

基因簇大片段克隆技术研究进展及挑战

高花花, 胡娟, 刘玲丽*

西北农林科技大学化学与药学院 天然产物化学生物学重点实验室, 陕西 杨凌 712100

高花花, 胡娟, 刘玲丽. 基因簇大片段克隆技术研究进展及挑战[J]. 微生物学通报, 2023, 50(1): 351-367.

GAO Huahua, HU Juan, LIU Lingli. Research progress and challenges in cloning large gene clusters[J]. Microbiology China, 2023, 50(1): 351-367.

摘要: 天然产物结构复杂、活性多样, 是新药开发的重要来源, 对天然产物生物合成途径的研究, 有利于探索酶催化的合成机制, 促进复杂天然产物的应用。天然产物的生物合成由其对应的基因簇调控, 其中大量天然产物生物合成基因簇(biosynthetic gene clusters, BGCs)在野生型菌株中无法表达或表达量低。对这些基因簇的研究, 需要进行克隆表达, 而如何克隆大片段基因簇并使其表达, 从而发现新型天然产物是一个具有挑战性的问题。其中构建基因组文库、转化关联重组(transformation-associated recombination, TAR)、Red/ET重组等是克隆大片段基因簇的重要技术。本文从克隆技术的策略和应用两个方面, 总结了这3种克隆技术目前的研究进展, 讨论了目前大片段基因簇克隆技术面临的挑战, 为研究大片段基因簇提供方法学借鉴。

关键词: 天然产物; 生物合成基因簇; 基因组文库; 转化关联重组; Red/ET; 基因簇异源表达

Research progress and challenges in cloning large gene clusters

GAO Huahua, HU Juan, LIU Lingli*

Key Laboratory of Natural Products Chemistry and Biology, College of Chemistry & Pharmacy, Northwest A&F University, Yangling 712100, Shaanxi, China

Abstract: Natural products, which feature complex structures and diverse activities, are important sources of new drugs. The research on the biosynthesis pathways of natural products is conducive to exploring the synthesis mechanism catalyzed by enzymes and promoting the application of complex natural products. The corresponding gene clusters regulate the biosynthesis of natural products, and a large number of biosynthetic gene clusters (BGCs) of natural products are not expressed or expressed at low levels in wild-type strains. Thus, these gene clusters are usually

资助项目: 国家重点研发计划(2018YFA0903200); 陕西省自然科学基金青年项目(2022JQ-142)

This work was supported by the National Key Research and Development Program of China (2018YFA0903200) and the Shaanxi Provincial Natural Science Foundation for Youths (2022JQ-142).

*Corresponding author. E-mail: leonie@nwsuaf.edu.cn

Received: 2022-04-02; Accepted: 2022-08-10; Published online: 2022-09-27

cloned before further research on them. It remains challenging to clone and express large gene clusters for the discovery of novel natural products. Construction of genomic library, transformation-associated recombination (TAR), and Red/ET recombination technology are essential techniques for cloning large gene clusters. This study summarizes the research on these three cloning techniques from the aspect of cloning strategy and application. Moreover, we discuss the challenges in cloning large gene clusters, providing a methodological reference for further study.

Keywords: natural product; biosynthetic gene cluster; genomic library; transformation-associated recombination (TAR); Red/ET; heterologous expression of gene clusters

天然产物在自然界中广泛存在,其来源包括动物、植物及微生物,大部分天然产物生物活性多样。目前世界上 60%以上的药物与天然产物有关^[1],在临床使用的抗生素和抗癌药物中,有 70%以上是天然产物,或者是受天然产物启发而形成的结构^[2-3]。天然产物本质为次级代谢产物,其化学结构复杂,在生物体中的合成途径主要由基因组中相关基因控制,该组基因被称为生物合成基因簇(biosynthetic gene clusters, BGCs)^[4],其包括编码核心骨架的合成基因、相关的调控基因及编码修饰酶的基因。研究生命体中的生物合成基因簇有利于充分利用基因组数据,从基因组的角度挖掘新颖天然产物,丰富天然产物结构库。

基因组挖掘是采用生物信息学手段对编码天然产物的基因簇进行分析,预测其可能产生的天然产物结构,为进一步激活或者异源表达基因簇等提供参考依据,最终获得目的天然产物^[5]。通过基因组挖掘手段来开展天然产物研究,其目的性更强、效率更高。同时,基因组挖掘手段有助于促进天然产物的生物合成研究,通过合成生物学解决天然产物进一步发展成先导化合物或者临床药物的瓶颈问题。基因组挖掘可以实现对多类型基因簇的分析,其目的是找出新颖的 BGCs,为获得其对应的新颖活性天然产物提供基础。例如, Malit 等^[6]对 162 672 个细菌基因组进行了全基因组挖掘,鉴定了 829 个含有细胞色

素 P450 基因的环二肽(cyclodipeptide, CDP)生物合成基因簇,通过直接克隆将挖掘出的 3 种菌株的 CDP 基因簇克隆到表达载体上,获得 2 种新的冠状 CDP,这 2 种 CDP 具有前所未有的复杂的大环结构,并表现出神经保护活性。目前,基因组挖掘的策略按照宿主的不同分为在本源菌中的同源表达和在异源宿主中的异源表达。同源表达即在原宿主菌中靶向激活生物合成基因簇,然后通过启动子和调控元件修饰将生物合成基因簇进行激活表达;异源表达则是将生物合成基因簇克隆至载体上,然后将其稳定地转移到一个合适的异源宿主中进行表达^[7]。

在异源表达操作过程中,关键技术是通过直接或间接克隆技术获得目的片段,然后将其转移到外源宿主细胞中进行表达,从而实现对目的片段的功能分析与验证^[8]。该异源表达方法对于那些自身产物代谢水平很低的菌株、环境基因组及难培养的微生物基因组的研究十分有利,基本能够使 BGCs 大量表达^[9]。然而,异源表达过程中的重点和难点是对大片段基因簇的克隆技术。目前,国内外已经发展出多种 BGCs 克隆策略,包括构建基因组文库、转化关联重组(transformation-associated recombination, TAR)、Red/ET 重组系统等用于微生物、动物、植物的基因表达(表 1)。这 3 种大片段 DNA 克隆技术各具有优缺点,其中构建基因组文库的方式能最大限度地保存基

基因组信息,但后期需要大量的筛选工作,同时还涉及片段之间的连接问题;而 TAR 克隆技术能实现目标基因簇的直接克隆,但这一技术遗传稳定性较差,而且 DNA 纯化产率较低,不利于后期的遗传操作;Red/ET 技术的效率较高且不依

赖于限制性内切酶,但操作步骤较为烦琐。我们列举了通过这 3 种方法成功表达的天然产物及其结构(图 1),研究人员可以通过基因簇的特点及实验条件选择合适的方法来开展大片段基因簇的克隆。

表 1 大片段 DNA 克隆技术的性质比较

Table 1 Comparison of properties of large DNA fragment cloning techniques

Cloning technology	The capacity of the carrier (kb)	Advantages	Disadvantages	Application area	Examples
Genomic library	100-300	Low chimerism rate; High genetic stability	Restriction endonuclease cleavage; DNA molecular connection	Gene cloning; Physical map construction; Plant functional genomics	Geldanamycin (1) ^[10] ; Daptomycin (2) ^[11]
TAR	<300	Assembly without enzyme; Direct isolation of BGCs	Genetic instability; A high degree of chimerism; Low yield of DNA purification products	Assembling gene clusters; Gene cluster cloning	Nataxazole (3) ^[12] ; Taromycin A (4) ^[13]
Red/ET	-	High efficiency Short homologous sequences; Restriction sites independent	Operation steps are relatively complex; The proportion of effective fragments is not high	Gene cluster cloning; Genetic modification	Syringolin G (5) ^[14] ; Syringolin H (6) ^[14]

-: Red/ET 重组技术对基因大小无限制

-: Red/ET recombination technology has no limit to gene size.

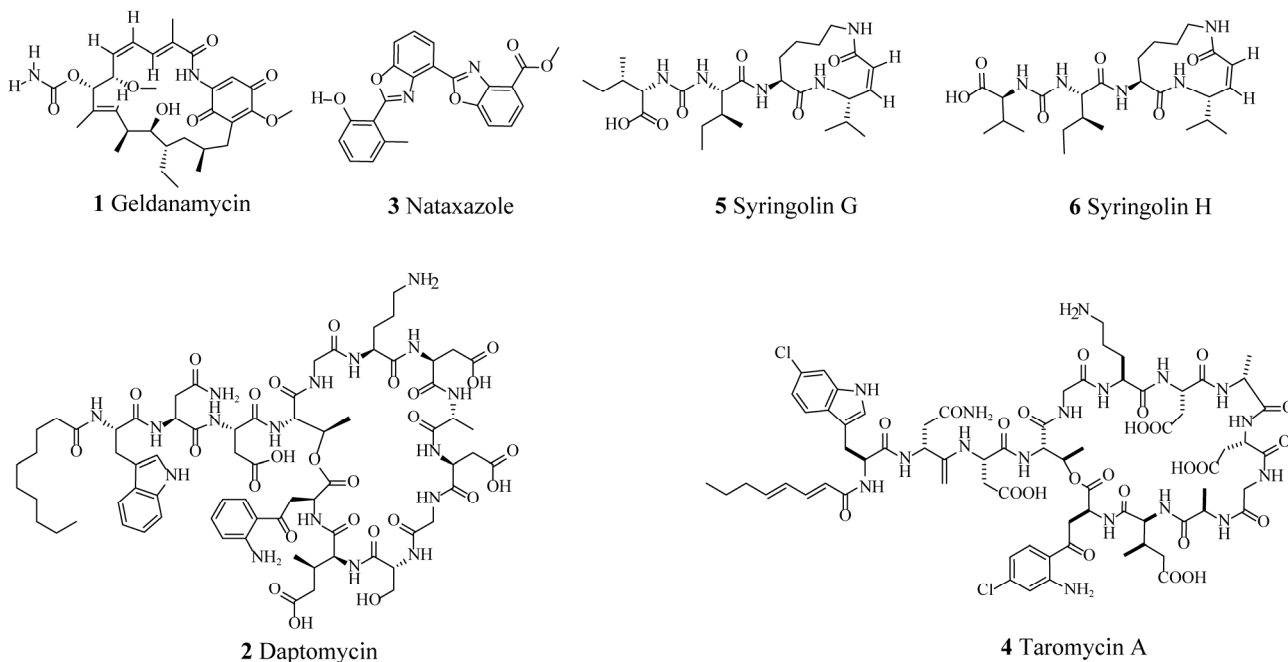


图 1 化合物 1-6 结构

Figure 1 Structures of compounds 1-6.

