

刺糖多孢菌鼠李糖和福乐糖胺合成基因的克隆和组装

郭航¹ 白亭丽¹ 陶美凤^{1,2}

1. 华中农业大学农业微生物学国家重点实验室, 武汉 430070;

2. 上海交通大学生命科学技术学院微生物代谢国家重点实验室, 上海 200030

摘要 鼠李糖和福乐糖胺是多杀菌素生物合成必需的2个脱氧糖结构单元,负责其合成的4种酶的基因(*gtt*、*epi*、*gdh*、*kre*)与多杀菌素生物合成基因簇并不在一起,而是分散分布在染色体的3个位点上。为克隆到完整的多杀菌素生物合成基因簇,采用构建基因组文库、PCR扩增等手段从刺糖多孢菌 NRRL18395 染色体分别克隆 *gtt*、*epi*、*gdh* 和 *kre* 基因,并将其克隆组装在同一整合型载体上,以便于构建用于表达多杀菌素糖单元合成基因的表达载体。

关键词 刺糖多孢菌; 多杀菌素; 鼠李糖; 福乐糖胺; cosmid 文库; 原位杂交

中图分类号 Q 78 **文献标识码** A **文章编号** 1000-2421(2012)03-0298-05

多杀菌素(spinosyns)是由放线菌刺糖多孢菌(*Saccharopolyspora spinosa*) NRRL18395 发酵产生的无公害高效杀虫剂,属大环内酯类化合物,制剂中主要成分为多杀菌素组份 A 和 D 的混合物,因此又称为 spinosad^[1]。多杀菌素最早于 1997 年在美国登记注册,开始应用于棉花和蔬菜;2005 年,多杀菌素经批准被用于 73 个国家的 250 多种农作物上,在我国商品名为菜喜(2.5%悬浮剂)和催杀(48%悬浮剂)^[2]。多杀菌素的母核是由 1 个 5,6,5-顺-反-顺三环和 1 个十二元的内酯环粘合而成,并通过糖苷键连接着 2 个不同的脱氧糖,其中 1 个是氨基糖二甲基福乐糖胺(dimethyl-forsamine),另 1 个是中性糖 α -L-鼠李糖(L-rhamnose)。多杀菌素生物合成的必需基因已经克隆测序,大部分基因聚集形成 1 个巨大的多杀菌素生物合成基因簇(*spn* 基因簇,80 kb)^[3]。但是,2 个脱氧糖单元合成所必需的 4 个基因不在基因簇中,而是分散位于染色体的 3 个不同位置,其中 NDP 葡萄糖脱水酶基因(*gdh*)和酮还原酶基因(*kre*)组成 1 个操纵子,NDP 葡萄糖合酶基因(*gtt*)和 3',5'-表异构酶基因(*epi*)分别独立存在^[4]。

鼠李糖和福乐糖胺合成过程中,*gtt*(编码 NDP-葡萄糖合成酶)、*gdh*(编码葡萄糖脱氢酶)催化葡萄

糖-1-磷酸生成 NDP-4-酮-6-脱氧-D-葡萄糖,作为鼠李糖和福乐糖胺合成的必需中间体,*epi*(编码 3',5'-表异构酶)和 *kre*(编码 4'-酮还原酶),将 NDP-4-酮-6-脱氧-D-葡萄糖转化为 α -L-鼠李糖。*epi*、*gtt*、*gdh*、*kre* 的缺失突变体不产多杀菌素,同时突变体生长情况也比野生型差^[4]。多杀菌素产生菌刺糖多孢菌菌株的遗传操作非常困难,使得旨在优化多杀菌素结构以及高产育种的基因工程改造困难重重^[5]。将生物合成基因转到异源宿主进行表达,并在异源表达宿主中开展组合生物合成和高产育种工作中成为理性的替代策略^[6-7]。本研究克隆了鼠李糖和福乐糖胺合成相关的 4 个基因,并将这 4 个基因组合到 1 个链霉菌整合型载体,构建得到合成 2 个脱氧糖单元的表达载体,为在异源宿主中表达多杀菌素及其基因工程改造奠定基础。

1 材料与方法

1.1 试验材料

- 1) 菌株和质粒。所用菌株和质粒见表 1。
- 2) 克隆引物设计。引物序列见表 2。
- 3) 培养基。大肠杆菌培养用 LB 培养基^[8];胰蛋白胨 10 g,酵母抽提物 5 g,NaCl 5 g,补加蒸馏水至 1 L,pH 7.0;刺糖多孢菌产孢用 YD 培养基^[12];

