA to phage DNA following mitomycin C induction[J]. *Journal of Molecular Biology*, 1968, 34(3): 429-437.
- [45] MCDONNELL GE, WOOD H, DEVINE KM, MCCONNELL DJ. Genetic control of bacterial suicide: regulation of the induction of PBSX in *Bacillus subtilis*[J]. *Journal of Bacteriology*, 1994, 176(18): 5820-5830.
- [46] BUXTON RS. Selection of *Bacillus subtilis* 168 mutants with deletions of the PBSX prophage[J]. *Journal of General Virology*, 1980, 46(2): 427-437.