

引用格式:

李鹏, 刘兰, 王通, 黄筱萍. 链霉菌中一类同源非典型端粒的克隆及分析[J]. 湖南农业大学学报(自然科学版), 2024, 50(4): 42–50.

LI P, LIU L, WANG T, HUANG X P. Cloning and analysis of a class of homologous atypical telomeres of *Streptomyces*[J]. Journal of Hunan Agricultural University(Natural Sciences), 2024, 50(4): 42–50.

投稿网址: <http://xb.hunau.edu.cn>



## 链霉菌中一类同源非典型端粒的克隆及分析

李鹏, 刘兰, 王通, 黄筱萍

(江西省科学院微生物研究所, 江西 南昌 330095)

**摘要:** 采用自连接-PCR 的方法克隆和测序了 *Streptomyces cattleya* DSM46488、*Streptomyces mobaraensis* DSM40847 及 *Streptomyces davawensis* JCM4913 的端粒, 序列比对发现, 这 3 种端粒具有较高的序列相似性, 前 60 个核苷酸的序列相似性达到了 80% 以上; 在二级结构水平, 这些端粒含有多个回文序列, 其中, 包含相同的回文序列 I 和 II, 但并不能通过折返形成超级发夹结构; *Streptomyces cattleya* DSM46488、*Streptomyces mobaraensis* DSM40847 及 *Streptomyces davawensis* JCM4913 的端粒与线性质粒 pSHJG1 端粒及 *Streptomyces albus* J1074 染色体端粒均具有较高的序列相似性; 系统发育分析显示它们与典型端粒及其他非典型端粒处于不同的进化分支, *S. cattleya* DSM46488、*S. mobaraensis* DSM40847、*S. davawensis* JCM4913、*S. albus* J1074 及 *S. hydropiscus* subsp. *Jinggangensis* 5008 都编码同源的潜在末端蛋白-端粒相关蛋白, 暗示了这类同源的非典型端粒系统广泛存在于链霉菌中。

**关键词:** 链霉菌; 线性复制子; 同源非典型端粒; 端粒系统; 回文序列; 末端蛋白-端粒相关蛋白

中图分类号: Q934.3

文献标志码: A

文章编号: 1007-1032(2024)04-0042-09

## Cloning and analysis of a class of homologous atypical telomeres of *Streptomyces*

LI Peng, LIU Lan, WANG Tong, HUANG Xiaoping

(Institute of Microbiology, Jiangxi Academy of Sciences, Nanchang, Jiangxi 330095, China)

**Abstract:** In this study, the telomeres of *Streptomyces cattleya* DSM46488, *S. mobaraensis* DSM40847 and *S. davawensis* JCM4913 were cloned and sequenced by using self-ligation-PCR method. High sequence similarity, especially within the first 60 nucleotides(identity>80%), was revealed by the sequence alignment. At the secondary structural level, these telomeres contained multiple palindromic sequences, including the same palindromic sequences I and II. However, super-hairpin structure could not be formed by folding back of palindrome I with the internal palindromes. Moreover, the three new telomeres cloned from *S. cattleya* DSM46488, *S. mobaraensis* DSM40847 and *S. davawensis* JCM4913 showed sequence similarity to telomeres of linear plasmid pSHJG1 and *S. albus* J1074 chromosome. Phylogenetic analysis based on the first 100 nucleotides of telomeres demonstrated that they were in different phylogenetic branches from the typical telomeres and other atypical telomeres. In addition, these strains encoded the homologous putative terminal proteins-telomere associated proteins, indicating that the class of homologous atypical telomere system was widely distributed within *Streptomyces*.

**Keywords:** *Streptomyces*; linear replicon; homologous atypical telomere; telomere system; palindrome; terminal protein-telomere associated protein

收稿日期: 2023-10-12

修回日期: 2024-08-15

基金项目: 国家自然科学基金项目(31900061); 江西省科学院科研开发专项基金项目(2017-YYB-03)

作者简介: 李鹏(1986—), 男, 江西南昌人, 博士, 助理研究员, 主要从事链霉菌遗传研究, pengli\_jx@126.com

链霉菌属(*Streptomyces*)是一类革兰阳性丝状细菌。与大多数细菌不同,链霉菌包含线性结构的染色体 DNA,并且多数链霉菌还包含大小不一的线性质粒<sup>[1-2]</sup>。链霉菌的线性复制子(包括线性染色体和线性质粒)的 5'末端均含有与末端 DNA 共价结合的末端蛋白<sup>[3-4]</sup>,与噬菌体 phi 29 相似。与噬菌体 phi 29 的复制不同<sup>[5]</sup>,链霉菌的线性复制子从位于复制子中央的起始位点起始复制,并以双向复制的方式进行延伸,但当复制延伸至 DNA 末端时,需补平滞后链中最后 1 个冈崎片段上 RNA 引物被去除后留下的 3'端单链区<sup>[6]</sup>。HUANG 等<sup>[7]</sup>对链霉菌线性复制子端粒进行克隆与分析,发现这些末端 DNA 都含有保守的回文序列,并且回文序列之间配对形成更复杂的折返结构,这种保守的端粒也被称为典型端粒。QIN 等<sup>[8]</sup>分析了线性质粒 pSLA2 维持线性复制必需的端粒 DNA 序列,证明了折返结构在典型端粒 DNA 复制过程中的重要性。