收稿日期: 2011-05-29

基金项目: 国家自然科学基金项目(30870073)

郭航,硕士研究生,研究方向:链霉菌分子生物学。E-mail: bobo-heao@126.com

通讯作者: 陶美凤,研究员,研究方向:链霉菌分子生物学。E-mail: tao_meifeng@sytu.edu.cn

表 1 菌株和质粒¹⁾

Table 1 Bacterial strains and plasmids

菌株或质粒 Strain or plasmid	特征 Characteristics	来源或文献 Source or reference
刺糖多孢菌 <i>S. spinosa</i>	多杀菌素产生菌 Producer of spinosad	NRRL
大肠杆菌 <i>Escherichia coli</i>		
DH5α	基因克隆宿主菌 Cloning host strain, F ⁻ <i>endA1 recA1 deoR gyrA96 relA</i>	[8]
XL1-Blue ^{MR}	Cosmid 文库构建宿主菌 Cosmid library construction host bacteria, Δ(<i>mcrA</i>) 183Δ(<i>mcrCB⁻ hsdSMR⁻ mrr</i>)173 <i>endA1supE44 thi-1 recA1 gyrA96 relA1 lac</i>	Epicenter
质粒 Plasmid		
pOJ446	Cosmid 文库构建载体 Cosmid library construction vector, 10.4 kb, <i>aac</i> (3) IV, <i>cos</i> sites	[9]
pMS82	大肠杆菌链霉菌穿梭载体 Shuttle plasmid vector of <i>Streptomyces</i> and <i>E. coli</i> , 6.7 kb, <i>hyg</i> , <i>int/attP</i> (ϕ BT1), <i>oriT</i>	[10]
pIJ2925	大肠杆菌载体 <i>E. coli</i> vector, 2.7 kb, <i>bla</i> , <i>lacZα</i>	[11]
pSET152	大肠杆菌链霉菌穿梭载体 Shuttle plasmid vector of <i>Streptomyces</i> and <i>E. coli</i> , 5.5 kb, <i>int/attP</i> (ϕ C31), <i>oriT</i> (RK2), <i>ori</i> (pUC18), <i>aac</i> (3) IV	[9]
pHL801	<i>gtt</i> 克隆在 pGEM-T Easy 载体 pGEM-T Easy containing <i>gtt</i> gene	本研究 This study
pHL802	<i>epi</i> 克隆在 pGEM-T Easy 载体 pGEM-T Easy containing <i>epi</i> gene	本研究 This study
pHL803	含 <i>gdh-kre</i> 的 ~8 kb 的 <i>Sph</i> I 片段连接在 pIJ2925 的 <i>Sph</i> I 位点 pIJ2925 containing a 8-kb <i>Sph</i> I fragment of <i>gdh-kre</i>	本研究 This study
pHL804	pHL801 中含 <i>gtt</i> 的片段连接在 pMS82 的 <i>Spe</i> I 位点 pMS82 containing <i>Spe</i> I fragment of <i>gtt</i> from pHL801	本研究 This study
pHL805	pHL802 中含 <i>epi</i> 的片段连接在 pHL804 的 <i>Spe</i> I 位点 pHL804 containing <i>Spe</i> I fragment of <i>epi</i> from pHL802	本研究 This study
pHL808	pHL803 中含 <i>gdh-kre</i> 的 <i>Nhe</i> I- <i>Hind</i> Ⅲ 片段连接在 pSET152 的 <i>Nhe</i> I- <i>Hind</i> Ⅲ 位点 pSET152 containing <i>Nhe</i> I - <i>Hind</i> Ⅲ fragment of <i>gdh-kre</i> from pHL803	本研究 This study
pHL809	pHL808 中含 <i>gdh-kre</i> 的 <i>Nhe</i> I - <i>Xba</i> I 片段连接在 pHL805 <i>Spe</i> I 位点 pHL805 containing <i>Nhe</i> I - <i>Xba</i> I fragment of <i>gdh-kre</i> from pHL808	本研究 This study

1)NRRL:美国农业研究菌种保藏中心 U. S. Agricultural Research Service Culture Collection.

