

链霉菌基因组中放线菌型整合性接合元件的识别

徐珍¹ 毕德玺² 李鹏² 谭之磊¹ 邓子新² 欧竑宇^{2*} 贾士儒^{1*}

(1. 天津科技大学 工业微生物教育部重点实验室 天津 300457)

(2. 上海交通大学 微生物代谢国家重点实验室 上海 200030)

摘要: 【目的】链霉菌染色体重组和外源 DNA 片段插入是影响其遗传多样性的主要因素。旨在考察放线菌型整合性接合元件(AICE)在链霉菌遗传多样性中所发挥的作用。【方法】基于 AICE 的特征性模块, 采用隐马尔科夫模型预测链霉菌基因组序列中的 AICEs。【结果】在已全测序的 12 条链霉菌染色体和 35 个质粒中, 共识别出 29 个 AICEs, 其中 12 个为首次报道。*Streptomyces coelicolor* 基因组中发现了 4 个 AICEs, 而其近缘的 *Streptomyces lividans* 却没有。【结论】AICEs 都整合在链霉菌染色体的核心区, 且都具有典型的整合环出、复制和接合转移等核心模块, 这些可自行转移的元件在链霉菌基因组可塑性中扮演了重要角色。

关键词: 链霉菌基因组, 放线菌型整合性接合元件, 隐马尔科夫模型

基金项目: 国家自然科学基金项目(No. 31170082, 21276197); 国家 973 计划项目(No. 2012CB721002)

*通讯作者: 欧竑宇: Tel: 86-21-62932943; ✉: hyou@sjtu.edu.cn

贾士儒: Tel: 86-22-60601598; ✉: jiashiru@tust.edu.cn

收稿日期: 2012-12-25; 接受日期: 2013-01-30

In silico identification of actinomycete integrative and conjugative elements in completely sequenced *Streptomyces* genomes

XU Zhen¹ BI De-Xi² LI Peng² TAN Zhi-Lei¹ DENG Zi-Xin²
OU Hong-Yu^{2*} JIA Shi-Ru^{1*}

(1. Key Laboratory of Industrial Microbiology, Ministry of Education, Tianjin University of Science and Technology, Tianjin 300457, China)

(2. State Key Laboratory of Microbial Metabolism, Shanghai Jiao Tong University, Shanghai 200030, China)

Abstract: [Objective] Chromosomal rearrangements and integration of alien DNA fragments contribute to the genome-wide genetic diversity of *Streptomyces*. This paper aims to study the role of actinomycete integrative and conjugative elements (AICEs) in *Streptomyces* genetic diversity. [Methods] AICEs were identified in the completely sequenced *Streptomyces* genomes using a profile Hidden Markov Model approach, based on their characteristic modules. [Results] Twenty-nine AICEs were found from 12 chromosome sequences and 35 plasmid sequences of *Streptomyces*, 12 of which were firstly reported. Four AICEs were identified in *S. coelicolor*, but none in the closely related *S. lividans*. [Conclusion] AICEs were found to be inserted into the core regions of the *Streptomyces* chromosomes. They typically comprised three core genetic modules, implementing chromosomal integration/excision, self-transfer by conjugation and replication, suggesting that they promote the plasticity of *Streptomyces* genomes.

Keywords: *Streptomyces* genome, Actinomycete integrative and conjugative elements, Profile Hidden Markov Model

链霉菌属细菌(*Streptomyces*)是一类革兰氏阳性、多细胞的丝状土壤微生物,是研究微生物遗传多样性和基因表达调控机制的良好模式材料。链霉菌具有复杂的生活周期和次生代谢途径,产生大量具有重要价值的天然代谢物,如大约 4 000 种以上抗生素(占自然界已知抗生素的 70%)^[1]。链霉菌形态分化、抗生素及色素合成等次生代谢过程受到其基因组遗传多样性的影响。链霉菌基因组大小约为 8 Mb, G+C 含量高达 70%,其染色体大部分呈线状,具有不稳定的末端^[2]。同源重组或外源 DNA 片段插入常导致链霉菌基

因组的高度不稳定性。此外,可移动遗传元件通过水平转移整合到链霉菌染色体上,所携带的新性状有助于宿主在特定生境下获得优势^[3]。

整合性接合元件(Integrative and conjugative elements, ICEs)是一种广泛存在于细菌中可自主移动的遗传元件,是基因水平转移的一种有效途径,在细菌进化中具有重要意义,但目前所发现的 ICEs 仅如冰山一角^[3]。典型的 ICE 通常编码整合酶、完整的接合系统和调控蛋白等。这些蛋白可控制 ICE 从供体菌染色体上切除和环出,以接合转移的方式进入受体菌并整合到其染色体上^[4-5]。放

线菌型 ICEs 一般用 AICEs (Actinomycete ICEs) 表示, 其核心模块如图 1A 所示。位点特异性整合发生在 ICE 的 *attB* 与染色体上 *attP* 之间, *attB* 和 *attP* 是两段短的高度同源序列。在革兰氏阴性菌和部分革兰氏阳性菌中, ICE 往往依赖四型分泌系统以单链的形式进行接合转移^[6-7]。而 AICE 的接合转移则是通过菌丝体之间的紧密接触完成的, 且在接合转移过程中直接以双链的形式进入受体菌。AICE 的独特之处还在于其具有自主复制功能^[8-9], 且这种自主复制功能与接合转移的过程紧密相关^[10-11]。AICE 的复制起始蛋白(Rep) 是 RepSA (命名于 pSAM2)和 RepAM (命名于 pMEA300)类型的蛋白, 也有一些是 Prim-Pol 类型的聚合酶(既可以介导 DNA 的复制起始又可以参与 DNA 的聚合)。一类包含有 FtsK/SpoIIIE 功能域的蛋白(Tra)负责双链 DNA 的接合转移^[12], 是 AICE 接合转移的核心模块。与这几个核心模块相关的特征蛋白质是预测典型 AICE 的重要依据。