随着越来越多的链霉菌线性复制子的端粒被克隆,发现并不是所有的链霉菌端粒 DNA 都具有这种保守的回文序列和折返结构,且较多的链霉菌的线性复制子包含全新的端粒,并将这种端粒统称为非典型端粒,如 *S. griseus* 的线性染色体<sup>[9]</sup>、线性质粒 SCP1<sup>[10]</sup>、线性质粒 pRL1 和 pRL2<sup>[11]</sup>、线性质粒 pSHJG1<sup>[12]</sup>、线性质粒 pLRS1366S<sup>[13]</sup>等。一方面,非典型端粒与典型端粒的序列差异较大;另一方面,虽然非典型端粒中也存在回文序列,但是往往不会形成类似典型端粒中的折返结构。有研究<sup>[14-15]</sup>表明,含有典型端粒的链霉菌都编码非常保守的末端蛋白,非典型端粒 DNA 的出现预示着新型的末端蛋白的存在。可见,链霉菌线性复制子的端粒系统即端粒 DNA 和末端蛋白-端粒相关蛋白存在着多样性;因此,对链霉菌线性复制子端粒的克隆将有助于发现更多的非典型端粒,以分离鉴定新的末端蛋白,揭示链霉菌线性复制子端粒的复制机制。

已报道的非典型端粒的序列各异,但同源非典型端粒还鲜有报道。本研究克隆了 *S. cattleya* DSM46488 的线性染色体和巨型质粒的端粒,在多种链霉菌中识别到了 DSM46488 巨型质粒端粒复制的末端蛋白-端粒相关蛋白的同源蛋白,基于端粒序列与末端蛋白-端粒相关蛋白之间共同进化的关系,猜想这类链霉菌应该包含同源的非典型端粒。

为了验证这种猜想,还克隆了另外 2 株编码同源潜在末端蛋白-端粒相关蛋白的链霉菌 *S. mobaraensis* DSM40847 和 *S. davawensis* JCM4913 的端粒,旨在挖掘同源非典型端粒,为非典型端粒系统的研究提供依据。

1 材料与方法

1.1 菌株、引物及主要试剂

*S. cattleya* DSM46488 及大肠埃希菌 TOP10 由江西省科学院微生物研究所保存;*S. mobaraensis* DSM40847 由江西师范大学生命科学学院杨慧林馈赠;*S. davawensis* JCM4913 购于日本微生物菌种保藏中心。本研究所用引物见表 1。生化试剂、常规抗生素购于生工生物工程(上海)股份有限公司;I5 高保真聚合酶购于北京擎科生物科技有限公司;限制性内切酶及 T4 DNA 连接酶购于 New England Biolabs 公司;胶回收试剂盒购于南京诺唯赞生物科技有限公司。

表 1 本研究中使用的引物

Table 1 Primers used in the study

引物	序列(5'-3')
cat-cF	AGAACTACCGGTACTCCACCA
cat-cR	CGGAGCGTCATGTCTGAGAG
cat-pF	CTGACCAGTGCTCCGCTTTC
cat-pR	TGTCTTTTTGCGGTGTGCTG
mot-p1F	AGGGTGAGTCGGACCTGC
mot-p1R	TGTGAGGCCTGAACAGG
mot-p2F	AGTCGCAAGTTCACGCTACC
mot-p2R	GAGAAAGGGCTACGGGTGAC
dav-pF	CGCCCTTTCGGACACTTCTAAC
dav-pR	TGCCTGGATCGTGCAGTACAC

1.2 方法

1.2.1 链霉菌培养及基本分子生物学操作

链霉菌的培养及基因组 DNA 的提取参考文献[16];大肠埃希菌 TOP10 的培养、转化及质粒提取等基本操作参考文献[17]。

1.2.2 链霉菌端粒的克隆

参照文献[18]克隆链霉菌线性复制子的端粒。首先提取 *S. cattleya* DSM46488 总 DNA,在提取过程中避免使用蛋白酶 K,以防止末端蛋白的降解。基于线性染色体的序列,在邻近末端位置挑选平端

内切酶 *Nru*I, 假定未知的末端序列不包含该酶的识别位点, 用 *Nru*I 对 15  $\mu$ g 基因组 DNA 进行完全酶切, 经过异丙醇沉淀及 2 次 70% 乙醇洗涤后, 用 90  $\mu$ L ddH<sub>2</sub>O 溶解 DNA。然后加入 10  $\mu$ L 1 mol/L NaOH, 在 37  $^{\circ}$ C 水浴锅中处理 1 h, 以达到降解末端蛋白的目的。加 2 mol/L HCl 和 1 mol/L 的 Tris-HCl 的混合液(pH 8.0, 终浓度为 0.1 mol/L)以中和 NaOH。最后, 加入 2 $\times$ SSC, 在 68  $^{\circ}$ C 下处理 1 h, 使单链 DNA 复性成双链 DNA。经过纯化, 溶解于 20  $\mu$ L ddH<sub>2</sub>O 中。配制 T4 DNA 连接酶的反应体系, 16  $^{\circ}$ C 下反应 12 h。根据已知 DNA 序列设计特异性引物(表 1), 以自连接 DNA 为模板进行 PCR 扩增反应。经过胶回收、克隆及测序, 最终获得 *S. cattleya* DSM46488 的端粒序列。

### 1.2.3 *S. mobaraensis* DSM40847 全基因组测序、组装及注释

全基因组序列测定由北京诺禾致源科技股份有限公司通过 PacBIO 及 Illumina NovaSeq PE150 测序平台完成。首先, 构建 10 kb 单分子实时测序文库, 产生 61 504 条平均读长为 7 947 bp 的 reads, 组装成 1 条完整的线性 contig, 然后与 Illumina 测序数据进行比对, 进而对组装结果进行校正。利用 GeneMarks 预测编码基因, 并对编码基因进行 COG 分类。利用 RNAmmer 及 tRNAscan-SE 分别预测 rRNA 及 tRNA<sup>[19-20]</sup>。利用 MUMmer 对 *S. mobaraensis* 及其他已测序链霉菌进行共线性分析<sup>[21]</sup>。将 *S. mobaraensis* DSM40847 的全基因组序列上传至 NCBI GenBank 数据库, 登录号为 CP072827。

