论著

文章编号:1000-5404(2013)15-1562-04

# 鲍曼不动杆菌噬菌体AB3的全基因组测序及生物信息学分析

张 劼,刘 茜,甘 丹 (400014 重庆,重庆市第三人民医院老年科)

[摘要] **目的** 对本研究小组分离的鲍曼不动杆菌噬菌体AB3进行测序和基因组生物信息学分析,阐明其亲缘关系。**方法** 采用鸟枪法和重叠群组装的策略对噬菌体AB3进行基因组测序,并通过EditSeq、tRNAscan-SE、TRF、FindTerm、ORF finder、BPROM、GeneMark™、Clustalx、phylip等软件对所获噬菌体AB3基因组的一般特性、编码基因的功能预测、RNA聚合酶基因系统的进化进行分析。**结果** 噬菌体AB3基因组为全长31 185 bp的双链DNA,G+C含量为39.18%,包含28个预测基因,1个转录终止子和4个可能的启动子序列。**结论** 基因分析和RNA聚合酶基因进化分析显示噬菌体AB3与噬菌体AB1类似,均属于phiKMV-like病毒属。

[关键词] 鲍曼不动杆菌;噬菌体;基因组;生物信息学
[中图法分类号] Q939.48; Q755 [文献标志码] A

# Sequencing and bioinformatic analysis of genome of *Acinetobacter baumannii* bacteriophage AB3

Zhang Jie, Liu Xi, Gan Dan (Department of Geriatrics, Third People's Hospital of Chongqing, Chongqing, 400014, China)

[**Abstract**] **Objective** To sequence the *Acinetobacter baumannii* bacteriophage AB3 separated by our team, and to perform bioinformatics analysis, so as to identify the classification of its phylogenetic relationship. **Methods** Shot-gun library and config package strategy were carried out for sequencing the genome of bacteriophage AB3. Such software as EditSeq, tRNAscan-SE, TRF, FindTerm, ORF finder, BPROM and GeneMark™ were applied to predict both general characteristics of the bacteriophage AB3 genome and the coding gene function. In addition, the evolution of RNA polymerase gene system was analyzed with the software of Clustalx and phylip. **Results** The genome of bacteriophage AB3 was a double-strand DNA with a full length of 31 185 bp, in which G+C mol% was 39.18% and 28 predicted genes, 1 transcription terminator, and 4 possible promoter sequences were included. **Conclusion** Genetic analysis and RNA polymerase gene evolution analysis indicate that bacteriophage AB3 is similar to bacteriophage AB1, and both of them belong to phiKMV-like virus.

[**Key words**] *Acinetobacter baumannii*; bacteriophage; genome; bioinformatics

Supported by the Project of Chongqing Municipal Health Bureau (2009-2-081, 2011-2-334, 2012-2-218). Corresponding author: Liu Xi, E-mail: Lixj960054@163.com

鲍曼不动杆菌对常用的抗菌药物具有广泛的耐药性,患者往往死于无药可治。根据2010年中国CHINET细菌耐药性监测网数据显示,鲍曼不动杆菌占临床分离革兰阴性菌的16.11%,仅次于大肠埃希菌与肺炎克雷伯菌。其对头孢哌酮/舒巴坦耐药率为30.7%、米诺环素为31.2%,其他药物如亚胺培南、头孢吡肟、哌拉西林/他唑巴坦、阿米卡星、庆大霉素、环丙沙星等耐药率均在50%以上。噬菌体作为一类细菌病毒,能在菌体内快速增殖并最终裂解细菌,从而达到抗菌作用。研究人员在利用噬菌体进行感染性疾病治疗方面做了许多探索[1]。本研究就课题组前期所分离的噬菌体AB3进行了全基因组测序,并做初步注释和分析。

