DOI:10.16016/j.1000-5404.2002.04.004
第24卷第4期 第三军医大学学报 Vol. 24, No. 4
2002年4月 ACTA ACADEMIAE MEDICINAE MILITARIS TERTIAE Apr. 2002

文章编号:1000-5404(2002)04-0385-04

论著

铜绿假单胞菌噬菌体 PaP3 的密码使用偏嗜性与 tRNA 基因

张克斌,金晓琳,陈志瑾,朱军民,饶贤才,胡福泉 (第三军医大学基础医学部微生物学教研室、重庆微生物工程重点实验室,重庆 400038)

提 要:目的 检索分析铜绿假单胞菌噬菌体 PaP3 的遗传密码使用偏嗜性及其 tRNA 基因的关系。方法 利用生物信息学软件和网上生物信息学工具,检索分析 PaP3 的遗传密码使用偏嗜性;并与铜绿假单胞菌 PAO1 的遗传密码使用偏嗜性进行对比分析;检索 PaP3 基因组中 tRNA 基因并分析其与遗传密码使用偏嗜性的关系。结果 遗传密码偏嗜性分析发现噬菌体 PaP3 和铜绿假单胞菌优势码一致的氨基酸有 F-uuc, I-auc, V-guc, A-gcc, Y-uac, H-cac, N-aac, K-aag 和 T-acc;PaP3 单独具有偏嗜性密码的氨基酸有 A-gcu, G-ggu, T-acu 和 V-gua;铜绿假单胞菌 PAO1 单独具有偏嗜性密码的氨基酸有 L-cug, V-gug, P-ccg, V-gcg, Q-cag, D-gac, E-gag, C-ugc, R-cgc, S-agc 和 G-ggc。tRNA 基因查询发现了 PaP3 基因组中存在三段 tRNA 编码序列,分别编码 $tRNA^{Asn}$, $tRNA^{Asp}$, $tRNA^{Pro}$;它们的反密码子分别为"gtt","gtc"和"tgg",分析结果没有表明这三个 tRNA 基因与优势转运氨基酸有关。结论 铜绿假单胞菌 PAO1 的遗传密码偏嗜性更为明显,而 PaP3 的遗传密码偏嗜性相对较弱,两者在遗传密码偏嗜性方面并未表现出一致性;噬菌体编码的 tRNA 基因与遗传密码偏嗜性之间有直接联系。

关键词:噬菌体 PaP3;基因组;密码偏嗜性;tRNA 基因
中图法分类号:Q755 文献标识码:A

Codon usage bias and tRNA genes analysis in genome of *Pseudomonas aeruginosa* phage PaP3

ZHANG Ke-bin, JIN Xiao-lin, CHEN Zhi-jin, ZHU Jun-min, RAO Xian-cai, HU Fu-quan (Department of Microbiology, Third Military Medical University, Chongqing 400038, China)

Abstract: Objective To analyze the relationship between codon usage bias and tRNA genes in *Pseudomonas areuginosa* phage PaP3 genome. Methods Codon usage bias analysis and tRNA genes searching in PaP3 genome and in the bacterial genome were performed respectively with bioinformatic software and tools in internet. And the relationships of the bias with rRNA and of the phage and host were analyzed. Results The common dominant codons in both PaP3 and P. aeruginosa PAO1 were F-uuc, I-auc, V-guc, A-gcc, Y-uac, H-cac, N-aac, K-aag and T-acc; Those only in PaP3 were A-gcu, G-ggu, T-acu and V-gua; While in PAO1 were L-cug, V-gug, P-ccg, V-gcg, Q-cag, D-gac, E-gag, C-ugc, R-cgc, S-agc and G-ggc. Three tRNA genes, coding for tRNAAsn, tRNAAsp, tRNAPro respectively, were found in the phage genome. Those tRNA genes were not correlated to the codon usage bias of the phage. Conclusion The P. aeruginosa PAO1 shows very evident codon usage bias, while the phage PaP3 doesn't. The codon usage bias of PaP3is is not kept in coincidence with that of PAO1 genome, and 3 tRNA genes coded by PaP3 are not correlated to codon usage bias of this phage.