1 天然产物 BGCs 克隆技术

1.1 基因组文库

随着测序技术的迅速发展,基于基因组信息的新型基因簇的深度挖掘将会带动结构新颖的天然产物不断出现。Malit 等^[15]以催化合成硫代酰胺键的 TufA-YcaO 蛋白为探针,通过 antibiotics & secondary metabolite analysis shell (antiSMASH)和 repeated incremental pruning to produce error reduction (RiPPER)分析确定了 613 个独特的 TufA 相关基因簇家族(gene cluster families, GCF)和 797 个前体肽家族。基因组挖掘出的这些独特基因簇,有些具有较多后修饰基因,使得整个基因簇较大,想要进一步通过实验验证其编码的产物较为困难。在直接克隆技术开发之前,获得大片段次级代谢产物的 BGCs 需要构建基因组文库。基于基因组文库的大片段异源表达方法作为有效的沉默基因簇激活策略,具有全局性和通用性。其理论上包含着目标物种的全部信息,可以使生物体的遗传信息以稳定的重组体形式储存起来。自 1974 年 Murray 等^[16]利用 λ

噬菌体载体成功构建基因组文库后,能够容纳更大片段、更稳定的载体逐步被研发出来(图 2),这些载体包括 λ 噬菌体、柯斯质粒(cosmid)、酵母人工染色体(yeast artificial chromosome, YAC)、P1 人工染色体(P1-derived artificial chromosome, PAC)、细菌人工染色体 (bacterial artificial chromosome, BAC)等,它们的结构和功能不断完善,各自具有其优缺点和适用领域。我们总结了其中能够容纳大片段 DNA 克隆载体的优缺点、载体目前的应用领域及应用相关技术成功表达的天然产物生物合成基因簇(表 2),以便根据载体的特点选择合适的载体对基因簇进行克隆。

克隆载体根据发展经历及特点等可划分为 3 个阶段^[17],第 1 阶段包括质粒、 λ 噬菌体和 cosmid,如 Li 等^[18]用引物对含有链丝菌素(streptothricins)抗生素生物合成相关基因的 cosmid 文库进行聚合酶链式反应(polymerase chain reaction, PCR)筛选,将筛选得到的基因片段 pTG6001 (41 kb)通过大肠杆菌-链霉菌接合法导入 *Streptomyces coelicolor* M145 中,生成了 streptothricin F、streptothricin D 及 2 个新的

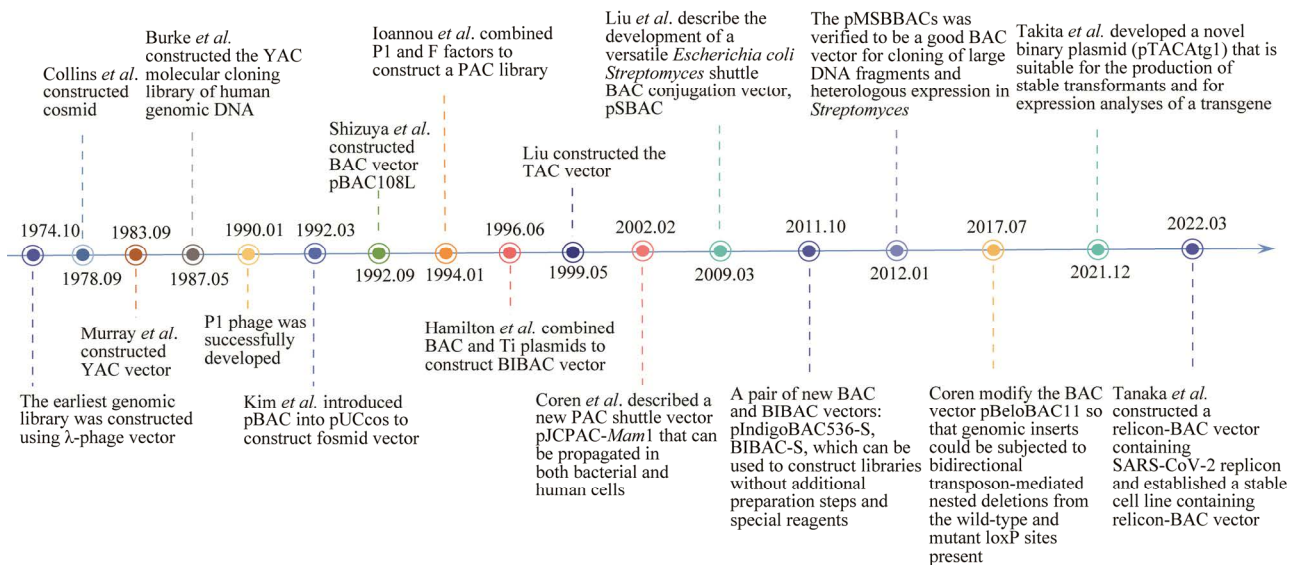


图 2 克隆载体的研发历程^[16,20-35]

Figure 2 The development process of cloning vector^[16,20-35].

表 2 几种克隆载体的性质比较^[21,25-28]Table 2 Comparison of properties of several cloning vectors^[21,25-28]

Vector	Constructi on years	Host	Capacity (kb)	Stability	Direct conversion	Advantages	Disadvantages	Application
YAC	1987	<i>Saccharomyces cerevisiae</i>	100–2 000	Poor stability	No direct conversion capability	Improvement of conversion capacity	Chimerism and recombination; Low conversion efficiency; Difficulty in separation	Gene cloning; Physical map construction; Genome sequencing
BAC	1992	<i>Escherichia coli</i>	100–300	Good stability	No direct conversion capability	Simple operation; Low chimerism rate; Great genetic stability	Complex procedures and time-consuming; False positives	Targeted marker development; Genome sequencing
PAC	1994	<i>Escherichia coli</i>	100–300	Good stability	No direct conversion capability	No obvious absence or chimerism; High genetic stability	Large carrier segment; Low efficiency	Physical map construction
BIBAC	1996	<i>Escherichia coli</i> / <i>Agrobacterium</i> sp.	100–300	Good stability	Have direct conversion capability	Direct transformation of plants; High conversion rate; Convenient selection	Large volume; Low copy number	Plant functional genomics
TAC	1999	<i>Escherichia coli</i> / <i>Agrobacterium</i> sp.	100–300	Good stability	Have direct conversion capability	Direct transformation of plants; High stability in <i>E. coli</i> and <i>Agrobacterium</i> sp.	Large carrier segment	Plant functional genomics

streptothricin (ST)类似物。蒋程恺等^[19]从林可链霉菌 NRRL 2936 的 fosmid 基因组文库中获得了含有完整的帕马霉素生物合成基因簇质粒 pJQK450 (32 kb), 并在 *S. coelicolor* M1154 中实现了异源表达。但质粒和 λ 噬菌体的插入片段大小为 20 kb, cosmid 和 fosmid 的插入片段极限为 40 kb, 这 4 种载体很难完成大量叠连序列的组装及哺乳动物基因组物理图谱的构建, 并且对于大片段 DNA 无法直接克隆, 从而限制了这一阶段载体的广泛应用。第 2 阶段包括 YAC、PAC 及 BAC 载体, 所能容纳的 DNA 片段明显增大, 稳定性增强, 其中 BAC 是目前应用最广泛的大片段 DNA 克隆载体。第 3 阶段包括二元细菌人工染色体 (binary bacterial artificial chromosome, BIBAC) 和可转化人工染色体 (transformation-competent artificial chromosome, TAC), 这类载体在大肠杆菌和农杆菌中稳定存在, 能够直接转化植物。在这 3 个阶段的载体中, 第 1、3 阶段载

体容纳的 DNA 片段较大, 因而常被用来构建大片段基因组文库。

1.1.1 YAC

λ 噬菌体文库和 cosmid 文库插入片段较小、稳定性差, 并不适用于复杂基因组的研究。为了容纳更大的 DNA 片段以便于研究, Murray 等^[21]构建了 55 kb 包括着丝粒、端粒、复制子和外源 DNA 在内的 YAC, 该染色体包含酵母有丝分裂繁殖所需的所有功能。