## 1 材料与方法

### 1.1 菌株和噬菌体

鲍曼不动杆菌为重庆市临床检验中心保存的临床分离株,噬菌体AB3以其为宿主菌从医院污水中分离得到。

[基金项目] 重庆市卫生局科研项目(2009-2-081,2011-2-334, 2012-2-218)
[通信作者] 刘 茜, E-mail:Lixj960054@163.com
[优先出版] http://www.cnki.net/kcms/detail/51.1095.R.20130329.1016.009.html(2013-03-29)

### 1.2 噬菌体 AB3 的分离及其基因组测序

取医院污水处理中心污水 1 000 mL,将 40 株鲍曼不动杆菌悬液和 LB 培养液 50 mL 加入到污水中混合培养 24 h;离心后过滤除菌;上清液中加 PEG8000 冰浴过夜后 4 ℃ 12 000 r/min 离心 10 min;重悬沉淀;10 μL 处理后的重悬液与 200 μL 宿主菌混合 20 min 后铺板。经反复纯化后,选取其中裂菌谱最宽的噬菌体 AB3 作为进一步研究的对象。噬菌体 AB3 可以形成圆形透亮的噬菌斑,具有典型裂解性噬菌体特性[2]。将噬菌体 AB3 悬液 20 μL 滴于铜网上,待其自然沉淀 20 min 后,加 1 滴 2% 磷钨酸于铜网上,染色 8 min,干燥后用透射电镜 Hitachi H-7500 观察。噬菌体 AB3 基因组测序按鸟枪法原理在华大基因公司完成,并上传 GenBank(登录号 KC311669)[3],序列组装用 Phred/Phrap/Consed 软件完成[4]。

### 1.3 噬菌体 AB3 基因组成分析

噬菌体 AB3 基因组成分析采用 Lasergene 软件的 EditSeq 软件完成,运用美国华盛顿大学编写的 tRNAscan-SE 软件对噬菌体 AB3 基因组中 tRNA 基因进行搜索,通过串联重复查找软件预测串联重复序列,在 softberry 网站上选择 FindTerm 软件对噬菌体 AB3 基因组中转录终止子进行预测,同时利用该网站中的 BPROM 软件预测噬菌体 AB3 基因组上的启动子。

### 1.4 噬菌体 AB3 编码基因预测分析

在美国国立生物技术信息中心(national center for biotechnology information,NCBI)网站上利用 ORF finder 软件进行开放阅读框(open reading frame,ORF)确认分析;运用美国佐治亚州技术研究所 Mark Borodovsky 实验室网站提供的 GeneMark™ 软件进行编码序列分析[5],同时用 Glimmer 软件进行验证;应用 NCBI 网站 BLAST 程序对噬菌体 AB3 基因组中推定基因进行逐个检索分析[6],采用两种查询方式:核酸序列比对核酸序列与蛋白质序列比对蛋白质序列;利用 Lasergene 软件的 EditSeq 软件对 AB3 基因组推算基因的平均长度和基因组中基因的密度进行分析。

### 1.5 噬菌体 AB3 RNA 聚合酶基因系统进化分析

取噬菌体 AB3 的 RNA 聚合酶基因同其同源基因进行比对,并做系统进化树分析。RNA 聚合酶基因的多序列比对采用 Clustalx 软件[7],进化树构建使用 Phylip 软件[8],用 TreeView 软件[9]观察结果。

## 2 结果

### 2.1 噬菌体 AB3 基因组序列组装和序列分析

透射电镜下可见噬菌体 AB3 头部为多面体立体结构,直径 60~70 nm(图 1)。其基因组可以被限制性内切酶切开,证实为双链 DNA(图 2)。测序完成经软件组装,噬菌体 AB3 基因组全长为 31 185 bp,G+C 含量为 39.18%,与其他 3 个鲍曼不动杆菌噬菌体基因组进行比较,结果见表 1。噬菌体 AP22、AB1[10]、YMC/09/02/B1251 ABA BP[11] 的 GenBank 登录号分别为 NC_017984、HM368260、JX403940。用 tRNAscan-SE 软件检索噬菌体 AB3 tRNA 基因未检测到,而其他 3 株噬菌体仅噬菌体 AB1 检测到 1 个 tRNA 基因。通过 TRF 软件预测噬菌体 AB3 串联重复序列,未检索到重复序列。