### 1.2.4 DNA 序列的测定和分析

DNA 测序由生工生物工程(上海)股份有限公司完成。运用在线工具 Clustal Omega(<https://www.ebi.ac.uk/Tools/msa/clustalo/>)<sup>[22]</sup> 比对 DNA 序列, 并运用在线工具 ESPript3([esprict.ibcp.fr/ESPript/ESPript/index.php](http://esprict.ibcp.fr/ESPript/ESPript/index.php))对 DNA 序列比对结果进行呈现<sup>[23]</sup>。采用 mfold 在线工具([http://www.unafold.org/DNA\\_form.php](http://www.unafold.org/DNA_form.php))分析 DNA 的二级结构<sup>[24]</sup>。运用软件 MEGA11<sup>[25]</sup>构建进化树, 并利用在线工具 iTOL<sup>[26]</sup>呈现进化树。*S. cattleya* DSM46488 染色体和巨型质

粒的端粒序列编号分别为 OP903022 及 OP903021, *S. mobaraensis* DSM40847 全基因组序列编号为 CP072827。

## 2 结果与分析

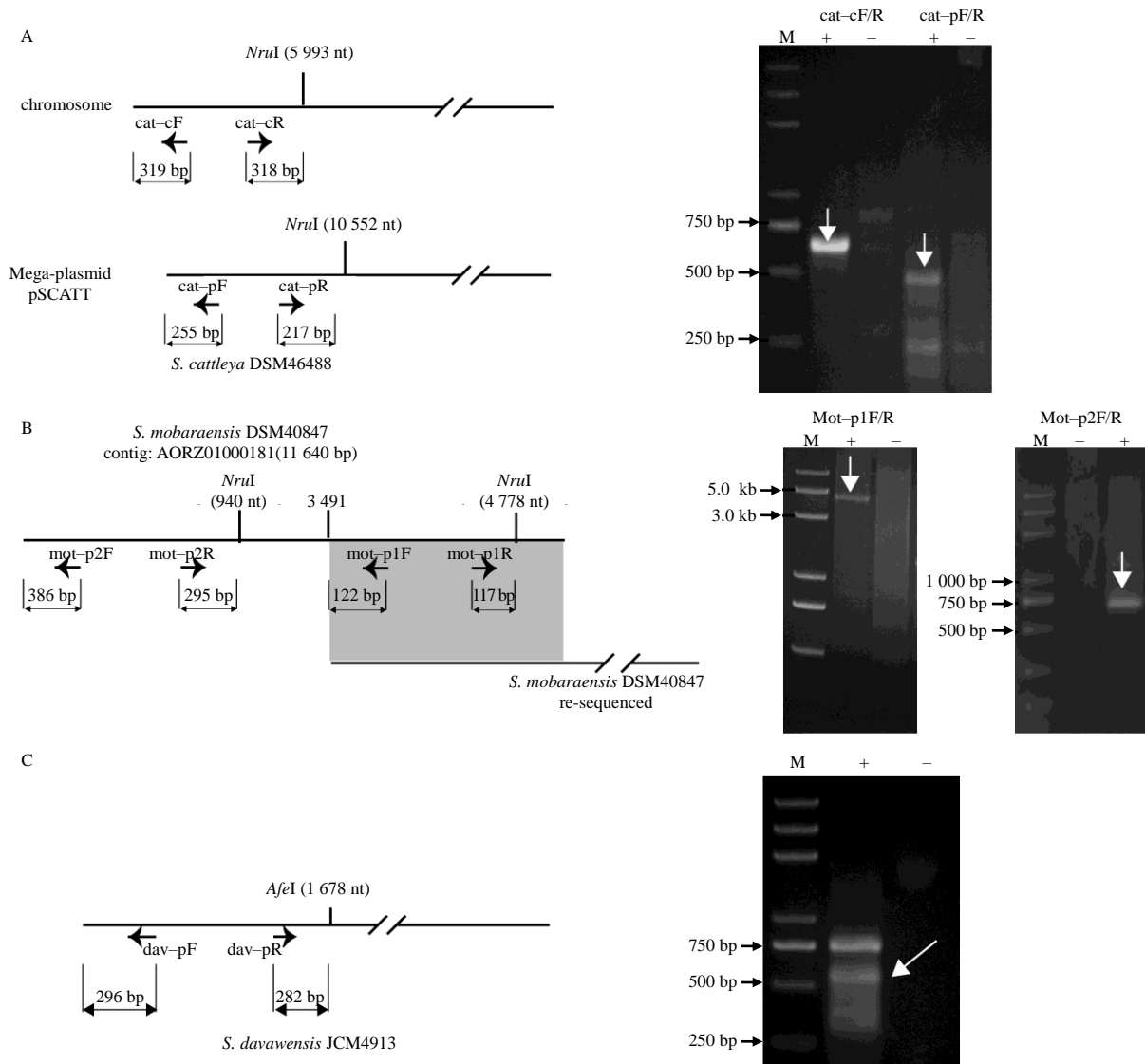
### 2.1 *S. cattleya* DSM46488 端粒的克隆及分析

基于 *S. cattleya* DSM46488 的染色体和巨型质粒的序列信息, 在邻近染色体末端 6 kb 及巨型质粒末端 10 kb 左右均有限制性内切酶 *Nru*I 的酶切位点(TCG↓CGA), 按照 1.2.2 所述方法, 经过 *Nru*I 酶解、去除末端蛋白、复性, 在 T4 DNA 连接酶的作用下, 形成自连接 DNA。以此 DNA 为模板, 根据已知序列设计特异性引物 cat-cF/cat-cR、cat-pF/cat-pR, 分别扩增包含染色体和巨型质粒端粒的 DNA 片段(图 1-A), 获得了目的条带; 经过克隆及测序, 最终获得 *S. cattleya* DSM46488 的染色体和巨型质粒的端粒序列。作为阴性对照, 以不加 T4 DNA 连接酶的 DNA 为模板, 以相同的引物在相同的条件下进行 PCR 扩增, 没有产生相应的条带(图 1-A)。测序结果显示, *S. cattleya* DSM46488 的线性染色体和线性巨型质粒具有相同的端粒序列, 为方便描述, 命名为 cpTelo。在核苷酸水平, cpTelo 与典型端粒无序列相似性(图 2); 在二级结构水平, cpTelo 也包含多个回文序列, 形成三核苷酸突出环, 但回文序列之间不能折返形成超级发夹结构。

