

以宏基因组技术探讨渤海秋冬季节病毒多样性*

夏 骏 汪 岷 宫 政 姜 勇 马 玉 汪 俭
邵红兵 王多兵 许志梦

(中国海洋大学海洋生命学院病毒实验室 青岛 266003)

摘要 为研究渤海海域病毒群落总体状况,本文以渤海的 B47 站表层海水为代表,对两个时间点(2010 年 9 月,2011 年 12 月)的病毒宏基因组做出全面分析。采集海水样品后过滤并经过切相流系统浓缩,提取 DNA 测序,再进行生物信息学分析,包括病毒组的种群分类、功能基因分析、系统进化分析等。分析结果指明,这两个时间点的病毒群落均由双链 DNA 病毒(97.75%, scaffold 百分比)占据优势地位,其中尤以有尾噬菌体(80.86%, scaffold 百分比)居多。按宿主区分,最大的一类病毒为聚球藻噬藻体(10.29%, scaffold 百分比)。渤海病毒群落最为丰富的功能基因为复制、结合和修复基因,且冬季(33.77%, ORF 百分比)大于秋季(16.39%, ORF 百分比)。结果表明:在渤海病毒组中,(1)存在理论上只出现在亚热带海域的原绿球藻噬藻体;(2)物种组成、物种多样性、功能基因比例呈现季节变化;(3)相比远洋病毒组有一定独特性;(4)可能存在未知的 T4-like 病毒分支。

关键词 宏基因组;病毒组;海洋病毒;病毒多样性;渤海

中图分类号 Q93 **doi:** 10.11693/hyhz20151100293

海洋环境中的病毒数量极为丰富,总共有将近 4×10^{30} 个,平均丰度约 3×10^9 ind./L (Suttle, 2005)。海洋病毒以微食物环、侵染微生物等形式影响海洋生物群落组成和碳氮循环(Wilhelm *et al*, 1999),浮游病毒优先杀死占优势宿主的特性能使海域富营养区的赤潮快速消散(Brusaard *et al*, 2005, 2008)。渤海的浮游病毒数量相当丰富,2010 年从春季至冬季数量变化范围是 $(6.40 \times 10^8 - 3.59 \times 10^{10})$ ind./L,并且呈现出中部海域丰度高于近岸海域的情况(王健等, 2013)。本文所研究的站位 B47(38.66°N, 118.97°E)处于渤海远离岸边的海域,受近岸的环境因素影响较小,2011 年 12 月以流式细胞仪测得的病毒丰度为 8.97×10^9 ind./L,因此选择这个点的数据来代表渤海浮游病毒群落的大致情况。

宏基因组学(metagenomics)作为研究环境总基因组的有效方法在 1998 年第一次提出(Handelsman *et al*, 1998),可以直接对环境样品中未经培养的微生物群

落复杂的总基因组进行分析,通过遗传物质的提取、测序、拼接、注释、统计分析等步骤(Thomas *et al*, 2012),了解环境微生物群落的物种组成、功能基因结构、进化地位、种间关系以及与环境因子的相关性(Handelsman, 2004)。

近十多年来,有多个探索海洋病毒宏基因组的研究成果被发表。有针对海水中病毒组的相关研究(Angly *et al*, 2006; Steward *et al*, 2011; Williamson *et al*, 2012; Hurwitz *et al*, 2013; Winter *et al*, 2014; Brum *et al*, 2015),从病毒物种组成和多样性、各大洋病毒宏基因组之间的关系、病毒种之间的进化关系、病毒群落结构与环境因子之间以及地域、水深、离岸距离的相关性等方面全面阐述了大洋海水中的病毒宏基因组特征;也有针对海洋沉积物病毒组的相关研究(Breitbart *et al*, 2004; Yoshida *et al*, 2013),对近岸海域以及深海的沉积物病毒组作出具体分析。

目前还没有针对渤海这一海域的病毒宏基因组

* 国家自然科学基金项目,41076088 号;国家重点基础研究发展计划(973)项目,2013CB29704 号;中国海洋大学中央高校基本科研业务费,201562018 号。夏骏,硕士研究生, E-mail: xiafrank@qq.com

通讯作者:汪岷,博士,教授, E-mail: mingwang@ouc.edu.cn

收稿日期:2015-11-30, 收修改稿日期:2016-03-25

的相关报道。本文所涉及的综合分析内容可为国内海域真光层病毒群落、微生物、海洋生态系统等各类研究提供相关数据支持。

1 材料与方法

1.1 采样和病毒宏基因组样品的预处理

于 2010 年 9 月和 2011 年 12 月东方红 2 号国家自然科学基金开放航次中, 用潜水泵在渤海 B47 站 (38.66°N, 118.97°E)(图 1)采集表层海水样各 150L。海水润洗塑料桶三次后泵入海水。先使用直径 300mm, 孔径 3 μ m 的混合纤维素膜进行第一次过滤, 再使用直径 300mm、孔径 0.22 μ m 的混合纤维素膜进行第二次过滤, 去除非病毒生物颗粒。滤过液经中型切相流系统(膜包: Pellicon® 2 Cassette, 材质: 聚醚砜, 孔径: 50kDa)浓缩至 0.5L, 再用小型切相流系统(膜包: Pellicon® XL Cassette, 材质: 聚醚砜, 孔径: 50kDa)浓缩至 0.01L。将高度浓缩后的样品置入液氮中速冻, 船上-20°C 储存, 实验室提取 DNA 前-80°C 储存。