表 2 引物及其特征

Table 2 Primers sequence and description

引物 Primer	序列 Sequence	用途 Application
pEpif1	5'-GATCTAGAGGATCCTAATACGACTCACTATAGGGATCAACAACAACCTCACCAGCA-3'	克隆 <i>epi</i> 基因
pEpirl	5'-GAACTAGTTGGAGGTGGATGTGAAATCCCTCGG-3'	Cloning of <i>epi</i> gene
pGttf1	5'-GATCTAGAGGATCCTAATACGACTCACTATAAAGGCCACCGCAAGGTCGTGCAGG-3'	克隆 <i>gtt</i> 基因
pGttr1	5'-GAACTAGTGCACCCGCCGATGGCCGACCGCATT-3'	Cloning of <i>gtt</i> gene
pGK1f	5'-GGATCCTGCTTCGTAGCTCG-3'	验证 <i>kre-gdh</i>
pGK1r	5'-GGATCCGCTTCCCCCACGG-3'	基因和制备探针
pGK2f	5'-TCCTGCTTCGTAGCTCGGTG-3'	Confirmation of
pGK2r	5'-TCAACGAAGCCCTGCACCAA-3'	<i>kre-gdh</i> gene and probe preparation

酵母抽提物 4 g, 麦芽糖 10 g, 葡萄糖 4 g, MgCl₂ TOYOBO 公司产品。

2 g, CaCl₂ 1.5 g, 青岛琼脂 15 g, 补加蒸馏水至 1 L, pH 7.2。

4) 抗生素与生化试剂。大肠杆菌培养中抗生素使用质量浓度为: 氨苄青霉素 100 mg/L、阿泊拉霉素 50 mg/L、潮霉素 B 50 mg/L。各种限制性内切酶、DNA 连接酶和碱性磷酸酶为 Fermentas 公司产品。DNA 聚合酶、dNTPs、DNA 分子质量 Marker 为 TaKaRa 公司产品。KOD 高保真聚合酶为

1.2 试验方法

1) 大肠杆菌 DH5α 化学法感受态制备、质粒转化、DNA 操作等参考文献[8]进行。

2) 链霉菌的培养与总 DNA 的提取。链霉菌菌丝体培养、总 DNA 提取参考文献[11]。

3) 基因组文库构建及 Southern 杂交筛选文库。基因组 cosmid 文库构建使用载体为 pOJ446^[9], 建库方法参考 SuperCos1 科斯质粒试剂盒 (Strata-

gene)和 CopyControl™ Fosmid 文库构建试剂盒说明书(Epicenter)。Southern blot 按照 DIG 杂交试剂盒说明书(Roche)。

4)Cosmid 文库基因组覆盖率计算^[8]。 $N = \ln(1 - P) / \ln(1 - f)$, P , 文库覆盖率, 分数表示; f , 插入片段大小占基因组的比例; N , 所需要的克隆子个数。

2 结果与分析

2.1 gtt、epi 的 PCR 扩增及克隆

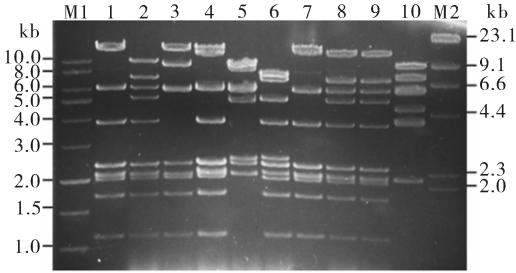
多杀菌素属大环内酯类化合物, 根据多杀菌素的化合物类型、合成途径以及宿主自身遗传背景, 克隆基因时选择保留其天然启动子; 根据异源表达的情况考虑是否使用其他链霉菌启动子。GenBank 中, *gtt*、*epi* 包括天然启动子在内的完整序列已知, 因此根据 GenBank 中的 *gtt* 基因及其上下游序列(AF355467)设计引物, 一端引物设置 *Xba* I 酶切位点, 另一端设置 *Spe* I 酶切位点, 以便于后续的基因克隆组装。以 *S. spinosa* NRRL18395 总 DNA 为模板, PCR 扩增 *gtt* 基因及其上游的天然启动子区。将 PCR 产物 TA 克隆到 pGEM-T Easy 载体, 得到 pHL801。同样, 根据 *epi* 序列(AF355466)设计引物, PCR 扩增并 TA 克隆得到 pHL802。测序验证 pHL801 和 pHL802 中的 *gtt*、*epi* 及上游启动子区序列正确。