与原核克隆系统相比, YAC 的优势在于插入片段增大(100–1 000 kb), 能够维持原核克隆系统中不稳定的序列^[36], 并广泛地应用于动物、植物、微生物的基因组研究。1987 年, Burke 等^[22]首次构建了人类基因组 DNA 的 YAC 分子克隆库。随后, YAC 载体逐渐应用于拟南芥^[37]等模式作物, 这些研究均已取得一定的成果。然而, 在传代培养时存在部分片段可能会缺失的现象或其中的片段重排, 这对物理图谱的构建及

基因分离不利。此外, YAC 容易发生嵌合、转化效率低、分离困难, 同样会对后期的研究造成障碍。

1.1.2 BAC

1992 年, Shizuya 等^[25]构建了第 1 个 BAC 载体 pBAC108L, 该载体由大肠杆菌中的致育因子(F 因子)演化而来, 具有操作简单、嵌合率低、遗传稳定性好等优点, 是目前应用最广泛的大型插入式 DNA 文库^[38]。

BAC 载体的插入片段长度可达 300 kb, 能够从单个基因组或混合微生物基因组样品中稳定克隆大片段 DNA。通过构建 BAC 文库后筛选获得的 BAC 克隆可用于组装重复的基因组区域, 重组 BAC 克隆可在大片段基因组区域中精确修饰调控 DNA, 并在严格的遗传控制下表达修饰的编码区^[39]。Liu 等^[30]描述了一种多功能大肠杆菌-链霉菌穿梭细菌人工染色体(BAC)接合载体 psBAC, 以促进放线菌次级代谢产物生物合成基因簇的克隆和异源表达。他们将 *Streptomyces* sp. NRRL 30748 的基因组 DNA 经 *Mfe* I 酶切后构建 psBAC 文库, 对该文库进行筛选后得到一个 95 kb 含美立霉素(meridamycin, *mer*)基因簇的克隆 pHLW30, 该克隆接合转移到 *S. lividans* K4-114 中, 并插入启动子 *ermE**得到菌株 E7, 通过液相色谱-质谱联用仪(liquid chromatograph mass spectrometer, LC-MS)在 E7 发酵提取物中成功检测到了 meridamycin, 但产量很低。Feng 等^[40]构建了异偏头痛素(*iso-migrastatin*, *iso-MGS*)产生菌 *Streptomyces platensis* NRRL18993 的细菌人工染色体文库, 从该文库中鉴定出 1 个 65 kb 含有完整的 *iso-MGS* BGC 的 BAC 克隆 pBS11001, 将 pBS11001 转入 5 种异源链霉菌宿主中, 得到重组菌株 SB11001、SB11002、SB11003、SB11004 和 SB11005。在标准的培养基和发酵条件下, 重

组菌株产生与 *S. platensis* NRRL18993 相同的 *iso-MGS* 图谱。Takao 等^[41]利用 BAC 文库进行蒽醌化合物 kinanthraquinone 基因簇的亚克隆, 通过原生质体转化在 *S. lividans* TK23 中进行异源表达, 并鉴定 kinanthraquinone 和 kinanthraquinone B 的产量。这些研究成果均验证了基于 BAC 载体的克隆技术在天然产物研究中表达大片段 BGCs 的有效性和灵活性。

BAC 载体可通过进一步修饰, 使其能够在特定的宿主菌株中复制和表达。Kakirde 等^[42]构建了一种新的革兰氏阴性穿梭 BAC 载体 pGNS-BAC, 将来自 IncP 质粒的 *oriT* 克隆到 pGNS-BAC 中, 使之带有结合转移位点, 从而使 pGNS-BAC 的 DNA 能够通过电穿孔等方式进入不同革兰氏阴性细菌宿主。该载体利用 F 复制子在大肠杆菌中进行稳定的单拷贝复制, 利用 RK2 微复制子在多种革兰氏阴性菌中进行高拷贝复制。

目前, BAC 载体已成功应用于 meridamycin^[30]、*iso-migrastatin*^[40]、格尔德霉素(geldanamycin)^[10]和达托霉素(daptomycin)^[11]等生物合成基因簇的外源表达, 其长度分别约为 95、65、85 和 100 kb。同时, BAC 载体还被应用于稻曲病菌 UV-2^[43]等大片段基因组文库的研究。基因工程中常用宿主天蓝色链霉菌的全基因组测序工作也是利用 BAC 载体构建基因组文库的方式完成的^[44]。在对微生物天然产物生物合成基因簇的研究中, BAC 文库被广泛使用, 目前已有多个生物合成基因簇通过 BAC 载体引入链霉菌菌株中。同时, BAC 载体还可以与其他克隆技术相结合, 高效地实现大量基因簇的克隆。Liang 等^[45]结合 CRISPR associated protein 12a (Cas12a)和 BAC 文库的特点, 开发了一种快速而高效的直接捕获大型 BGCs 的体外平台, 命名为 CAT-FISHING (CRISPR/Cas12a 介导的快速直接生物合成基因

簇克隆), CAT-FISHING 能够有效地从不同放线菌基因组 DNA 样品中捕获多个大型 BGC, 最大捕获量为 145 kb, GC 含量为 75%。通过该方法, Liang 等^[45]从 *Micromonospora* sp. 181 中直接克隆了 110 kb 长的聚酮化合物合酶 BGC, 随后在 *S. albus* J1074 中异源表达, 鉴定和表征了一种新的大环内酰胺化合物 马里诺霉素 A (marinolactam A), 其具有很好的抗癌活性。另一类大环内酰胺的活性次级代谢产物安莎霉素主要由链霉菌和芽孢杆菌产生, 其中还包含一些海洋来源的细菌^[46]。这类化合物的显著特征是一个中型或大型大环内酰胺“手柄”融合到 1 个单环或双环芳香核心。其中一类重要的苯安莎霉素化合物是三烯霉素。我们课题组从秦岭太白山北坡苔藓中 1 株放线菌 *S. cacaoi* subsp. *asoensis* H2S5 中获得了一系列三烯霉素类化合物, 并预测对应的 BGC 为一个 89 kb 的 *tym* 基因簇^[47-49]。基于 BAC 文库构建特别适合于 GC 含量高的大片段 DNA 的克隆^[45], 因此, 我们针对 *S. cacaoi* subsp. *asoensis* H2S5 基因组数据构建 BAC 文库, 进而展开对安莎霉素生物合成基因簇的异源表达研究。

1.1.3 BIBAC

由于 BAC、YAC 文库需要随后的亚克隆步骤, 不能直接转化植物, 使得植物转化和互补实验非常费力。为了规避这一限制, Hamilton^[27]研发了 BIBAC 载体, 其在结构上同时具有 BAC 的复制系统和农杆菌的 root inducing (Ri) 复制子、转移 DNA (transfer DNA, T-DNA) 的左右边界, 可以在大肠杆菌和农杆菌中穿梭复制。

作为大片段 DNA 克隆载体, BIBAC 构建的基因组文库, 一方面能够在农杆菌介导下转化植物, 简化了大片段 DNA 转化过程, 提高了植物基因定位克隆的效率, 以较小的库容保存植物体完整的基因组信息; 另一方面, 由于克隆所携带

的目的 DNA 片段大, 在后续的转化试验中, BIBAC 载体能够实现基因的成簇转化, 而且外源基因在受体中的表达不易产生位置效应和基因沉默现象, 可避免单基因转化的弊端^[50], 促进基因发现和功能研究。He 等^[51]构建了第一个野生稻基因组 BIBAC 文库, 利用 2 个限制性片段长度多态性 (restriction-fragment length polymorphism, RFLP) 标记 R288 和 C820 对文库进行筛选, 分别鉴定出 7 和 8 个阳性克隆, 通过农杆菌介导, 将含有 120 kb 插入片段的 BIBAC 克隆成功转入水稻基因组。

在培育转基因植物时, 目标之一是实现转基因的稳定表达。通过 T-DNA 的单拷贝整合可以避免转基因沉默。BIBAC 载体在根癌杆菌和大肠杆菌(T-DNA 组装的生物体)中都是低拷贝, 虽然低拷贝数对大肠杆菌中大片段 DNA 的稳定维持很重要, 但其会导致 DNA 产量降低, 阻碍载体制备。此外, BIBAC 系统用于克隆的限制性酶切位点数量有限, 而且载体本身骨架庞大, 使得设计和克隆一个 T-DNA 结构变得复杂。Anggoro 等^[52]构建了 BIBAC-GW, 通过引入 gateway 特异性重组位点增强了其功能, 从而实现快速、简单的载体修改; BIBAC-GW 载体可以像原始 BIBAC 载体一样, 高效地产生单一完整的 T-DNA 整合的转基因拟南芥植株。使用 BIBAC-GW, 已经衍生出许多带有荧光标记基因的结构, 当转入植物时, 可以通过荧光的读数来研究和优化位点特异性基因。

虽然 BIBAC 在大肠杆菌和根癌农杆菌中体积大、拷贝数低, 使得克隆实验会比其他现有的二元载体系统更加困难^[53]。但是 BIBAC 的植物转化能力推动了许多功能基因组学、分子育种和分子农业研究, 包括重要经济基因和数量性状位点 (quantitative trait locus, QTL) 的克隆和基因工程。对这样一个插入的 BIBAC 文库进行转化,

可以转移单个途径或生物过程中涉及的基因、基因簇和调控元件,减少基因沉默,促进功能研究。目前, BIBAC 已成功应用于人参^[54]等物种基因组文库的构建。而且插入大片段 DNA 的 BIBAC 载体已成功转移到番茄等,引入的 T-DNA 稳定维持并遗传了数代,未观察到基因沉默^[55]。