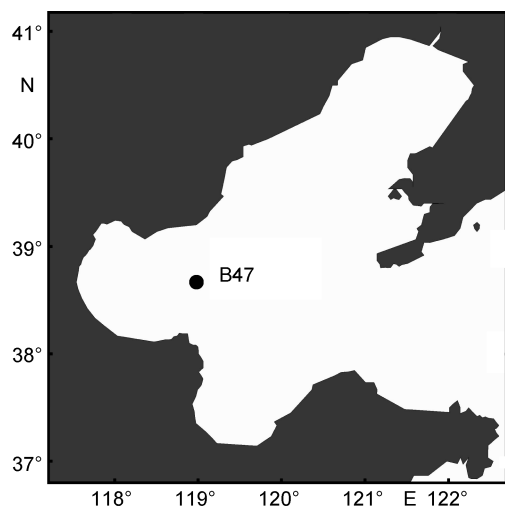


图 1 渤海病毒宏基因组采样站位图

Fig.1 Sites map for sampling viral metagenome in Bohai Sea

1.2 DNA 的提取

样品 37°C 水浴解冻, 加入聚乙二醇(终浓度 10%) 和 NaCl(终浓度 0.6%), 4°C 避光静置 24h。在 4°C 下以 13500r/min 离心 40min, 去除上清液, 让沉淀悬浮于 300 μ L 钠-镁离子缓冲液。加入 100 μ L KCl, 冰浴 30min。4°C 12000r/min 离心 10min, 去除沉淀。加入 10 μ L 蛋白酶 K, 混匀后再加入 20 μ L 10%十二烷基硫酸钠溶液, 56°C 金属浴 1h(每 10min 震荡摇匀一次)。连续两次用 215 μ L 的平衡酚和 215 μ L 的氯仿-异戊醇

萃取, 取上清液 20°C 12000r/min 离心 4min, 去除有机相。加入 43 μ L 醋酸钠, 860 μ L 无水乙醇, -20°C 静置 3h。4°C 12000r/min 离心 15min, 去除上清液, 加入 0.001L 70%乙醇, 使沉淀悬浮, 4°C 12000r/min 离心 5min。去除上清液, 加入 0.001L 无水乙醇, 吹打悬浮, 4°C 12000r/min 离心 5min。吹干, 使沉淀溶于 30 μ L Tris-EDTA 溶液, 4°C 静置 30min, -80°C 保存。

1.3 建库、测序和拼接

利用 Hiseq 2000(Illumina)为样品进行高通量测序, 2 $\times$ 100bp 双末端测序, 插入片段长度为 170bp, 测序公司为深圳华大基因科技有限公司(BGI)。首先对样品进行检测, 然后根据 Illumina 标准建库测序流程进行测序: (1) 电泳回收主要的 DNA 片段; (2) 用 T₄ DNA Polymerase、Klenow DNA Polymerase 和 T₄ 多核苷酸激酶将打断形成的黏性末端修复成平末端; (3) 通过 3'端加碱基“A”, 使得 DNA 片段能与 3'端带有“T”碱基的特殊接头连接; (4) 用合格的文库进行 cluster 制备和测序。对得到的所有 reads 序列进行进一步处理: (1) 去除含 N 碱基数目总和达到 3 个或以上的 reads; (2) 去除与 reads 序列有 15bp 的重叠区的接头序列污染; (3) 去除质量值连续(默认 reads 中质量值小于 Q20 的碱基数目大于 36, 设置为 84、36); (4) 去除重复污染; (5) 对数据进行去宿主污染的分析, 一致性大于等于 90%的条件下, 与宿主比对上的被认为是宿主污染 reads。最终得到干净数据。用 SOAP denovo (1.06 版)软件对 reads 进行拼接, 设置多个 K 值选择质量最高的结果, 2010 年 9 月样品选择 K 值为 59, 2011 年 12 月样品选择 K 值为 47, 参数设置为“SOAP denovo-63mer all-K*-p8-F-M2-d1-R-u-k-o”。拼接完成后过滤掉长度小于 500bp 的 scaffold。

1.4 病毒组物种组成分析

上传拼接好的 scaffold 序列, 利用 MetaVir 网站平台(<http://metavir-meb.univ-bpclermont.fr/>)在病毒基因组参考序列蛋白库(Refseq complete viral genomes protein sequences database)中进行 Blastp 比对 (e-value<10⁻³)。用 GAAS(Genome relative Abundance and Average Size)软件进行病毒物种分类统计(Roux *et al*, 2014)。

1.5 功能基因分析

使用 MetaGeneMark(2.10 版)软件预测 scaffold 中的开放阅读框(ORF)(Zhu *et al*, 2010), 再用 CD-HIT (4.6 版)软件去除冗余的 ORF。将非冗余的 ORF 利用 Blastp(e-value<10⁻³)在 WebMGA 网站上进行 COG 功

能基因数据库比对(Wu *et al.*, 2011)。

1.6 系统进化分析

两个样品中的病毒超过 97%属于双链 DNA 病毒, 所以选择以双链 DNA 病毒基因的保守结构域 (conserved domain)AVS、G20、GP23、MCP、PhoH、Polβ、Polβ2、T7gp17、TerL 作为待建树的标记基因, 用 Blastp(e-value<10⁻³) 在 ORF 中查找。其中属于 Polβ 和 TerL 保守结构域中的各一条基因在样品的 ORF 和 GenBank NR 库均有最好的比对结果(e-value<10⁻¹⁰), 被挑选出来进行系统进化分析, 分别是 DNA 聚合酶 Polβ 中的 CyPHP-SSM4(pfam 编号: PF00136)和 DNA

末端酶大亚基 TerL 中的 Q58LJ7_BPPRS(pfam 编号: 03237)。在用 CLASTALW 进行两两比对和多序列比对后, 用 MEGA(6.06 版)软件建立邻接树(Hall, 2013)。

2 结果

2.1 病毒物种分类

2010 年 9 月和 2011 年 12 月的样品中各拼接得到 19875 条、31319 条 scaffold, 总长度分别为 28.29Mbp、33.46Mbp, 拼接序列其它数据见表 1。两个样品预测到非冗余 ORF 48227 个(2010 年 9 月)、63660 个(2011 年 12 月)(表 1)。