Key words: bacteriophage PaP3; genome; codon usage bias; tRNA tRNA coding sequence

迄今为止的研究表明,地球上的生物使用的是同一套遗传密码,但不同的生物在遗传密码的使用上往往表现出偏嗜性(Codon usage bias),即如果某个氨基酸具有几个密码的话,有些生物表现出偏好使用其中某一个或两个密码,有的生物甚至使用由稀有碱基构成的特殊密码。这种密码使用的偏嗜性是生物物种的特征。我们最近完成基因组测序的 PaP3 噬菌体是铜绿假单胞菌的胞内寄生物,作为病毒它没有自己的细胞器核糖体,因此推测它在遗传密码使用的偏嗜性上可能与宿主菌保持一致性。本研究试图通过对噬菌体 PaP3 及其宿主铜绿假单胞菌的遗传密码使用偏嗜性进行分析,以期考察噬菌体与其宿主菌的密码使用偏嗜性是否存在差异性?同时对噬菌体的基因组中 tRNA 基因进行检索,并进一步分析遗传密码偏嗜性是否与噬菌体自己编码的 tRNA 基因有关?从而为揭示噬菌体 PaP3 自己编码的这 3 个 tRNA 基因的功能提供线索。

1 方法

1.1 遗传密码子偏嗜性分析

挑选 10 个在 Genemark 分析中认定的 PaP3 的推定基因(同时在随后的 BLAST 分析中认定它们极可能是真正的编码基因,其中包括蛋白 N-端测序证实为 PaP3 主要结构蛋白的基因),利用 DNAStar 软件包中的 EditSeq 软件对其遗传密码子的使用频率进行统计计算,再对其偏嗜性进行分析。同时在 Genbank 中随机挑选 10 个确定的铜绿假单胞菌 PAO1 的基因,作同样的偏嗜性分析。并将结果做对照分析,以便考察寄生物噬菌体与宿

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)
作者简介:张克斌(1968-),男,四川省岳池县人,博士,主治医师,讲师,主要从事微生物基因组和生物信息学方面的研究,发表论文 10 篇,现在附属新桥医院检验科,重庆 400037。电话:(023)68755000-74654
通信作者:胡福泉,电话:(023)68752240, E-mail: hufuquan@yahoo.com
收稿日期:2002-02-03;修回日期:2002-03-08

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

主菌之间在遗传密码使用方面有无一致性。

1.2 噬菌体 PaP3 tRNA 基因的检索

在法国 Pasteur 实验室网站上进行，网址：http://bioweb.pasteur.fr/seqanal/，选择 Fast tRNA Analysis Method，以简易格式或 *.seq 格式输入 PaP3 全长 44 139 bp 序列。结果以 e-mail 形式发回。

1.3 tRNA 二级结构折叠模拟

使用 DNAStar 中 GeneQuest 软件分析。

2 结果

2.1 遗传密码偏嗜性分析

PaP3 及 PAO1 遗传密码使用频率见表 1 所示。从表中可见：①噬菌体 PaP3 和铜绿假单胞菌优势码一致的氨基酸有 F-uuc，I-auc，V-guc，A-gcc，Y-uac，H-cac，N-aac，K-aag 和 T-acc。②PaP3 单独具有偏嗜性密码的氨基酸有 A-gcu，G-ggu，T-acu 和 V-gua。③铜绿假单胞菌 PAO1 单独具有密码偏嗜性的氨基酸有 L-cug，V-gug，P-ccg，V-gcg，Q-cag，D-gac，E-gag，G-ugc，R-cgc，S-agc 和 G-ggc。总体来看，绿脓杆菌 PAO1 在遗传密码偏嗜性上表现非常明显，除 Glu(E)和 Tyr(W)原本只有一个遗传密码外，其余各氨基酸均有一种或两种使用率优势密码。PaP3 也具有遗传密码使用的偏嗜性，但密码偏嗜性相对较弱，只在部分氨基酸中存在，且优势密码使用百分率相对偏低。噬菌体与宿主菌二者之间虽有一定共性，但仍存在不小差异。