目前报道最多的 ICEs 多位于革兰氏阴性菌和硬壁菌门的部分革兰氏阳性菌中, 它们常赋予宿主新的性状, 如耐药性和致病性等。但关于链霉菌中 AICE 的系统性研究相对较少。随着 DNA 测序成本和时间的剧减, 多个链霉菌全基因组序列已被测定, 急需比较分析其基因组多样性和深度挖掘其代谢潜力。本文采用了隐马尔科夫模型预测链霉菌基因组序列中的 AICE, 系统探讨了在已全测序链霉菌基因组中 AICEs 的分布、核心模块和进化关系, 为链霉菌基因组多样性的研究提供了新方法和思路。

1 材料与方法

1.1 链霉菌基因组序列来源

从 NCBI 收集整理了以下 3 类序列数据: (1) 目前全基因组测序的 11 条链霉菌染色体, *S. griseus* NBRC 13350 (NC_010572), *S. avermitilis*

MA-4680 (NC_003155), *S. scabiei* 87.22 (NC_013929), *S. bingchengensis* BCW-1 (NC_016582), *S. flavogriseus* ATCC 33331 (NC_016114), *S. venezuelae* ATCC 10712 (NC_018750), *S. violaceusniger* Tu 4113 (NC_015957), *Streptomyces* sp. SirexAA-E (NC_015953), *S. cattleya* DSM 46488 (NC_017586), *S. hygrosopicus* jinggangensis 5008 (NC_017765), *S. coelicolor* A3(2) (NC_003888); (2) 已完成全基因组拼接的 *S. lividans* TK24 (NZ_GG657756); (3) 35 个全测序的链霉菌质粒, *S. avermitilis* MA-4680 plasmid SAP1 (NC_004719), *S. flavogriseus* ATCC 33331 pSFLA02 (NC_016115)和 pSFLA01 (NC_016110), *S. violaceusniger* Tu 4113 pSTRVI01 (NC_015951)和 pSTRVI02 (NC_01595), *S. cattleya* DSM 46488 pSCATT (NC_017585), *S. hygrosopicus* jinggangensis 5008 pSHJG1 (NC_017766)和 pSHJG2 (NC_016972), *S. coelicolor* A3(2) SCP1 (NC_003903) 和 SCP2 (NC_003904), *S. albulus* pNO33 (NC_006571), *S. clavuligerus* ATCC 27064 pSCL (NC_001738), *S. coelicolor* pSCP2* (NC_003319), *S. cyaneus* pSA1.1 (NC_002112), *S. flavovirens* pSN22 (NC_001425), *S. ghanaensis* pSG5 (NC_008792), *S. laurentii* pSLS (NC_008441), *S. lavendulae* pSLG33 (NC_010097), *S. lividans* 1326 pSLP2 (NC_004933), *S. natalensis* pSNA1 (NC_002149), *S. phaeochromogenes* pJV1 (NC_001759), *S. rochei* pSLA2-L (NC_004808), *Streptomyces* sp. 44030 pRL1 (NC_010849), *Streptomyces* sp. 44414 pRL2 (NC_007927), *Streptomyces* sp. EN27 pEN2701 (NC_004931), *Streptomyces* sp. F11 pFP11 (NC_006911), *Streptomyces* sp. FQ1 pFP1 (NC_006912), *Streptomyces* sp. FR1 pFRL1 (NC_010851), *Streptomyces* sp. HK1 pSHK1 (NC_010311), *Streptomyces* sp. W9 pCQ3 (NC_013449), *Streptomyces* sp. Y27 pWTY27 (NC_013667), *Streptomyces* sp. ZL12 pZL12 (NC_013420), *Streptomyces* sp. x3 pTSC2 (NC_013417), *S. venezuelae* pSVH1 (NC_007431), *S. violaceoruber* pSV2 (NC_004934)。

1.2 AICE 预测与进化分析

整体预测方法如图 1B 所示。与整合、转移

和复制相关的 3 个核心模块用于识别典型的 AICE。本文使用隐马尔科夫模型来表征与这些核心模块相关的特征蛋白的序列保守性(表 1)。采用 HMMer 3.0^[14]软件来查找链霉菌基因组序列中的 AICE 特征蛋白, 设定阈值 $E\text{-value} \leq 0.000\ 05$ 。编码这些特征蛋白的基因距离设定为 60 kb 之内(目前已知的同 AICE 中特征蛋白之间的距离小于 50 kb)^[5]。用 tRNAscan-SE 软件^[15]预测常用作外源 DNA 片段插入位点的 tRNA 基因。采用 Vmatch (<http://vmatch.de/>)来识别 AICE 两侧的正向重复序列。最后结合正向重复序列和 tRNA 基因位点进行人工校验, 确定 AICE 插入位点、两侧正向重复序列及整合、复制和转移等核心模块。蛋白序列的系统发生树用 MEGA 5.0 软件^[16]生成, 采用邻接法, 自展值取 1 000。以模式菌 *S. coelicolor* A3(2)为参照, 使用 MUMmer 3.0 软件^[17]确定已测序链霉菌染色体的核心区及左右两臂。

2 结果与讨论

2.1 AICE 整合在链霉菌染色体的核心区

从已完成全基因组测序或拼装的链霉菌基因组中, 共收集了 47 个复制子。我们使用隐马尔科夫模型来捕捉与整合、转移和复制等 AICE 核心模块相关的特征蛋白, 共识别出了 29 个 AICEs (图 2)。其中, 17 个 AICEs 与之前的文献报道相一

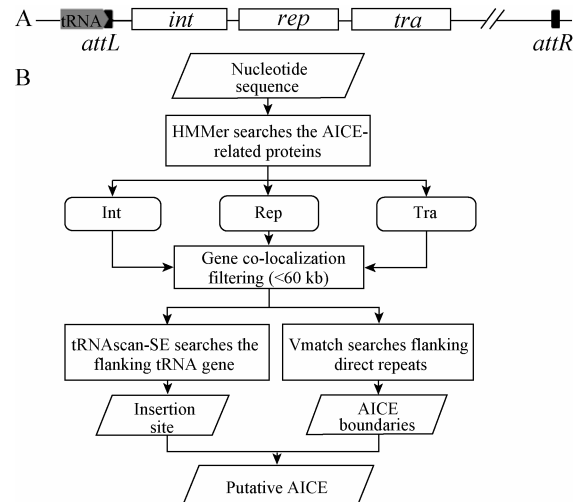


图 1 典型 AICE 的核心模块(A)与本文使用的 AICE 预测方法(B)