2.2 刺糖多孢菌 NRRL18395 基因组文库构建及 *gdh-kre* 的克隆

刺糖多孢菌总 DNA 用 *Mbo* I 进行部分酶切, 跑胶回收 40 kb 以上的部分酶切总 DNA, CIAP 去磷酸化处理; 载体 pOJ446 用 *Hpa* I 完全酶切, CIAP 去磷酸化, 纯化回收后再用 *Bam*H I 完全酶切, 回收 8.6 kb 片段。将总 DNA 和载体连接、包装、感染 XL1-Blue^{MR} 感受态细胞, 挑取阿泊拉霉素抗性的克隆 2 400 个, 得到刺糖多孢菌 NRRL18395 的基因组科斯质粒文库。随机挑选 20 个克隆酶切检查文库质量, 插入片段平均 30 kb。

根据 *gdh-kre* 序列(AF355468)设计引物, 以 PCR 扩增产物作为分子探针进行 Southern 杂交筛选基因组文库, 得到 10 个科斯质粒克隆子 7G10、8E9、8E12、11A8、13C8、15C7、18C5、18E8、18G9 和 24G6, 用 *Bam*H I 酶切制作酶谱, 发现这些阳性克隆互相重叠, 形成一个重叠连锁群(Contig), 如图 1 和图 2 所示。用 *Sph* I 酶切 7G10, 回收包含 *gdh-*

kre 的 8.5 kb 片段, 克隆到 pIJ2925 的 *Sph* I 位点, 得到 pHL803, PCR 扩增验证含有 *gdh-kre* 完整基因。分别对两侧未知序列进行测序, pHL803 上插入的总 DNA 片段中 *gdh* 基因上游 DNA > 1 kb, 显然包含天然启动子区。



M1: 1 kb ladder; M2: λ/*Hind*III; 1~10: 阳性克隆 *Bam*HI digested positives clones, 1: 7G10; 2: 8E9; 3: 8E12; 4: 11A8; 5: 13C8; 6: 15C7; 7: 18C5; 8: 18E8; 9: 18G9; 10: 24G6.

图 1 Cosmid 文库 *gdh-kre* 基因阳性克隆的 *Bam*H I 酶切图谱

Fig. 1 *Bam*H I restriction map of the *gdh-kre* gene positive clones from cosmid bank

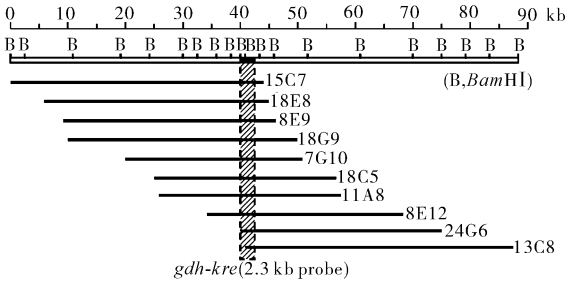


图 2 根据 *gdh-kre* 基因阳性克隆的 *Bam*H I 酶切图谱排列得到的重叠连锁群

Fig. 2 Contig of the positive clones from the *Bam*H I restriction map

2.3 鼠李糖和福乐糖胺合成基因的组装

利用 *Xba* I、*Nhe* I 和 *Spe* I 等 3 个限制酶产生相同粘性末端的特点, 依次将含有 *gtt*、*epi* 和 *gdh-kre* 完整基因及启动子区的 DNA 片段插入到链霉菌整合性载体 pMS82 上, 得到 pHL809。即将 pHL801 中含 *gtt* 的 *Xba* I-*Spe* I 片段连接在 pMS82 的 *Spe* I 位点, 得到 pHL804; 将 pHL802 中含 *epi* 的 *Xba* I-*Spe* I 片段连接在 pHL804 的 *Spe* I 位点, 得到 pHL805; 由文库 cosmid 7G10 亚克隆得到含 *gdh-kre* 的 ~8 kb 的 *Sph* I 片段连接在 pIJ2925 的相应位点得到 pHL803, 测序分析完整的 *gdh-kre* 包含在 *Nhe* I-*Hind*III 片段(~3.5 kb)中, 连接在 pSET152 的相应位点, 得到 pHL808; 最后

将 pHL808 中含 *gdh-kre* 基因的 *Nhe* I-*Xba* I 片段连接在 pHL805 的 *Spe* I 位点,得到 pHL809 (图 3)。pHL809 经酶切、PCR 及部分测序验证正确。