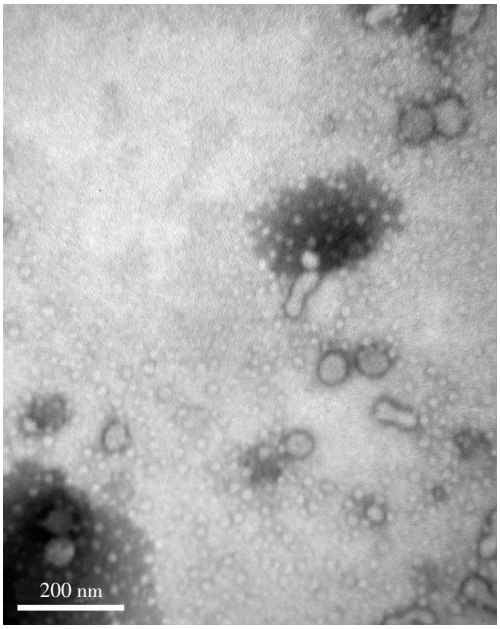
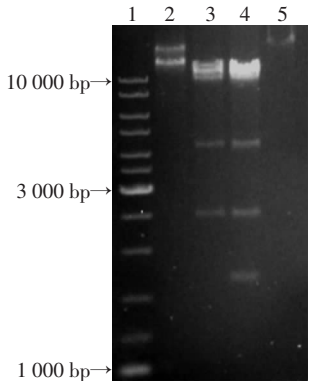



**图 1 透射电镜下观察噬菌体 AB3 经 2%磷钨酸负染后形态变化**




1:DNA 标准; 2: *Eco*RV 酶切; 3:*Sph* Ⅰ酶切; 4: *Eco*RV 和 *Sph* Ⅰ双酶切; 5:未消化

**图 2 噬菌体 AB3 基因组单双酶切电泳结果**

### 2.2 噬菌体 AB3 预测编码基因

噬菌体 AB3 基因组经 ORF finder 软件检索共有 97 个大于 100 bp 的开放阅读框,但是开放阅读框仅具有基因序列特征并不一定是真正的基因。用 GeneMark™ 软件对噬菌体 AB3 基因组分析,共找到 28 个可能的编码序列,同时用 Glimmer 软件进一步验证,结果是一致的。28 个预测基因及其注释见表 2。

### 2.3 噬菌体 AB3 基因长度和基因密度分析

噬菌体 AB3 基因组中最长的基因为 orf5(3 099 bp),最短的基因为 orf23(120 bp),基因组的 5′端有 1 477 bp 长未检索到编码基因。分析显示整个编码序列均位于负链上,占基因组序列全长的 90.8%。大于 2 000 bp 的基因最多,共有 6 个。

### 2.4 噬菌体 AB3 启动子和转录终止子分析

在 softberry 网站上选择 FindTerm 软件对噬菌体 AB3 基因组中转录终止子进行预测,仅发现 1 个可能的不依赖 ρ-因子的转录终止子的位点。同时利用该网站中的 BPROM 软件预测噬菌体 AB3 基因组上的启动子,对比噬菌体的基因序列位置,共发现 4 个可能的启动子序列(表 3)。

**表1 不同鲍曼不动杆菌噬菌体基因组中碱基组成的差异**

| 噬菌体 | 序列长度(nt) | 碱基G+C含量(%) | 碱基A含量(%) | 碱基G含量(%) | 碱基T含量(%) | 碱基C含量(%) | 预测基因 | tRNA基因 |
|---|---|---|---|---|---|---|---|---|
| 噬菌体AB3 | 31 185 | 39.18 | 29.25 | 17.19 | 31.57 | 21.99 | 28 | 0 |
| 噬菌体AB1 | 45 159 | 37.69 | 30.62 | 16.66 | 31.69 | 21.03 | 85 | 1 |
| 噬菌体AP22 | 46 387 | 37.74 | 31.72 | 21.17 | 30.54 | 16.57 | 89 | 0 |
| 噬菌体YMC/09/02/B1251 ABA BP | 45 364 | 39.05 | 31.57 | 21.24 | 29.37 | 17.81 | 62 | 0 |