表 1 铜绿假单胞菌 **PAO1** 与噬菌体 **PaP3** 遗传密码使用频率

Tab 1 Codon usage frequency of PAO1 and bacteriophage PaP3

	U			C			A			G		
U	F	**22.0**	7.4	S	**21.7**	2.1	Y	**30.1**	19.1	C	**57.1**	12.5	U
	F	**78.0**	92.6	S	**28.0**	27.1	Y	**69.9**	80.9	C	**42.9**	87.5	C
	L *	**5.4**	0.9	S	**6.7**	1.4	Stop	**40.0**	0	Stop	**50.0**	75.0	A
	L *	**12.6**	7.0	S	**9.8**	23.6	Stop	**10.0**	25.0	W	**100**	100	G
C	L	**19.4**	4.1	P	**29.3**	5.2	H	**30.9**	24.0	R	**26.3**	16.0	U
	L	**24.8**	20.1	P	**32.0**	22.0	H	**69.1**	76.0	R	**18.0**	62.4	C
	L	**14.3**	2.2	P	**23.3**	1.2	Q	**57.5**	15.2	R	**18.4**	4.1	A
	L	**23.5**	65.7	P	**15.4**	71.6	Q	**42.5**	84.8	R	**12.7**	13.1	G
A	I	**25.7**	13.5	T	**31.9**	6.1	N	**23.8**	13.2	S	**16.6**	5.4	U
	I	**67.8**	83.8	T	**35.8**	75.8	N	**76.2**	86.8	S	**17.2**	40.4	C
	I	**6.5**	2.7	T	**14.7**	4.0	K	**22.8**	18.4	R	**5.7**	1.8	A
	M	**100**	100	T	**17.6**	14.1	K	**77.2**	81.6	R	**18.9**	2.5	G
G	V	**19.7**	7.7	A	**41.1**	5.1	D	**50.8**	23.4	G	**42.7**	13.9	U
	V	**32.1**	36.7	A	**30.4**	56.6	D	**49.2**	76.6	G	**18.1**	66.9	C
	V	**30.3**	8.2	A	**20.4**	4.8	E	**40.8**	40.4	G	**18.8**	6.1	A
	V	**17.9**	47.4	A	**8.1**	33.6	E	**59.2**	59.6	G	**20.4**	13.1	G

Note: The left-hand of the table column contains the first nucleotides of the codons, the second nucleotides at the top, and the third in the right-hand column. The amino acids or functions are given on the left. The numbers represent percentages, which of black body represent percentages of PaP3, and those with and another corresponding percentage of PAO1, those with represent obvious dominant codons; *: Here the first nucleotide of codons for Leu is C

2.2 PaP3 中 tRNA 编码序列分析及二级结构模拟

将 PaP3 全序列发送到法国巴斯德实验室网站进行 tRNA 基因查询，结果发现该基因组中有三段 tRNA 编码序列，均位于负链，其起始位点、序列长度和序列如下：