1.1.4 TAC

1999 年, Liu 等^[28]结合 PAC 载体和 BIBAC 载体的特点,构建了多个植物可转化人工染色体载体。该系列载体具有 P1 复制子和 Ri 质粒、pRiA 复制子、Hind III 和 BamH I 等一系列的酶切位点、P1 裂解复制子、重组筛选标记以及植物选择标记,这些特征使 TAC 载体能够以单拷贝形式在大肠杆菌和农杆菌中穿梭复制,有利于克隆 DNA 的提取和文库筛选,减少假阳性的发生。

TAC 文库适合快速定位克隆,比如,某些 TAC 克隆的位置行走和互补突变揭示了异常表型的基因;同时, TAC 文库可用于含有目的基因的基因簇克隆,并用于基因组行走定位克隆^[56]。一旦分离到目标 TAC 克隆,可以立即用农杆菌系统将其转入植物基因组中进行突变体的功能实验,从而加快研究进程。Liu 等^[57]以植物胁迫相关 NAC 转录因子的保守序列为探针,采用 Southern 印迹杂交(Southern blotting)方法筛选 *Oryza officinalis* 基因组 TAC 文库;通过农杆菌介导法,利用含 NAC 基因的 TAC 阳性克隆对水稻品种 HJX74 进行转化,从而实现将 *O. officinalis* 基因组片段整合到 HJX74 基因组中,得到的转基因株系对于旱胁迫表现出较高的耐受性。目前, TAC 文库已成功应用于黑籽南瓜^[58]等物种的基因组文库构建。

1.2 TAR

截至目前,通过构建基因组文库获取大片段 DNA 的方法仍存在部分缺陷。例如,随机建库

克隆需要依靠高通量筛选;筛选结果存在假阳性;PCR 难以扩增 10 kb 以上片段,从小片段拼装费时费力且突变率高;基因片段的完整性不确定;基于限制性酶切连接难以找到片段两端适宜的酶切位点等。作为全新的大片段 DNA 克隆技术, TAR 能够直接从基因组中分离特定基因或基因簇,无需构建基因组文库,克服了大片段基因难以被克隆到同一宿主菌株的缺陷,是目前唯一能够选择性分离长达 300 kb DNA 片段的克隆方法。

1.2.1 TAR、ARS 和 DNA 序列信息

TAR 克隆通过酵母体内重组,准确、选择性地从复杂的基因组 DNA 中分离出完整的基因簇或大的染色体片段^[59]。通过在 DNA 5'端和 3'端设计带有目的基因簇的特异性同源臂的线性化载体,再将该载体与目的基因片段共同转移到酿酒酵母中发生同源重组,实现目的基因簇的直接克隆。

原始的 TAR 克隆方法需要克隆的 DNA 片段至少携带一个自主复制序列(autonomous replication sequences, ARS),该序列可以作为酵母复制的起点,这种策略适用于大多数染色体区域。然而,在具有多个重复元件的染色体区域,如着丝粒和端粒,ARS 频率可能会降低,因而无法通过上述方法进行分离^[60]。Noskov 等^[61]构建了新的 TAR 克隆系统,可以分离缺乏酵母 ARS 序列的基因组区域;在该系统中,ARS 与 URA3 一起作为负选择标记插入 TAR 载体中,消除了酵母转化过程中由于末端连接导致的载体再循环现象,从而能够从原核和真核基因组中分离任何染色体区域,无论是否存在 ARS,这一系统极大地扩展了 TAR 克隆的应用领域。Kouprina 等^[62]研发了一种直接从基因组 DNA 中分离单拷贝基因的方法,该方法使用具有特异性靶向序列和一个重复序列(分别为人类或小鼠

DNA 的 Alu 序列和 B1 重复序列)的载体, 从人类总 DNA 中分离出完整的单拷贝基因 *HPRT*; 因为载体中一个端点是固定的, 所以这种方法被称为径向 TAR 克隆, 径向 TAR 克隆通过改变唯一的靶向序列的排列方式, 可以从特定的靶向序列中克隆出沿染色体两侧方向延伸的大片段区域; 经径向克隆获得的 DNA 片段大小为 40–250 kb, 阳性克隆的产量与使用两个独特的靶向序列克隆的产量大致相同。

当 3'端和 5'端序列信息都可用时, TAR 可以根据目标基因簇两端序列, 设计一定长度的同源臂, 并将这 2 个独特同源臂序列连接到线性载体两端, 从而获得具有定位功能的线性载体。使用该载体可以从总基因组 DNA 中定位目标基因簇。将基因组 DNA 和载体 DNA 共转染到酵母细胞后, 载体和目标基因簇末端之间的重组形成了环状质粒, 用于进一步的 DNA 分离、测序和分析。Wu 等^[63]利用 TAR 克隆技术, 从 1 株高产 6-demethylchlortetracycline (6-DCT) 菌株 *Streptomyces. aureofaciens* DM-1 基因组 DNA 中克隆了 6-DCT BGC (43.5 kb), 该 BGC 经 *Nsi* I/*Mfe* I 酶切后克隆到一个整合载体 pCAP01 中, 得到的质粒被整合到 *S. aureofaciens* DM-1 的染色体中, 与亲本菌株相比, 重组菌株的 6-DCT 的产量提高了 34%。

1.2.2 CRISPR/Cas9 耦连 TAR 克隆技术

CRISPR/Cas9 系统不依赖于特定的酶切位点, 只需识别长度为 20 nt 的特异性序列。Lee 等^[64]通过 CRISPR/Cas9 偶联 TAR 克隆技术, 使目的 DNA 片段两端发生双链断裂, 将基因的阳性克隆率提高至 32%。Kang 等通过结合 CRISPR/Cas9 和 TAR 建立 mCRISTAR 系统, 可以在一个转化步骤中同时替换多个启动子位点, 该系统通过两轮 mCRISTAR 使沉默的 *Tam* 基因簇被 8 种特征化的合成启动子重构, 然后在 *S.*

albus 中异源表达得到 tetarimycin A^[65]。在此基础上, Kim 等^[66]构建了 miCASTAR 系统, 该系统首先在体外利用 CRISPR/Cas9 在操纵子区域分解基因簇, 通过 PCR 扩增的短 DNA 进行重组, 然后携带合成启动子, 在酵母中利用 TAR 进行重组, 克服了 CRISTAR 系统需要多个质粒的缺点; 之后在放线菌 *Amycolatopsis tolypomycina* NRRL B-24205 基因组中克隆了 *atolypene (ato)* 基因簇, 利用 miCASTAR 激活了 *ato* 生物合成基因簇, 在 miCASTAR 的启动子重构后, 发现 *ato* 基因簇编码 atolypenes A 和 B 对人癌细胞具有细胞毒作用。然而单个 CRISPR 阵列中存在多个直接重复序列, 使得 DNA 合成变得尤为困难, 并且由于所有间隔序列都被整合到一个 CRISPR 阵列中, 因此, 如果需要修改启动子位点的组合以进行多路重构, 就需要合成一个全新的 CRISPR 阵列, 这显著增加了整体成本和时间。Kim 等^[67]通过多质粒共转化的方法克服了这些局限性, 每个质粒都含有一个或两个独特的引导 RNA (s) 和一个独特的营养缺陷标记, 将多个 CRISPR 质粒共转化为酵母实现对 BGCs 的多重重构, 该方法命名为 mpCRISTAR, 与以往基于 CRISPR 阵列的方法相比, 显著提高了启动子工程的复用能力; 另外, 他们用 actinorhodin (ACT) 生物合成基因簇作为模型簇, 证明了 mpCRISTAR 的多路复用能力。

1.2.3 TAR 缺点与改进

TAR 克隆在 NPs 研究中的首次应用是从土壤衍生的 eDNA cosmid library 中组装大型 NPs BGCs。近年来, 随着克隆技术的飞速发展, TAR 被广泛应用于微生物次级代谢产物生物合成基因簇的异源表达和次级代谢产物的异源生物合成。基于 TAR 克隆技术的异源表达策略可实现绝大多数基因簇的完整异源表达。Yamanaka 等^[13]利用基于 pCAP01 质粒的 TAR 克隆技术捕获海

洋放线菌 *Saccharomonospora* sp. CNQ-490 中长达 67 kb 的 *NRPS* 沉默生物合成基因簇, 并对其调控基因进行改造, 然后将改造过的基因簇异源表达于模式链霉菌 *S. coelicolor*, 从中获得该基因簇编码的二氯化脂肽类抗生素 taromycin A。由于大 DNA 片段容易断裂, 因此 TAR 克隆技术只能用于克隆小于 300 kb 的 DNA 片段。若想克隆大于 300 kb 的 DNA 片段, 首先可以通过不断优化 TAR 克隆技术来实现, 其次还可通过 TAR 克隆组装基因片段从而合成生物合成基因簇。Hu 等^[68]利用 TAR 技术将 3 个大小差异显著的 DNA 片段(分别为 38.4、0.3 和 0.8 kb)组装后得到 plipastatin 的生物合成基因簇并在枯草芽孢杆菌中成功表达。

此外, 当克隆区域富含重复序列时 TAR 不稳定, 在酵母有丝分裂时容易发生片段缺失。与高 AT 含量的 DNA 在细菌细胞中难于克隆一样, 富含 GC 的 DNA 区域在酿酒酵母中也难于准确克隆, 但通过对酵母的改造, Noskov 等^[69]仍然克隆到了大于 100 kb 的高 GC 含量区域(GC 含量质量分数为 55%)。