Fig. 1 Schematic representation of AICEs core modules (A) and the proposed prediction approach (B)

致^[12-13], 另外 12 个 AICEs 是本文首次报道的(图 2, 星号标注)。在 12 个已测序的链霉菌基因组中, 其中 10 个链霉菌整合了一个或多个不同类型的 AICEs, 其中含有 AICEs 最多的是 *S. coelicolor* A3(2); 而与 *S. coelicolor* 近缘的 *S. lividans* TK24 则没有发现 AICE, *Streptomyces* sp. SirexAA-E 也没有识别出 AICE。AICEs 大小分布在 9–24 kb, 基本都具备 4 个核心模块: 整合/切除模块(图 2 橙色与浅蓝色)、复制模块(图 2 黄色)、接合转移

表 1 识别与 AICE 核心模块相关的特征蛋白所使用的隐马尔可夫模型

Table 1 HMM profiles of conserved proteins related to AICE core modules used in this study

核心模块 Core module	蛋白家族或域 Protein family/domain (Pfam ID)	描述 Description
Integration	Phage_integrase (PF00589)	Phage integrase family (Tyrosine recombinase)
	Recombinase (PF07508)	Recombinase (Serine recombinase)
Replication	RepSA ^[13]	RepSA, Replication initiator protein, pSAM2
	DUF3631 (PF12307)	RepAM, Replication initiator protein, pMEA300
	Prim-Pol (PF09250)	Bifunctional DNA primase/polymerase, N-terminal
Translocation	FtsK_SpoIIIE (PF01580)	Intercellular chromosomal DNA transfer

模块(图 2 绿色)和调控模块(图 2 蓝色)。识别出的 29 个 AICEs 中, 有 26 个位于染色体上(表 2); 另外 3 个 AICEs 是质粒, 这 3 个质粒自身都携带有整合酶、复制相关蛋白、控制接合转移的基因以及相应的调控基因, 符合 AICE 的典型结构。

识别出的 AICEs 都位于链霉菌染色体的核心保守区(图 3)。值得注意的是, *S. violaceusniger* Tu 4113 染色体是环形的(NC_015957), 其复制起始位点(*ori*)位于 2 kb 处(图 3 SVI)。与 *S. coelicolor* A3(2)染色体进行比对可以区分其核心保守区与不稳定区域, 与其它链霉菌线性染色体有所不同。但从 *S. violaceusniger* Tu 4113 染色体识别出的 3 个 AICEs 也同样都位于保守区域。我们推测, 虽然不同链霉菌染色体的核心区域较为保守, 但外源 AICEs 在核心区的整合是链霉菌基因组多样性的重要原因之一。

2.2 天蓝色链霉菌与阿维链霉菌中的 AICEs

S. coelicolor A3(2)是目前所发现含有 AICEs 最多的链霉菌。其染色体大小为 8.6 Mb, 核心区上有 4 个完整的 AICEs, 和一个包含有部分 AICE 特征的片段(长度为 11.3 kb, 数据未给出)^[5,12]。其中, SLP1 是研究最早最详细的 AICE 之一。它的大小为 17 kb, 整合在染色体的 tRNA-Tyr 基因位点, 编码来自噬菌体的酪氨酸重组酶家族的整合酶, 同时还有切除酶辅助其从染色体环出, 接合转移依赖于其包含有 FtsK-SpoIIIE 功能域的 Tra 蛋白, 整套系统可以保证 SLP1 从 *S. coelicolor* 高效率的转移到 *S. lividans* 中^[18]。而另外 3 个 AICEs 的研究都是基于生物信息学分析, 目前还没有实验报道其功能。pSAM2 类型的 AICESco3250 和 AICESco5349 都整合到 tRNA-Arg 基因位点, AICESco3250 位于钙离子依赖型抗生素生物合成基因簇的附近, 而 AICESco5349 位于孢子灰色色素(一种聚酮类化合物)合成基因簇(*whiE*)的附近。te Poele 等推测这两个 AICEs 可能与 *S. coelicolor*

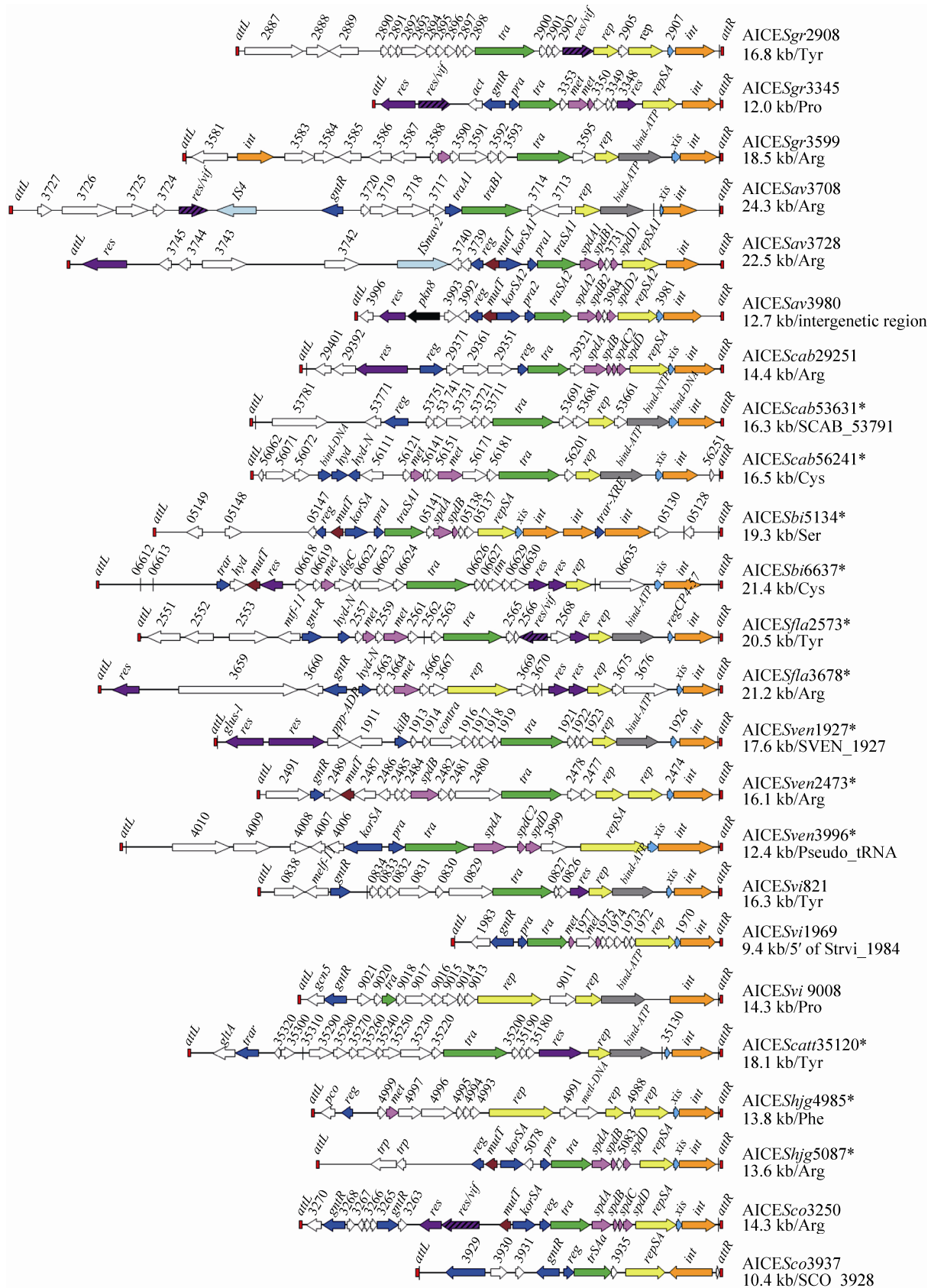
A3(2)通过水平转移获得次级代谢物合成基因簇相关^[12]。AICESco3937 则插入硫胺素(Thiamine)生物合成基因 *thiC* (SCO_3928), 介导此特异性插入的整合酶属于丝氨酸重组酶家族, 这在 AICEs 中是比较罕见的(详见 2.3 部分)。

