

文章编号: 1000-0615(2018)02-0216-10

DOI: 10.11964/jfc.20161210667

## 基于长片段PCR扩增的牡蛎疱疹病毒基因组高通量测序

史杰<sup>1,2</sup>, 白昌明<sup>2</sup>, 李晨<sup>2</sup>, 蔡生力<sup>1</sup>, 王崇明<sup>2\*</sup>

(1. 上海海洋大学水产与生命学院, 上海 201306;

2. 中国水产科学研究院黄海水产研究所,

青岛海洋科学与技术国家实验室海洋渔业科学与食物产出过程功能实验室,

农业部海水养殖病害防治重点实验室,

青岛市海水养殖流行病学与生物安保重点实验室, 山东 青岛 266071)

**摘要:** 为获取2001年低温冻存栉孔扇贝感染牡蛎疱疹病毒(OsHV-1)变异株(ZK2001)基因组序列, 并分析ZK2001与其他OsHV-1变异株的序列差异和系统发育关系, 利用基于长片段PCR的基因组DNA的扩增和富集技术, 获取2001年栉孔扇贝感染OsHV-1变异株的基因组DNA; 再使用Illumina Hiseq 2500 PE250高通量测序平台对其测序。最后分析ZK2001与OsHV-1其他变异株基因组的序列差异和系统发育关系。测序数据组装后获得8个Scaffold。基因组变异分析结果显示, ZK2001与参考基因组相比存在328个SNP位点, SNP和序列插入/缺失变异是导致OsHV-1基因组序列变异的主要变异类型。系统发育分析结果显示, ZK2001变异株与分离自我国的OsHV-1变异株亲缘关系最近, 与分离自欧洲的OsHV-1 $\mu$ var及其相关变异株的亲缘关系最远, 说明中国和欧洲分布OsHV-1间存在因地理隔离导致的遗传分化。研究表明, 基于长片段PCR的DNA富集技术, 可以有效地扩增和富集冷冻样本中OsHV-1基因组DNA, 并应用于高通量测序。OsHV-1不同变异株基因组序列数据的获取和积累, 将为其基因组尺度的基因变异、株系演化和系统发育关系等研究提供重要基础。

**关键词:** 栉孔扇贝; 高通量测序; 长片段PCR; 牡蛎疱疹病毒

**中图分类号:** S 968.3

**文献标志码:** A

疱疹病毒是一类具有独特构造的双链DNA病毒, 在脊椎动物和无脊椎动物中广泛流行。完整疱疹病毒由内至外由核心、衣壳、被膜(tegument)及囊膜组成。疱疹病毒核心为双链基因组DNA, 大小为124~241 kb; 衣壳具有正二十面体结构( $T=16$ ), 大小为115~130 nm; 被膜由位于核衣壳与囊膜之间的基质蛋白组成; 囊膜来源于宿主细胞(磷脂层和膜蛋白), 同时含有来源于病毒编码的糖蛋白<sup>[1]</sup>。牡蛎疱疹病毒(*Ostreid herpesvirus 1*, OsHV-1)是疱疹病毒目(Herpesvirales), 软体动物疱疹病毒科(Malacoherpesviridae), 牡蛎疱疹病毒属(*Ostreavirus*)下的唯一一种;

也是最早被发现可以感染无脊椎动物的疱疹病毒种。与脊椎动物疱疹病毒较高的宿主特异性不同, OsHV-1具有跨种属感染和流行的致病特点, 已知可以感染牡蛎、扇贝、蛤仔和魁蚶(*Scapharca broughtonii*)等12种海洋双壳类<sup>[2-5]</sup>。OsHV-1的这一表型特点可能与其基因组比脊椎动物疱疹病毒基因组表现出更高的多态性有关。自2008年以来, 一种具有高致病力的变异株(OsHV-1 $\mu$ var)开始在法国海岸出现, 并导致长牡蛎(*Crassostrea gigas*)的大规模死亡<sup>[6]</sup>。流行病学调查结果显示, OsHV-1 $\mu$ var在法国的检出频率逐年增高, 至2009年完全取代OsHV-1, 成为造成

收稿日期: 2016-12-30 修回日期: 2017-05-09

资助项目: 国家自然科学基金(31502208, 31302233); 现代农业产业技术体系建设专项(CARS-48); 青岛海洋科学与技术国家实验室鳌山科技创新计划(2015ASKJ02)

通信作者: 王崇明, E-mail: wangcm@ysfri.ac.cn

法国长牡蛎死亡的主要变异株。随后, OsHV-1  $\mu$ var在意大利、西班牙、澳大利亚等国家相继出现, 并引起当地长牡蛎的大规模死亡<sup>[2]</sup>。除了该变异株之外, 其他OsHV-1变异株也被相继鉴定出来, 包括在栉孔扇贝(*Chlamys farreri*)中检测出的急性病毒性坏死病毒(acute viral necrosis virus, AVNV)<sup>[3]</sup>、魁蚶中检出的OsHV-1-SB<sup>[5]</sup>, 长牡蛎、菲律宾蛤仔(*Ruditapes philippinarum*)及欧洲扇贝(*Pecten maximus*)幼虫中检出的OsHV-1var等<sup>[7]</sup>。其中OsHV-1参考株、AVNV和OsHV-1-SB的全基因组序列已经通过Sanger测序技术成功获得, 基因组大小分别为207 439、210 993和199 354 bp; OsHV-1参考株与AVNV和OsHV-1-SB基因组DNA的相似度分别为95.2%和97.3%<sup>[8]</sup>。

由于Sanger测序技术通量低, 因此基于该技术进行OsHV-1基因组测序时, 存在覆盖深度低、费用高和耗时费力的缺点, 不适合大规模地开展多变异株基因组重测序和变异检测研究。近年来随着测序技术的发展, 高通量测序技术(next-generation sequencing, NGS)被越来越多地应用到包括疱疹病毒在内的多种病毒基因组测序相关的研究领域<sup>[9]</sup>。与Sanger技术相比, NGS通量大、费用低、周期短的优点显而易见<sup>[10]</sup>。但NGS技术对病毒核酸的纯度和数量(达到 $\mu$ g级)要求很高, 而无脊椎动物缺乏稳定细胞系, 目前主要使用密度梯度离心法从组织病料中纯化病毒粒子, 由于病毒回收效率太低, 所提取的DNA很难满足NGS的要求。