表 1 reads 拼接以及开放阅读框预测结果
Tab.1 Result of reads assembly and predicted ORF

样品	序列数	总长 (Mbp)	N50 (bp)	N90 (bp)	最长 (bp)	最短 (bp)	序列定位			开放阅读框数量	开放阅读框平均长度
							双末端拼接	单末端拼接	%		
B47_Sep.2010	19875	28.29	1820	614	73314	500	4407626	1227695	37.74	48227	529
B47_Dec.2011	31319	33.46	1115	568	38650	500	1926379	756354	30.73	63660	484

在 NR 库(non-redundant protein database, 非冗余蛋白库)中比对上(blastx, e-value<10⁻³)病毒序列的分别有 9716 条(48.89%)和 13599 条(43.42%)。两个样品共比对上 1240 个病毒种。与太平洋深海沉积物中大多数序列属于单链 DNA 病毒不同(Mitsuhiro *et al.*, 2013), 在两个样品比对上病毒基因组的序列中, 双链 DNA 病毒占据主要地位, 2010 年 9 月的样品中占 97.65%, 2011 年 12 月的样品中占 97.51%。以病毒目来划分, 有尾噬菌体目占据主导地位(81.19%、80.63%)。以病毒科来划分, 除了有尾噬菌体目下的肌尾病毒科(2010 年 9 月为 27.26%, 2011 年 12 月为 21.08%)、长尾病毒科(27.83%和 25.34%)和短尾病毒科(23.64%和 30.90%)外, 藻类 DNA 病毒科(3.04%和 3.28%)和米米病毒科(1.38%和 0.83%)也占据一定的比例, 这五个科的病毒序列数量占到总病毒序列数的 83.15%(2010 年 9 月)、81.43%(2011 年 12 月)(图 2)。

两个样品中总序列数排在前十的病毒种(图 3)为 *Puniceispirillum* phage HMO-2011(5.26%和 6.97%)、*Pelagibacter* phage HTVC008M(3.58%和 2.77%)、*Pelagibacter* phage HTVC010P(2.33%和 3.01%)、cyanophage KBS-S-2A(1.82%和 1.15%)、*Cellulophaga* phage phi38:1 (0.54%和 1.97%)、*Prochlorococcus* phage P-SSM2(1.17%和 1.52%)、*Synechococcus* phage S-SM2(1.86%和 0.51%)、*Pelagibacter* phage HTVC011P(0.60%和 1.38%)、*Pelagibacter* phage

HTVC019P(0.70%和 1.17%)、*Idiomarinaceae* phage 1N2-2(0.98%和 0.93%)。这十个种的病毒序列占到总序列数的 20.32%。

序列按照主要宿主来划分病毒大类(图 4), *Synechococcus* phage (共 33 个种, 10.28%)、*Pelagibacter* phage (共 4 个种, 7.86%)和 *Puniceispirillum* phage (共 1 个种, 6.26%)为数量最多的三大类, 以这 12 类微生物为宿主的病毒序列占到两个样品总序列数的 52.01%。而其中在两个样品中差距最大的为 *Synechococcus* phage, 2010 年 9 月样品中所占比例为 12.75%, 而 2011 年 12 月的为 8.52%。原绿球藻噬藻体(*Prochlorococcus* phage)所占比例为 3.09%至 3.73%。

另外, 对于存在于渤海病毒组中不占优势的病毒大类, 共有 27 条序列比对上单链 DNA 病毒的 8 个种, 5 条比对上单链 RNA 病毒 5 个种。

物种多样性指数采用 Shannon-Wiener 指数 H' , 以序列的个数作为病毒个体数进行计算, 2010 年 9 月病毒组为 5.87nats, 2011 年 12 月的为 5.83nats。Pielou's 均匀度 J' 冬季为 0.612, 秋季为 0.639。

2.2 病毒功能基因多样性

将所有预测到的 ORF 在 COG(Clusters of Orthologous Groups of proteins, 直系同源蛋白簇)数据库中进行比对(blastp, e-value<10⁻³)。共有 20.85%(2010 年 9 月)和 9.39%(2011 年 12 月)的 ORF 在 COG 数据库中比对上。在具有特定功能的基因中,