同时, 载体自连接和共转化效率低下等问题同样限制了 TAR 的应用。科研人员针对这一局限从多个角度对 TAR 克隆技术进行了改进, 主要为 CRISPR/Cas9 系统偶联 TAR 克隆技术。除该偶联技术外, 科研人员通过对 TAR 的优化降低了载体的自连接等问题。苏会娟^[70]在 TAR 克隆的基础上构建了诱导型酵母转化重组系统, 即 Itar (inducible transformation-associated recombination) 克隆系统, 该系统利用半乳糖诱导 *I-Sel* 内切酶表达使载体线性化, 通过同源重组将目的 DNA 片段组装至捕捉载体, 实现大片段 DNA 的直接克隆; 通过该系统将基因组上的 *polymyxin* 基因簇转移至 *Paenibacillus polymyxa*, 成功合成了多黏菌素(*polymyxin*)。有研究在 TAR

克隆载体 pCAP01 的基础上, 引入强启动子 *pADHI* 基因和反向选择性标记 *URA3* 基因, 使阳性克隆率达到 80%^[71]。此外, TAR 还被广泛应用于等位基因分离与远程分子单倍型分析、同源基因的分离、HACs、细菌载体中不可克隆的染色体区域分离及合成生物学研究等。

1.3 Red/ET 重组系统

“Red/ET 重组”是近年发展起来的一种基于同源重组原理在大肠杆菌中直接修饰各类 DNA 分子的新技术。该克隆技术利用重组酶 Red α / β 或 Rec E/T 介导的同源重组对 DNA 精确修饰, 不受酶切位点和基因大小限制, 可以简单、快速地对任意大小的 DNA 分子进行插入、敲除、突变等多种修饰。同时能免于建库, 只需要 40–50 bp 的同源臂就能实现高效的同源重组, 从基因组中直接克隆大型基因簇。因此, 可通过合成引物的方式将短序列同源臂加在 PCR 产物的两端, 并实现与任意位置的目标序列进行同源重组。这一点突破了酶切位点和长片段 PCR 扩增的局限性对传统基因操作技术的限制, 大大促进了微生物次级代谢产物的异源表达研究。

Red/ET 重组技术省略了标准限制/连接基因修饰的许多步骤, 极大地促进了复杂生物合成途径的遗传操作^[14]。结合 BAC 或 cosmid 文库的构建和筛选, 该技术已经通过重新组装实现了几条完整的生物合成途径。随后 Red/ET 重组技术介导了生物合成途径的拼接和修饰, 以便在技术上更适宜于微生物系统中的异源生物合成。

对野生型菌株中相对较小的基因簇(<30 kb) 展开研究时, 一方面, 可以构建基因组文库, 从单个 cosmid 中检索到基因簇, 再通过 Red/ET 技术对目标基因簇进行重组。Binz 等^[72]利用 Red/ET 重组技术, 从 2 个 cosmid 中重新构建了 *pla* 生物合成基因簇, 随后在 3 株链霉菌中成功表达了完整的基因簇, 并从表达结构中删除一个

修饰基因, 得到一个新的非 glycosylated phenalin 内酯衍生物。另一方面, 可以通过 Red/ET 重组技术直接从基因组中克隆目标基因簇。Huang 等^[73]通过 Red/ET 重组技术直接克隆得到 *syl* 基因簇, 将含有的 *sylA-sylE* 基因簇片段克隆到 pASK 载体中, 异源表达得到 syringolin。

然而, 很多天然产物生物合成基因簇的大小都大于普通的 cosmid 载体, 克服这种限制的策略目前有 3 种。(1) 构建 BAC 文库, BAC 载体可以容纳 100–300 kb DNA 片段的插入。Red/ET 同源重组能够对 BAC 进行有效的目的性修饰, 如基因的剪切、突变、克隆等^[74]。Tu 等^[75]利用 Red/ET 重组技术从 BACs 中重组出 58 kb 的 *dis* 核心基因簇, 当该基因簇转座到异源宿主 *Myxococcus xanthus* DK1622 的染色体上时, 产生了几种二恶唑衍生物; 并且随后的基因缺失实验证明, 双萨拉唑(disorazol)的生产只需要 *disA–D* 基因, 用合成的启动子取代编码离散 AT 蛋白的 *disD* 基因的天然启动子, 改进了异源宿主 *M. xanthus* DK1622 的 disorazol 产量; Red/ET 还可在 BAC 上高效构建各类载体, 包括基因打靶载体。卢智勇等^[76]利用基于 BAC 的细菌内同源重组构建新型 CCR5Δ32 打靶载体; 张亮等^[77]基于 Red/ET 同源重组技术构建温敏型敲除载体, 通过双交换同源重组技术获得短短芽胞杆菌 X23 全局性转录负调控因子缺失的突变株, 然后采用平板抑菌、高效液相色谱-质谱联用技术(high performance liquid chromatography-mass spectrometry, HPLC-MS)和实时荧光定量 PCR (quantitative real-time PCR, qPCR)研究了全局性转录负调控因子敲除对短短芽胞杆菌 X23 不同生长期发酵液的抑菌活性、edeines 产量、*ede* 操纵子转录水平的影响。(2) 在单个高拷贝数质粒上重组基因簇, 但耗时较长且质粒插入片段较小, 限制了其使用。为了解决上述问题, 可以通

过构建 2 个相容的载体, 然后逐步引入并在合适的宿主菌中共表达, 方便了大片段 DNA 插入的操作。Su 等^[78]构建了 2 个稳定共存的载体(p184 和 pMT), 它们包含兼容的复制源(p15A 或 ColE), 用于表达完整的 PKS 基因簇; 2 个载体不仅具有稳定性和兼容性等特性, 而且具有相似的拷贝数, 因此表达水平可以达到化学计量平衡, 并且其稳定性和兼容性得到了证实。(3) 利用 Red/ET 介导的线性+线性同源重组(linear plus linear homologous recombination, LLHR)应用于直接从基因组 DNA 中克隆大基因簇, 从而避免了基因组文库的构建和筛选^[79]。Bian 等^[14]利用 Red/ET 介导的 LLHR 从 *Pseudomonas syringae* 基因组 DNA 中直接克隆 syringolin 基因簇, 通过 *sylCDE* 在大肠杆菌 GB05-MtaA 中的异源表达, 得到了 2 个新的 syringolin 成员 syringolin G 和 H 并进行了表征, 通过添加 *sylAB* 组装完整的丁香脂生物合成途径, 形成与天然生产者相同的基因结构。

在上述策略中, Red/ET 重组可结合反向筛选技术对 DNA 进行无痕敲除、定点突变和模块替换等修饰。常用的正筛选标记包括各种抗生素抗性基因等, 负筛选标记包括基因 *sacB*、*rspl*、*galK* 和 *ccdB* 等, 但这些标记具有较高的突变频率, 显著增加了负选择后的背景, 并且在重组效率低的条件下还会促进假阳性克隆的发生。Liu 等^[80]利用 Red/ET 重组技术和 *ccdB* 反向筛选标记, 使用一个带有枯草芽胞杆菌 *amyE* 基因和 *ccdB* 负选择标记的 5.2 kb 特异性直接克隆载体, 利用两步法, 分别从 *Brevibacillus brevis* X23 和 *B. amyloliquefaciens* FZB42 中直接克隆了两种大型的 edeine (48.3 kb)和 bacillomycin (37.2 kb) 的生物合成途径, 在 1 周内整合到枯草芽胞杆菌的染色体中并在异源宿主中成功表达。此外, Red/ET 在研究化合物合成机理上也有突出的表

现, Red/ET 技术通过对 colibactin 的合成后修饰过程进行阻断, 确定了 colibactin 的前体物质^[81]。

2 挑战与展望

随着基因测序技术和生物信息学的快速发展, 通过基因组挖掘发现天然产物生物合成基因簇, 最终获得新颖天然产物的策略, 具有快速、高效的特点, 并逐渐成为研究热点。基于该策略开展的天然产物研究往往依赖于其生物合成途径在异源宿主中的表达, 通过激活编码其生物合成途径的基因簇来发现天然产物。基因组测序数据表明, 微生物产生生物活性 NPs 的潜力比目前报道的要大得多^[82], 但截至目前, 只有少数 BGCs 被鉴定出来, 有些 BGCs 受到严格的转录调控, 在野生型宿主中表达量低或不表达; 有些 BGCs 存在环境基因组中难以表达; 还有些 BGCs 存在于实验室难培养微生物中, 无法在实验条件获得其对应的天然产物。在本文综述的 3 种主要的大片段 DNA 克隆技术中, 传统的分子克隆方法依赖于限制性内切酶的裂解和 DNA 分子的连接, 需要良好的酶切位点, 这限制了其实际应用; TAR 克隆能够分离小于 300 kb 的片段, 但对于富含 GC 的 DNA 片段在酵母中的克隆效率较低; Red/ET 克隆策略存在操作步骤相对复杂、合成效率不高、合成有效片段的比例不高等问题。

综上所述, 已有克隆技术虽然仍具有缺点, 但是针对这些劣势, 科研人员逐渐研发了新的克隆技术攻克或改进了原有克隆方法。这些新型克隆技术的使用可以有效促进天然产物生物合成基因簇的克隆与表达, 从而促进新颖天然产物的发现, 但对大片段基因簇的克隆仍然是一个具有挑战性的问题。相信随着基因挖掘技术和测序技术的不断发展, 将会挖掘出更多的基因簇, 结合新型大片段克隆表达技术, 从而可以获得活性多