长片段PCR(Long-range PCR或者Long PCR)通常可以扩增3~15 kb的片段, 近来被用于包括病毒基因组和人类致病相关基因等目的基因的扩增、富集和高通量测序研究中<sup>[11-13]</sup>。本团队已成功建立了基于长片段PCR技术的OsHV-1基因组DNA富集技术, 利用该技术得到的OsHV-1基因组DNA可以用于高通量测序。本研究尝试利用该技术对长期冻存栉孔扇贝病料样本中, 不具备感染性且形态可能不完整的OsHV-1进行基因组扩增富集和高通量测序, 并分析其与已知OsHV-1几个变异株的基因组序列差异和系统发育关系。

## 1 材料与方法

### 1.1 实验材料

栉孔扇贝病料为2001年青岛胶州湾海区发生大规模死亡的养殖个体, 保存于-40 °C低温冰

箱中。OsHV-1参考株、AVNV和OsHV-1-SB(GenBank登录号分别为AY509253、GQ153938和KP412538)全基因组序列均下载自美国国家生物技术信息中心(National Center of Biotechnology Information, NCBI) DNA序列数据库。DNeasy<sup>®</sup>血液&组织DNA提取试剂盒、QIAquick<sup>®</sup>胶回收试剂盒购自QIAGEN, PrimeSTAR<sup>®</sup>GXL DNA聚合酶、LA taq购自宝生物工程(上海)股份有限公司, FastStart Essential DNA Probes Master实时定量PCR试剂盒购自Roche。所有引物、探针和质粒均由生工生物工程(上海)股份有限公司合成。

### 1.2 样品预处理

剪取1 g病贝外套膜组织, 使用剪刀剪碎, 经0.22  $\mu$ m滤膜过滤的灭菌海水洗涤3次后, 加入9 mL过滤海水。使用组织匀浆机(23 000 r/min, 匀浆3次, 每次10 s)搅拌获得组织匀浆液, 1000 $\times$ g (4 °C)离心5 min取上清液。吸取上清液180  $\mu$ L为靶物质, 按照DNeasy<sup>®</sup>血液&组织DNA提取试剂盒的操作步骤进行DNA的提取。阴性对照取自健康栉孔扇贝样本, 并使用相同处理获得DNA。

### 1.3 引物设计

长片段PCR引物(表1)根据GenBank中已发表的OsHV-1参考变异株(GenBank登录号为AY509253)序列, 使用在线引物设计软件GenoFrag设计(<http://genoweb1.irisa.fr/Serveur-GPO/outils/generationAmorces/GENOFRAG/indexen.php>)。

### 1.4 PCR反应

实时定量PCR 牡蛎疱疹病毒的实时定量PCR使用Martenot等<sup>[14]</sup>建立的探针法进行检测。每个PCR反应体系包括2 $\times$ Mix 12.5  $\mu$ L、上下游引物(10  $\mu$ mol/L)各1  $\mu$ L、探针(10  $\mu$ mol/L) 0.5  $\mu$ L、模板2  $\mu$ L、超纯水补充至25  $\mu$ L。PCR反应程序: 95 °C预变性10 min; 95 °C变性10 s, 60 °C 20 s, 40个循环。定量所用的质粒为10<sup>7</sup>~10<sup>1</sup>拷贝/ $\mu$ L、10倍梯度稀释的人工合成质粒。所用的引物和探针为BF: 5'-GTCGCATCTTTGGATTTAACAA-3', B4: 5'-ACTGGGATCCGACTGACAAC-3', BP: 6FAM-TGCCCTGTCATCTTGAGG TATAGACAATC-BHQ-1。

长片段PCR 长片段PCR采用25  $\mu$ L体系, 包括5 $\times$ Mix 5  $\mu$ L、dNTP 2  $\mu$ L、GXL DNA聚合酶0.5  $\mu$ L、上下游引物(10  $\mu$ mol/L)各0.5  $\mu$ L、模板

表 1 牡蛎疱疹病毒(OsHV-1)长片段PCR引物信息

Tab. 1 Primer sequences for Long-range PCR of OsHV-1

引物名称 primer ID	引物序列 primer sequences	引物名称 primer ID	引物序列 primer sequences
1F	CCC GCACACATACGACTACATAAA	12R	CGGCATTACCAAATATAGGCACACG
1R	GGTTCCAGTAGGGTGTTAAGAGC	13F	GGCACGCTCATTCTTACAACCTCTG
2F	GCAATCCAGTTCCCAAACCAATAGG	13R	CAACCAATCAGATCGACGAGACTCA
2R	ATCGCTTCCTATCACCTGTGGTCT	14F	GCGATGCCTTAATTGTTGCCAGAGT
3F	CCGTGAAATATCTGCCAAGGTGTTG	14R	TCCTGTGGAATGGTTGTTGGTGATG
3R	CGACCAGGAGAACATGAACGACTTT	15F	GCAAACGAAAGAGCGGCTATAACAG
4F	GCCCTCCTATTGGTACAAGATTGCT	15R	CTCCGTCATCGGTGTTATTACTAGG
4R	GAGCACAAACACTACCGCATACATG	16F	GTCGGTGTGGGTTTGAAATGTAG
5F	GCTTGTTTCTGGTGTCTGAGGTCA	16R	CTGGAAGCGAGTGTCAAGGTTAAAC
5R	ATGGCAGAAATAGAAACCCGAGGTC	17F	GCAGTTTGATTCATGTGTGGCAGAG
6F	ATCCAGTCTGTCAAATGCTCTCTC	17R	GCCATCCACCTCATATCCATTTCTC
6R	CCAGATATGAAGAGGAAGGGATGTC	18F	GGATGATGGATTGTTGGACGAGAGA
7F	ATGCCTGGGCGTAATTGTCTCTTGA	18R	GCTGCGGTCAGTACATGGTCATTTA
7R	CCAACCTTTCATCGTCACTCATCTC	19F	CGTGAAGACGCCATGAAGAGAAGTT
8F	GGGCGTCACTTTAGACTTCCAATC	19R	TCTGCCAGCCTCTGTGAACTTGTA
8R	TTCCCTGGCGATACTCTCATAGACA	20F	TTGGCAGATGAGGACACCTTATACC
9F	GGTCCGTCAACATCGAGAAAGAGAA	20R	TTCCTGATTCTCCACGCCATAACA
9R	CACGATAAATATGCTGCCTGGGTCA	21F	GGCAGCTAGTAAGGTCAATCTCAAC
10F	AGGGCGAGCATGGTCACATTCAAA	21R	TGGTTCCTGGGCGACGTTTACATAA
10R	GGGATATTCTGAGGGTGTGTGGA	22F	CGAAACGACAGGTTGAAGTGAGG
11F	GCCCAATAAACCTACAGAGGATGAG	22R	CAATGAATCGCCAATTAAGGAGG
11R	ATGGCAGATTCAGGAGAGGGTTGTA	23F	CCATTGTCAATCTCGGTTCTGC
12F	TGGCTTCTGTGGTGGTAGTTGTTGT	23R	GGAGGTGGGGTTTGAATACGAAG

