DOI:10.16016/j.1000-5404.2004.01.019
第26卷第1期 第三军医大学学报 Vol. 26, No. 1
2004年1月 ACTA ACADEMIAE MEDICINAE MILITARIS TERTIAE Jan. 2004

文章编号:1000-5404(2004)01-0051-03

论著

PaP3噬菌体53×10³蛋白编码基因的确定

朱军民,金晓琳,饶贤才,黄建军,张克斌,胡福泉 (第三军医大学基础医学部微生物学教研室,重庆市微生物工程实验室,重庆 400038)

提 要:目的 确认PaP3噬菌体53×10³蛋白的编码基因。方法 用PEG沉淀结合CsCl梯度密度离心分离、纯化噬菌体颗粒,通过SDS-PAGE电泳分析该噬菌体的结构性蛋白带,转印PVDF膜后,对53×10³蛋白进行N-端氨基酸测序,进而从PaP3全基因组的252个ORFs中确认该蛋白质的编码基因及其对应的氨基酸序列。结果 PaP3噬菌体的结构性蛋白共发现有8个,其中53×10³蛋白是由ORF29893编码的。53×10³蛋白编码基因全长1 524 bp,G+C=54.92%,编码508个氨基酸。该蛋白分子质量为53.7×10³,等电点(PI)为7.207。结论 53×10³蛋白是由噬菌体基因组中ORF29893编码的,共由508个氨基酸组成。该蛋白可能是一种结构性蛋白。

关键词:PaP3噬菌体;结构性蛋白;编码基因
中图法分类号:Q755;Q782 文献标识码:A

Identification of 53×10³ protein gene of bacteriophage PaP3

ZHU Jun-min, JIN Xiao-lin, RAO Xian-cai, HUANG Jian-jun, ZHANG Ke-bin, HU Fu-quan (Department of Microbiology, College of Medicine, Third Military Medical University, Chongqing 400038, China)

Abstract: Objective To identify the coding gene of 53×10³ protein of bacteriophage PaP3. Methods Bacteriophage PaP3 particles were isolated and purified by using PEG-sedimentation combined with CsCl gradient centrifugation, and then the total proteins were isolated on SDS-PAGE to figure out how many proteins making up of the virion particle. The protein bands on SDS-PAGE were transferred onto PVDF membrane by electronic transfer method. Afterwards, the 6 amino acid sequence of the N-terminal of the 53×10³ protein was determined by using Edman degradation method. The coding gene of 53×10³ protein from 252 ORFs of genome PaP3 was clarified based on the sequence result. Results A total of at least 8 structural proteins were found in the phage. The 6 amino acid sequence of the N-terminal of the 53×10³ protein was "MASMAY". Based on this sequence, it might be confirmed that 53×10³ protein with molecular weight (MW) of 53.7×10³ was encoded by ORF29893, consisting of 1 524 base pairs, with G+C content of 54.92%, and coding 508 amino acids (isoelectric point, PI=7.207). The theoretical MW of ORF29893 coding protein was 53.7×10³, very close to the actually detected value. Conclusion Protein of 53×10³ is encoded by ORF29893 of genome PaP3, and the whole amino acid sequence of 53×10³ protein is clarified. This protein may be one of capsomeres making up of the phage capsid protein.

Key words: bacteriophage PaP3; structural protein; coding gene

噬菌体(bacteriophage)是寄生于细菌等微生物体内的一类病毒,它是一种比原核生物细菌更小、更简单、在进化阶梯上更为原始的非细胞形态的生命现象。因为噬菌体生活在细菌体内,而细菌又生活在人体内,这就构成了双重环境嵌套的特殊环境。因此,它是研究噬菌体、细菌及人体三者相互作用关系的一种极好的模式生物。对噬菌体的深入研究,有助于揭示由噬菌体介导的细菌毒力因子和耐药性因子的横向转移等医学重大问题。PaP3噬菌体是我室自行分离的一株铜绿假单胞菌噬菌体,其全基因组测序和初步注释业已全部完成[1,2]。已经弄清PaP3噬菌体全基因组由44 249 bp组成,经ORF finder搜索表明PaP3基因组中具有252个(大于100 bp)的ORFs,基因预测分析提示有69个编码基因(coding sequences, CDS)。应该注意的是,上述基因的预测,都是根据现行常用生物信息学分析软件的分析结果。我们知道,这些计算机分析结果最终还是需要实验证据来证实。本研究旨在确认