阿维链霉菌 *S. avermitilis* MA-4680 染色体大小为 9 Mb, 共识别出 3 个 AICEs。其中 AICESav3708 和 AICESav3728 以串联的形式整合在同一个 tRNA-Arg 位点。两个 AICEs 分别具有独立而完整的整合、复制、接合转移以及调控模块。这一现象说明, AICESav3708 和 AICESav3728 编码的整合酶具有特异性插入 tRNA-Arg 位点的特征。我们推测, 可能是 AICESav3728 先插入 tRNA-Arg 位点, 在整合完成之后, 此 tRNA-Arg 的 3'端 *attP* 恢复原始状态; 随后, AICESav3708 识别新的 *attP* 并插入到同一位置(图 4)。此外, 两者复制和接合转移模块相关蛋白的进化距离相差较远(图 5B 和 5C), 考虑到移动遗传元件的不相容性, 推测这是 AICESav3708 和 AICESav3728 能够稳定存在于同一位置的重要原因。另一个 AICESav3980 插入的是基因间区, 没有明显的特征, 但是从其基因构成上看, 与 AICESav3728 很相似(图 2), 并且组成相关模块的蛋白进化距离也很近(图 5)。AICESav3980 和 AICESav3728 极有可能来自同一个祖先, 在进化中 AICESav3980 的整合酶发生了改变, 从而导致插入的位点也发生了变化。

2.3 链霉菌基因组中 AICE 具有整合环出、复制和接合转移等核心模块

在已全测序的链霉菌基因组中识别出的 29 个 AICEs 都具有典型的整合环出、复制和接合转移等核心模块, 我们推测这些 AICEs 可自行转移。

在整合环出模块中, 27 个 AICEs 都是依赖于酪氨酸重组酶家族的整合酶进行切除和整合, 只



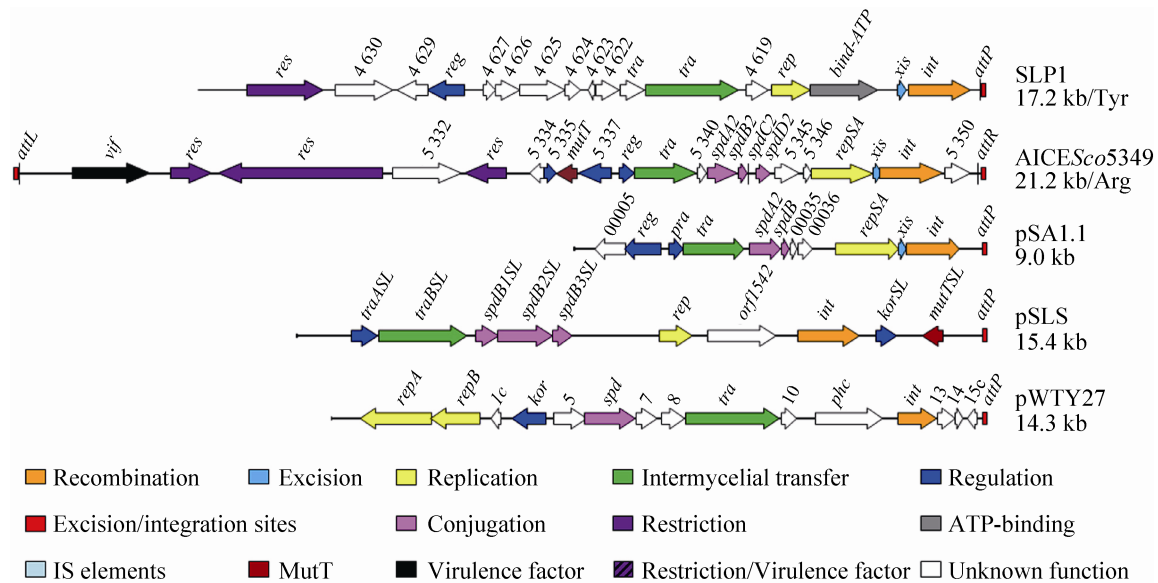


图 2 29 个链霉菌 AICEs 的核心模块、插入位点及两侧正向重复序列

Fig. 2 Genetic organization of 29 identified AICEs in the sequenced *Streptomyces* genomes