1  $\mu$ L, 超纯水15.5  $\mu$ L。PCR反应程序: 94  $^{\circ}$ C预变性1 min; 98  $^{\circ}$ C变性10 s, 50  $^{\circ}$ C退火15 s, 68  $^{\circ}$ C延伸10 min, 35个循环。

### 1.5 长片段PCR产物检测

取4.5  $\mu$ L长片段PCR扩增产物, 加入1.5  $\mu$ L的上样缓冲液(GeneFinder与6 $\times$ loading buffer按1:9体积混合)混合后, 在浓度为0.8%的琼脂糖凝胶中以110 V电压电泳50 min; 使用凝胶成像系统观察并拍照, 存在约10 kb的特异性条带, 则确定成功扩增。

### 1.6 长片段PCR产物纯化和定量混合

对成功扩增的PCR产物, 按照QIAquick<sup>®</sup>胶

回收试剂盒的操作步骤进行产物纯化, 然后使用NanoDrop 2000超微量分光光度计对回收的产物进行定量。最后根据PCR产物定量浓度和大小, 将23对长片段产物按等摩尔数进行混合。混合后的长片段PCR产物送北京诺禾致源生物信息科技有限公司, 构建DNA小片段文库, 并使用Illumina Hiseq 2500 PE250高通量测序平台进行测序。

### 1.7 全基因组组装和分析

首先使用FASTX-Toolkit对原始数据进行过滤(去接头序列和低质量reads), 得到clean reads。序列组装以OsHV-1参考株(GenBank登录号为AY509253)作为参考, 使用基于de-Bruijn算法构图的SOAPdenovo与Velvet软件来完成。编写

perl脚本统计本研究中测序变异株(ZK2001)及其他3株变异株的GC含量。以OsHV-1为参考基因组,使用比对软件BWA(0.7.12)采用mem算法将clean reads比对到参考基因组上,利用Integrative Genomics Viewer(IGV)软件绘制reads在参考基因组上的分布图(测序深度)。

### 1.8 基因组序列变异分析

使用picard(1.129)对ZK2001变异株与OsHV-1参考株比对之后的结果进行变异位点(Variant)分析,Variant指在基因组水平上由单个核苷酸或者几个核苷酸的插入或缺失所引起的DNA序列多态性。然后使用软件GATK(3.4-46)对处理好的比对文件进行单个样本的Variant检测,检测到的变异使用Variant Filtration进行过滤。AVNV、OsHV-1-SB与OsHV-1的比对和SNP分析使用MUMmer 3.23软件,默认参数。最后通过SVG软件绘制SNP在基因组中的分布图。

使用ANNOVAR软件对检测出的Variant进行功能注释,将其中的SNP注释成同义突变和非同义突变,用perl脚本统计数据,最后使用SVG软件绘制每个开放阅读框(open reading frame, ORF)中氨基酸发生变异的个数和概率。

### 1.9 系统发育关系分析

从ZK2001基因组中抽取32个开放阅读框(ORF)核苷酸序列,与已公布的其他10个OsHV-1变异株的同源序列,使用最大似然法(maximum likelihood, ML)进行系统发育分析。先使用MAFFT选择合适的参数进行多重比对,再用Bioedit 7.0.0对比对结果进一步校正。使用jModelTest 2.1.4中的Akaike Information Criterion(AIC)标准,选择最优进化模型。PAUP 4.0b10构建ML树,选择启发式搜索(heuristic search),逐步添加法随机加入序列,参数设置为100。构树方法采用树二等分再连接(tree bisection reconnection, TBR),所有数据均未加权,获得50%一致性的树,支持率由1000次自展分支检验获得。

## 2 结果

### 2.1 PCR扩增结果

**实时定量PCR结果** 经实时定量PCR检测,本研究所用2001年采集的栉孔扇贝样本OsHV-1感染强度为 $8.7 \times 10^5$ 拷贝/mg组织,达到进

行长片段PCR扩增所要求的DNA感染强度( $10^4$ 拷贝/mg组织以上)。

**长片段PCR扩增结果** 23对长片段PCR引物均能成功扩增片段大小在10 kb左右的产物,与预期结果相吻合;而阴性对照均不能被扩增(图1)。

### 2.2 高通量测序结果

对高通量原始数据过滤后得到clean reads。使用SOAP denovo软件对clean reads进行拼接,参数设置:-R-d1-u-K#,其中-K参数设置多个不同的值(45、55、73、87、95),比较多个组装参数的N50长度,当-K为87时组装结果最优,保留200 bp以上的Scaffold序列并对其进行长度统计。共得到8个Scaffold,长度分别为170 095、18 675、1929、1805、1395、834、825和600 bp。

