

研究论文

DOI: 10.12211/2096-8280.2021-082

ExoCET-BAC策略高效抓取和组装高AT含量基因组大片段

姜婵娟^{1,2}, 崔天琦¹, 孙洪变¹, 焦念志², 符军¹, 张友明¹, 王海龙¹

(¹ 山东大学微生物技术研究院, 微生物技术国家重点实验室, 山东 青岛 266237; ² 山东大学海洋研究院, 山东 青岛 266237)

摘要: 基因克隆是解析基因功能的重要手段, 但仍有很多基因难以克隆, 比如高AT含量 (>60%) 基因组来源的DNA。ExoCET克隆技术通过联合核酸外切酶介导的体外同源重组和大肠杆菌 RecET重组酶介导的细胞内同源重组, 不仅能从微生物基因组中靶向抓取>100 kb的大片段, 而且能高效组装>13个DNA片段, 是基因克隆的有力工具, 迄今未有利用ExoCET技术从AT含量>63%的基因组克隆大片段的报道。本研究以AT含量为69%的海洋单细胞光合蓝细菌原绿球藻MIT 9301菌株的基因组为研究对象, 探究了利用ExoCET技术进行高AT含量基因组大片段克隆的最佳条件。结果显示: ①在核酸外切酶介导的体外同源重组时使用Gibson体系较T4聚合酶体系能获得更高的克隆效率; ②载体应选择单拷贝的细菌人工染色体(BAC), 多拷贝质粒载体会导致克隆失败; ③ExoCET可以从原绿球藻基因组上抓取>80 kb的大片段, 并且能以100%的正确率组装11个3 kb的DNA片段; ④可以一步同时抓取4个7~20 kb的基因组大片段。大规模基因组测序显示高AT含量生物占比超过30%, 该研究建立的ExoCET-BAC策略将为高AT含量生物的基因组功能研究提供高效使能技术。

关键词: 难克隆DNA; 高AT含量; 基因组; 同源重组; 直接克隆; DNA组装; ExoCET

中图分类号: Q781 文献标志码: A

Efficient capture and assembly of AT-rich genomic fragments using ExoCET-BAC strategy

JIANG Chanjuan^{1,2}, CUI Tianqi¹, SUN Hongluan¹, JIAO Nianzhi², FU Jun¹, ZHANG Youming¹,
WANG Hailong¹

(¹State Key Laboratory of Microbial Technology, Institute of Microbial Technology, Shandong University, Qingdao 266237, Shandong, China; ²Institute of Marine Science and Technology, Shandong University, Qingdao 266237, Shandong, China)

Abstract: Gene cloning is a routine experiment to elucidate its functions, however, many genes are difficult to be cloned, such as those from AT-rich (>60%) fragments. Exonuclease *in vitro* assembly combined with RecET recombination for highly efficient direct DNA (ExoCET) cloning combines *in vitro* homologous recombination

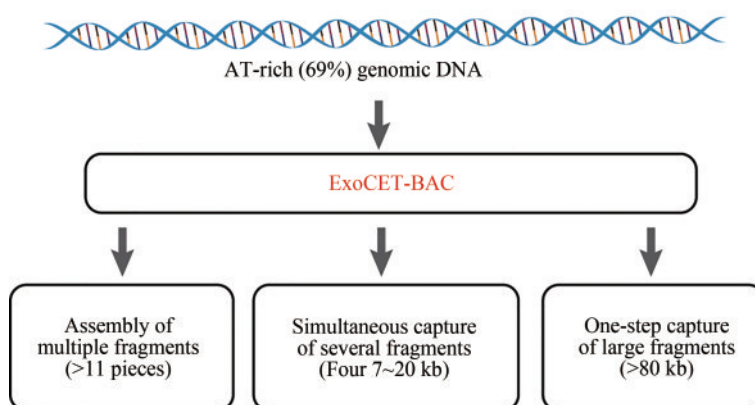
收稿日期: 2021-08-06 修回日期: 2021-08-28

基金项目: 山东省泰山学者项目 (tsqn201812008); 山东大学齐鲁青年学者项目

引用本文: 姜婵娟, 崔天琦, 孙洪变, 焦念志, 符军, 张友明, 王海龙. ExoCET-BAC策略高效抓取和组装高AT含量基因组大片段[J]. 合成生物学, 2022, 3(1): 238-251