根据端粒序列与末端蛋白之间的共同进化关系, 推测 *S. cattleya* DSM46488 可能会编码新的末端蛋白。基于此推测, 在 *S. cattleya* DSM46488 的巨型质粒中识别并初步鉴定到与 cpTelo 复制相关的新末端蛋白-端粒相关蛋白 SCATT\_p03450-p03440。通过 BLASTp 比对, 发现该末端蛋白及端粒相关蛋白的同源蛋白存在于 *S. albus* J1074、*S. hygroscopicus* subsp. *Jinggangensis* 5008 线性质粒 pSHJG1、*S. mobaraensis* DSM40847 及 *S. davawensis* JCM4913 链霉菌中(图 3), 序列相似性>55%。其中, *S. albus* J1074、*S. hygroscopicus* subsp. *Jinggangensis* 5008 线性质粒 pSHJG1 的端粒已经被克隆, 这 2 个端粒与 *S. cattleya* DSM46488 端粒 cpTelo 存在一定的序列相似性(二者前 100 个核苷酸与 cpTelo 的相似性分别为 49%、60%), 尤其是 5'端前 60 个核苷酸的序列相

似性大于 80%，推测这些链霉菌可能包含同源的非常端粒系统。为了验证这个推测，分别克隆了 *S.*

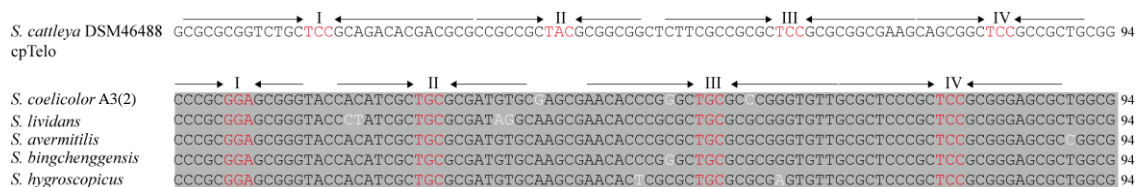
*mobaraensis* DSM40847 及 *S. davawensis* JCM4913 的端粒。



M 为 2 kb plus DNA ladder; +/- 为加/未加 T4 DNA 连接酶; 白色箭头示预期目标条带。A 为 *S. cattleya* DSM46488 染色体和巨型质粒端粒克隆; B 为 *S. mobaraensis* DSM40847 染色体端粒的克隆; C 为 *S. davawensis* JCM4913 染色体端粒的克隆。

图 1 自连接方法克隆链霉菌端粒

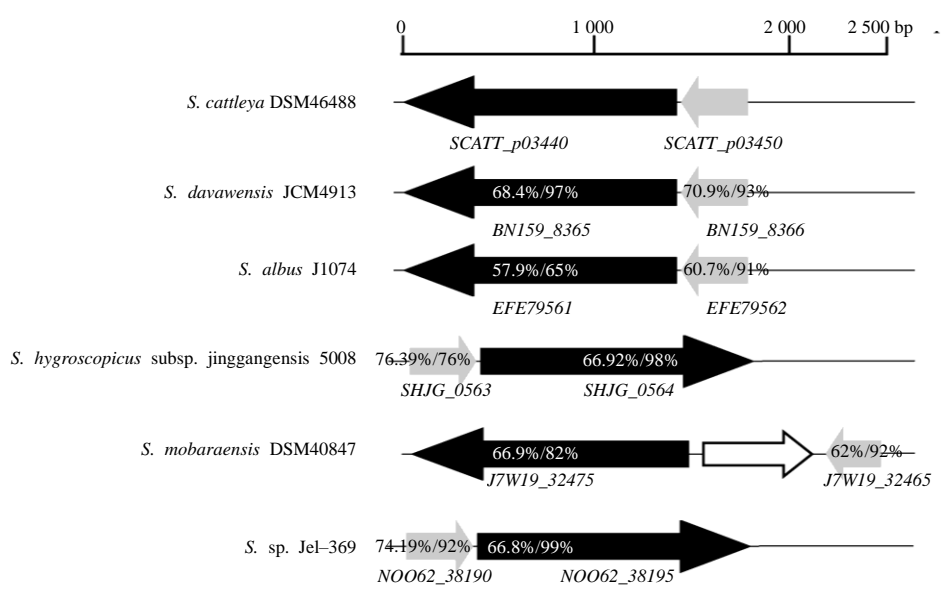
Fig.1 Cloning of *Streptomyces* telomeres by self-ligation method



箭头表示回文序列, 红色核苷酸表示 3 碱基 loop。

图 2 *Streptomyces cattleya* DSM46488 端粒 cpTelo 与典型端粒序列比对结果

Fig.2 Sequence alignments of *Streptomyces cattleya* DSM46488 telomere cpTelo against *Streptomyces* typical telomeres



第 1 个百分数表示氨基酸序列一致性；第 2 个百分数表示序列覆盖度。

图 3 SCATT\_p03440-p03450 在链霉菌基因组中的分布

Fig. 3 Distribution of SCATT\_p03440-p03450 among the completely sequenced *Streptomyces*