以OsHV-1参考株(AY509253)作为参考序列,统计测序深度,以及ZK2001和其他3个变异株GC含量分布图(图2)。ZK2001序列GC含量为38.58%,与AVNV(GQ153938)及OsHV-1-SB株(KP412538)GC含量一致,略低于OsHV-1参考株的38.72%。相邻长片段的overlap测序深度有明显增加,基因组末尾重复区组装质量较差。

### 2.3 基因组序列变异分析

将组装好的ZK2001、AVNV(GQ153938)、OsHV-1-SB(KP412538)与参考变异株基因组进行比对,查看不同变异株间SNP(图3),以及SNP引起各开放阅读框(ORF)氨基酸变异的个数和概率(图4)。ZK2001与参考基因组相比存在328个SNP位点,几乎遍布ZK2001整个基因组,但也存在基因组局部区间位点分布的差异。AVNV和OsHV-1-SB与参考基因组相比,各有465和577个SNP,明显高于ZK2001的SNP数量,出现这一现象的原因还不清楚。SNP引起氨基酸变异个数最多的几个ORF依次为ORF70、69、9、38和59,这些ORF分别编码环指蛋白(ORF9和38)、膜蛋白(ORF59)和2个未知功能的蛋白质(ORF69和70)。SNP引起氨基酸变异频率最快的几个ORF依次为ORF15、12、14和16等,这些编码蛋白均不是疱疹病毒的核心功能蛋白,这也许是其能够表现出高变异率的原因。

### 2.4 系统发育关系分析

系统发育关系分析显示ZK2001与2009年分

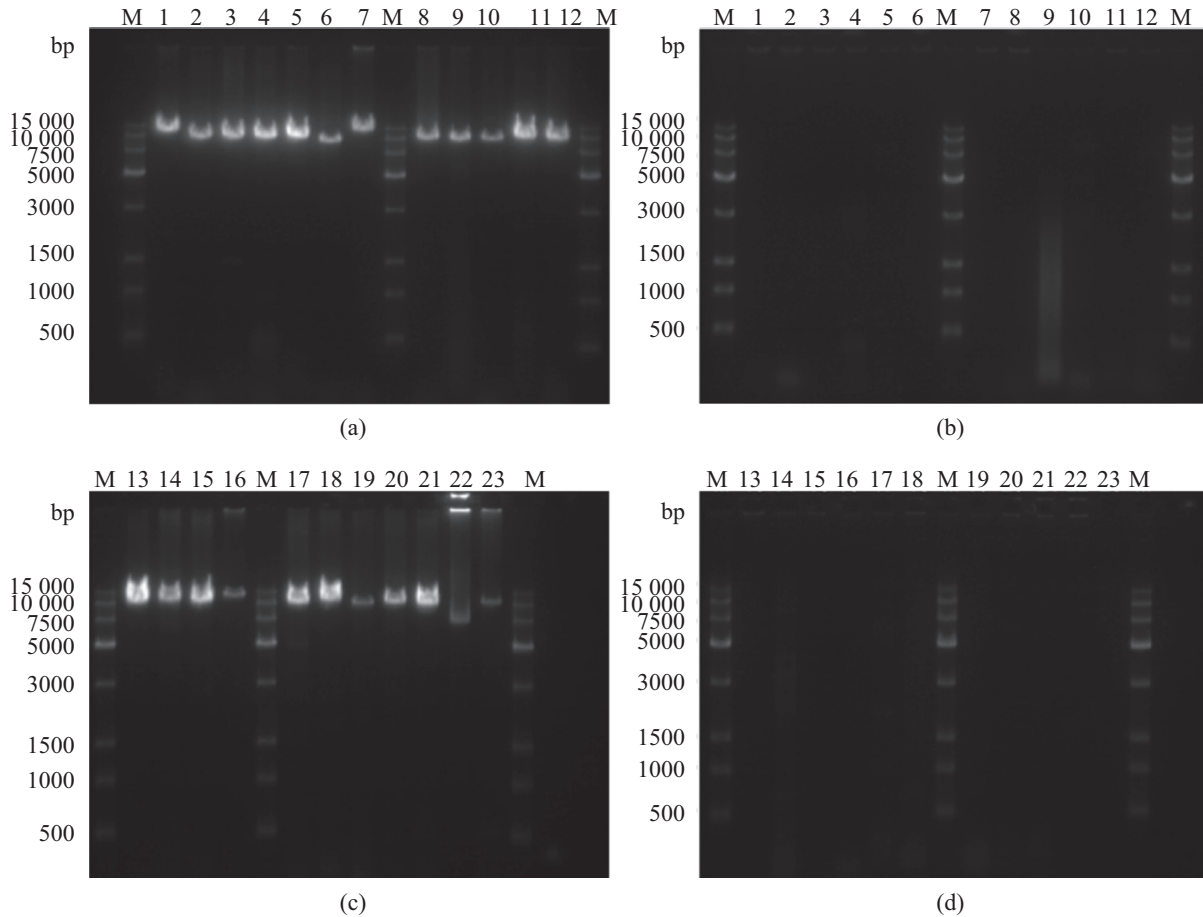


图 1 PCR产物琼脂糖凝胶电泳结果

(a)、(c) 23个PCR产物的电泳结果；(b)、(d) 阴性对照电泳结果；1~23. 23个对应引物的PCR产物；M. *Trans* 15 kb DNA Marker

Fig. 1 Agarose gel electrophoresis results of PCR amplicons

(a), (c) 23 PCR amplicons result ; (b), (d) negative controls result; 1-23.23 PCR amplicons result of corresponding primers; M. *Trans* 15 kb DNA Marker

离自我国栉孔扇贝的AVNV亲缘关系最近，其次为2012年分离自我国魁蚶的OsHV-1-SB和20世纪90年代分离自欧洲长牡蛎的OsHV-1参考株；而与2008年以后出现的OsHV-1 $\mu$ var及其相关变异株的亲缘关系最远(图5)。