Citation: JIANG Chanjuan, CUI Tianqi, SUN Hongluan, JIAO Nianzhi, FU Jun, ZHANG Youming, WANG Hailong. Efficient capture and assembly of AT-rich genomic fragments using ExoCET-BAC strategy[J]. Synthetic Biology Journal, 2022, 3(1): 238-251

mediated by exonucleases and intracellular homologous recombination mediated by *E. coli* RecET recombinases. ExoCET is a powerful gene cloning technique because it can not only capture fragments larger than 100 kb from genomes but also assemble more than 13 pieces of DNA fragments. So far, there has been no report on the use of ExoCET to clone large fragments from genomes with AT content more than 63%. In this study, the genome of the marine cyanobacteria *Prochlorococcus* MIT 9301 strain with 69% AT content was used to optimize the ExoCET conditions for cloning of large AT-rich DNA fragments. The results indicate that: (1) Compared to T4 polymerase method, higher cloning efficiency was obtained when Gibson assembly method was used for the *in vitro* homologous recombination, since the covalently stitched DNA molecule by Gibson assembly is much more stable than the DNA molecule with gap stitched by T4 DNA polymerase; (2) The single-copy bacterial artificial chromosome (BAC) vector should be used, and the multi-copy plasmid vector cannot clone those fragments, because the single-copy BAC can minimize the expression of cloned genes to avoid their toxic effect on the host; (3) The ExoCET-BAC strategy can not only capture fragments larger than 80 kb from the *Prochlorococcus* genome, but also assemble 11 pieces of DNA fragments with 100% accuracy; (4) ExoCET-BAC can capture 4 pieces of genomic fragments of 7-20 kb simultaneously in one step. Genome sequencing revealed that AT-rich organisms account for more than 30%, and thus the ExoCET-BAC strategy established in this study provides efficient enabling technology for the genome functional research of AT-rich organisms.



Keywords: difficult cloning DNA; high AT content; genome; homologous recombination; direct cloning; DNA assembly; ExoCET

随着测序技术的不断进步，越来越多的微生物基因组被测序，其中蕴含的众多未知生物合成资源亟待开发。为此人们陆续开发了靶向抓取基因组大片段的方法，包括：基于酵母内源同源重组的 TAR 技术^[1-2]，基于大肠杆菌噬菌体同源重组酶的 RecET 技术^[3]，基于 CRISPR/Cas9 体外酶切和 Gibson 体外组装的 CATCH 技术^[4-5]，基于核酸外切酶介导的体外同源重组和大肠杆菌 RecET 蛋白介导的细胞内同源重组联用的 ExoCET 技术^[6]，基于 CRISPR/Cas9 体外酶切和体外 λ 包装相结合的

方法^[7]，基于 Cas12a 体外酶切和大肠杆菌细胞内 Cre-lox 位点特异性重组的 CAPTURE 技术^[8]。除了基因组大片段直接抓取外，人们还开发了一些将大型基因簇拆分为若干 DNA 片段，然后先利用 PCR 扩增这些片段，再将它们拼接起来形成完整基因簇的克隆技术，比如 DiPaC^[9-10]、CPEC^[11]、DNA Assembler^[12-13] 等。然而，近年来人们大多针对次级代谢产物丰富的放线菌等高 GC 含量微生物对这些技术进行条件优化，却忽略了对高 AT 含量大型基因簇克隆技术的开发，导致高 AT 含量基因