Note: *: New identified AICEs/Insertion site. Gene abbreviation: *rep*: Replication initiator; *int*: Integrase; *xis*: Excisionase; *met*: Mobile element transfer protein; *kor*: Regulatory protein; *tra*: Transfer protein; *res*: Restriction protein; *mut*: MutT-like protein; *act*: Acetyltransferase; *reg*: Regulator; *hyd*: Hydrolase; *pra*: Replication activator protein; *digc*: Diguanylate cyclase; *kil*: Protein kilB; *gltA*: GltA; *itm*: Integral membrane protein; *mf*: Methyltransferase; *trp*: Transposase; *vif*: Virulence factor; *glus*: Glutamine synthetase; *contra*: Putative conjugal transfer protein; *gcn*: N-acetyltransferase GCN5; *pco*: Polycistronic operon; *bind-NTP*: NTP-binding protein; *bind-DNA*: DNA-binding protein; *trar*: Transcriptional regulator protein; *spd*: Mobile element transfer protein; *bind-ATP*: ATP-binding protein; *gntR*: GntR family transcriptional regulator; *ppp-ADP*: ADP-ribose pyrophosphatase. Italic letters in AICE and full name of host strains: *Sgr*: *S. griseus* subsp. *griseus* NBRC 13350; *Sav*: *S. avermitilis* MA-4680; *Scab*: *S. scabiei* 87.22; *Sbi*: *S. bingchengensis* BCW-1; *Sfla*: *S. flavogriseus* ATCC 33331; *Sven*: *S. venezuelae* ATCC 10712; *Svi*: *S. violaceusniger* Tu 4113; *Scatt*: *S. cattleya* DSM 46488; *Sco*: *S. coelicolor* A3(2); *Shjg*: *S. hygroscopicus* subsp. *jinggangensis* 5008.

表 2 已测序链霉菌基因组中识别出的 AICEs
Table 2 Identified AICEs in the sequenced *Streptomyces* genomes

	全测序数量 Sequenced number	复制子 Replicon	复制子数量 Replicon number	AICE 数量 AICE number	本文首次报道 Firstly reported
Complete genome	12	Chromosome	12	26	12
		Plasmid	10	0	0
Plasmid sequenced alone	25	Plasmid	25	3	0
Total			47	29	12

有 2 个 AICEs 编码丝氨酸重组酶家族的整合酶 (图 5A)。这一现象与革兰氏阴性菌 ICEs 和温和噬菌体是一致的。在目前所发现的具有类似机制的可移动遗传元件当中, 由丝氨酸重组酶家族整合酶所介导的特异性插入 tRNA 位点是非常罕见的, 这一现象目前只在温和噬菌体 Ψ RSM1 中得

到了验证^[19-20], 其它都是基于生物信息学的预测。从另外 27 个整合酶的系统发育树(图 5A)可以看出, 这些 AICEs 的整合酶同源性较为接近, 根据序列相似性可以将其简单的归类成 5 个亚家族(图 5 中每条竖线代表一个亚家族, 竖线右侧为典型的代表 AICE) (图 2 橙色, 图 5A)。同一株链

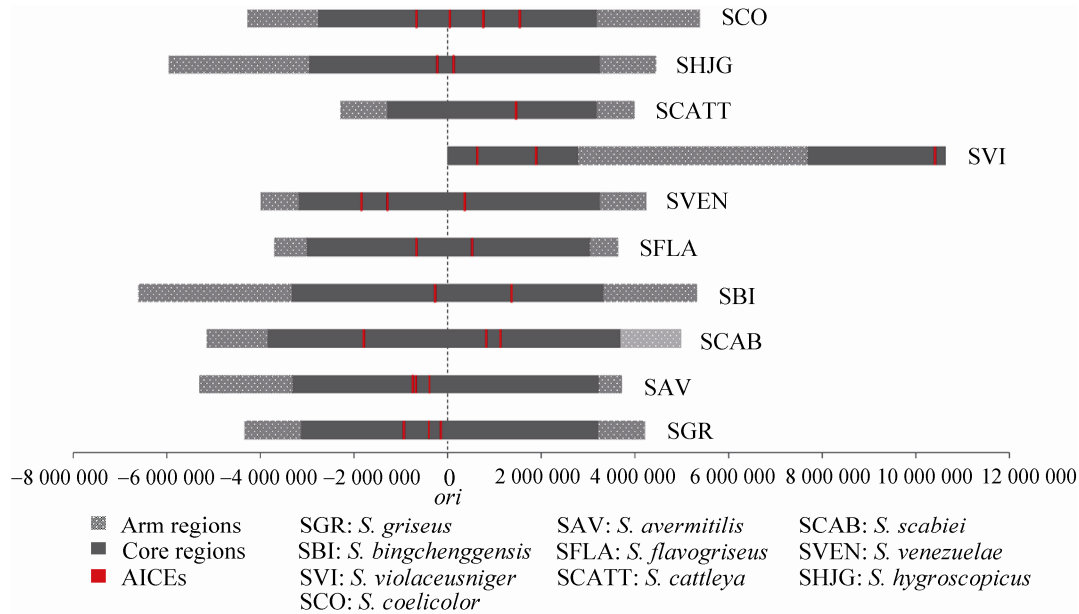


图3 AICEs 在已测序链霉菌染色体上的分布

Fig. 3 Distribution of identified AICEs in the sequenced *Streptomyces* chromosomes

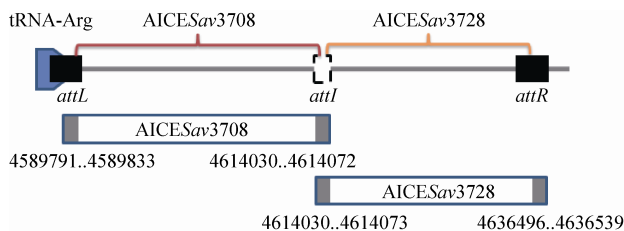


图4 阿维链霉菌染色体上 AICESav3708 与 AICESav3728 整合在同一个 tRNA-Arg 基因位点

Fig. 4 AICESav3708 and AICESav3728 insert into the same tRNA-Arg gene site in the *S. avermitilis* MA-4680 chromosome