①转运 Asn 的 tRNA：该基因位于 42 967～43 043，全长 76 个核苷酸。其反密码子(Anticodon)是“gtt”，编码 tRNA 序列为

tgggatgtagctcagttggtagagcagggagctgttaactctc

aggtcgtagattcgagccctaccgtcccagcca

反密码子臂(anticodon arm)的分值为 29，全域分值(Global score，SG)＝9，二级结构预测见图 1。

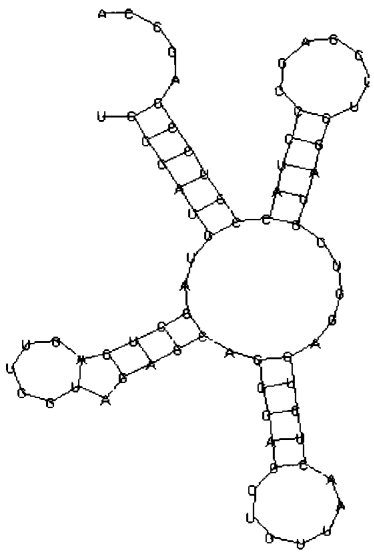


图 1 **tRNA^{Asn}** 二级结构模拟图

Fig 1 Secondary structure remodeling of tRNA^{Asn}

② 转运 Asp 的 tRNA：该基因位于 42 887～42 963，全长 76 个核苷酸。其反密码子是“gtc”，编码的 tRNA 序列是

ggcccattaacttagcggttagagtgctcccctgtctagggag

aagccgtcgggttcgaatccgacatgggtcgcca

反密码子臂的分值为 33，全域分值＝7，其二级结构预测见图 2。

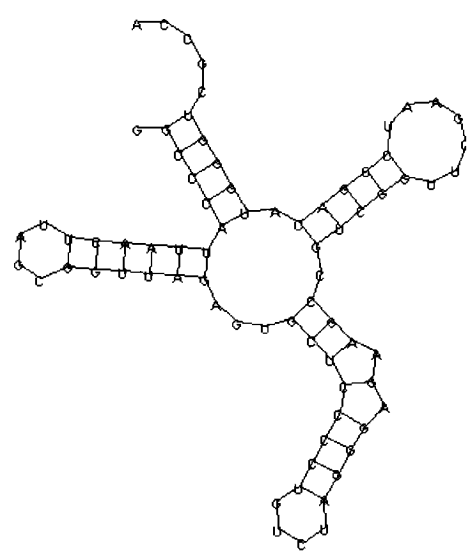


图 2 **tRNA^{Asp}** 二级结构模拟图

Fig 2 Secondary structure remodeling of tRNA^{Asp}

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

③转运 Pro 的 tRNA：该基因位于 42 715 ~ 42 793，全长 78 个核苷酸。其反密码子是"tgg"，其编码的 tRNA 序列是

ctcctcgtagctcagtctggttagagtgcctgcctggaagcag

gaggtcgtaggttcgaatcctaccgggtgacat

反密码子臂分值为 33，全域分值＝8，其二级结构预测见图 3。

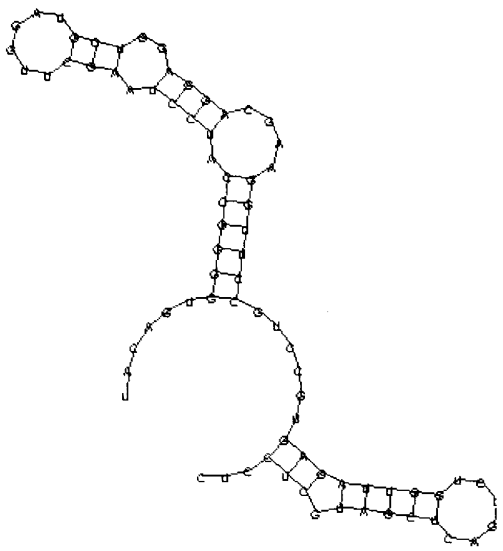


图 3　$tRNA^{Pro}$二级结构模拟图

Fig 3　Secondary structure remodeling of $tRNA^{Pro}$

3　讨论

不同生物对遗传密码的使用有不同的偏嗜性，这一现象早为人们所知并在设计兼并引物时加以运用。我们在核苷酸水平对 PaP3 基因组各 ORF 做 Blastn 检索时未发现同源检出物，但在氨基酸水平做 BLASTx 检索时却发现了四种同源蛋白，提示噬菌体 PaP3 与其它噬菌体或物种间存在遗传密码子偏嗜性差异。故本研究选择 10 种 PaP3 的推定基因，对其遗传密码的偏嗜性进行统计分析。同时，从 Genbank 中随意挑选 10 个铜绿假单胞菌 PAO1 株基因，同步分析其遗传密码使用偏嗜性。结果发现，二者在遗传密码使用频率上存在较大差异。尽管噬菌体在基因转录和蛋白翻译过程中使用宿主菌的许多工具，但迄今尚无研究表明噬菌体在遗传密码偏嗜性上必须与宿主菌保持一致。Carolyn 等[1] 的研究发现鼠伤寒沙门氏菌噬菌体 P22 在遗传密码偏嗜性上与其宿主菌有较大差异；本研究中噬菌体 PaP3 与铜绿假单胞菌 PAO1 在遗传密码偏嗜性上也并不完全重叠，可见遗传密码使用的偏嗜性是生物物种的特征。