基金项目:国家自然科学基金资助项目(30070037);重庆市攻关课题资助项目(2000)
Supported by the National Natural Science Foundation of China (30070037) and the Key Sci & Tech Research Project of Chongqing (2000)
作者简介:朱军民(1975-),男,山东省淄博市人,硕士,助教,主要从事噬菌体基因组学方面的研究。电话:(023)68752239
通信作者:胡福泉,电话:(023)68752240,E-mail:hufuquan@yahoo.com
收稿日期:2003-04-04;修回日期:2003-07-10

?1994-2016 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

PaP3 噬菌体 53× 10^3 蛋白的编码基因。

1 材料和方法

1.1 铜绿假单胞菌 PaP3 噬菌体

从临床标本中分离到其宿主菌铜绿假单胞菌，然后在医院下水道中分离到该噬菌体[3]。其全基因组测序和初步注释已完成[1,2]。

1.2 PaP3 噬菌体颗粒的纯化

按照本室常规方法，参照文献[4]λ 噬菌体纯化方法，配备 CsCl 3 层梯度液，密度分别为 1.45、1.50、1.70 g/ml。将培养的 PaP3 和宿主菌混合悬液经 DNase Ⅰ及 RNase A 消化后，饱和氯化钠沉淀细菌。离心取上清液，用 PEG 8000 冰浴沉淀噬菌体颗粒。将沉淀的噬菌体颗粒用 2 ml 的 TM 液溶解稀释后，加在 3 层 CsCl 梯度液上，22 000 r/min 高速离心 2 h，收集其中相应的灰色区带即为铜绿假单胞菌 PaP3 噬菌体颗粒纯化物。

1.3 PaP3 噬菌体 SDS-PAGE 蛋白电泳

参照文献[4]方法执行。配备 12% 分离胶和 5% 积层胶，将 PaP3 噬菌体纯化颗粒和中分子质量标准蛋白质，加入等量的蛋白加样缓冲液混合，煮沸变性处理后上样，恒压 80 V 电泳约 4 h 待溴酚蓝指示剂到达凝胶下缘时，停止电泳。电泳后，切下含中分子质量标准蛋白质的凝胶，用考马斯亮蓝 R-250 染色观察。其余各泳道的蛋白凝胶切去积层胶部分供电转印用。

1.4 PaP3 噬菌体蛋白转印 PVDF 膜

参照 http://www.expasy.ch/ch2d/protocols/protocols.fm8.html/ 网站上所列方法进行。将余下的 PaP3 蛋白凝胶切去积层胶部分，与 PVDF 膜一起，在电转印缓冲液中平衡 30 min。然后在夹板中制备三明治夹层，由正极到负极依次为：3 层浸湿的 Whatman 滤纸、PVDF 膜、稍小于 PVDF 膜的凝胶块、稍小于凝胶块的 3 层 Whatman 滤纸。整个操作在电转印缓冲液中进行，注意各层间不能有气泡。将夹板置于电转印槽中，置 4 ℃冰箱中恒压 30 V 电转印过夜，将 PVDF 膜取出用考马斯亮蓝染色，脱色，直至出现明显的蛋白条带而背景色脱净为止。

1.5 PaP3 噬菌体蛋白 N-端测序

将转至 PVDF 膜上 53× 10^3 左右的蛋白样品邮寄北京大学生命科学院进行蛋白 N-端测序，根据 Edman 降解法，完成 N-端的 5 个氨基酸的测序。