### 3 讨论

自20世纪90年代末以来，OsHV-1引起包括我国在内的全球14个国家和地区海水养殖双壳贝类的大规模死亡。目前受影响的种类包括牡蛎(6种)、扇贝(3种)、蛤类(2种)和魁蚶在内的12个物种；其中受影响最严重的是欧洲、澳大利亚和美国养殖的长牡蛎，以及我国养殖的栉孔扇贝和魁蚶<sup>[2-5, 7, 15-16]</sup>。基于OsHV-1基因组序列变异的研究显示，该病毒发生了较强的株系分化，且其不同变异株的分布可能具有较强的时空特

异性和宿主偏好。但目前对OsHV-1的分型都建立在从该病毒基因组序列中随机选取的单个或少数几个基因位点上<sup>[17]</sup>。而对人和脊椎动物疱疹病毒的研究显示，少数基因位点很难反映疱疹病毒全基因组序列变异和病毒演化的真实情况。

在不同流行病学特征OsHV-1病毒株全基因组序列的获取和差异分析的基础上，筛选更多具有代表性的基因位点，并以此为基础进行病毒多样性、株系演化和系统发育关系研究，是解决上述问题的关键。由于OsHV-1随时间的发展不断发生演化，对不同年代来源OsHV-1变异株基因组序列的差异分析，有助于了解病毒随时间发展的演化情况，从而更加全面地认识病毒的进化过程。但是由于历史病料样本经过长时间的冻存，病毒失去了感染能力，病毒粒子形态也受到严重破坏，使用常规方法很难富集。

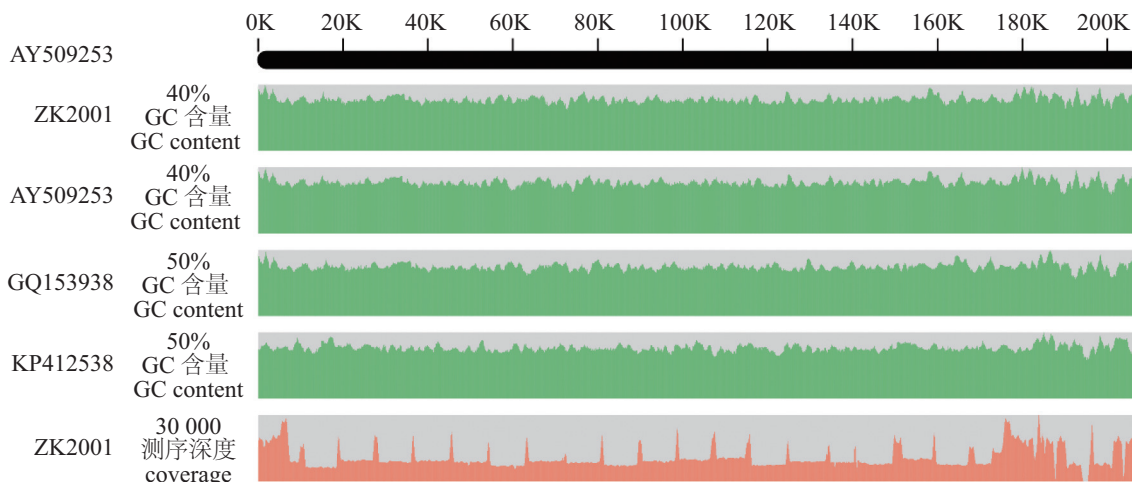


图2 测序基因组(ZK2001)深度和GC含量分布图

Fig. 2 Distribution of coverage and GC content of the sequenced ZK2001 genome

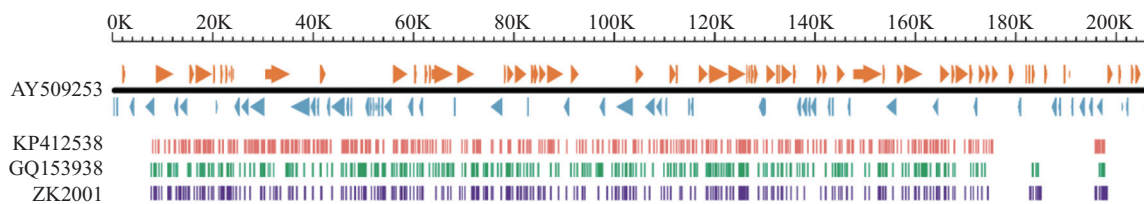


图3 测序基因组(ZK2001)和参考基因组相比SNP分布图

Fig. 3 SNP distribution of the sequenced ZK2001 compared to the reference genome

虽然利用RNA探针的技术可以从病料DNA中捕获病毒基因组DNA<sup>[18]</sup>,但目前市场上仅有针对感染人的几种疱疹病毒基因组富集RNA探针可供使用。随着高通量测序费用的降低,也有学者直接提取感染疱疹病毒的病料组织DNA进行高通量测序;通过提高测序数据量,达到深度覆盖占总DNA比例很低的病毒基因组的目的<sup>[19]</sup>。例如有研究对感染疱疹病毒的亚洲象(*Elephas maximus*)组织病料DNA进行高通量测序,完成了占宿主基因组比例极低的象疱疹病毒(*Elephantid herpesvirus 1*)基因组的测序。该研究结果表明,EEHV1A和EEHV1B基因组数据仅占全部数据的0.169%和0.038%。由于病毒基因组数据占总数据比例太低,只有将宿主基因组数据通过与参考基因组比对、删除后,才能完成这2个病毒基因组的组装<sup>[19]</sup>。目前OsHV-1的十几种贝类宿主,除了长牡蛎外,对其他宿主基因组信息都不了解<sup>[20]</sup>,因此直接测序的方法也不适用于OsHV-1基因组测序。