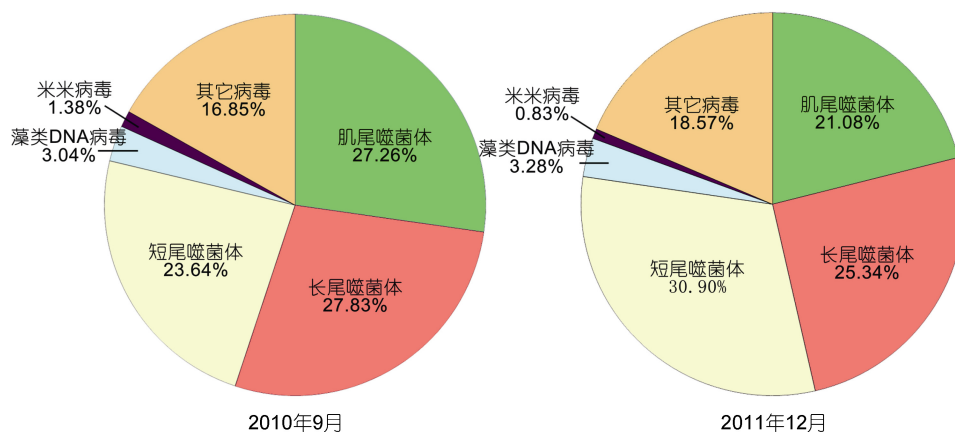


图2 病毒物种科级分类
Fig.2 Viral taxonomy in family

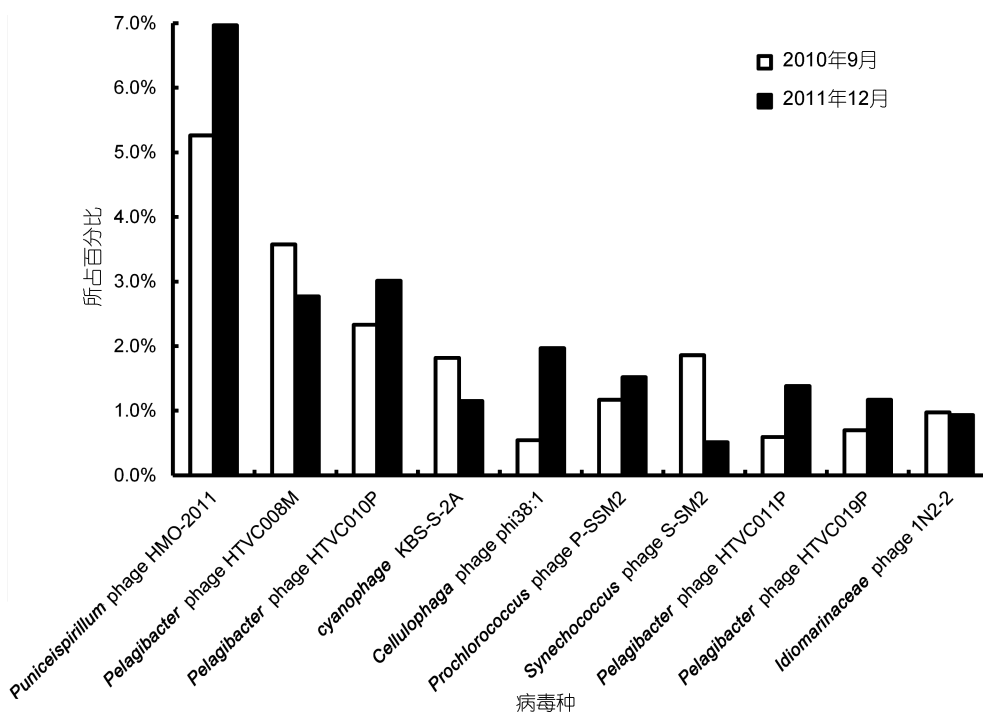


图3 序列数量丰富度前十的病毒种
Fig.3 The top 10 virus species in sequence abundance

以复制、结合和修复蛋白最为丰富,其次为细胞壁/细胞膜/包膜合成蛋白,以及转录蛋白等(表 2)。渤海冬季病毒群落复制、结合和修复蛋白所占比例远高于秋季,在北黄海的 9 个病毒组中呈现相同的现象(汪俭, 2015)。

2.3 系统发育关系

将全部 ORF 与各标记基因比对(blastp, e-value< 10^{-3}), 有多个 ORF 被比对上, 结果见表 3, 其中以 DNA 末端酶大亚基 TerL 最多, 其次为 T7 噬菌体尾丝蛋白 gp17 和 DNA 聚合酶 Polβ。选取比对结果最

好的两个标记基因 Polβ 和 TerL 建立 N-J 系统进化树(图 5, 6)。结果显示, 渤海病毒组的 ORF 在已分离的病毒基因中, 亲缘关系与聚球藻噬藻体(*Synechococcus* phage)、原绿球藻噬藻体(*Prochlorococcus* phage)、根瘤菌噬菌体(*Sinorhizobium* phage)、远洋杆菌噬菌体(*Pelagibacter* phage)较近。

在以 Polβ 建的系统进化树中, 一支 Cluster I 与已知的蓝藻噬藻体亲缘关系较远, 与之最近的一个种为 *Sinorhizobium* phage phiM12。以 TerL 建的系统进化树中, 渤海病毒组的序列较为分散, 其中 Cluster II

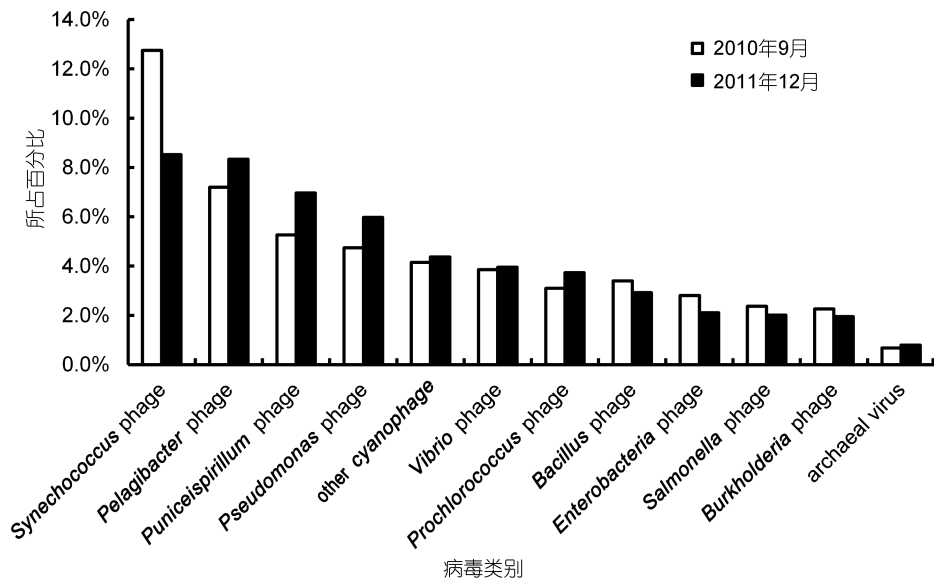


图 4 以主要宿主分类的前 12 大类病毒
Fig.4 The top 12 major viral groups classified in by hosts

表 2 COG 功能基因注释结果
Tab.2 Functional gene annotation by COG

类别	描述	2010 年 9 月	2011 年 12 月
L	复制、结合和修复蛋白	16.39%	33.77%
M	细胞壁/细胞膜/包膜合成蛋白	6.19%	7.46%
K	转录蛋白	5.23%	4.01%
O	翻译后修饰、蛋白转换和分子伴侣蛋白	4.02%	4.92%
F	核酸转运和代谢蛋白	3.26%	3.65%
E	氨基酸转运和代谢蛋白	4.64%	0.84%
C	能量产生和转化蛋白	4.09%	0.50%
G	碳水化合物转运和代谢蛋白	3.24%	0.47%
J	翻译、核糖体结构和合成蛋白	2.85%	1.07%
P	无机盐转运和代谢蛋白	2.88%	0.30%