2.2 *S. mobaraensis* DSM40847 全基因组序列的测定及端粒克隆

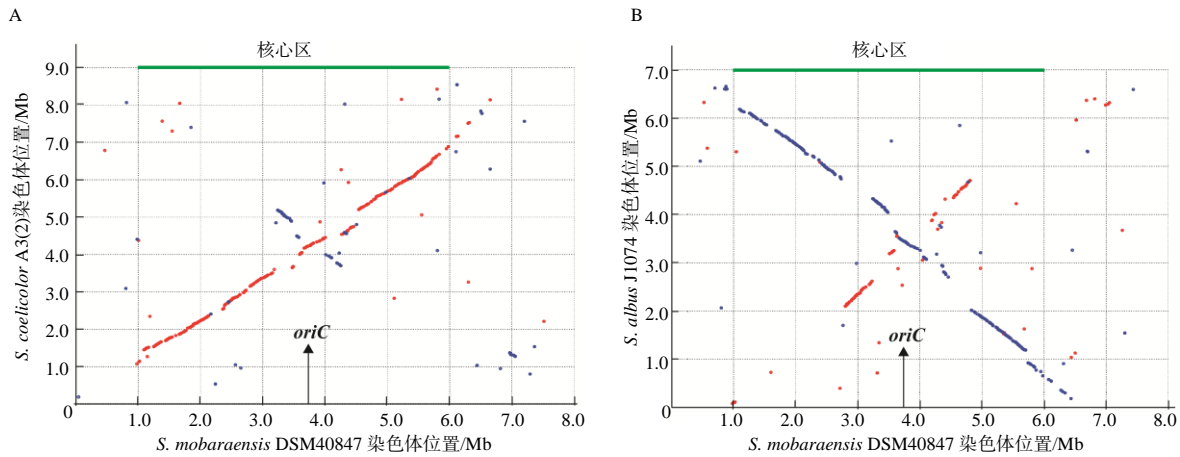
为克隆 *S. mobaraensis* DSM40847 的端粒，重新测定了该菌株的全基因组序列，获得了 1 条完整的线性染色体，大小为 7 574 085 bp。染色体的基

本特征如表 2 所示。该染色体包含 6 451 个编码基因，BLASTp 比对显示 2 018 个基因编码假定蛋白。复制起始区 *oriC*(3 776 948 ~ 3 777 818 nt)在 *dnaA* 及 *dnaN* 基因的中间，位于接近染色体中央的位置(图 4)。另外，*S. mobaraensis* DSM40847 的染色体

表 2 *Streptomyces mobaraensis* DSM40847 染色体的基本特征

Table 2 General features of *Streptomyces mobaraensis* DSM40847 chromosome

菌株	拓扑结构	染色体长度/bp	G+C 含量/%	开放阅读框数量/个	ORF 平均长度/bp	编码密度/%	rRNA 数量/个	tRNA 数量/个	参考文献
<i>S. mobaraensis</i> DSM40847	线性	7 574 085	73.2	6 451	1 017	86.6	7	66	
<i>S. albus</i> J1074	线性	6 841 649	73.3	5 832	1 011	86.8	7	66	[27]
<i>S. coelicolor</i> A3(2)	线性	8 667 507	72.1	7 825	991	88.9	6	63	[28]



绿色线及 *oriC* 分别表示 *S. mobaraensis* DSM40847 染色体的核心区及复制起始区。A 为 *S. mobaraensis* DSM40847 染色体与 *S. coelicolor* A3(2)染色体的序列比对；B 为 *S. mobaraensis* DSM40847 染色体与 *S. albus* J1074 染色体的序列比对。

图 4 3 种链霉菌染色体的结构比对

Fig.4 Comparison of chromosome structure for *Streptomyces mobaraensis* versus *S. coelicolor* and *S. mobarsenis* versus *S. albus*

还包含 66 个 tRNA 编码基因。与大多数链霉菌不同的是, *S. mobaraensis* DSM40847 包含 7 个 rRNA 操纵子, 类似的情况只在 *S. albus* J1074 中报道过, 而 *S. albus* J1074 作为 1 种通用的异源表达宿主, 具有作为链霉菌底盘的潜力<sup>[27]</sup>。 *S. mobaraensis* DSM40847 是否也有相似的通用异源表达宿主的特性还有待进一步研究。与其他链霉菌一样, *S. mobaraensis* DSM40847 的染色体也包含明显的核心区(1.0 ~ 6.0 Mb)和两臂区的结构(图 4)。

通过与现有的 *S. mobaraensis* DSM40847 (AORZ01000000)的序列进行比对, 181 号 contig 匹配到了线性染色体的末端, 显示线性染色体的末端尚有 3.5 kb 左右的缺口(图 1-B)。为了克隆染色体的端粒, 利用平端内切酶 *NruI* 处理基因组 DNA, 该酶在线性染色体的末端 900 bp 位置有 1 个酶切位点, 而在 181 号 contig 中还存在另 1 个 *NruI* 的酶切位点(图 1-B); 因此, 设计了 2 对特异性引物 mot-p1F/R 及 mot-p2F/R, 分别位于 2 个 *NruI* 的识别位点的邻近位置, 如图 1-B 所示。当利用 mot-p1F/R 为扩增引物时, 获得了 1 条 3 kb 左右的目的条带, 其长度与 2 个 *NruI* 的识别位点中间条带的长度一致。测序结果显示序列与预期的一致。以 mot-p2F/R 为扩增引物时, 也获得了特异性目的条带(图 1-B), 阴性对照中均无任何条带产生。经过克隆及测序, 最终获得了 *S. mobaraensis* DSM40847 线性染色体的端粒 mobTelo(图 5)。