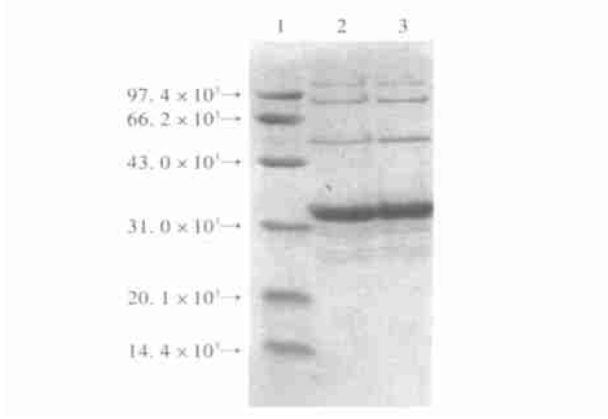
2 结果

2.1 PaP3 噬菌体蛋白的 SDS-PAGE 电泳结果

将用 PEG 8 000 沉淀和 CsCl 梯度密度离心法提取纯化的 PaP3 噬菌体颗粒，做 SDS-PAGE 蛋白电泳，结果如图 1 所示。在中分子质量标准蛋白范围内，凝胶上共可以出现 8 条蛋白条带。分子质量依次可能为 114× 10^3，90× 10^3，53× 10^3，35× 10^3，33× 10^3，29× 10^3，27× 10^3，25× 10^3（其中 35× 10^3，33× 10^3 两者相距较近）。根据这一结果，我们可以肯定，PaP3 噬菌体中至少有 8 个 ORFs 是可表达出蛋白质的。

2.2 PaP3 噬菌体蛋白 N-端测序

通过对 SDS-PAGE 电泳条件的优化，分离到 8 条蛋白带。并将其中的 53× 10^3 蛋白条带送往北京大学生命科学院进行蛋白 N-端测序。本次蛋白质 N-端测序前 5 个氨基酸结果依次为：A（丙氨酸）S（丝氨酸）M（蛋氨酸）A（丙氨酸）Y（酪氨酸），加上第一位的 M（蛋氨酸），排列起来顺序为：MASMAY。



1：marker；2，3：frotein of bacteriophage Pap3

图 1 PaP3 噬菌体蛋白 SDS-PAGE 图谱

Fig 1 SDS-PAGE analysis of bacteriophage PaP3

2.3 PaP3 噬菌体 53× 10^3 蛋白编码基因的确认

确认 53× 10^3 蛋白编码基因的依据有两个：一是 N-末端氨基酸的序列，二是分子量的大小。首先，从 252 个 ORFs 中，将编码接近 53× 10^3 蛋白产物的 ORF 全部列出来，并同时将它们的 N-末端氨基酸序列出来，与 53× 10^3 蛋白的实际测序结果进行比对，以确认哪一个 ORF 是 53× 10^3 蛋白的编码基因。

表 1 为 PaP3 噬菌体编码近 53× 10^3 蛋白的 ORFs 基因产物的 N-末端序列比较表。由上述测序结果可知，我们所测的结果加上其前面的蛋氨酸为 MASMAY，具有惟一性，对应的是 ORF29893 的 N-端的前 6 个氨基酸。对 ORF29893 进行序列分析，其理论计算分子质量为 53.7× 10^3，与我们用中分子质量标准蛋白所测的结果 53× 10^3 最为接近。因此，根据 ORF29893 的 N-端氨基酸序列，理论计算分子量大小，可以肯定 53× 10^3 的蛋白条带所对应的 ORF 即为 ORF29893。且在整个噬菌体基因序列中，ASMAY 的出现也具有惟一性，只出现在 ORF29893 的 N-末端。

表 1 PaP3 噬菌体所有近 53× 10^3 推定蛋白的 N-端氨基酸序列比较

Tab 1 N-terminal amino acid sequences coded by PaP3 genome

ORF	N-terminal amino acid sequence	Theoretic MW(× 10^3)
ORF20768	MKMENY	37.1
ORF37917	MDLENQ	37.6
ORF4999	MARLKY	38.1
ORF8296	MSALYS	44.5
ORF29893	MASMAY	53.7
Terminase gene	MDTQER	54.6
ORF9505	MCGLVG	56.8
ORF33864	MALERQ	57.4
DNA polymerase2	VDWRKS	61.2
Primase/helicase gene	MKKIIG	64.2
ORF40279	MAKRRK	80.9