目前人们对于不同遗传密码偏嗜性机制尚缺乏深入的研究。文献[2] 认为遗传密码偏嗜性与终产物蛋白质的二级结构有关，也有研究表明，在哺乳类蛋白中(如免疫球蛋白的可变区和恒定区)，密码使用偏嗜性也与蛋白的二级结构有一定相关性，但在原核生物蛋白中未发现这种相关性[3]。在这个问题上作者更倾向于认为不同物种在遗传密码使用偏嗜性上决定于两个因素：一是该种生物不同 tRNA 的可利用度(tRNA availability)，即哪一种 tRNA 可供使用的数量多？而这一数量又可能与 tRNA 基因的拷贝数有关；二是 tRNA 二级结构与核糖体结合位点的适配性，即具有哪一种反密码子的 tRNA 所携带的氨基酸更容易与核糖体结合位点相结合。

基于这种认识，我们想进一步分析是否在 PaP3 基因组中存在编码 tRNA 的基因？这些基因编码的 tRNA 如果作为氨基酰的搬运工具是否与 PaP3 的遗传密码子偏嗜性有关？于是我们对 PaP3 基因组中的 tRNA 基因进行了检索。

基因组中 tRNA 基因的编码产物不是蛋白质，而是 RNA 分子，用 ORF 查询为基础的检索方法发现不了 tRNA 基因。由于噬菌体没有自己的核糖体，很容易使人想到它可能也没有 tRNA，然而噬菌体有自己的 tRNA 基因却是事实[4]。

过去几十年来，人们对于 tRNA 分子进行了深入的研究，并对积累的许多 tRNA 编码序列进行比对分析[5]，在比对中发现了一些 tRNA 共有的保守序列。早期在计算机上对 tRNA 序列的分析主要基于对其保守区域的搜索，通过序列中核苷酸权值(Weight)矩阵或同义群(Consensus)矩阵计算分析得出结果[6]。第一个可靠的 tRNA 分析软件 tRNAscan，由 Fichant &Burks (1991)完成[7]。随后 Nadia El-Mabrouk 等开发了更为迅速的分析软件 Fast tRNA Analysis Method[8]，本研究采用了这种分析软件。其分析的基本思路是：首先，从序列中搜索和鉴定潜在的 TΨC 区，设定 7nt 形成 UΨC 环，5 个碱基对形成 TΨC 臂；其次是搜索和鉴定 D 区，分析该区域与 TΨC 区的相对距离和二者之间可能存在的可变区和内含子序列；接下来是分析目的序列的氨基酰臂；至此，由每步的得分相加得到一总体相似性分值(Global similarity score)，若大于设定阈值，则进行最后一步，既搜索和分析反密码子臂和反密码子环形成的可能性；如大于设定阈值，则将此段序列鉴定为 tRNA 编码序列[8]。