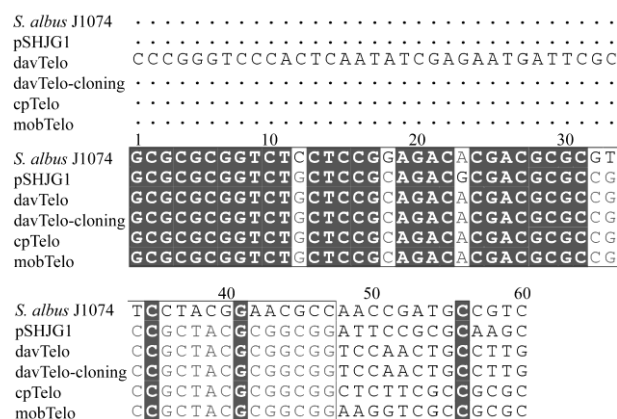


图 5 链霉菌同源非典型端粒序列比对

Fig.5 Sequence alignments of homologous atypical telomeres of *Streptomyces*

### 2.3 *S. davawensis* JCM4913 染色体端粒的克隆

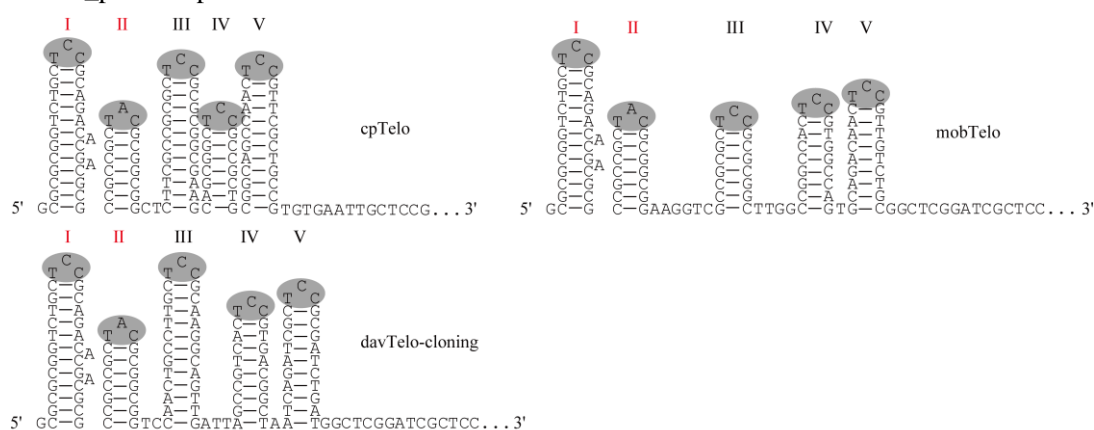
BLASTn 比对结果表明 *S. davawensis* JCM4913

的末端序列 *davTelo* 在 5'端比 *cpTelo* 及 *mobTelo* 多了 33 个碱基(图 5)。为了验证该末端序列的正确性, 利用平端内切酶 *AfeI* 并根据已知序列设计特异性引物 *dav-pF/R*, 以自连接 DNA 为模板, 获得了条带明显的 DNA 片段(图 1-C), 经过胶回收及克隆测序, 最终获得了 *S. davawensis* JCM4913 线性染色体的端粒序列, 命名为 *davTelo-cloning*。与原始的端粒序列相比, 重新克隆获得的端粒 *davTelo-cloning* 并不包含原始序列中 5'端的 33 个核苷酸(图 5)。

### 2.4 同源非典型端粒序列分析

已报道的链霉菌端粒, 不论是典型端粒还是非典型端粒, 均含有多个回文序列。虽然在核苷酸水平, 端粒 *cpTelo*、*mobTelo* 及 *davTelo-cloning* 均没有典型端粒末端 13 个核苷酸序列, 即回文序列 I, 但 *cpTelo*、*mobTelo* 及 *davTelo-cloning* 均含有多个回文序列(图 6), 并形成 TCC 及 TAC 2 种不同的三核苷酸突出环(图 6)。这些端粒的回文序列 I 均不能与回文序列 III 或 IV 配对, 形成超级发夹结构, 说明折返结构并不是端粒复制的必需结构。通过序列比对发现, *cpTelo*、*mobTelo*、*davTelo-cloning*、*S. hygroscopicus* subsp. *Jinggangensis* 5008 中线性质粒 pSHJG1 及 *S. albus* J1074 的端粒存在较高的序列相似性, 尤其是前 60 个核苷酸更是如此(图 5), 在二级结构水平, 包含相同的回文序列 I 及 II。提取端粒的前 100 个核苷酸序列进行进化分析, 结果显示非典型端粒与典型端粒位于 2 个不同的分支(图 7), 在非典型端粒分支中, *cpTelo*、*mobTelo*、*davTelo-cloning*、*S. hygroscopicus* subsp. *Jinggangensis* 5008 中线性质粒 pSHJG1、*S. albus* J1074 及 *S. sp.* Je1-369 的端粒聚类构成 1 个亚分支(图 7), 暗示它们组成一类同源的非典型端粒序列。此外, 在 *S. mobaraensis* DSM40847、*S. davawensis* JCM4913、*S. hygroscopicus* subsp. *Jinggangensis* 5008、*S. albus* J1074 及 *S. sp.* Je1-369 基因组中均识别到 *S. cattleya* DSM46488 中 *cpTelo* 复制必需的末端蛋白及端粒相关蛋白 SCATT\_p03450-p03440 的同源蛋白(图 3), 暗示 *S. mobaraensis* DSM40847、*S. davawensis* JCM4913、*S. hygroscopicus* subsp. *Jinggangensis* 5008、*S. albus* J1074 及 *S. sp.* Je1-369 均编码与 *S. cattleya* DSM46488 端粒系统同源的非典型端粒系统

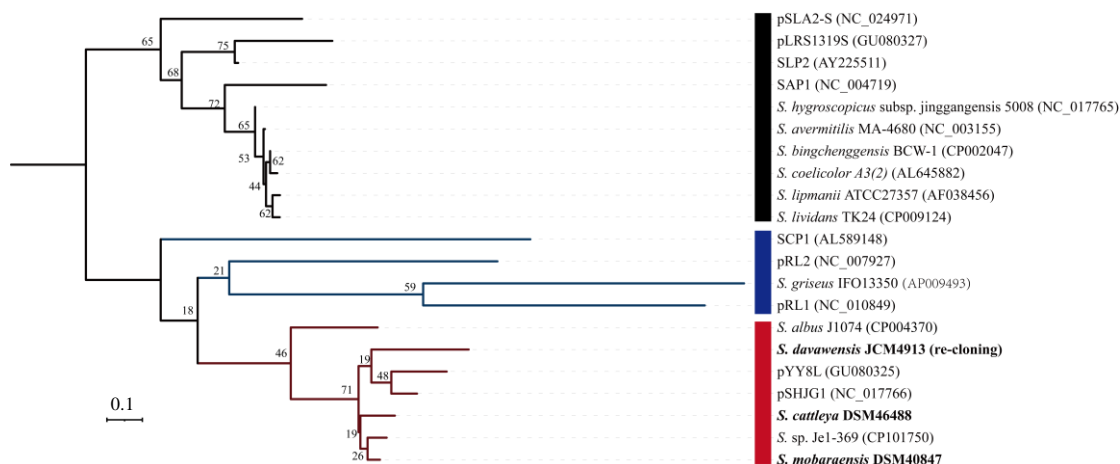
cpTelo-SCATT\_p03450-p03440。



I、II、III、IV 及 V 表示回文序列；红色标记的为回文序列 I 及 II；灰色标记的核苷酸表示 3 碱基 loop。

图 6 端粒 cpTelo、mobTelo 及 davTelo-cloning 的二级结构预测

Fig.6 Putative secondary structure of telomeres cpTelo, mobTelo and davTelo-cloning