霉菌中的几个 AICEs 具有不同亚家族的整合酶, 可以使更多的 AICEs 整合到染色体不同位点, 还可以保证不同 AICEs 在染色体上的遗传稳定性。因为不同的移动遗传元件携带相同的整合酶或整合在相同的位点容易发生重组, 且元件之间的不相容性会使得某些元件容易丢失^[21-23]。除了 pSLS 亚家族的整合酶(图 5A 紫色竖线)之外, 其它酪氨酸重组酶家族的整合酶都是介导特异性整合于染色体的 tRNA 位点。整合酶进化关系越近, 所介导的特异整合位点也趋于一致, 如携带

编码 SLP1 亚家族整合酶的 AICEs 趋于整合到染色体的 tRNA-Tyr 位点, 携带编码 pSA1.1 亚家族整合酶的 AICEs 趋于整合到染色体的 tRNA-Arg 位点(图 5A)。

AICEs 的复制模块是其区别于革兰氏阴性菌 ICEs 的重要特征之一。AICEs 的复制起始位点 (*ori*)和复制起始蛋白与其接合转移过程是紧密相关的, 实验证明, pSAM2 的这两个重要位点的任何一个发生缺失, pSAM2 都不能接合转移到受体细胞中^[10-11]。链霉菌 AICEs 的复制起始蛋白主要有 3 种, 分别是 RepSA、RepAM 和 Prim-Pol 家族的蛋白(图 2 黄色, 表 1, 图 5C)。链霉菌 AICEs 的复制模块具有多态性, 某些 AICEs 还具有多个复制蛋白(图 2 黄色), 而且没有种特异性, 这之前所观察到的整合酶的特性相一致, 保证了 AICEs 在染色体上的相对稳定。

AICEs 接合模块最大的特点是介导双链 DNA 的转移, 与链霉菌的接合质粒类似, 这与链霉菌本身的遗传体系有很大的关系。在接合转移

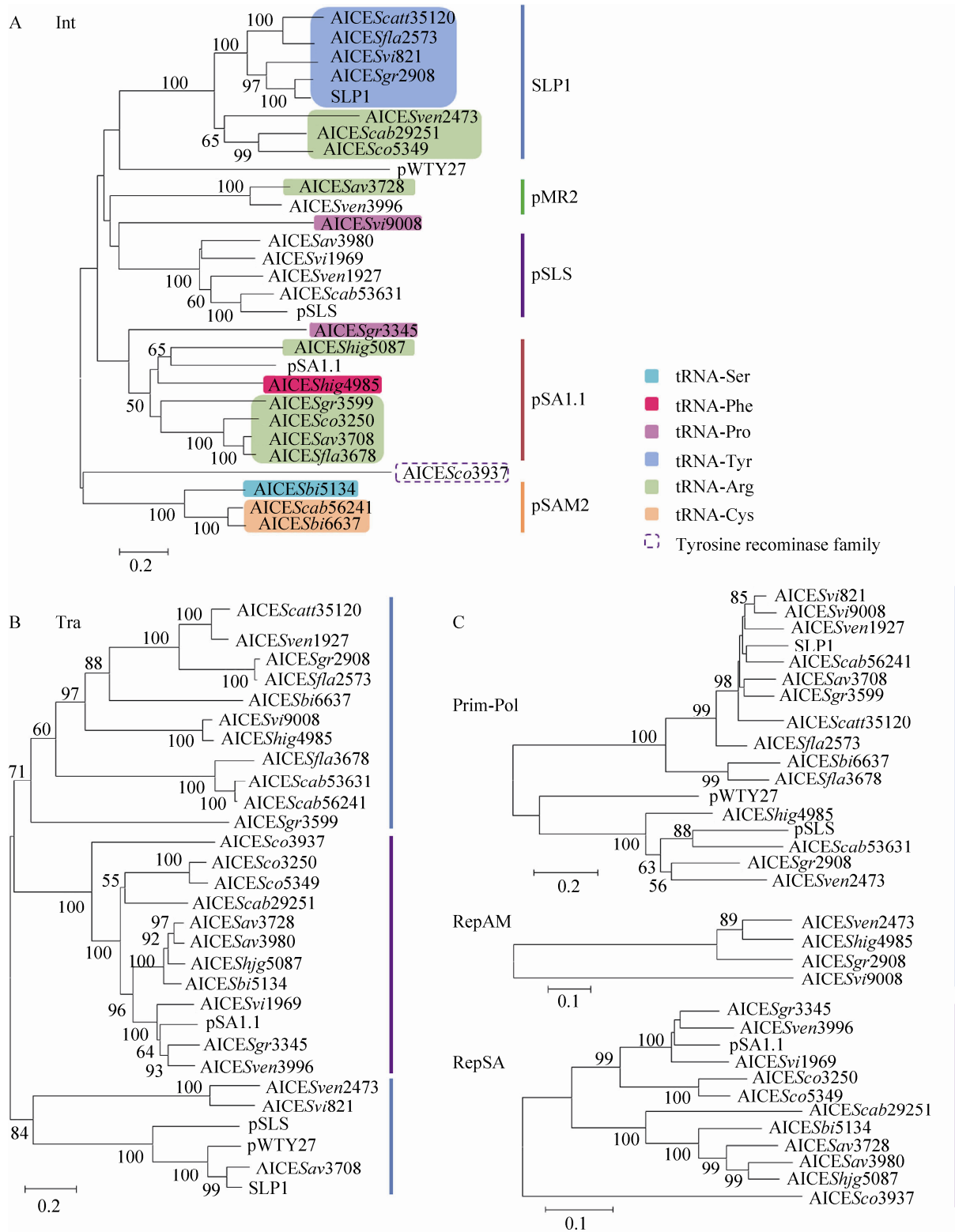


图5 AICEs 核心模块相关特征蛋白的系统发生树

Fig. 5 Phylogenetic analyses of conserved proteins involving in diverse core modules of AICEs

Note: Int: Integrase; Tra: Translocation; Prim-Pol: Bifunctional DNA primase/polymerase; RepAM: Replication initiator protein, pMEA300; RepSA: Replication initiator protein, pSAM2.