注: 只列出有特定功能的、比对上的 ORF 总数排在前十位的基因

一支的 bootstrap 值为 100%, 聚类极为可信, 其进化地位同样在 T4-like virus 的 HTVC008M 与 phiM12 之间, 但不与 phiM12 在同一分支上。

3 讨论

在渤海表层海水 2010 年 9 月和 2011 年 12 月的

两个病毒宏基因组中, 大部分的病毒序列属于双链 DNA 病毒, 其中肌尾病毒科、长尾病毒科、短尾病毒科、藻 DNA 病毒科占据主导地位。类似结果也出现在近期的印度洋和太平洋东北部真光层海水病毒组中, 渤海各级分类与太平洋物种分类结果接近一致, 原因可能是均没有对样品 DNA 进行扩增, 最大程度地还原了病毒群落结构的真实情况(Steward *et al*, 2011)。而印度洋虽然同样以双链 DNA 病毒—有尾噬菌体目占绝对主导地位, 但其扩增后的病毒组中肌尾病毒科占据 54.3% 与渤海病毒组相差较大(Williamson *et al*, 2012)。

噬藻体(cyanophage)能够侵染蓝细菌中聚球藻属和原绿球藻属两个重要的初级生产者, 在自然水体中数量极为丰富(Bergh *et al*, 1989), 其对宿主的致死率在 1% 至 8%(Garza *et al*, 1998), 是微食物环(microbial loop)中的重要成员。噬藻体序列在病毒组序列中占有相当大的比例, 通过系统进化分析指明, 渤海噬藻体中存在未知的物种分支 Cluster I(图 5)和 Cluster II(图 6), 且很有可能属于肌尾噬菌体—T4-like 病毒。对未知物种分支的探索, 可为将来的病

表 3 保守结构域在样品开放阅读框中比对的匹配结果
Tab.3 The alignment between conserved domains and ORFs of samples

样品	保守结构域								
	AVS	G20	GP23	MCP	PhoH	Polβ	Polβ2	T7gp17	TerL
2010 年 9 月	23	41	41	0	27	49	3	68	529
2011 年 12 月	14	13	14	0	26	45	23	94	747

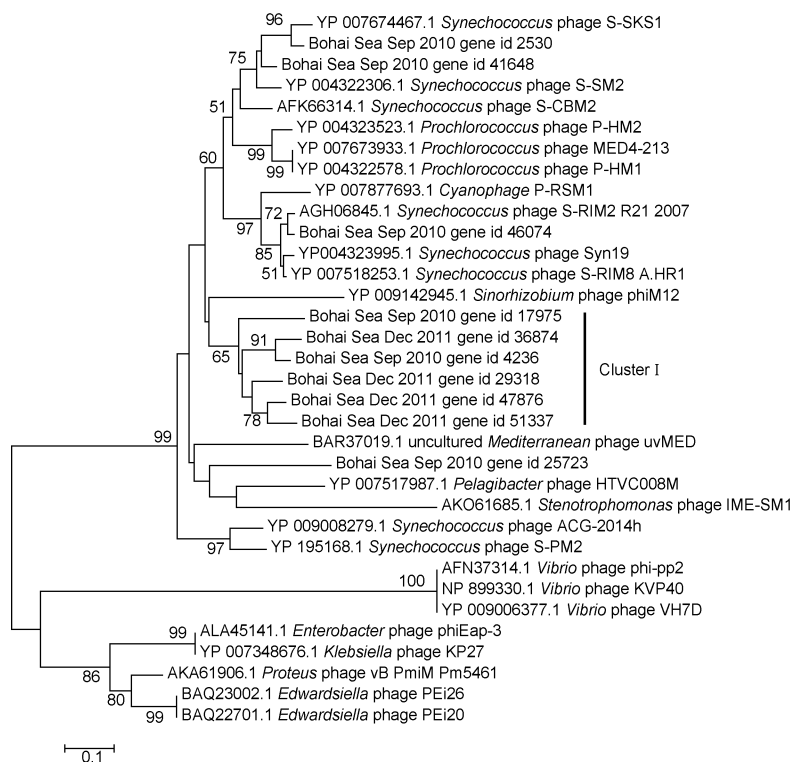


图 5 渤海病毒与已知病毒 Polβ(DNA 聚合酶 β)保守结构域基因的 N-J 系统进化树

Fig.5 Neighbor-joining tree based on conserved domain Polβ (DNA polymerase β) from known virus
以 *Vibrio* phage、*Enterobacter* phage、*Proteus* phage 和 *Edwardsiella* phage 为外类群