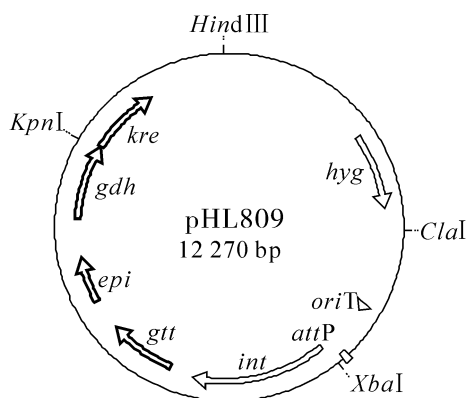


图3 鼠李糖和福乐糖胺合成基因表达载体 pHL809

Fig. 3 pHL809, the expression vector of the forosamine and rhamnose biosynthetic pathway

3 讨论

自首次从刺糖多孢菌发酵液中分离出来至今,多杀菌素及其类似物由于杀虫效果好、环境友好、对非靶标昆虫无毒等特点,在医疗卫生、农业生产、畜牧业等方面得到广泛应用,效果显著^[13-14]。多杀菌素兼具生物农药安全性和化学合成农药速效性特点,使其作为环保、高效的新型杀虫剂具有广阔的应用前景,相关产品在美国分别于1999年和2008年2次获得美国“总统绿色化学品挑战奖”。多杀菌素专利已过期,在我国多杀菌素的研发及产业化生产将带来巨大的经济效益和社会效益^[2]。因此,我国相关科研和产业界也对多杀菌素开发一直有着浓厚的兴趣,但由于菌种低产及遗传操作困难等原因,对多杀菌素的研究仍处于实验室阶段,尚不具备工业化生产的条件^[2]。

考虑到刺糖多孢菌的遗传学操作存在巨大困难,尝试在异源宿主内表达多杀菌素,通过对异源宿主的代谢工程和遗传操作以实现高产多杀菌素,乃至通过组合生物合成技术优化结构、合成新的结构类似物,将是一个合理有效的策略。但是,由于多杀菌素生物合成的部分必需基因与 *spn* 基因簇不在一起,这些基因也必须克隆、组装并引入到异源宿主,才能保障异源表达的成功。

根据基因序列特点,本研究分别采用构建基因组文库和 PCR 手段克隆鼠李糖和福乐糖胺合成的

4个基因。分别克隆到4个基因后,利用 *Spe* I、*Nhe* I 和 *Xba* I 酶切后产生相同粘性末端的特点依次组装得到表达2个糖单元合成基因的表达载体 pHL809。在 GenBank 给出的 *kre-gdh* 序列中,结构基因上游只有88 bp 非编码序列^[4],因而不能肯定是否包含基因转录和翻译及其调控所必需的功能序列,因此采用建库的方法克隆包括上游必需功能区的完整基因,以确保目的基因能够采用自身的天然启动子进行表达。pHL809 还含有噬菌体 BT1 的整合位点 *attP* 和整合酶基因 *int*、潮霉素抗性基因 *hyg* 和广宿主接合质粒 RK2 的接合转移起始区 *oriT*,从而可以方便地接合转移到各种候选的放线菌异源表达宿主,并将组装后的2个糖单元合成基因一起稳定整合到宿主染色体中的 *attB* 整合位点上,为多杀菌素的异源表达奠定基础。

要实现多杀菌素的异源表达,还需要将多杀菌素合成基因簇 *spn* 导入表达宿主中。*spn* 基因簇较大,共有80 kb,至少2~3个科斯质粒才能囊括完整的基因簇,在我们的基因组文库中也筛选到多个分别包含部分 *spn* 基因簇的阳性克隆,进一步工作需要采用 RED/ET 重组技术^[15-16]将位于不同科斯质粒的部分基因簇拼接起来。同时,我们也在尝试采用细菌人工染色体(BAC)载体将该 *spn* 基因簇克隆下来。最后,分别将糖合成基因表达载体和完整 *spn* 基因簇一起转到异源宿主,可望得到多杀菌素的异源表达。在异源宿主的选择方面,模式菌株在遗传背景、抗生素合成及其调控以及遗传操作等方面研究得清楚^[17],有较多成功经验可以借鉴。