簇克隆困难。

在DNA双螺旋结构中，配对碱基之间的氢键通过持续断裂和再生的动态平衡而发生DNA呼吸作用^[14]。由于AT之间的氢键相互作用与GC相比较弱，因此高AT含量DNA片段的呼吸作用较强，导致其中含有较多的单链泡状结构，基因表达相关的蛋白更容易结合在这些区域，使其中基因的转录活性较强。这是导致高AT含量DNA较不稳定和克隆困难的重要原因。为此，Godiska等^[15]建立了利用线性pJAZZ载体在大肠杆菌中克隆高AT含量DNA的方法，但是此载体的最大容量仅为30 kb。前述CATCH等克隆技术虽然能从基因组上抓取大于30 kb的基因组大片段，但是它们没有针对高AT含量DNA进行优化，因此在克隆高AT含量基因簇时遇到了困难，比如CATCH技术无法从AT含量为63%的变形链球菌（*Streptococcus mutans*）基因组中克隆DNA大片段^[16]。中国科学院微生物研究所陈义华教授课题组^[16]针对高AT含量厌氧菌来源次级代谢产物生物合成基因簇克隆的瓶颈，利用变形链球菌UA159菌株能够天然转化外源DNA和具有较高同源重组活性的特点开发的NabLC大片段克隆技术，能将40 kb以内的高AT含量厌氧菌来源的基因簇整合到UA159基因组的特定位置上，进而成功克隆了变形链球菌（AT含量63%）、表皮葡萄球菌（AT含量68%）和梭菌（AT含量72%）的基因簇，并在UA159菌株中进行了这些基因簇的异源表达研究。但是利用NabLC技术克隆的基因簇整合在变形链球菌UA159菌株的基因组上，因此，只能在UA159菌株中进行表达研究。如果基因簇在UA159菌株中不表达，则无法将其转移到其他宿主中进行表达优化。

ExoCET技术是通过联合核酸外切酶（exonuclease）介导的体外同源重组和大肠杆菌RecET蛋白介导的细胞内同源重组，将基因组大片段从限制性酶切或CRISPR/Cas9酶切的基因组中一步抓取到载体上的克隆技术^[6]。ExoCET是在RecET^[3]的基础上发展起来的基因组大片段克隆技术。噬菌体蛋白RecE是5'→3'双链DNA核酸外切酶，RecT是单链DNA退火蛋白。双链线状DNA在RecE的作用下暴露出3'单链末端，然后

RecT介导单链末端通过碱基互补配对进行退火，使DNA片段首尾相接串联起来^[3]。基于上述原理开发的RecET技术可以从基因组上将50 kb以内的目的DNA片段抓取到表达载体上^[3]。RecET技术要求克隆载体和目的DNA片段必须同时进入一个宿主细胞内并相遇才能发生同源重组。受到当前基因组制备方法的限制，基因组DNA在制备过程中容易断裂的特点致使越大的目的DNA片段在制备的基因组中含量越低，转化时其与克隆载体进入同一细胞的概率也就越小，因此克隆效率也就越低。所以RecET技术不能从细菌基因组中克隆大于50 kb的DNA片段^[3]。ExoCET技术靶向抓取基因组大片段分为两个过程，首先体外同源重组将载体和目的DNA片段串联起来提高其一起进入克隆宿主的概率，然后细胞内表达的RecET蛋白将只完成了一端退火的重组DNA分子（占体外反应产物的85%）通过另一端的同源臂进行退火，形成环状重组质粒^[6]。这种体外同源重组和细胞内同源重组相结合的方法，使基因组大片段的克隆效率比RecET技术高80多倍。ExoCET能以超过80%的正确率将>50 kb的非核糖体多肽合成酶基因从发光杆菌基因组克隆到质粒载体上，也能将106 kb的盐霉素聚酮合酶基因簇以4%~8%的正确率从限制性酶切或CRISPR/Cas9酶切的白色链霉菌基因组克隆到BAC载体上^[6]。而且，利用该ExoCET技术克隆的基因簇还能在大肠杆菌中通过丰富的Red/ET重组^[17-18]工具箱进行修饰，以插入基因转移元件^[19-20]、替换启动子^[21-22]、删除/插入调控基因^[23]等实现在多种不同宿主的异源表达，而且还可以进行结构域点突变或模块替换进行组合生物合成以获得更多衍生物^[24-26]。目前尚未有利用ExoCET技术从AT含量>63%的基因组中抓取大片段的报道。由于ExoCET技术通过联合体外同源重组和细胞内同源重组，能够最大限度地促使克隆载体和目的基因组片段同时进入一个宿主细胞并发生同源重组，因此具有很高的克隆效率。如果对ExoCET技术进行优化，使其能够高效克隆高AT含量基因，将非常有利于高AT含量生物的基因组功能研究。