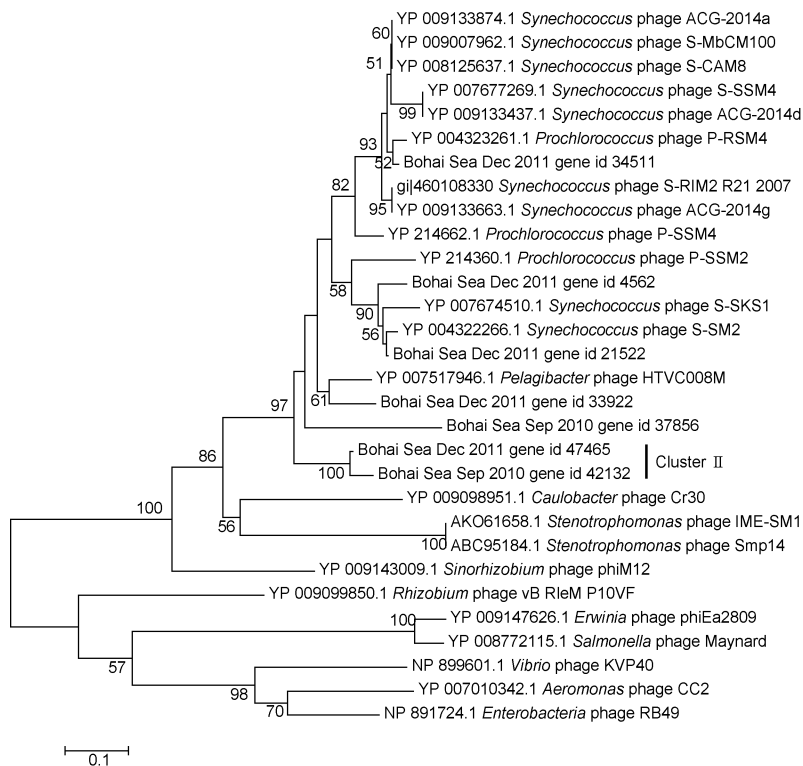


图 6 渤海病毒与已知病毒 TerL(DNA 末端酶大亚基)保守结构域基因的 N-J 系统进化树

Fig.6 Neighbor-joining tree based on conserved domain TerL (DNA terminase large subunit) from known virus
以 *Rhizobium* phage、*Vibrio* phage、*Enterobacter* phage、*Erwinia* phage、*Aeromonas* phage 和 *Salmonella* phage 为外类群

毒分离提供指导,并完善病毒宏基因组中的“Unknown”物种部分。

聚球藻(*Synechococcus*)对于海洋初级生产力的贡献为 25%(Li, 1994; Smith *et al*, 2001),是海洋中重要的微生物类群。在黄海海域,由于水温降低,聚球藻冬季的数量小于秋季数量(Zhao *et al*, 2011)。聚球藻噬藻体呈现出与宿主相同的变化趋势,2011 年 12 月所占比例小于 2010 年 9 月。另一个重要的大类,原绿球藻噬藻体(*Prochlorococcus* phage)所占比例为 3.09%至 3.73%。但相关研究指出,原绿球藻仅分布于夏季水温较高的亚热带和热带海域(焦念志, 2006),对中国海域原绿球藻生长的实验研究也表明,其最适生长水温范围在 25—30°C(冯宪栋等, 2007)。渤海 B47 站 2011 年夏季表层海水水温为 18.0°C,且本实验室并没有以流式细胞仪测出原绿球藻的存在。可能的原因是原绿球藻噬藻体宿主范围较广,能够侵染其他属蓝细菌,或是对马暖流分支将部分聚球藻从亚热带海域携带至渤海。

在冬季和秋季渤海病毒组中,4 个病毒种 *Puniceispirillum* phage HMO-2011、*Pelagibacter* phage HTVC010P、*Pelagibacter* phage HTVC008M 和 *Synechococcus* phage S-SM2 占据优势地位。海洋真光层细菌其中一个主要分支为 SAR116(*Candidatus* *Puniceispirillum* 类群),它对于甲基营养和光合异养代谢有着潜在影响(Grote *et al*, 2011; Giovannoni *et al*, 2012)。*Puniceispirillum* phage HMO-2011 序列在印度洋的 3 个病毒组以及太平洋的 4 个病毒组中占到病毒序列总数的 10.3%—25.3%,为丰富度第一或第二的病毒种(Kang *et al*, 2013),与之前报道的太平洋病毒组(POV)中开阔海域和近岸的 SAR11 噬菌体 *Pelagibacter* phage HTVC010P 总数结合来看(Zhao *et al*, 2013),大洋真光层水体中最占优势的病毒种为 HMO-2011 和 HTVC010P。其次为 *Pelagibacter* phage HTVC008M 和 *Synechococcus* phage S-SM2(Kang *et al*, 2014)。因此处于太平洋西岸的渤海,样品中这四种已分离纯化病毒的基因序列比例占据相当大的优势应为正常现象。但渤海病毒组与其它海域不同的是, cyanophage KBS-S-2A(1.43%)、*Cellulophaga* phage phi38:1(1.38%)和 *Idiomarinaceae* phage 1N2-2(0.95%)为序列数排在前十位的优势种。

仅有极少数的单链 DNA 病毒存在于渤海病毒组中,这与 MetaVir 病毒宏基因组数据库中大部分海域的病毒组相似,但也与部分海域差异很大,如 2005

年马尾藻海的 ssDNA 序列占到了总病毒序列的 22.43%。27 条单链 DNA 序列中最丰富的病毒科为丝杆噬菌体(*Inoviridae*)。有研究表明,海洋中存在大量单链 DNA 病毒(Labonté *et al*, 2013),但由于其高突变率难以被分离纯化,数据库中仅有少数病毒种的全基因组序列。在病毒组研究中,将以发掘海水中未知单链 DNA 病毒种为未来的一个方向。

在冬季病毒种数(1093)高于秋季(1010)、病毒序列数(13599)也高于秋季(9716)的情况下,冬季病毒物种多样性低于秋季,原因是冬季优势种的优势度更大。从 Pielou's 均匀度 J' (冬季为 0.612,秋季为 0.639)可以看出,秋季各物种序列数更为平均。冬季功能基因中的复制、结合和修复基因所占比例却是秋季的两倍多,说明渤海病毒群落(特别是劣势种)在冬季水温较低的情况下代谢活性减弱,但以能够适应低温条件的微生物为宿主的病毒则更具优势。例如 *Puniceispirillum* phage 和 *Pelagibacter* phage 两个主要的大类,在冬季的比例明显高于秋季。聚球藻噬藻体在冬季数量降低,是由于宿主聚球藻的数量与水温呈显著相关,在中国海域数量有冬季<秋季的规律(赵苑, 2010)。