样的天然产物, 丰富天然产物化学结构和生物活性数据库。

REFERENCES

- [1] KIM LJ, OHASHI M, ZHANG Z, TAN D, ASAY M, CASCIO D, RODRIGUEZ JA, TANG Y, NELSON HM. Prospecting for natural products by genome mining and microcrystal electron diffraction[J]. *Nature Chemical Biology*, 2021, 17(8): 872-877.
- [2] 张琳, 宫春杰. 微生物来源的天然产物研究进展[J]. *生物化工*, 2020, 6(6): 121-125.
ZHANG L, GONG CJ. Research progress of natural products from microorganisms[J]. *Biological Chemical Engineering*, 2020, 6(6): 121-125 (in Chinese).
- [3] NEWMAN DJ, CRAGG GM. Natural products as sources of new drugs over the last 25 years[J]. *Journal of Natural Products*, 2007 70(3): 461-77.
- [4] VIJU N, PUNITHA SMJ, SATHEESH S. An analysis of biosynthesis gene clusters and bioactivity of marine bacterial symbionts[J]. *Current Microbiology*, 2021, 78(7): 2522-2533.
- [5] 杨谦, 程伯涛, 汤志军, 刘文. 基因组挖掘在天然产物发现中的应用和前景[J]. *合成生物学*, 2021, 2(5): 697-715.
YANG Q, CHENG BT, TANG ZJ, LIU W. Applications and prospects of genome mining in the discovery of natural products[J]. *Synthetic Biology Journal*, 2021, 2(5): 697-715 (in Chinese).
- [6] MALIT JIL, LIU WC, CHENG AF, SAHA S, LIU LL, QIAN PY. Global genome mining reveals a cytochrome P450-catalyzed cyclization of crownlike cyclodipeptides with neuroprotective activity[J]. *Organic Letters*, 2021, 23(17): 6601-6605.
- [7] 李瑞娟, 赵晓雨, 杨润雨, 刘洋, 颜富, 王海龙, 张友明, 符军. 噬菌体重组酶介导的 DNA 同源重组工程[J]. *微生物学通报*, 2021, 48(9): 3230-3248.
LI RJ, ZHAO XY, YANG RY, LIU Y, YAN F, WANG HL, ZHANG YM, FU J. Recombineering mediated by bacteriophage recombinases[J]. *Microbiology China*, 2021, 48(9): 3230-3248 (in Chinese).
- [8] TANG L, SHAH S, CHUNG L, CARNEY J, KATZ L, KHOSLA C, JULIEN B. Cloning and heterologous expression of the epothilone gene cluster[J]. *Science*, 2000, 287(5453): 640-642.

- [9] 付博, 樊泽正, 杜毅涛, 郭子晟, 李忠玲. 微生物基因组挖掘的方法和研究策略[J]. 基因组学与应用生物学, 2018, 37(6): 2451-2458.
FU B, FAN ZZ, DU YT, GUO ZS, LI ZL. Approaches and strategies of genome mining in microbes[J]. Genomics and Applied Biology, 2018, 37(6): 2451-2458 (in Chinese).
- [10] RASCHER A, HU ZH, VISWANATHAN N, SCHIRMER A, REID R, NIERMAN WC, LEWIS M, HUTCHINSON CR. Cloning and characterization of a gene cluster for geldanamycin production in *Streptomyces hygroscopicus* NRRL 3602[J]. FEMS Microbiology Letters, 2003, 218(2): 223-230.
- [11] PENN J, LI X, WHITING A, LATIF M, GIBSON T, SILVA CJ, BRIAN P, DAVIES J, MIAO V, WRIGLEY SK, BALTZ RH. Heterologous production of daptomycin in *Streptomyces lividans*[J]. Journal of Industrial Microbiology and Biotechnology, 2006, 33(2): 121-128.
- [12] CANO-PRieto C, GARCÍA-SALCEDO R, SÁNCHEZ-HIDALGO M, BRAÑA AF, FIEDLER HP, MÉNDEZ C, SALAS JA, OLANO C. Genome mining of *Streptomyces* sp. Tü 6176: characterization of the nataxazole biosynthesis pathway[J]. ChemBiochem: a European Journal of Chemical Biology, 2015, 16(10):1461-1473.
- [13] YAMANAKA K, REYNOLDS KA, KERSTEN RD, RYAN KS, GONZALEZ DJ, NIZET V, DORRESTEIN PC, MOORE BS. Direct cloning and refactoring of a silent lipopeptide biosynthetic gene cluster yields the antibiotic taromycin A[J]. Proceedings of the National Academy of Sciences of the United States of America, 2014, 111(5): 1957-1962.
- [14] BIAN XY, HUANG F, STEWART FA, XIA LQ, ZHANG YM, MÜLLER R. Direct cloning, genetic engineering, and heterologous expression of the syringolin biosynthetic gene cluster in *E. coli* through Red/ET recombineering[J]. ChemBiochem: a European Journal of Chemical Biology, 2012, 13(13): 1946-1952.
- [15] MALIT JLL, WU CH, LIU LL, QIAN PY. Global genome mining reveals the distribution of diverse thioamidated RiPP biosynthesis gene clusters[J]. Frontiers in Microbiology, 2021, 12: 635389.
- [16] MURRAY NE, MURRAY K. Manipulation of restriction targets in phage lambda to form receptor chromosomes for DNA fragments[J]. Nature, 1974, 251(5475): 476-481.
- [17] 于洋, 蒋世翠, 王康宇, 薛菲, 张美萍, 王义. 大片段DNA克隆载体的研究进展[J]. 黑龙江农业科学, 2015(2): 147-151.
YU Y, JIANG SC, WANG KY, XUE F, ZHANG MP, WANG Y. Research progress of large fragment DNA cloning vector[J]. Heilongjiang Agricultural Sciences, 2015(2): 147-151 (in Chinese).
- [18] LI JE, GUO ZY, HUANG W, MENG XX, AI GM, TANG GL, CHEN YH. Mining of a streptothricin gene cluster from *Streptomyces* sp. TP-A0356 genome via heterologous expression[J]. Science China Life Sciences, 2013, 56(7): 619-627.
- [19] 蒋程恺, 孟思童, 张菲, 胡晓婧, 谢守锋, 陶美凤, 梁晶丹, 康前进, 白林泉, 邓子新. 林可链霉菌 NRRL 2936 中帕马霉素生物合成基因簇的异源表达及调控基因功能研究[J]. 微生物学通报, 2018, 45(2): 334-346.
JIANG CK, MENG ST, ZHANG F, HU XJ, XIE SF, TAO MF, LIANG JD, KANG QJ, BAI LQ, DENG ZX. Heterologous expression of the biosynthetic gene cluster of pamamycin and investigation on the functions of regulatory genes from *Streptomyces lincolnensis* NRRL 2936[J]. Microbiology China, 2018, 45(2): 334-346 (in Chinese).
- [20] COLLINS J, HOHN B. Cosmids: a type of plasmid gene-cloning vector that is packageable *in vitro* in bacteriophage lambda heads[J]. Proceedings of the National Academy of Sciences of the United States of America, 1978, 75(9): 4242-4246.
- [21] MURRAY AW, SZOSTAK JW. Construction of chromosomes in yeast[J]. Nature, 1983, 305: 189-193.
- [22] BURKE DT, CARLE GF, OLSON MV. Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors[J]. Science, 1987, 236(4803): 806-812.
- [23] STERNBERG N. Bacteriophage P1 cloning system for the isolation, amplification, and recovery of DNA fragments as large as 100 kilobase pairs[J]. Proceedings of the National Academy of Sciences of the United States of America, 1990, 87(1): 103-107.
- [24] KIM UJ, SHIZUYA H, de JONG PJ, BIRREN B, SIMON MI. Stable propagation of cosmid sized human DNA inserts in an F factor-based vector[J]. Nucleic Acids Research, 1992, 20: 1083-1085.