**表2 噬菌体AB3基因组预测基因及注释**

| 开放阅读框编号 | 起始位点 | 终止位点 | 核酸链 | 长度(bp) | 相对分子质量($\times10^3$) | 功能预测 |
|---|---|---|---|---|---|---|
| orf01 | 1 478 | 1 786 | - | 309 | 11.09 | DNA maturase A |
| orf02 | 1 795 | 2 352 | - | 558 | 21.07 | endolysin |
| orf03 | 2 339 | 2 674 | - | 336 | 11.91 | holin |
| orf04 | 2 951 | 5 506 | - | 2 556 | 92.29 | tail fiber |
| orf05 | 5 513 | 8 611 | - | 3 099 | 114.07 | internal virion core protein |
| orf06 | 8 621 | 11 506 | - | 2 886 | 105.85 | structural protein |
| orf07 | 11 519 | 12 193 | - | 675 | 23.44 | internal virion protein B |
| orf08 | 12 193 | 14 484 | - | 2 292 | 84.07 | tail tubular protein B |
| orf09 | 14 493 | 15 119 | - | 627 | 24.36 | tail tubular protein A |
| orf10 | 15 256 | 15 441 | - | 186 | 6.43 | hypothetical protein phiAB1_gp34 |
| orf11 | 15 497 | 16 528 | - | 1 032 | 38.41 | capsid protein |
| orf12 | 16 544 | 17 404 | - | 861 | 30.53 | scaffolding protein |
| orf13 | 17 413 | 18 969 | - | 1 557 | 58.88 | Head-tail connector protein |
| orf14 | 18 978 | 19 229 | - | 252 | 9.08 | structural protein |
| orf15 | 19 226 | 19 423 | - | 198 | 7.69 | hypothetical protein phiAB1_gp29 |
| orf16 | 19 527 | 21 944 | - | 2 418 | 90.81 | RNA polymerase |
| orf17 | 21 953 | 22 603 | - | 651 | 24.69 | dNMP kinase |
| orf18 | 22 603 | 23 538 | - | 936 | 35.61 | hypothetical protein phiAB1_gp26 |
| orf19 | 23 542 | 23 982 | - | 441 | 16.65 | DNA endonuclease VII |
| orf20 | 23 979 | 24 548 | - | 570 | 21.71 | hypothetical protein |
| orf21 | 24 538 | 25 440 | - | 603 | 34.02 | DNA exonuclease |
| orf22 | 25 484 | 25 819 | - | 336 | 13.33 | hypothetical protein |
| orf23 | 25 816 | 25 935 | - | 120 | 4.41 | hypothetical protein |
| orf24 | 25 992 | 26 882 | - | 891 | 32.45 | hypothetical protein phiAB1_gp22 |
| orf25 | 26 886 | 27 125 | - | 240 | 9.19 | hypothetical protein |
| orf26 | 27 115 | 29 448 | - | 2 334 | 88.39 | DNA polymerase A domain-like protein |
| orf27 | 29 814 | 30 800 | - | 987 | 36.91 | ATP-dependent DNA ligase |
| orf28 | 30 797 | 31 162 | - | 366 | 13.53 | DNA helicase |

**表3 噬菌体AB3基因组上启动子的预测**

| 编号 | 启动子位置 | -10 box位置 | -10 box序列 | -35 box位置 | -35 box序列 | 可能相关基因 |
|---|---|---|---|---|---|---|
| 1 | 2 875 | 2 890 | ATAATAGCA | 2 911 | TTACAG | orf03 |
| 2 | 15 226 | 15 241 | AGTTTACCA | 15 263 | ACTCAG | orf09 |
| 3 | 29 591 | 29 606 | AGTATAAAT | 29 625 | TATTAC | orf26 |
| 4 | 30 501 | 30 516 | ATTGTAGGC | 30 538 | TGTAAA | orf27 |