OsHV-1为双链DNA病毒,基因组相对稳

定,种内变异株间基因组序列相似度在95%左右。基于上述特性,本实验团队基于长片段PCR扩增方法,建立了OsHV-1全基因组扩增和富集技术,并成功应用第三代高通量测序平台对富集的基因组进行了测序。本研究首次尝试使用长片段PCR和第二代高通量测序技术,成功完成了长期冻存病料样本感染OsHV-1变异株的基因组DNA的富集和高通量测序。预实验结果表明,直接使用组织样本(剪碎或液氮研磨)提取DNA进行长片段PCR,会出现扩增失败和结果不稳定等问题。分析出现这一问题的原因是由于宿主DNA的大量存在,干扰了长片段PCR引物与病毒基因组的结合,从而导致扩增效率较低。为了解决这一问题,实验中设计了一套通过去除大部分宿主组织和细胞来消减宿主DNA比重的方法(如材料与方法部分所述)。利用该方法对病料组织样本处理后,再进行核酸提取得到的DNA样本可以被长片段PCR稳定扩增。但由于本研究采用了PCR扩增后的产物建立测序文库,虽然采用具有高保守性的聚合酶,仍然存在不均

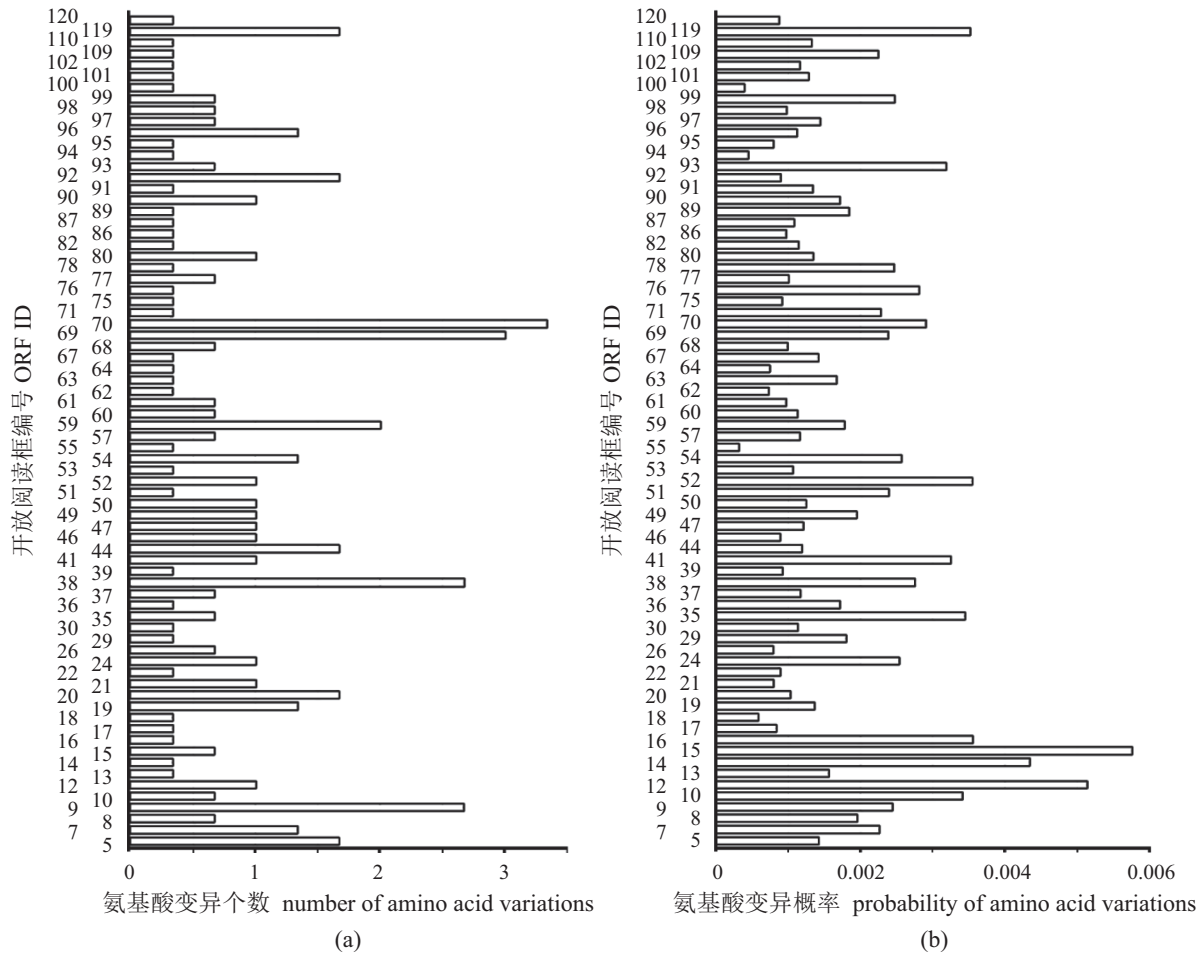


图4 开放阅读框中氨基酸变异个数和概率分布

(a)开放阅读框中发生变异的氨基酸数量；(b)开放阅读框中氨基酸发生变异的概率

Fig. 4 Number and probability of amino acid variations in each open reading frame

(a) number of amino acid variations in ORF; (b) probability of amino acid variations in ORF

匀放大导致的一些序列比例过高，以及核苷酸的错误插入2个主要问题。另外，使用长片段PCR产物进行高通量测序得到的数据，无法获取样本内病毒准种群基因组可能存在的基因多样性信息。

由于牡蛎疱疹病毒基因组末端包含2个特殊的反向重复区( $TR_L$ 和 $IR_L$ 以及 $IR_S$ 和 $TR_S$ )<sup>[21]</sup>，使用二代测序技术获得的小片段文库序列进行组装时，重复区不能很好地完成组装，这也是导致本研究最终不能得到完整的全基因组的主要原因。类似的情况在人疱疹病毒高通量测序过程中也有报道<sup>[22]</sup>，但这并不影响对变异株基因组序列差异等相关分析的开展。本研究结果也显示，虽然OsHV-1基因组末端重复区影响了完整基因组的组装，但对后续基因组变异分析并没

有太大影响。第三代单分子测序技术具有读长长的特性，单个读长可以达到10 kb，能跨过整个重复区，因此使用第三代测序平台对所富集的基因组DNA进行测序，可以得到完整的全基因组。但第三代测序技术成本过高，因此，需要根据研究的目的选择合适的实验平台。