- [25] SHIZUYA H, BIRREN B, KIM UJ, MANCINO V, SLEPAK T, TACHIIRI Y, SIMON M. Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector[J]. Proceedings of the National Academy of Sciences of the United States of America, 1992, 89(18): 8794-8797.
- [26] IOANNOU PA, AMEMIYA CT, GARNES J, KROISEL PM, SHIZUYA H, CHEN C, BATZER MA, JONG PJD. A new bacteriophage P1-derived vector for the propagation of large human DNA fragments[J]. Nature Genetics, 1994, 6: 84-89.
- [27] HAMILTON CM. A binary-BAC system for plant transformation with high-molecular-weight DNA[J]. Gene, 1997, 200(1/2): 107-116.
- [28] LIU YG, SHIRANO Y, FUKAKI H, YANAI Y, TASAKA M, TABATA S, SHIBATA D. Complementation of plant mutants with large genomic DNA fragments by a transformation-competent artificial chromosome vector accelerates positional cloning[J]. Proceedings of the National Academy of Sciences of the United States of America, 1999, 96(11): 6535-6540.
- [29] COREN JS, STERNBERG N. Construction of a PAC vector system for the propagation of genomic DNA in bacterial and mammalian cells and subsequent generation of nested deletions in individual library members[J]. Gene, 2001, 264(1): 11-18.
- [30] LIU HB, JIANG H, HALTLI B, KULOWSKI K, MUSZYNSKA E, FENG XD, SUMMERS M, YOUNG M, GRAZIANI E, KOEHN F, Carter GT, He M. Rapid cloning and heterologous expression of the meridamycin biosynthetic gene cluster using a versatile *Escherichia coli-Streptomyces* artificial chromosome vector, pSBAC[J]. Journal of Natural Products, 2009, 72(3): 389-395.
- [31] SHI X, ZENG H, XUE Y, LUO M. A pair of new BAC and BIBAC vectors that facilitate BAC/BIBAC library construction and intact large genomic DNA insert exchange[J]. Plant Methods, 2011, 7: 33.
- [32] 黄胜,李娜,周俊,何璟. 适用于链霉菌大片段基因组 DNA 克隆和异源表达的细菌人工染色体(BAC)载体的构建及应用[J]. 微生物学报, 2012, 52(1): 30-37.
HUANG S, LI N, ZHOU J, HE J. Construction of a new bacterial artificial chromosome (BAC) vector for cloning of large DNA fragments and heterologous expression in *Streptomyces*[J]. Acta Microbiologica Sinica, 2012, 52(1): 30-37 (in chinese).
- [33] COREN JS. Retrofitting the BAC cloning vector pBeloBAC11 by the insertion of a mutant *loxP* site[J]. BMC Research Notes, 2017, 10(1): 344.
- [34] TAKITA E, YOSHIDA K, HANANO S, SHINMYO A, SHIBATA D. Development of the binary vector pTACA_{tg1} for stable gene expression in plant: reduction of gene silencing in transgenic plants carrying the target gene with long flanking sequences[J]. Plant Biotechnology Journal (Tokyo), 2021, 38(4): 391-400.
- [35] TANAKA T, SAITO A, SUZUKI T, MIYAMOTO Y, TAKAYAMA K, OKAMOTO T, MORIISHI K. Establishment of a stable SARS-CoV-2 replicon system for application in high-throughput screening[J]. Antiviral Research, 2022, 199: 105268.
- [36] SANTRA DK, SANDHU D, TAI T, BHATTACHARYYA MK. Construction and characterization of a soybean yeast artificial chromosome library and identification of clones for the *Rps6* region[J]. Functional & Integrative Genomics, 2003, 3(4): 153-159.
- [37] CREUSOT F, FOUILLOUX E, DRON M, LAFLEURIEL J, PICARD G, BILLAULT A, LE PASLIER D, COHEN D, CHABOUTÉ ME, DURR A. The CIC library: a large insert YAC library for genome mapping in *Arabidopsis thaliana*[J]. Plant Journal: for Cell and Molecular Biology, 1995, 8(5): 763-770.
- [38] ŠAĀRĀ J, ŠIMKOVÁ H, DOLEŽEL J. Construction of BAC libraries from flow-sorted chromosomes[J]. Methods in Molecular Biology: Clifton, N J, 2016, 1429: 135-149.
- [39] BUCKLEY KM, DONG P, CAMERON RA, RAST JP. Bacterial artificial chromosomes as recombinant reporter constructs to investigate gene expression and regulation in echinoderms[J]. Briefings in Functional Genomics, 2018, 17(5): 362-371.
- [40] FENG ZY, WANG LY, RAJSKI SR, XU ZN, Coeffet-LeGal MF, SHEN B. Engineered production of iso-migrastatin in heterologous *Streptomyces* hosts[J]. Bioorganic & Medicinal Chemistry, 2009, 17(6): 2147-2153.
- [41] TAKAO RS, SAKAI K, KOSHINO H, OSADA H, TAKAHASHI S. Identification of the kinanthraquinone biosynthetic gene cluster by expression of an atypical

- response regulator[J]. *Bioscience, Biotechnology, and Biochemistry*, 2021, 85(3): 714-721.
- [42] KAKIRDE KS, WILD J, GODISKA R, MEAD DA, WIGGINS AG, GOODMAN RM, SZYBALSKI W, LILES MR. Gram negative shuttle BAC vector for heterologous expression of metagenomic libraries[J]. *Gene*, 2011, 475(2): 57-62.
- [43] 刘庆丽, 王晓明, 王革娇, 罗朝喜, 谭新球, 罗美中. 稻曲病菌 UV-2 菌株细菌人工染色体文库构建及分析[J]. *微生物学通报*, 2013, 40(9): 1715-1722.
- LIU QL, WANG XM, WANG GJ, LUO CX, TAN XQ, LUO MZ. Construction of a bacterial artificial chromosome library of *Villosioclava virens* UV-2 genome[J]. *Microbiology China*, 2013, 40(9): 1715-1722 (in Chinese).
- [44] BENTLEY SD, CHATER KF, CERDEÑO-TÁRRAGA AM, CHALLIS GL, THOMSON NR, JAMES KD, HARRIS DE, QUAIL MA, KIESER H, HARPER D, BATEMAN A, BROWN S, CHANDRA G, CHEN CW, COLLINS M, CRONIN A, FRASER A, GOBLE A, HIDALGO J, HORNSBY T, et al. Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2)[J]. *Nature*, 2002, 417(6885): 141-147.
- [45] LIANG MD, LIU LS, XU F, ZENG XQ, WANG RJ, YANG JL, WANG WS, KARTHIK L, LIU JK, YANG ZH, ZHU GL, WANG SL, BAI LQ, TONG YJ, LIU XT, WU M, ZHANG LX, TAN GY. Activating cryptic biosynthetic gene cluster through a CRISPR-Cas12a-mediated direct cloning approach[J]. *Nucleic Acids Research*, 2022, 50(6): 3581-3592.
- [46] FAN YQ, WANG C, WANG LP, CHAIROUNGDU A, PIYACHATURAWAT P, FU P, ZHU WM. New Ansamycins from the Deep-Sea-Derived Bacterium *Ochrobactrum* sp. OUCMDZ-2164[J]. *Marine Drugs*, 2018, 16(8): 282.
- [47] 唐丹. 黑蛋巢菌与秦岭链霉菌中活性成分挖掘及生物合成初探[D]. 杨凌: 西北农林科技大学博士学位论文, 2019.
- TANG D. Bioactive components isolated from *Cyathus* and Qinling-mountain-derived *Streptomyces* and their preliminary biosynthesis study[D]. Yangling: Doctoral Dissertation of Northwest A&F University, 2019 (in Chinese).
- [48] LIU LL, CHEN ZF, LIU Y, TANG D, GAO HH, ZHANG Q, GAO JM. Molecular networking-based for the target discovery of potent antiproliferative polycyclic macrolactam ansamycins from *Streptomyces cacaoi* subsp. *asoensis*[J]. *Organic Chemistry Frontiers*, 2020, 7(24): 4008-4018.
- [49] TANG D, LIU LL, HE QR, YAN W, LI D, GAO JM. Ansamycins with antiproliferative and antineuroinflammatory activity from moss-soil-derived *Streptomyces cacaoi* subsp. *asoensis* H2S5[J]. *Journal of Natural Products*, 2018, 81(9): 1984-1991.
- [50] 李娥贤, 殷富有, 张敦宇, 陈越, 余腾琼, 雷涌涛, 肖素勤, 程在全, 柯学. 云南药用野生稻高质量染色体 DNA 的制备[J]. *作物杂志*, 2020(5): 103-109.
- LI EX, YIN FY, ZHANG DY, CHEN Y, YU TQ, LEI YT, XIAO SQ, CHENG ZQ, KE X. Extraction of high-quality chromosome DNA in Yunnan wild rice *Oryza officinalis*[J]. *Crops*, 2020(5): 103-109 (in Chinese).
- [51] HE RF, WANG YY, SHI ZY, REN X, ZHU LL, WENG QM, HE GC. Construction of a genomic library of wild rice and *Agrobacterium*-mediated transformation of large insert DNA linked to BPH resistance locus[J]. *Gene*, 2003, 321: 113-121.
- [52] ANGGORO DT, TARK-DAME M, WALMSLEY A, OKA R, de SAIN M, STAM M. BIBAC-GW-based vectors for generating reporter lines for site-specific genome editing in planta[J]. *Plasmid*, 2017, 89: 27-36.
- [53] CODY JP, GRAHAM ND, BIRCHLER JA. BiBAC modification and stable transfer into maize (*Zea mays*) hi-II immature embryos via *Agrobacterium*-mediated transformation[J]. *Current Protocols in Plant Biology*, 2017, 2(4): 350-369.
- [54] ZHAI JF, WANG Y, SUN CY, JIANG SC, WANG KY, ZHANG Y, ZHANG HB, ZHANG MP. A plant-transformation-competent BIBAC library of ginseng (*Panax ginseng* C. A. Meyer) for functional genomics research and characterization of genes involved in ginsenoside biosynthesis[J]. *Molecular Breeding*, 2013, 31(3): 685-692.
- [55] FRARY A, HAMILTON CM. Efficiency and stability of high molecular weight DNA transformation: an analysis in tomato[J]. *Transgenic Research*, 2001, 10(2): 121-132.
- [56] HIROSE Y, SUDA K, LIU YG, SATO S, NAKAMURA Y, YOKOYAMA K, YAMAMOTO N, HANANO S, TAKITA E, SAKURAI N, SUZUKI H, NAKAMURA Y, KANEKO T, YANO K, TABATA S, SHIBATA D. The *Arabidopsis* TAC position viewer: a high-resolution map