### 2.5 噬菌体AB3 RNA聚合酶基因系统进化分析

取噬菌体AB3的RNA聚合酶基因同其相关的10个同源基因进行比对,并用Phylip软件构建系统进化树,可见噬菌体AB3与噬菌体phiAB1(即噬菌体AB1)RNA聚合酶基因最为相似,见图3。

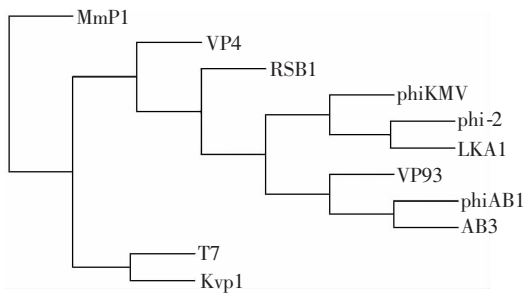



**图3 噬菌体AB3 RNA聚合酶基因系统进化树**

## 3 讨论

噬菌体AB3基因组用限制性内切酶*Eco*RⅤ、*Sph* Ⅰ均可切开,用DNase Ⅰ可完全消化,证实为双链DNA。基因组全长为31 185 bp,小于已完成测序的3个鲍曼不动杆菌噬菌体基因组[10-11],其中有大量序列与噬菌体AB1相似。G+C含量为39.18%,均大于已完成测序的其他3个鲍曼不动杆菌噬菌体(表1),但与噬菌体AB1最为相似。软件检索噬菌体AB3中tRNA基因未检测到,而与之序列相似的噬菌体AB1检测到1个tRNA基因。用TRF软件检索噬菌体AB3串联重复序列也未检测到,这与其他3个噬菌体是一致的。开放阅

读框是满足编码基因序列条件的必备条件,噬菌体 AB3 经 ORF finder 软件检索共有 97 个大于 100 bp 的开放阅读框,但是开放阅读框仅具有基因序列特征并不一定是真正的基因。用 GeneMark™ 软件对噬菌体 AB3 基因组进行分析,共找到 28 个可能的编码序列,同时用 glimmer 软件验证结果是一致的。噬菌体 AB3 基因组的序列长度短于其他 3 个同宿主噬菌体,造成了其编码基因明显少于其他 3 个同宿主噬菌体(表 1)。同一宿主菌可以有不同大小长度的噬菌体[12],说明噬菌体具有多样性。

噬菌体 AB3 的基因排列很有规律性,基本上按早中晚期基因排序[13]。而发现的转录终止子的位置非常重要,是噬菌体 AB3 中晚期基因的分割点,转录终止子的 5'端为噬菌体的晚期基因,基本为噬菌体的结构蛋白,而 3'端为噬菌体的早中期基因,为噬菌体的非结构蛋白[14]。

噬菌体 AB3 基因组中基因的长度为 120 bp(orf23) ~ 3 099 bp(orf5),平均长度为 1 011 bp,与其他噬菌体基本是一致的。分析显示整个编码序列均位于负链上,占基因组序列全长的 90.8%。在预测的 28 个基因中,有 24 个基因与噬菌体 AB1 具有同源性,其中 20 个基因有明确的功能,说明噬菌体 AB3 可能与噬菌体 AB1 具有共同的祖先。RNA 聚合酶基因是 T7 噬菌体家族的分类的一个金指标。将噬菌体 AB3 的 RNA 聚合酶基因同其相关的 10 个同源基因利用 Clustalx 软件进行序列比对,进而用 Phylip 软件最大距离法构建系统进化树,用 TreeView 软件[15-16] 观察结果(图 3)。可见噬菌体 AB3 的 RNA 聚合酶与噬菌体 AB1 高度同源,与先前基因的分析是一致的,故噬菌体 AB3 在分类上属于 phiKMV-like 病毒属。