过程中,链霉菌的菌丝体紧密接触并缠绕在一起形成聚合体,从染色体上环出的双链 AICE 在 Tra 蛋白的辅助下即可从供体菌进入受体细胞中。此机制类似于链霉菌在孢子形成和细胞分裂过程中的染色体分裂过程。而革兰氏阴性菌和硬壁菌中部分革兰氏阳性菌的 ICEs 大多依赖于四型分泌系统进行传播。Tra 蛋白是一类包含有 FtsK-SpoIIIE 功能域的蛋白,含有 FtsK-SpoIIIE 功能域的蛋白往往参与芽孢的形成和细胞的分裂过程,主要负责染色体的分区并转移双链染色体 DNA 到新的细胞^[24-25]。在识别出的 29 个链霉菌 AICEs 中, Tra 蛋白可以分为 3 个亚家族(图 2 绿色,图 5B)。有意思的是,所有与 pSAM2 的 Tra 蛋白类似的 AICEs,它们的复制起始蛋白都属于 RepSA 家族(图 5B 和 C 紫色),其它 AICEs 的复制起始蛋白分属于 RepAM 和 Prim-Pol 家族(图 5B 和 C 蓝色)。这与前面提到的 AICEs 复制与接合转移过程的紧密关联是吻合的,这种紧密的关联必然需要负责相关过程的蛋白在功能上相匹配。

此外, AICEs 还具有调控模块,但目前的研究并不深入。AICEs 具有温和噬菌体和链霉菌质粒的双重特性,既可以整合在染色体上与染色体一起进行复制,又可以从染色体上环出然后像质粒那样进行自我复制。因此,参与温和噬菌体和质粒调控的相关因子都可能与 AICEs 的调控相关。AICEs 在不同的条件诱导下环出的频率也不同,培养基的营养成分及其降解的中间产物、抗生素、紫外线以及群体感应等等都是 AICEs 潜在的调控因素^[3,26], AICEs 的部分调控因子还可以与其它可移动遗传元件相互作用^[27],因此,参与 AICEs 调控的蛋白种类和数量也非常庞大。目前所研究的对 AICEs 起调控作用的蛋白主要有以下几大类: GntR 家族的 KorSA、具有 Nudix 结构的水解酶、具有螺旋-转角-螺旋结构的 DNA 接合蛋白、*kil-kor* 系统、XRE 家族蛋白等,还有一些

与以上几个家族蛋白具有高度同源性或结构相似性的蛋白也参与了 AICEs 的调控。在识别出的 29 个 AICEs 中,基本都携带以上几个家族的调控蛋白(图 2 蓝色)。目前我们没有找到这些调控蛋白的分布规律,尚不能作为 AICEs 的识别依据,这也是 AICEs 调控机制较难研究的原因之一。

2.4 链霉菌 AICE 编码的限制修饰系统

与通常的 ICEs 类似,除了与自行转移相关的核心模块之外, AICEs 还携带其它的重要基因。这些基因还会赋予宿主新性状,如耐药性、重金属抗性、固氮能力等^[13]。这些特性极大地促进了宿主对环境的适应性,增加了宿主在自然选择过程中的竞争力。在链霉菌基因组识别出的 29 个 AICEs 中,我们发现 16 个 AICEs 携带与限制修饰系统相关的基因(图 2 紫色)。我们推测,这些基因的引入对宿主防范外来 DNA 的侵入有一定的作用。*S. colicolor* 中 SLP1 携带的 SCO_4631 就编码一个限制修饰系统相关蛋白 ScoA3McrA^[28]。ScoA3McrA 是一种四型甲基依赖型限制性内切酶,它在体外既可以切割硫修饰的 DNA 也可以切割 DCM-甲基化 DNA,因此,硫修饰的 DNA 不能进入含有 SLP1 的 *S. colicolor*^[28]。相似地,把携带有控制 DNA 硫修饰的基因 *dndA-E* 导入 *S. colicolor* 之后,观察不到菌的生长,因为 *dndA-E* 的表达会使宿主的 DNA 被硫修饰化,SCO_4631 切割硫修饰的染色体导致了宿主的死亡。除了 SLP1,其它 AICEs 携带的限制系统还没有得到实验证实。细菌的限制修饰系统往往位于可移动遗传元件上^[29], AICEs 这种较高频率携带限制系统的现象在链霉菌的进化中具有重要的意义。限制修饰系统是细菌自身的一种保护机制,既可以防范外来 DNA 的入侵,又可以防范外来 DNA 内切酶对自身 DNA 的破坏。在长期的进化过程中,限制系统与修饰系统的共同作用保证了群体在遗传中的相对稳定性。

3 小结

AICEs 是具有典型特征和代表性的可移动遗传元件, 但关于链霉菌中 AICEs 的研究较少。对当前已全测序的链霉菌基因组, 本文采用了隐马尔科夫模型来表征与 AICE 核心模块相关的特征蛋白, 进行 AICEs 的预测。在 12 个链霉菌染色体上共识别了 26 个 AICEs (包括 12 个新的 AICEs), 有 3 个质粒本身也具有 AICEs 的典型特征; 并确定了这些 AICE 的插入位点、正向重复序列及整合、复制和转移等核心模块。比较分析发现, 染色体上的 26 个 AICEs 都位于核心区, 这一分布特征表明可自行移动的 AICEs 是链霉菌基因组多样性的重要影响因素之一。有 16 个 AICEs 携带编码限制系统的基因, 这有利于宿主防范外来 DNA 的入侵, 增加了宿主在自然选择过程中进一步进化的竞争力。随着链霉菌新物种的发现与基因组测序数据的增多, 越来越多具有更多样性特征的 AICEs 将会涌现出来。关于 AICEs 功能研究的逐渐深入, 也将会为链霉菌甚至是其它放线菌的遗传操作提供强大的技术支持, 从而进一步加快通过基因工程手段改良工业生产抗生素菌株的研究进程。