参考文献

- [1] THOMPSON G D, DUTTON R, SPARKS T C. Spinosad—a case study: an example from a natural products discovery programme[J]. Pest Manag Sci, 2000, 56(8): 696-702.
- [2] 蔡恒, 王燕, 万红贵, 等. 刺糖多孢菌生产多杀菌素的研究进展[J]. 中国生物工程杂志, 2011, 31(2): 124-129.
- [3] WALDRON C, MATSUSHIMA P, ROSTECK P R J, et al. Cloning and analysis of the spinosad biosynthetic gene cluster of *Saccharopolyspora spinosa* [J]. Chem Biol, 2001, 8: 487-499.
- [4] MADDURI K, WALDRON C, MATSUSHIMA P, et al. Genes for the biosynthesis of spinosyns: applications for yield improvement in *Saccharopolyspora spinosa* [J]. J Ind Microbiol Biotechnol, 2001, 27: 399-402.
- [5] MATSUSHIMA P, BALTZ R H. Transformation of *Saccharopolyspora spinosa* protoplasts with plasmid DNA modified in

vitro to avoid host restriction[J]. Microbiology, 1994, 140: 139-143.

[6] BALTZ R H. Streptomyces and *Saccharopolyspora* hosts for heterologous expression of secondary metabolite gene clusters [J]. J Ind Microbiol Biotechnol, 2010, 37: 759-772.

[7] NGUYEN K T, RITZ D, GU J Q, et al. Combinational biosynthesis of novel antibiotics related to daptomycin [J]. Proc Natl Acad Sci, 2006, 103: 17462-17467.

[8] SAMBROOK J, RUSSELL D. Molecular cloning: a laboratory manual [M]. 3rd. New York: Spring Harbor Laboratory Press, 2001.

[9] BIERMAN M, LOGAN R, O'BRIEN K, et al. Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp. [J]. Gene, 1992, 116: 43-49.

[10] GREGORY M A, TILL R, SMITH M C. Integration site for *Streptomyces* phage phiBT1 and development of site-specific integrating vectors [J]. J Bacteriol, 2003, 185(17): 5320-5323.

[11] KIESER T, BIBB M J, BUTTNER M J, et al. Practical streptomyces genetics [M]. Norwich: John Innes Foundation, 2000.

[12] FOOR F, ROBERTS G, MORIN N, et al. Isolation and characterization of the *Streptomyces cattleya* temperate phage TG1 [J]. Gene, 1985, 39: 11-16.

[13] HUANG K, XIA L, ZHANG Y, et al. Recent advances in the biochemistry of spinosyns [J]. Appl Microbiol Biotechnol, 2009, 82: 13-23.

[14] KIRST H A. The spinosyn family of insecticides: realizing the potential of natural products research [J]. The Journal of Antibiotics, 2010, 63: 101-111.

[15] WENZEL S C, GROSS F, ZHANG Y, et al. Heterologous expression of a myxobacterial natural products assembly line in pseudomonads *via* red/ET recombineering [J]. Chem Biol, 2005, 12(3): 349-356.

[16] 季思思, 朱义广, 彭东海, 等. 苏云金芽胞杆菌质粒 p26 复制功能区的克隆 [J]. 华中农业大学学报, 2011, 30(5): 558-562.

[17] BENTLEY S D, CHATER K F, CERDENO-TARRAGA A M, et al. Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3 (2) [J]. Nature, 2002, 417: 141-147.

Cloning and reconstitution of rhamnose and forosamine biosynthetic gene of *Saccharopolyspora spinosa* NRRL18395

GUO Hang¹ BAI Ting-li¹ TAO Mei-feng^{1,2}

1. State Key Laboratory of Agricultural Microbiology, Huazhong Agricultural University, Wuhan 430070, China;
2. State Key Laboratory of Microbial Metabolism and School of Life Science and Biotechnology, Shanghai Jiaotong University, Shanghai 200030, China

Abstract Rhamnose and forosamine are two essential deoxysugars components of spinosad. Four genes involved in the biosynthesis of rhamnose and forosamine, i. e. *gtt* (encoding NDP-glucose synthase), *gdh* (encoding NDP-glucose dehydratase), *epi* (encoding epimerase) and *kre* (encoding 4'-ketoreductase), are not located together with the spinosad biosynthetic gene cluster, but distributed in 3 scattered loci in the chromosome of *Saccharopolyspora spinosa* NRRL18395. To get an integral spinosad biosynthetic pathway, the four genes were cloned by constructing a genomic library and PCR amplifying from *S. spinosa* NRRL18395. The 4 genes were successively assembled into an integrative vector to construct an expression vector for biosynthesis of rhamnose and forosamine in heterologous hosts.

Key words *Saccharopolyspora spinosa*; spinosad; rhamnose; forosamine; cosmid library; *in situ* hybridization

(责任编辑: 陆文昌)