综上所述,对本课题组前期研究分离得到的鲍曼不动杆菌噬菌体 AB3 进行全基因组测序,基因组全长 31 185 bp,G + C 含量为 39.18%,包含 28 个预测基因,1 个转录终止子和 4 个可能的启动子序列,基因分析和 RNA 聚合酶基因进化分析显示噬菌体 AB3 与噬菌体 AB1 类似,均属于 phiKMV-like 病毒属。

**参考文献:**


[1] Gupta R, Prasad Y. Efficacy of polyvalent bacteriophage P-27/HP to control multidrug resistant Staphylococcus aureus associated with human infections[J]. Curr Microbiol, 2011, 62(1): 255 - 260.

[2] 张劼, 罗永艾. 耐碳青霉烯类鲍曼不动杆菌噬菌体的分离鉴定及其生物学特性[J]. 上海交通大学学报: 医学版, 2012, 32(5): 589 - 594.

[3] Dong H, Lin W, Zhang C K, *et al*. Genomic sequence and expression analyses of human chromatin assembly factor 1 p150 gene[J]. Gene, 2001, 264(2): 187 - 196.

[4] de-la-Bastide M, McCombie W R. Assembling genomic DNA sequences with PHRAP[J]. Curr Protoc Bioinformatics, 2007, Chapter 11: Unit11 - 4.

[5] Besemer J, Borodovsky M. Heuristic approach to deriving models for gene finding[J]. Nucleic Acids Res, 1999, 27(19): 3911 - 3920.

[6] Altschul S F, Madden T L, Schaffer A A, *et al*. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs[J]. Nucleic Acids Res, 1997, 25(17): 3389 - 3402.

[7] Jeanmougin F, Thompson J D, Gouy M, *et al*. Multiple sequence alignment with Clustal X[J]. Trends Biochem Sci, 1998, 23(10): 403 - 405.

[8] Retief J D. Phylogenetic analysis using PHYLIP[J]. Methods Mol Biol, 2000, 132: 243 - 258.

[9] Page R D. Visualizing phylogenetic trees using TreeView[J]. Curr Protoc Bioinformatics, 2002, Chapter 6: Unit 6.2.

[10] Li P, Chen B, Song Z, *et al*. Bioinformatic analysis of the Acinetobacter baumannii phage AB1 genome[J]. Gene, 2012, 507(2): 125 - 134.

[11] Jeon J, Kim J W, Yong D, *et al*. Complete genome sequence of the podoviral bacteriophage YMC/09/02/B1251 ABA BP, which causes the lysis of an OXA-23-producing carbapenem-resistant Acinetobacter baumannii isolate from a septic patient[J]. J Virol, 2012, 86(22): 12437 - 12438.

[12] Romero P, Croucher N J, Hiller N L, *et al*. Comparative genomic analysis of ten Streptococcus pneumoniae temperate bacteriophages[J]. J Bacteriol, 2009, 191(15): 4854 - 4862.

[13] Rokyta D R, Wichman H A. Genic incompatibilities in two hybrid bacteriophages[J]. Mol Biol Evol, 2009, 26(12): 2831 - 2839.

[14] Williams T M, Loman N J, Ebruke C, *et al*. Genome analysis of a highly virulent serotype 1 strain of Streptococcus pneumoniae from West Africa[J]. PLoS One, 2012, 7(10): e26742.

[15] Du P, Zhang W, Wang H, *et al*. Comparative genomic analysis of Escherichia coli O104: H4 stx2 prophage reveals a potential new method to identify virulence factors[J]. Genome, 2012, 55(9): 697 - 700.

[16] Dorscht J, Klumpp J, Bielmann R, *et al*. Comparative genome analysis of Listeria bacteriophages reveals extensive mosaicism, programmed translational frameshifting, and a novel prophage insertion site[J]. J Bacteriol, 2009, 191(23): 7206 - 7215.


(**收稿**: 2013-01-06;**修回**: 2013-03-15)

(编辑 邓强庭)