黑色、蓝色及红色分别表示典型端粒、非同源的非典型端粒及同源非典型端粒的进化分支。

图 7 链霉菌端粒序列进化分析结果

Fig.7 Phylogenetic analysis based on the first 100 nucleotides of *Streptomyces* telomeres

### 3 结论与讨论

端粒复制是链霉菌线性复制子复制中的重要组成部分，保证了链霉菌线性复制子的完整性，而端粒序列决定了端粒复制所需的末端蛋白及端粒相关蛋白。研究<sup>[7]</sup>发现，许多链霉菌线性复制子包含保守的端粒序列，这些端粒都富含回文序列，回文序列 I 与回文序列 IV 配对形成超级发卡结构，且回文序列含有 3-nt 的突出环。也有研究<sup>[9-13]</sup>发现，链霉菌线性复制子包含非典型端粒，如 *S. griseus* IFO13350 染色体、线性质粒 SCP1、pRL1、pRL2、pSHJG1 等。但目前有关同源的非典型端粒的相关报道较少。

本研究利用自连接-PCR 方法成功克隆了 *S.*

*cattleya* DSM46488 的端粒 cpTelo，通过与典型端粒的一级结构及二级结构比对，cpTelo 与典型端粒无序列相似性，与典型端粒一样，cpTelo 也包含多个回文序列，并形成 3-nt 的突出环。然而回文序列之间不能折返形成超级发卡结构，进一步佐证了折返结构不是端粒复制，尤其是非典型端粒复制的必需结构。基于蛋白序列相似性，BLASTp 比对发现，与 cpTelo 复制相关的末端蛋白-端粒相关蛋白(TP-TAP)的同源蛋白广泛存在于其他链霉菌中，包括 *S. albus* J1074、*S. hygroscopicus* subsp. *Jinggangensis* 5008、*S. sp.* Je1-369、*S. mobaraensis* DSM40847 及 *S. davawensis* JCM4913 等，其中 *S. albus* J1074 及 *S. hygroscopicus* subsp. *Jinggangensis* 5008 的线性质粒 pSHJG1 的端粒已经被克隆。*S. albus* J1074 和线



性质粒 pSHJG1 的端粒与 cpTelo 具有较高的序列相似性, 尤其是前 60 个核苷酸更是如此, 而 *S. davawensis* JCM4913 的末端序列除了 5'突出的 33 个核苷酸外, 与 cpTelo 的前 60 个核苷酸也具有很高的相似性。根据端粒序列与末端蛋白-端粒相关蛋白之间相互适应并共同进化的关系, 推测这类链霉菌应该都包含与 cpTelo 相似的端粒。对 *S. mobaraensis* DSM40847 端粒的克隆及对 *S. davawensis* JCM4913 端粒的重新克隆结果显示, mobTelo、davTelo-cloning 均与 cpTelo 高度相似, 尤其是前 60 个核苷酸更是如此, 进一步验证了上述推测。除了上述链霉菌外, 以 cpTelo 为查询序列, 通过 BLASTn 搜寻其相似序列, 结果在许多链霉菌的线性染色体中找到了相似序列(均位于序列的末端, 尚不能确定是不是真正的端粒), 相似性也主要集中在前 60 个核苷酸。推测 *S. cattleya* DSM46488 端粒系统 cpTelo-SCATT\_p03450-p03440 可能代表着一类广泛存在于链霉菌的同源非典型端粒系统。初步验证了 cpTelo 端粒复制所需的末端蛋白及端粒相关蛋白 SCATT\_p03450-p03440, 但还需要进一步试验来鉴定其同源蛋白是否也参与了 cpTelo 相似端粒的复制。此外, 由于这类同源端粒序列的相似性主要集中于前 60 个核苷酸, 在二级结构水平包含相同的回文序列 I 及 II, 暗示在这类非典型端粒复制过程中回文序列 I 及 II 的重要性。

综上所述, 本研究首先克隆了 *S. cattleya* DSM46488 的端粒 cpTelo, 在 *S. mobaraensis* DSM40847 及 *S. davawensis* JCM4913 等多个链霉菌基因组中识别到了与 cpTelo 复制相关的末端蛋白及端粒相关蛋白 SCATT\_p03450-p03440 的同源蛋白, 根据端粒序列与末端蛋白-端粒相关蛋白间的趋同性, 推测这类链霉菌包含同源的非典型端粒, 对 *S. mobaraensis* DSM40847 及 *S. davawensis* JCM4913 端粒进行克隆及测序, 结果证实了该推测, 暗示 *S. cattleya* DSM46488 的端粒系统 cpTelo-SCATT\_p03450-p03440 代表了一类广泛存在于链霉菌中的同源非典型端粒系统。

#### 参考文献:

[1] LIN Y S, KIESER H M, HOPWOOD D A, et al. The chromosomal DNA of *Streptomyces lividans* 66 is linear[J]. *Molecular Microbiology*, 1994, 14(5): 1103.