进化关系中,与 Cluster I 最近的一个种为 *Sinorhizobium* phage phiM12,宿主 *Sinorhizobium meliloti* 1021 能在土壤中与根系共生,达到固氮的作用(Stroupe ME *et al*, 2014),属于有尾噬菌体目-肌尾噬菌体科-T4-like 病毒属。T4-like 噬菌体有着可收缩尾部因而宿主范围广,从肠杆菌至蓝细菌噬藻体,广泛分布于海水和淡水中(Mann *et al*, 2005; Sullivan *et al*, 2005; Weigele *et al*, 2007; Dreher *et al*, 2011)。Cluster I 可能代表海洋真光层中新的一类病毒,从 Pol β 进化地位在同属于 T4-like 病毒的 *Pelagibacter* phage HTVC008M 与 *Sinorhizobium* phage phiM12 之间可以推断,Cluster I 所代表的一类病毒很可能属于 T4-like 病毒。Cluster II 极为可信,进化地位在 HTVC008M 与 phiM12 之间,可能属于 Cluster I 以外的另一支 T4-like 病毒类群。

在将来的海洋病毒宏基因组研究中,采样和实验方法需要联系最新的研究进展不断改善,如最近的 Fe 离子沉降病毒法(John *et al*, 2011)可使病毒宏基因组采样步骤简化。数据分析方面,除了对群落的基础分析,更应在结合环境因子和宿主序列,拓宽研究病毒大类领域,明确病毒群落的海洋生态学地位方面有所突破。

4 结论

综上所述, 渤海 B47 站的表层海水病毒宏基因组中双链 DNA 病毒占据主要地位, 前五位的病毒科依次为肌尾病毒科、长尾病毒科、短尾病毒科、藻类 DNA 病毒科、米米病毒科, 病毒种类与太平洋、大西洋、印度洋等远洋相比, 有一定的独特性。渤海秋季病毒物种多样性和均匀性都比冬季更高, 且功能基因分析结果表明秋季病毒代谢活性更强。原绿球藻一般情况下只存在于水温较高的亚热带海域, 但在秋季和冬季水温较低的渤海海域, 病毒组中却存在原绿球藻噬藻体, 值得进一步研究探讨原因。在系统进化分析结果中发现噬藻体中可能属于未知 T4-like 噬菌体的病毒分支, 可为噬菌体分离提供指导。

参 考 文 献

- 王 健, 汪 岷, 刘 哲等, 2013. 渤海浮游病毒的时空分布. 海洋与湖沼, 44(6): 1597—1603
- 冯宪栋, 蒋霞敏, 符尧尧, 2007. 理化因子对原绿球藻生长及其色素含量的影响. 水产科学, 26(12): 643—647
- 汪 俭, 2015. 北黄海浮游病毒群落的宏基因组学研究. 青岛: 中国海洋大学硕士学位论文, 36—37
- 赵 苑, 2010. 黄海和东海微小型浮游生物分布研究. 青岛: 中国海洋大学博士学位论文, 101—102
- 焦念志, 2006. 海洋微生物生态学. 北京: 科学出版社, 77—78
- Angly F E, Felts B, Breitbart M *et al*, 2006. The marine viromes of four oceanic regions. PLoS Biology, 4(11): e368
- Bergh Ø, Børsheim K Y, Bratbak G *et al*, 1989. High abundance of viruses found in aquatic environments. Nature, 340(6233): 467—468
- Breitbart M, Felts B, Kelley S *et al*, 2004. Diversity and population structure of a near-shore marine—sediment viral community. Proceedings of the Royal Society B: Biological Sciences, 271(1539): 565—574
- Brum J R, Ignacio-Espinoza J C, Roux S *et al*, 2015. Patterns and ecological drivers of ocean viral communities. Science, 348(6237): 1261498
- Brussaard C P D, Kuipers B, Veldhuis M J W, 2005. A mesocosm study of *Phaeocystis globosa* population dynamics: I. Regulatory role of viruses in bloom control. Harmful Algae, 4(5): 859—874
- Brussaard C P D, Martínez J M, 2008. Algal Bloom Viruses. Plant Viruses, 2(1): 1—13
- Dreher T W, Brown N, Bozarth C S *et al*, 2011. A freshwater cyanophage whose genome indicates close relationships to photosynthetic marine cyanomyoviruses. Environmental Microbiology, 13(7): 1858—1874
- Garza D R, Suttle C A, 1998. The Effect of cyanophages on the mortality of *Synechococcus* spp. and selection for UV resistant viral communities. Microbial Ecology, 36(3—4): 281—292
- Giovannoni S J, Vergin K L, 2012. Seasonality in ocean microbial communities. Science, 335(6069): 671—676
- Grote J, Bayindirli C, Bergauer K *et al*, 2011. Draft genome sequence of strain HIMB100, a cultured representative of the SAR116 clade of marine *Alphaproteobacteria*. Standards in Genomic Sciences, 5(3): 269—278
- Hall B G, 2013. Building phylogenetic trees from molecular data with MEGA. Molecular Biology and Evolution, 30(5): 1229—1235
- Handelsman J, 2004. Metagenomics: application of genomics to uncultured microorganisms. Microbiology and Molecular Biology Reviews, 68(4): 669—685
- Handelsman J, Rondon M R, Brady S F *et al*, 1998. Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. Chemistry & Biology, 5(10): R245—R249
- Hurwitz B L, Sullivan M B, 2013. The Pacific Ocean Virome (POV): a marine viral metagenomic dataset and associated protein clusters for quantitative viral ecology. PLoS One, 8(2): e57355
- John S G, Mendez C B, Deng L *et al*, 2011. A simple and efficient method for concentration of ocean viruses by chemical flocculation. Environmental Microbiology Reports, 3(2): 195—202
- Kang I, Cho J C, 2014. Depth-specific distribution of the SAR116 phages revealed by virome binning. Journal of Microbiology and Biotechnology, 24(5): 592—596
- Kang I, Oh H M, Kang D *et al*, 2013. Genome of a SAR116 bacteriophage shows the prevalence of this phage type in the oceans. Proceedings of the National Academy of Sciences of the United States of America, 110(30): 12343—12348
- Labonté J M, Suttle C A, 2013. Previously unknown and highly divergent ssDNA viruses populate the oceans. The ISME Journal, 7(11): 2169—2177
- Li W K W, 1994. Primary production of prochlorophytes, cyanobacteria, and eucaryotic ultraphytoplankton: measurements from flow cytometric sorting. Limnology and Oceanography, 39(1): 169—175
- Mann N H, Clokie M R J, Millard A *et al*, 2005. The genome of S-PM2, a “photosynthetic” T4-type bacteriophage that infects marine *Synechococcus* strains. Journal of Bacteriology, 187(9): 3188—3200
- Roux S, Tournayre J, Mahul A *et al*, 2014. Metavir 2: new tools for viral metagenome comparison and assembled virome analysis. BMC Bioinformatics, 15(1): 76
- Smith E M, Kemp M W, 2001. Size structure and the production/respiration balance in a coastal plankton community. Limnology and Oceanography, 46(3): 473—485
- Steward G F, Preston C M, 2011. Analysis of a viral metagenomic library from 200 m depth in Monterey Bay, California constructed by direct shotgun cloning. Virology Journal, 8(1): 287
- Stroupe M E, Brewer T E, Sousa D R *et al*, 2014. The structure of *Sinorhizobium meliloti* phage ΦM12, which has a novel T=19l triangulation number and is the founder of a new group of T4-superfamily phages. Virology, 450—451:

- 205—212
- Sullivan M B, Coleman M L, Weigele P *et al*, 2005. Three *Prochlorococcus* cyanophage genomes: signature features and ecological interpretations. *PLoS Biology*, 3(5): e144
- Suttle C A, 2005. Viruses in the sea. *Nature*, 437(7057): 356—361
- Thomas T, Gilbert J, Meyer F, 2012. Metagenomics—a guide from sampling to data analysis. *Microbial Informatics and Experimentation*, 2(1): 3
- Weigele P R, Pope W H, Pedulla M L *et al*, 2007. Genomic and structural analysis of Syn9, a cyanophage infecting marine *Prochlorococcus* and *Synechococcus*. *Environmental Microbiology*, 9(7): 1675—1695
- Wilhelm S W, Suttle C A, 1999. Viruses and nutrient cycles in the sea: viruses play critical roles in the structure and function of aquatic food webs. *Bioscience*, 49(10): 781—788
- Williamson S J, Allen L Z, Lorenzi H A *et al*, 2012. Metagenomic exploration of viruses throughout the Indian Ocean. *PLoS One*, 7(10): e42047
- Winter C, Garcia J A L, Weinbauer M G *et al*, 2014. Comparison of deep-water viromes from the Atlantic Ocean and the Mediterranean Sea. *PLoS One*, 9(6): e100600
- Wu S T, Zhu Z W, Fu L M *et al*, 2011. WebMGA: a customizable web server for fast metagenomic sequence analysis. *BMC Genomics*, 12(1): 444
- Yoshida M, Takaki Y, Eitoku M *et al*, 2013. Metagenomic analysis of viral communities in (Hado) pelagic sediments. *PLoS One*, 8(2): e57271
- Zhao Y L, Temperton B, Thrash J C *et al*, 2013. Abundant SAR11 viruses in the ocean. *Nature*, 494(7437): 357—360
- Zhao Y, Zhao L, Xiao T *et al*, 2011. Spatial and temporal variation of picoplankton distribution in the Yellow Sea, China. *Chinese Journal of Oceanology and Limnology*, 29(1): 150—162
- Zhu W H, Lomsadze A, Borodovsky M, 2010. *Ab initio* gene identification in metagenomic sequences. *Nucleic Acids Research*, 38(12): e132

METAGENOMIC STUDY ON VIRAL DIVERSITY IN AUTUMN AND WINTER IN BOHAI SEA

XIA Jun, WANG Min, GONG Zheng, JIANG Yong, MA Yu, WANG Jian, SHAO Hong-Bing, WANG Duo-Bing, XU Zhi-Meng

(Laboratory of Virology, College of Marine Life Science, Ocean University of China, Qingdao 266003, China)

Abstract To study the viral community, we collected two surface water samples in September, 2010 and December, 2011 in Bohai Sea, China. Virome study included taxonomic analysis, functional gene analysis, and phylogenetic analysis. Samples were filtered and concentrated by TFF (Tangential Flow Filtration) for DNA extraction. Results show that dsDNA virus (97.75%, scaffold percentage), especially the most abundant order Caudovirales (80.86%, scaffold percentage), was dominant in both samples. Classified in the host, *Synechococcus* phage (10.29%, scaffold percentage) was the main component. Furthermore, the most common functional genes of viroplankton in Bohai Sea were those for replication, recombination, and repair, and the proportion in winter (33.77%, ORF percentage) was more than that in autumn (16.39%, ORF percentage). Moreover, phylogenetic analysis indicated that the viral genes are clustered with conserved domains of *Synechococcus* phage, *Pelagibacter* phage, and *Prochlorococcus* phage. The virome included few *Prochlorococcus* that existed theoretically in subtropical waters. The species diversity and functional gene proportion varied seasonally in own unique characteristics to that of pelagic ones, and a possible branch of unknown T4-like phage may exist.

Key words metagenomics; virome; marine virus; viral diversity; Bohai Sea