为了解析法国养殖贝类感染7个主要OsHV-1变异株与参考变异株间的系统发育关系，Martenet等<sup>[23]</sup>利用PCR扩增和Sanger测序的方法，对这些变异株的39个编码不同功能蛋白质ORF的DNA序列进行测序。随后使用多基因联合的方法，充分解析了这些变异株与参考株之间的系统发育关系。本研究在上述7个变异株的基础上，从已知3个OsHV-1全基因组和本研究获取的基因组中抽提相关ORF的DNA序列，在去除过短

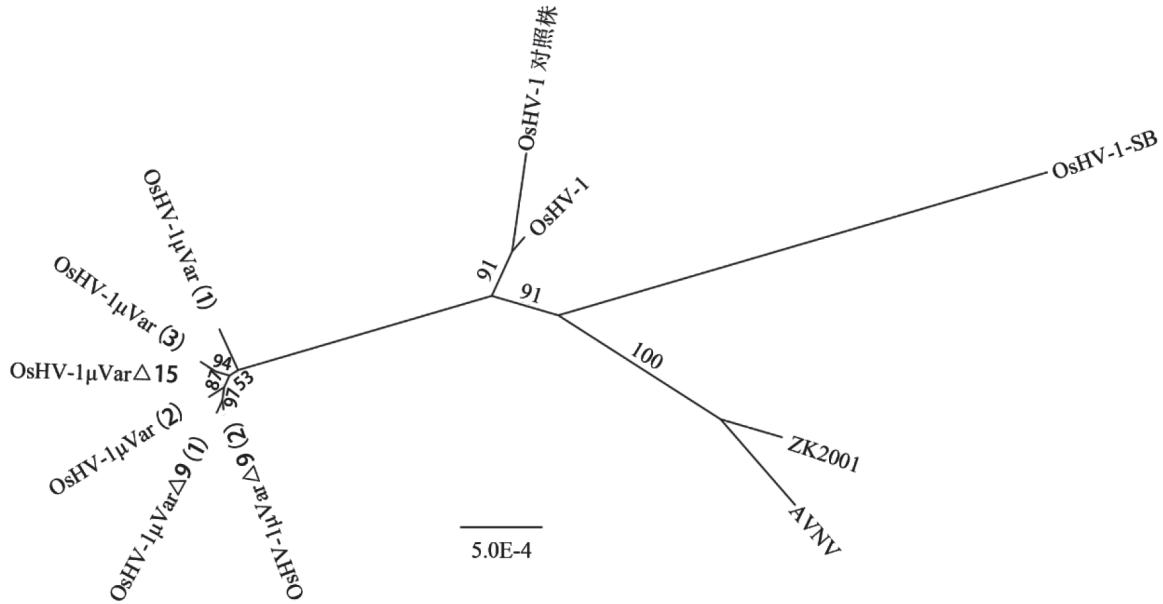


图 5 ZK2001变异株与其他10个变异株系统发育关系树

Fig. 5 Phylogenetic tree of the sequenced ZK2001 and the other ten OsHV-1 variants

的ORF序列后, 得到11个OsHV-1变异株共有的32个ORF序列。基于这些序列信息构建的系统发育树结构显示, 分布于我国的几个变异株间亲缘关系较近, 但又以相同宿主的2个变异株(AVNV和ZK2001)间亲缘关系最近。而欧洲分布的OsHV-1μvar及其相关变异株彼此间的亲缘关系很近, 但与其他所有变异株的亲缘关系都较远。结果表明, OsHV-1变异株间的亲缘关系受到宿主和时空分布范围等多个因素的共同影响。

参考文献:

[ 1 ] McGeoch D J, Rixon F J, Davison A J. Topics in herpesvirus genomics and evolution[J]. *Virus Research*, 2006, 117(1): 90-104.

[ 2 ] Barbosa S V, Renault T, Travers M A. Mass mortality in bivalves and the intricate case of the Pacific oyster, *Crassostrea gigas*[J]. *Journal of Invertebrate Pathology*, 2015, 131: 2-10.

[ 3 ] 王崇明, 王秀华, 宋晓玲, 等. 栉孔扇贝一种球形病毒的分离纯化及其超微结构观察[J]. *水产学报*, 2002, 26(2): 180-184.

Wang C M, Wang X H, Song X L, *et al.* Purification and ultrastructure of a spherical virus in cultured scallop *Chlamys farreri*[J]. *Journal of Fisheries of China*, 2002, 26(2): 180-184(in Chinese).

[ 4 ] Bai C M, Wang C M, Xia J Y, *et al.* Emerging and endemic types of *Ostreid herpesvirus 1* were detected in bivalves in China[J]. *Journal of Invertebrate Pathology*, 2015, 124: 98-106.

[ 5 ] Bai C M, Gao W H, Wang C M, *et al.* Identification and characterization of *Ostreid herpesvirus 1* associated with massive mortalities of *Scapharca broughtonii* broodstocks in China[J]. *Diseases of Aquatic Organisms*, 2016, 118: 65-75.

[ 6 ] Segarra A, Pépin J F, Arzul I, *et al.* Detection and description of a particular *Ostreid herpesvirus 1* genotype associated with massive mortality outbreaks of Pacific oysters, *Crassostrea gigas*, in France in 2008[J]. *Virus Research*, 2010, 153(1): 92-99.

[ 7 ] Arzul I, Nicolas J L, Davison A J, *et al.* French scallops: a new host for *Ostreid herpesvirus-1*[J]. *Virology*, 2001, 290(2): 342-349.

[ 8 ] Xia J Y, Bai C M, Wang C M, *et al.* Complete genome sequence of *Ostreid herpesvirus-1* associated with mortalities of *Scapharca broughtonii* broodstocks[J]. *Virology Journal*, 2015, 12: 110.

[ 9 ] Shendure J, Ji H. Next-generation DNA sequencing[J]. *Nature Biotechnology*, 2008, 26(10): 1135-1145.