2.4 PaP3 噬菌体 ORF29893 基因序列分析

ORF29893 位于 PaP3 噬菌体基因组的负链上，位置是从 29 893～28 370 号碱基，全长 1 524 bp，% A＝22.84，% G＝

?1994-2016 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

30. 38, % T= 22. 24, % C= 24. 54, % A+ T= 45. 08, % G+ C= 54. 92. 该基因编码产物为 508 个氨基酸, 分子质量为 53. 7× 10^3, 等电点(PI)为 7. 207, 氨基酸全序列为: MASMAYEGSPIRPSILRAAQNELDMARIARNKLPLAVEGSGIPDRVRRAAQAALENPERWSRAAQEVAPAAEATGRGALGRIAGILGGPVGVGVQAAVTPADIGDQERTRAEEMAQAAQAVENMGPEVAQSANQWAQGVGQRAAQKATGGPQGAGLLSYGVTPNQPSIEPEIAPEVASEAGAAVAEEEEANRQVIQQGAIEGLRSGAVSRPEMAQAVVEADAQREGVELKPEELKTRVNEELTQMRTMDNDDLSRYVSYALIGTGLLASALDKTGKAGDMFAASYERQLDRNLQAGINQQKMAAAAADRQIKEKDLERKVAKDKSDERLGIGNLEVKRGTLEETSRKNEGLLSKWEKQAAVERAKLALTQRGQDLATQRAQLQAETARRGQDMSQENAQLSSAVRLKTAKISAAARKAAASASRGEAITTKDANELVKGVSGSDLVGGAKIGEQDQSAISQKLRNHLRQNPGADPVTFVQQEAAKLQKQPSGLFGWGSGILSRPAPK。

3　讨论

本研究室于 1999 年开始了对铜绿假单胞菌噬菌体的研究, 自行分离到一株铜绿假单胞菌噬菌体, 命名为 PaP3。目前已经完成了 PaP3 的全基因组测序, 已经弄清 PaP3 噬菌体全基因组由 44 249 bp 组成。经 ORF finder 搜索表明 PaP3 基因组中具有 252 个(大于 100 bp)的 ORFs, 基因预测分析提示有 69 个 ORFs 可能是编码基因。经 BLAST ORF 序列同源性比对, 在蛋白质氨基酸水平上, 只发现有 6 个 ORFs 编码产物有同源性序列存在, 其余均是功能尚不清楚的新基因。

基因组测序与注释完成后, 对于基因组功能的研究甚为重要。作者将注意力首先集中到了噬菌体的结构性蛋白上。在前期工作中, 仅分离到 4 个结构蛋白条带, 且对其中的 35× 10^3 蛋白进行了鉴定。作者考虑到, 对于一个 4 万多碱基的噬菌体基因组, 其结构蛋白不应该仅有 4 个。在本研究中, 作者采用了网上介绍的改良方法, 对整个实验条件进行了优化, 结果从 PaP3 噬菌体颗粒中分离到 8 个蛋白条带, 它们的分子质量依次为 114× 10^3, 90× 10^3, 53× 10^3, 35× 10^3, 33× 10^3, 29× 10^3, 27× 10^3, 25× 10^3。目前, 已经知道的与 PaP3 基因组大小相差不多的几个噬菌体, 其噬菌体颗粒在 SDS-PAGE 电泳后的凝胶上, 通常可出现几个到十几个不同的条带[5,6], 而对于与其相近的铜假单胞菌噬菌体如约 36 kb 的铜假单胞菌噬菌体 phi CTX 在 SDS-PAGE 电泳后的凝胶上共出现 10 条带[7], 56 kb 的噬菌体 D3 在 SDS-PAGE 凝胶电泳后显示有 6 条蛋白条带[8], 而本实验的 PaP3 噬菌体共有 8 条带, 与它们都是相类似的, 可能已经接近了 PaP3 噬菌体在 SDS-PAGE 凝胶上的结构蛋白数。在这 8 个蛋白条带中, 由图可见, 53× 10^3 蛋白量达到了可进行 N-末端氨基酸测序的水平, 故本研究对该蛋白进行了研究。