- of transformation-competent artificial chromosome (TAC) clones aligned with the *Arabidopsis thaliana* Columbia-0 genome[J]. *The Plant Journal: for Cell and Molecular Biology*, 2015, 83(6): 1114-1122.
- [57] LIU R, ZHANG HH, CHEN ZX, SHAHID MQ, FU XL, LIU XD. Drought-tolerant rice germplasm developed from an *Oryza officinalis* transformation-competent artificial chromosome clone[J]. *Genetics and Molecular Research: GMR*, 2015, 14(4): 13667-13678.
- [58] 杨正安, 丁玉梅, 张应华, 张兴国. 黑籽南瓜基因组可转化人工染色体(TAC)文库的构建及分析[J]. *农业生物技术学报*, 2012, 20(2): 211-217.
- YANG ZA, DING YM, ZHANG YH, ZHANG XG. The construction and analysis of genomic library based on transformation-competent artificial chromosome (TAC) vector in *Cucurbita ficifolia*[J]. *Journal of Agricultural Biotechnology*, 2012, 20(2): 211-217 (in Chinese).
- [59] KOUPRINA N, KIM JH, LARIONOV V. Highly selective, CRISPR/Cas9-mediated isolation of genes and genomic loci from complex genomes by TAR cloning in yeast[J]. *Current Protocols*, 2021, 1(8): e207.
- [60] KOUPRINA N, NOSKOV VN, LARIONOV V. Selective isolation of large segments from individual microbial genomes and environmental DNA samples using transformation-associated recombination cloning in yeast[J]. *Nature Protocols*, 2020, 15(3): 734-749.
- [61] NOSKOV VN, KOUPRINA N, LEEM SH, OUSPENSKI I, BARRETT JC, LARIONOV V. A general cloning system to selectively isolate any eukaryotic or prokaryotic genomic region in yeast[J]. *BMC Genomics*, 2003, 4(1): 16.
- [62] KOUPRINA N, ANNAB L, GRAVES J, AFSHARI C, BARRETT JC, RESNICK MA, LARIONOV V. Functional copies of a human gene can be directly isolated by transformation-associated recombination cloning with a small 3' end target sequence[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 1998, 95(8): 4469-4474.
- [63] WU NX, HUANG H, MIN TL, HU HF. TAR cloning and integrated overexpression of 6-demethylchlortetracycline biosynthetic gene cluster in *Streptomyces aureofaciens*[J]. *Acta Biochimica et Biophysica Sinica*, 2017, 49(12): 1129-1134.
- [64] LEE NCO, LARIONOV V, KOUPRINA N. Highly efficient CRISPR/Cas9-mediated TAR cloning of genes and chromosomal loci from complex genomes in yeast[J]. *Nucleic Acids Research*, 2015, 43(8): e55.
- [65] KANG HS, CHARLOP-POWERS Z, BRADY SF. Multiplexed CRISPR/Cas9- and TAR-mediated promoter engineering of natural product biosynthetic gene clusters in yeast[J]. *ACS Synthetic Biology*, 2016, 5(9): 1002-1010.
- [66] KIM SH, LU WL, AHMADI MK, MONTIEL D, TERNEI MA, BRADY SF. Atolypenes, tricyclic bacterial sesterterpenes discovered using a multiplexed *in vitro* Cas9-TAR gene cluster refactoring approach[J]. *ACS Synthetic Biology*, 2019, 8(1): 109-118.
- [67] KIM H, JI CH, JE HW, KIM JP, KANG HS. mpCRISTAR: multiple plasmid approach for CRISPR/Cas9 and TAR-mediated multiplexed refactoring of natural product biosynthetic gene clusters[J]. *ACS Synthetic Biology*, 2020, 9(1): 175-180.
- [68] HU YM, NAN F, MAINA SW, GUO J, WU SL, XIN ZH. Clone of plipastatin biosynthetic gene cluster by transformation-associated recombination technique and high efficient expression in model organism *Bacillus subtilis*[J]. *Journal of Biotechnology*, 2018, 288: 1-8.
- [69] NOSKOV VN, KARAS BJ, YOUNG L, CHUANG RY, GIBSON DG, LIN YC, STAM J, YONEMOTO IT, SUZUKI Y, ANDREWS-PFANNKOCHE C, GLASS JI, SMITH HO, HUTCHISON CA 3RD, VENTER JC, WEYMAN PD. Assembly of large, high G+C bacterial DNA fragments in yeast[J]. *ACS Synthetic Biology*, 2012, 1(7): 267-273.
- [70] 苏会娟. 诱导型酵母转化重组系统的构建及其应用[D]. 南京: 南京理工大学硕士学位论文, 2018.
- SU HJ. Construction and application of the inducible transformation-associated recombination in *Saccharomyces cerevisiae*[D]. Nanjing: Master's Thesis of Nanjing University of Science and Technology, 2018 (in Chinese).
- [71] TANG XY, LI J, MILLÁN-AGUIÑAGA N, ZHANG JJ, O'NEILL EC, UGALDE JA, JENSEN PR, MANTOVANI SM, MOORE BS. Identification of thiotetronic acid antibiotic biosynthetic pathways by target-directed genome mining[J]. *ACS Chemical Biology*, 2015, 10(12): 2841-2849.
- [72] BINZ TM, WENZEL SC, SCHNELL HJ, BECHTHOLD A, MÜLLER R. Heterologous expression and genetic engineering of the phenalinolactone biosynthetic gene

- cluster by using Red/ET recombineering[J]. *ChemBioChem: a European Journal of Chemical Biology*, 2008, 9(3): 447-454.
- [73] HUANG F, TANG JL, HE L, DING XZ, HUANG SY, ZHANG YM, SUN YJ, XIA LQ. Heterologous expression and antitumor activity analysis of syringolin from *Pseudomonas syringae* pv. *syringae* B728a[J]. *Microbial Cell Factories*, 2018, 17(1): 31.
- [74] 杨朝霞, 杨霄旭, 李苗苗, 石小亚, 向双林. 通过减毒和 Red/ET 同源重组构建跨界基因沉默菌株[J]. *生命科学研究*, 2016, 20(3): 208-213.
- YANG ZX, YANG XX, LI MM, SHI XY, XIANG SL. Construction of tTrans-kingdom RNAi bacteria by attenuation and Red/ET homologous recombination[J]. *Life Science Research*, 2016, 20(3): 208-213 (in Chinese).
- [75] TU Q, HERRMANN J, HU SB, RAJU R, BIAN XY, ZHANG YM, MÜLLER R. Genetic engineering and heterologous expression of the disorazol biosynthetic gene cluster via Red/ET recombineering[J]. *Scientific Reports*, 2016, 6: 21066.
- [76] 卢智勇, 杨卓顺, 丁妍, 袁雅红, 王小莉, 于莉, 李东升. 利用 Red/ET 同源重组构建新型 CCR5 Δ 32 打靶载体[J]. *生物技术通讯*, 2017, 28(4): 415-421.
- LU ZY, YANG ZS, DING Y, YUAN YH, WANG XL, YU L, LI DS. Construction of CCR5 Δ 32 targeting vector with Red/ET homologous recombination[J]. *Letters in Biotechnology*, 2017, 28(4): 415-421 (in Chinese).
- [77] 张亮, 吕翠, 段彩琛, 王运生, 黄军, 杜杰, 戴良英, 黎定军, 刘清术, 陈武. 转录调控因子 AbrB 对生防短芽孢杆菌 X23 中抗生素 edeines 生物合成的影响[J]. *中国生物防治学报*, 2020, 36(4): 564-574.
- ZHANG L, LÜ C, DUAN CC, WANG YS, HUANG J, DU J, DAI LY, LI DJ, LIU QS, CHEN W. Effect of transcription factor AbrB on the biosynthesis of antibiotic edeines in biocontrol bacteria *Brevibacillus brevis* X23[J]. *Chinese Journal of Biological Control*, 2020, 36(4): 564-574 (in Chinese).
- [78] SU C, ZHAO XQ, WANG HN, QIU RG, TANG L. Seamless stitching of biosynthetic gene cluster containing type I polyketide synthases using Red/ET mediated recombination for construction of stably co-existing plasmids[J]. *Gene*, 2015, 554(2): 233-240.
- [79] FU J, BIAN XY, HU S, WANG HL, HUANG F, SEIBERT PM, PLAZA A, XIA LQ, Müller R, STEWART AF, ZHANG Y. Full-length RecE enhances linear-linear homologous recombination and facilitates direct cloning for bioprospecting[J]. *Nature Biotechnology*, 2012, 30(5): 440-446.
- [80] LIU QS, SHEN QY, BIAN XY, CHEN HN, FU J, WANG HL, LEI P, GUO ZH, CHEN W, LI DJ, ZHANG YM. Simple and rapid direct cloning and heterologous expression of natural product biosynthetic gene cluster in *Bacillus subtilis* via Red/ET recombineering[J]. *Scientific Reports*, 2016, 6: 34623.
- [81] BIAN XY, FU J, PLAZA A, HERRMANN J, PISTORIUS D, STEWART AF, ZHANG YM, MÜLLER R. *In vivo* evidence for a prodrug activation mechanism during colibactin maturation[J]. *ChemBioChem: a European Journal of Chemical Biology*, 2013, 14(10): 1194-1197.
- [82] RUTLEDGE PJ, CHALLIS GL. Discovery of microbial natural products by activation of silent biosynthetic gene clusters[J]. *Nature Reviews Microbiology*, 2015, 13(8): 509-523.