本研究共发现 3 段 tRNA 编码序列，其中 $tRNA^{Asn}$和 $tRNA^{Asp}$经二级结构模拟分析后发现为典型的三叶草结构，各臂和环均完整无缺，且氨基酰臂 3′端碱基为 CCA。但 $tRNA^{Pro}$经二级结构模拟后为不规则的茎环结构，除反密码子环外，其余各结构均不典型，且氨基酰臂 3′-端碱基并非其它 tRNA 所共有的 CCA，而是 CAU。目前，我们尚未发现此类非典型 tRNA 的相关资料，这

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

有待于我们进一步的学习和探讨。

进一步逐一将这 3 个 tRNA 基因的反密码子与 PaP3 的遗传密码偏嗜性联系分析。$tRNA^{Asn}$ 的反密码子为“gtt”，对应密码子为“aac”，恰好为 PaP3 中 Asn(N)的使用优势码，使用频率为 76.2%；而 $tRNA^{Asp}$ 的反密码子为“gtc”，其对应密码子为“gac”，在 PaP3 中 Asp(D)的“gac”使用率为 49.2%，而另一密码“gau”的使用率为 50.8%，二者之间无显著差异；$tRNA^{Pro}$ 的反密码子为“tgg”，其对应密码子为“cca”，在 PaP3 中 Pro(P)的“cca”使用率为 23.3%，与其它密码“ccc”32.0%，“ccg”15.3%，“ccu”29.3%的使用率相比，也非优势码。这些分析提示，噬菌体 PaP3 中的三个 tRNA 基因与其对应的遗传密码偏嗜性之间并无直接联系。这些 tRNA 基因是否真正转录为 tRNA，抑或具有其它的生物学意义，值得我们进一步研究和探讨。

近来有一些研究表明，tRNA 基因与溶原性噬菌体的整合位点密切相关。如根瘤菌 Meliloti 41 株(Rm41)的溶原性噬菌体 16-3 的整合位点便覆盖了一个推测的脯氨酸 tRNA 基因的 3′端，该基因与放线菌中某一整合性质粒的相应的脯氨酸 tRNA 样基因组靶序列有 79%的相似性[9]；噬菌体 phiCTX 的基因组可整合进入铜绿假单胞菌的丝氨酸 tRNA 基因，该基因图谱位于铜绿假单胞菌染色体的 43～45 min 区域[10]；温和噬菌体 mv4 将其 DNA 整合进德氏乳杆菌 subsp. bulgaricus 株的染色体中，细菌附着位点(AttB)的序列分析表明，其中心区域与丝氨酸的 tRNA 基因有重叠，mv4 DNA 整合进丝氨酸 tRNA，将一个完整的丝氨酸 tRNA 基因保留在噬菌体附着位点(attL)位点上[11]。温和噬菌体 phiSfi21 将其 DNA 整合进嗜热链球菌染色体中，附着位点的序列分析鉴定出 40 bp 的中心区，该区与整合酶基因末端的 18 bp 和宿主精氨酸 tRNA 基因末端的 11 bp 高度重叠[12]。在所有研究过的噬菌体与 15 株酒明串珠菌的整合作用中，attB 均位于亮氨酸的 tRNA 基因内[13]。温和噬菌体 VWB 插入委内瑞拉链霉菌 ETH14630 中，AttP、attB、attL 及 attR 的序列分析揭示了一个 45 bp 的通用中心区；在 attB 中 45 bp 序列由精氨酸 tRNA 基因的 3′末端组成[14]。在新近完成测序的 *E. coli* 基因组中，发现了两个新的前噬菌体，研究者发现其中之一的 cp4-6 的溶原位点在 *E. coli* 的 tRNA 基因 thrW(60)上，因为 thrW 的 3′端在相隔 34 242 bp 的下游有重复序列，该序列与几种噬菌体整合酶(Integrase)的同源基因 b0281 相比邻[15]。噬菌体经常利用 tRNA 基因作为插入位点，可能是利用反密码子茎环序列可作为替代的中心位点。