分离纯化了蛋白条带, 就有条件确定该蛋白的编码基因和编码产物的氨基酸序列。在本研究中, 采用的是 N-末端氨基酸测序结合分子量大小这两个指标来确认 53× 10^3 蛋白的编码基因。测序结果表明 53× 10^3 蛋白的 N-末端氨基酸序列为"ASMAY"。检查 PaP3 噬菌体的全部 252 个 ORFs, 只有处在负链上的 ORF29893-28370(根据起始位点命名为 ORF29893)的编码产物其 N-端的 5 个氨基酸序列为"ASMAY"。随后, 对 ORF29893 进行了序列分析, 根据计算, 其分子质量为 53. 7× 10^3。而且, 分析显示在 PaP3 噬菌体所有 ORFs 中只有这个 ORF 编码的基因产物分子质量与所测到的 53× 10^3 蛋白的分子质量最为接近, 因此可以确认 ORF29893 就是 53× 10^3 蛋白的编码基因。

噬菌体基因组所编码的蛋白中, 结构性蛋白通常表达较稳定, 且表达拷贝数较高。而调节性蛋白的转录和翻译水平相对很低, 甚至只在噬菌体生活周期的某些阶段一过性的表达[9,10]。由于提取的是噬菌体的颗粒, 根据实验分析 53× 10^3 蛋白其表达拷贝数较高。因此, 本次研究的 53× 10^3 蛋白可能是 PaP3 噬菌体的结构蛋白之一, 很可能是噬菌体衣壳蛋白中的一种壳微粒。至于该蛋白是位于头部、尾部或其它位置, 有待进一步研究。

参考文献:

[1] 张克斌, 金晓琳, 朱军民, 等. 铜绿假单胞菌 PaP3 噬菌体基因组测序[J]. 第三军医大学学报, 2002, 24(4): 378－380.

[2] 张克斌, 金晓琳, 朱军民, 等. 铜绿假单胞菌 PaP3 噬菌体基因组的初步注释[J]. 第三军医大学学报, 2002, 24(4): 381－384.

[3] 张克斌, 陈志瑾, 金晓琳, 等. 铜绿假单胞菌噬菌体分离鉴定及耐噬菌体突变频率测定[J]. 微生物学通报, 2002, 29(1): 40－45.

[4] J 萨姆布鲁克, E F 弗里奇, T 曼尼阿蒂斯. 分子克隆实验指南[M]. 第 2 版. 北京: 科学出版社, 1996. 136－137.

[5] Tremblay D M, Moineau S. Complete genomic sequence of the lytic Bacteriophage DT1 of Streptococcus thermophilus[J]. Virology, 1999, 255(1): 63－76.

[6] Volkmar B J R, Stefan H, Horst N, *et al*. Taxonomic differentiation of bacteriophages of Lactococcus lactis by electron microscopy, DNA-DNA hybridization, and protein profiles[J]. J Gen Microb, 1989, 135: 2551－2560.

[7] Elsabhagh H, Xiong G, Lutz F. Nucleotide sequence of attP and cos sites of phage CTX and expression of cytotoxin in Pseudomonas aeruginosa PA158[J]. Mol Gen Genet, 1993, 237(3): 421－428.

[8] Gilakjan Z A, Kropinski A M. Cloning and analysis of the capsid morphogenesis genes of Pseudomonas aeruginosa bacteriophage D3: another example of protein chain mail[J]. J Bacteriol, 1999, 181(23): 7221－7227.

[9] 司稚东, 何晓青, 汪家驯, 等. 噬菌体学[M]. 北京: 科学出版社, 1996. 63－88.

[10] Calender R. The Bacteriophages[M]. New York: Plenum Press, 1988. 117－167.

(编辑　薛国文)

?1994-2016 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net