[ 10 ] Stensvold C R, Lebbad M, Verweij J J. The impact of genetic diversity in protozoa on molecular diagnostics[J]. *Trends in Parasitology*, 2011, 27(2): 53-58.

[ 11 ] Ozcelik H, Shi X J, Chang M C, *et al.* Long-range PCR



- and next-generation sequencing of *BRCA1* and *BRCA2* in breast cancer[J]. *Journal of Molecular Diagnostics*, 2012, 14(5): 467-475.
- [12] Uribe-Convers S, Duke J R, Moore M J, *et al.* A long PCR-based approach for DNA enrichment prior to next-generation sequencing for systematic studies[J]. *Applications in Plant Sciences*, 2014, 2(1): 1300063.
- [13] Kvisgaard L K, Hjulsgaard C K, Fahnøe U, *et al.* A fast and robust method for full genome sequencing of porcine reproductive and respiratory syndrome virus (PRRSV) type 1 and type 2[J]. *Journal of Virological Methods*, 2013, 193(2): 697-705.
- [14] Martenot C, Oden E, Travaillé E, *et al.* Comparison of two real-time PCR methods for detection of *Ostreid herpesvirus 1* in the Pacific oyster *Crassostrea gigas*[J]. *Journal of Virological Methods*, 2010, 170(1-2): 86-89.
- [15] Friedman C S, Estes R M, Stokes N A, *et al.* Herpes virus in juvenile Pacific oysters *Crassostrea gigas* from Tomales Bay, California, coincides with summer mortality episodes[J]. *Diseases of Aquatic Organisms*, 2005, 63(1): 33-41.
- [16] Jenkins C, Hick P, Gabor M, *et al.* Identification and characterisation of an *Ostreid herpesvirus-1* microvariant (OsHV-1  $\mu$ -var) in *Crassostrea gigas* (Pacific oysters) in Australia[J]. *Diseases of Aquatic Organisms*, 2013, 105(2): 109-126.
- [17] Renault T, Moreau P, Faury N, *et al.* Analysis of clinical *Ostreid herpesvirus 1* (*Malacoherpesviridae*) specimens by sequencing amplified fragments from three virus genome areas[J]. *Journal of Virology*, 2012, 86(10): 5942-5947.
- [18] Depledge D P, Palser A L, Watson S J, *et al.* Specific capture and whole-genome sequencing of viruses from clinical samples[J]. *PLoS One*, 2011, 6(11): e27805.
- [19] Wilkie G S, Davison A J, Watson M, *et al.* Complete genome sequences of elephant endotheliotropic herpesviruses 1A and 1B determined directly from fatal cases[J]. *Journal of Virology*, 2013, 87(12): 6700-6712.
- [20] Zhang G F, Fang X D, Guo X M, *et al.* The oyster genome reveals stress adaptation and complexity of shell formation[J]. *Nature*, 2012, 490(7418): 49-54.
- [21] Davison A J, Trus B L, Cheng N Q, *et al.* A novel class of herpesvirus with bivalve hosts[J]. *Journal of General Virology*, 2005, 86(1): 41-53.
- [22] Kolb A W, Adams M, Cabot E L, *et al.* Multiplex sequencing of seven ocular herpes simplex virus type-1 genomes: phylogeny, sequence variability, and SNP distribution[J]. *Investigative Ophthalmology & Visual Science*, 2011, 52(12): 9061-9073.
- [23] Martenot C, Travaillé E, Lethuillier O, *et al.* Genome exploration of six variants of the *Ostreid herpesvirus 1* and characterization of large deletion in OsHV-1 $\mu$ Var specimens[J]. *Virus Research*, 2013, 178(2): 462-470.

## Long-range PCR and next generation sequencing of *Ostreid herpesvirus* genome

SHI Jie<sup>1,2</sup>, BAI Changming<sup>2</sup>, LI Chen<sup>2</sup>, CAI Shengli<sup>1</sup>, WANG Chongming<sup>2\*</sup>

(1. College of Fisheries and Life Science, Shanghai Ocean University, Shanghai 201306, China;

2. Key Laboratory of Maricultural Organism Disease Control, Ministry of Agriculture,

Laboratory for Marine Fisheries Science and Food Production Processes, Qingdao National Laboratory for Marine Science and Technology,

Qingdao Key Laboratory of Mariculture Epidemiology and Biosecurity; Yellow Sea Fisheries Research Institute,

Chinese Academy of Fishery Sciences, Qingdao 266071, China)

**Abstract:** To obtain the genome sequence of *Ostreid herpesvirus* 1 (OsHV-1) infected Zhikong scallop (*Chlamys farreri*) in 2001 (ZK2001), and its phylogenetic relationship with the other reported variants, the genome DNA of the ZK2001 variant was firstly enriched through long-range PCR, and sequenced with Illumina Hiseq 2500 PE250 platform. Then the variation and phylogenetic relationship between ZK2001 and the other reported variants were analyzed. 8 scaffolds were obtained after assembly. Genome variation analysis indicated that there were 328 SNPs between ZK2001 and the reference genome. SNPs and insert/deletion were the main cause of genome differentiation. Phylogenetic inference indicated that ZK2001 variant was more closely related to OsHV-1 variants isolated from China than those from Europe. These results indicated that genetic differentiation had occurred as a result of geographic isolation. This study indicated genomic DNA of OsHV-1, which had been frozen for a long time, could be enriched with long-range PCR and used for high-throughput sequencing. The genomic data of different OsHV-1 variants will be necessary for further study of the genetic variation, evolution and phylogenetic analysis of the virus.

**Key words:** *Chlamys farreri*; next-generation sequencing; long-range PCR; *Ostreid herpesvirus* 1

**Corresponding author:** WANG Chongming. E-mail: wangcm@ysfri.ac.cn

**Funding projects:** National Natural Science Foundation of China (31502208, 31302233); China Agriculture Research System (CARS-48); Scientific and Technological Innovation Project Financially Supported by Qingdao National Laboratory for Marine Science and Technology (2015ASKJ02)