目前，噬菌体 PaP3 的溶原性位点的鉴定工作正在进行之中，基因组中的三个 tRNA 基因与溶原性位点之间的关系将得到揭示，有关的研究结果将另文发表。

参考文献：

[1] Carolyn V B, Andrew M K. Sequence of the Genome of Salmonella bacteriophage P22[J]. J Bacterio, 2000, 182(22): 6 472－6 481.

[2] Xie T, Ding D, Tao X, *et al*. The relationship between synonymous codon usage and protein structure[J]. FEBS Lett, 1998, 434(1－2): 93－96.

[3] Miyata T, Hayashida H, Yasunaga T, *et al*. The preferential codon usages in variable and constant regions of immunoglobulin genes are quite distinct from each other[J]. Nucleic Acids Res, 1979, 7(8): 2 431－2 438.

[4] Kropinski A M, Sibbald M J. Transfer RNA genes and their significance to codon usage in the Pseudomonas aeruginosa lamboid bacteriophage D3[J]. Can J Microbiol, 1999, 45(9): 791－796.

[5] Sprinzl M, Steegborn C, Huebel F, *et al*. Complilation of tRNA sequences and sequences of tRNA genes[J]. Nucl Acids Res, 1992, 24: 68－72.

[6] Staden R. Methods to define and locate patterns of motifs in sequences[J]. Comput Appl Biol sci, 1988, 4: 53－60.

[7] Fichant G A, Burks C. Identifying potential tRNA genes in genomic DNA sequences[J]. J Mol Biol, 1991, 220: 659－671.

[8] El-Mabrouk N, Lisacek F. Very fast identification of RNA motifs in genomic DNA. Application to tRNA search in the yeast genome[J]. J Mol Biol, 1996, 264: 46－55.

[9] Papp I, Dorgai L, Papp P, *et al*. The bacterial attachment site of the temperate Rhizobium phage 16－3 overlaps the 3' end of a putative proline tRNA gene[J]. Mol Gen Genet, 1993, 240(2): 258－264.

[10] Hayashi T, Matsumoto H, Ohnishi M, *et al*. Molecular analysis of a cytotoxin-converting phage, phi CTX, of Pseudomonas aeruginosa: structure of the attP-cos-ctx region and integration into the serine tRNA gene[J]. Mol Microbiol, 1993, 7(5): 657－667.

[11] Dupont L, Boizet B B, Coddeville M, *et al*. Characterization of genetic elements required for site-specific integration of Lactobacillus delbrueckii subsp. bulgaricus bacteriophage mv4 and construction of an integration-proficient vector for Lactobacillus plantarum[J]. J Bacteriol, 1995, 177(3): 586－95.

[12] Bruttin A, Foley S, Brussow H. The site-specific integration system of the temperate Streptococcus thermophilus bacteriophage phiSfi21[J]. Virology, 1997, 237(1): 148－158.

[13] Gindreau E, Torlois S, Lonvaud Funel A, *et al*. Identification and sequence analysis of the region encoding the site-specific integration system from Leuconostoc oenos (OEnococcus oeni) temperate bacteriophage phi 10MC[J]. FEMS Microbiol Lett, 1997, 147(2): 279－285.

[14] Van Mellaert L, Mei L, Lammertyn E, *et al*. Site-specific integration of bacteriophage VWB genome into Streptomyces venezuelae and construction of a VWB-based integrative vector[J]. Microbiology, 1998, 144(Pt 12): 3 351－3 358.

[15] Blattner F R, Plunkett G III, Bloch C A, *et al*. The complete genome sequence of Escherichia coli K-12[J]. Science, 1997, 277(5 331): 1 453－1 474.

（编辑 冷怀明）

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