参 考 文 献

- [1] Omura S. The expanded horizon for microbial metabolites—a review[J]. *Gene*, 1992, 115(1/2): 141–149.
- [2] Ikeda H, Ishikawa J, Hanamoto A, et al. Complete genome sequence and comparative analysis of the industrial microorganism *Streptomyces avermitilis*[J]. *Nature Biotechnology*, 2003, 21(5): 526–531.
- [3] Wozniak RA, Waldor MK. Integrative and conjugative elements: mosaic mobile genetic elements enabling dynamic lateral gene flow[J]. *Nature Reviews Microbiology*, 2010, 8(8): 552–563.
- [4] Burrus V, Waldor MK. Shaping bacterial genomes

- with integrative and conjugative elements[J]. *Research in Microbiology*, 2004, 155(5): 376–386.
- [5] Bi D, Xu Z, Harrison EM, et al. ICEberg: a web-based resource for integrative and conjugative elements found in Bacteria[J]. *Nucleic Acids Research*, 2012, 40(Database issue): D621–D626.
- [6] Bi D, Liu L, Tai C, et al. SecReT4: a web-based bacterial type IV secretion system resource[J]. *Nucleic Acids Research*, 2013, 41(D1): D660–D665.
- [7] Christie PJ, Atmakuri K, Krishnamoorthy V, et al. Biogenesis, architecture, and function of bacterial type IV secretion systems[J]. *Annual Review of Microbiology*, 2005, 59: 451–485.
- [8] Sezonov G, Duchêne AM, Friedmann A, et al. Replicase, excisionase, and integrase genes of the *Streptomyces* element pSAM2 constitute an operon positively regulated by the *pra* gene[J]. *Journal of Bacteriology*, 1998, 180(12): 3056–3061.
- [9] Sezonov G, Hagège J, Pernodet JL, et al. Characterization of *pra*, a gene for replication control in pSAM2, the integrating element of *Streptomyces ambofaciens*[J]. *Molecular Microbiology*, 1995, 17(3): 533–544.
- [10] Hagège J, Boccard F, Smokvina T, et al. Identification of a gene encoding the replication initiator protein of the *Streptomyces* integrating element, pSAM2[J]. *Plasmid*, 1994, 31(2): 166–183.
- [11] Hagège J, Pernodet JL, Friedmann A, et al. Mode and origin of replication of pSAM2, a conjugative integrating element of *Streptomyces ambofaciens*[J]. *Molecular Microbiology*, 1993, 10(4): 799–812.
- [12] te Poele EM, Bolhuis H, Dijkhuizen L. Actinomycete integrative and conjugative elements[J]. *Antonie Van Leeuwenhoek*, 2008, 94(1): 127–143.
- [13] Ghinet MG, Bordeleau E, Beaudin J, et al. Uncovering the prevalence and diversity of integrating conjugative elements in actinobacteria[J]. *PLoS One*, 2011, 6(11): e27846.
- [14] Finn RD, Clements J, Eddy SR. HMMER web server: interactive sequence similarity searching[J]. *Nucleic Acids Research*, 2011, 39(Web Server issue): W29–W37.
- [15] Schattner P, Brooks AN, Lowe TM. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs[J]. *Nucleic*

- Acids Research, 2005, 33(Web Server issue): W686–W689.
- [16] Tamura K, Peterson D, Peterson N, et al. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods[J]. *Molecular Biology and Evolution*, 2011, 28(10): 2731–2739.
- [17] Kurtz S, Phillippy A, Delcher AL, et al. Versatile and open software for comparing large genomes[J]. *Genome Biology*, 2004, 5(2): R12.
- [18] Bibb MJ, Ward JM, Kieser T, et al. Excision of chromosomal DNA sequences from *Streptomyces coelicolor* forms a novel family of plasmids detectable in *Streptomyces lividans*[J]. *Molecular & General Genetics: MGG*, 1981, 184(2): 230–240.
- [19] Askora A, Kawasaki T, Fujie M, et al. Resolvase-like serine recombinase mediates integration/excision in the bacteriophage ϕ RSM[J]. *Journal of Bioscience and Bioengineering*, 2011, 111(2): 109–116.
- [20] Williams KP. Integration sites for genetic elements in prokaryotic tRNA and tmRNA genes: sublocation preference of integrase subfamilies[J]. *Nucleic Acids Research*, 2002, 30(4): 866–875.
- [21] Burrus V, Waldor MK. Formation of SXT tandem arrays and SXT-R391 hybrids[J]. *Journal of Bacteriology*, 2004, 186(9): 2636–2645.
- [22] Hochhut B, Beaber JW, Woodgate R, et al. Formation of chromosomal tandem arrays of the SXT element and R391, two conjugative chromosomally integrating elements that share an attachment site[J]. *Journal of Bacteriology*, 2001, 183(4): 1124–1132.
- [23] Pembroke JT, Murphy DB. Isolation and analysis of a circular form of the IncJ conjugative transposon-like elements, R391 and R997: implications for IncJ incompatibility[J]. *FEMS Microbiology Letters*, 2000, 187(2): 133–138.
- [24] Begg KJ, Dewar SJ, Donachie WD. A new *Escherichia coli* cell division gene, *ftsK*[J]. *Journal of Bacteriology*, 1995, 177(21): 6211–6222.
- [25] Wu LJ, Lewis PJ, Allmansberger R, et al. A conjugation-like mechanism for prespore chromosome partitioning during sporulation in *Bacillus subtilis*[J]. *Genes & Development*, 1995, 9(11): 1316–1326.
- [26] Vrijbloed JW, Jelínková M, Hessels GI, et al. Identification of the minimal replicon of plasmid pMEA300 of the methylotrophic actinomycete *Amycolatopsis methanolica*[J]. *Molecular Microbiology*, 1995, 18(1): 21–31.
- [27] Daccord A, Ceccarelli D, Burrus V. Integrating conjugative elements of the SXT/R391 family trigger the excision and drive the mobilization of a new class of Vibrio genomic islands[J]. *Molecular Microbiology*, 2010, 78(3): 576–588.
- [28] Liu G, Ou HY, Wang T, et al. Cleavage of phosphorothioated DNA and methylated DNA by the type IV restriction endonuclease ScoMcrA[J]. *PLoS Genetics*, 2010, 6(12): e1001253.
- [29] Furuta Y, Abe K, Kobayashi I. Genome comparison and context analysis reveals putative Mobile forms of restriction-modification systems and related rearrangements[J]. *Nucleic Acids Research*, 2010, 38(7): 2428–2443.