- [2] HIROCHIKA H, SAKAGUCHI K. Analysis of linear plasmids isolated from *Streptomyces*: association of protein with the ends of the plasmid DNA[J]. *Plasmid*, 1982, 7(1): 59-65.
- [3] BAO K, COHEN S N. Terminal proteins essential for the replication of linear plasmids and chromosomes in *Streptomyces*[J]. *Genes & Development*, 2001, 15(12): 1518-1527.
- [4] BAO K, COHEN S N. Recruitment of terminal protein to the ends of *Streptomyces* linear plasmids and chromosomes by a novel telomere-binding protein essential for linear DNA replication[J]. *Genes & Development*, 2003, 17(6): 774-785.
- [5] BLANCO L, BERNAD A, LÁZARO J M, et al. Highly efficient DNA synthesis by the phage phi 29 DNA polymerase:symmetrical mode of DNA replication[J]. *The Journal of Biological Chemistry*, 1989, 264(15): 8935-8940.
- [6] CHANG P C, KIM E S, COHEN S N. *Streptomyces* linear plasmids that contain a phage-like, centrally located, replication origin[J]. *Molecular Microbiology*, 1996, 22(5): 789-800.
- [7] HUANG C H, LIN Y S, YANG Y L, et al. The telomeres of *Streptomyces* chromosomes contain conserved palindromic sequences with potential to form complex secondary structures[J]. *Molecular Microbiology*, 1998, 28(5): 905-916.
- [8] QIN Z, COHEN S N. Replication at the telomeres of the *Streptomyces* linear plasmid pSLA2[J]. *Molecular Microbiology*, 1998, 28(5): 893-903.
- [9] GOSHI K, UCHIDA T, LEZHAVA A, et al. Cloning and analysis of the telomere and terminal inverted repeat of the linear chromosome of *Streptomyces griseus*[J]. *Journal of Bacteriology*, 2002, 184(12): 3411-3415.
- [10] KINASHI H, SHIMAJI-MURAYAMA M, HANAFUSA T. Nucleotide sequence analysis of the unusually long terminal inverted repeats of a giant linear plasmid, SCP1[J]. *Plasmid*, 1991, 26(2): 123-130.
- [11] ZHANG R, YANG Y, FANG P, et al. Diversity of telomere palindromic sequences and replication genes among *Streptomyces* linear plasmids[J]. *Applied and Environmental Microbiology*, 2006, 72(9): 5728-5733.
- [12] WU H, QU S, LU C Y, et al. Genomic and transcriptomic insights into the thermo-regulated biosynthesis of validamycin in *Streptomyces hygroscopicus* 5008[J]. *BMC Genomics*, 2012, 13: 337.
- [13] 杨勇, 代玉梅, 陈振华, 等. 五个链霉菌线型质粒端粒的克隆和分析[J]. *微生物学报*, 2010, 50(8): 1008-1013.
- [14] HUANG C H, TSAI H H, TSAY Y G, et al. The telomere



- system of the *Streptomyces* linear plasmid SCP1 represents a novel class[J]. *Molecular Microbiology*, 2007, 63(6): 1710–1718.
- [15] SUZUKI H, MARUSHIMA K, OHNISHI Y, et al. A novel pair of terminal protein and telomere-associated protein for replication of the linear chromosome of *Streptomyces griseus* IFO13350[J]. *Bioscience, Biotechnology, and Biochemistry*, 2008, 72(11): 2973–2980.
- [16] KIESER T, BIBB M J, CHATER K F, et al. *Practical Streptomyces Genetics*[M]. Norwich: The John Innes Foundation, 2000.
- [17] SAMBROOK J, RUSSELL D W. *Molecular Cloning: A Laboratory Manual*[M]. 3rd ed. New York: Cold Spring Harbor Laboratory Press, 2001.
- [18] FAN Y, DAI Y M, CHENG Q X, et al. A self-ligation method for PCR-sequencing the telomeres of *Streptomyces* and *Mycobacterium* linear replicons[J]. *Journal of Microbiological Methods*, 2012, 90(2): 105–107.
- [19] LAGESSEN K, HALLIN P, RØDLAND E A, et al. RNAmmer: consistent and rapid annotation of ribosomal RNA genes[J]. *Nucleic Acids Research*, 2007, 35(9): 3100–3108.
- [20] CHAN P P, LIN B Y, MAK A J, et al. tRNAscan-SE 2.0: improved detection and functional classification of transfer RNA genes[J]. *Nucleic Acids Research*, 2021, 49(16): 9077–9096.
- [21] KURTZ S, PHILLIPPY A, DELCHER A L, et al. Versatile and open software for comparing large genomes[J]. *Genome Biology*, 2004, 5(2): R12.
- [22] MADEIRA F, PEARCE M, TIVEY A R N, et al. Search and sequence analysis tools services from EMBL-EBI in 2022[J]. *Nucleic Acids Research*, 2022, 50(W1): W276–W279.
- [23] ROBERT X, GOUET P. Deciphering key features in protein structures with the new ENDscript server[J]. *Nucleic Acids Research*, 2014, 42(Web Server issue): W320–W324.
- [24] ZUKER M. Mfold web server for nucleic acid folding and hybridization prediction[J]. *Nucleic Acids Research*, 2003, 31(13): 3406–3415.
- [25] TAMURA K, STECHER G, KUMAR S. MEGA11: molecular evolutionary genetics analysis version 11[J]. *Molecular Biology and Evolution*, 2021, 38(7): 3022–3027.
- [26] LETUNIC I, BORK P. Interactive Tree of Life(iTOL) v5: an online tool for phylogenetic tree display and annotation[J]. *Nucleic Acids Research*, 2021, 49(W1): W293–W296.
- [27] ZABURANNYI N, RABYK M, OSTASH B, et al. Insights into naturally minimised *Streptomyces albus* J1074 genome[J]. *BMC Genomics*, 2014, 15: 97.
- [28] BENTLEY S D, CHATER K F, CERDEÑO-TÁRRAGA A M, et al. Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2)[J]. *Nature*, 2002, 417: 141–147.

责任编辑: 毛友纯  
英文编辑: 柳 正