本研究以AT含量为69%的海洋单细胞光合蓝细菌原绿球藻（*Prochlorococcus*）MIT 9301菌株的

基因组为研究对象，探索了利用 ExoCET 技术进行高 AT 含量基因组大片段克隆的可行性，并对相关操作条件进行了优化，成功建立了基于 ExoCET 和单拷贝 BAC 载体的 ExoCET-BAC 策略，实现了大于 80 kb 基因组大片段的目标抓取、11 个片段的高效组装和 4 个基因组大片段的一步目标抓取，为高 AT 含量生物的基因组功能研究提供了高效克隆技术。

1 材料和方法

1.1 材料

1.1.1 菌株、质粒和引物

本研究所用的菌株和质粒见表 1。寡核苷酸引物（由睿博兴科生物技术有限公司合成）见表 2。

表 1 本研究所用的菌株和质粒

Tab. 1 Strains and plasmids used in this study

	Characteristics	Source
Strains		
<i>E. coli</i> GB2005	<i>E. coli</i> DH10B derivatives in which <i>fhuA</i> ::IS2, <i>ArecET</i> , <i>AybcC</i> ; endogenous <i>recET</i> and DLP12 phage <i>ybcC</i> genes are deleted, and <i>fhuA</i> is mutated to make it resistant to T1 phage	Laboratory stock
<i>E. coli</i> GB05-dir	GB05 derivatives expressing full-length RecE/RecT under the arabinose-inducible P _{BAD} promoter	Laboratory stock
Plasmids		
pSC101-BAD-ETgA-tet	Full-length RecE/RecT inducible (P _{BAD} promoter) expression plasmid contains a temperature-sensitive pSC101 origin and a tetracycline resistance gene.	Laboratory stock
pBR322-amp-ccdB-rpsL	A pBR322 plasmid with the ampicillin resistance gene, kanamycin resistance gene, and counterselection genes <i>ccdB</i> and <i>rpsL</i> .	Laboratory stock
pBeloBAC11-hygecdB	A bacterial artificial chromosome (BAC) with the chloramphenicol resistance gene, hygromycin resistance gene and <i>ccdB</i> gene	Laboratory stock

表 2 本研究所用寡核苷酸序列

Tab. 2 Oligonucleotide sequences used in this study

Oligos	Sequence (5'→3')	Purpose	Template
pBR322-5pieces-1	TTTTAGTTTATTTTAAAGACTTTTAAATACTGGATCATTC	Amplification of the pBR322 vector for 5-piece DNA assembly	pBR322-amp-ccdB-rpsL
pBR322-2	GGGAACTGTGGAATCTTAAATTAATAACCTTGTGCGAG		
	GTGTTTAAACGGTGTGGTAGCTCGCGTATT		
F1-1	ACCTCGACAAGGTATTTAAT	Amplification of fragments for DNA assembly	MIT 9301 genomic DNA
F1-2	CAACATTCTTCTTGCGATT		
F2-1	CATACCAATATTTATTATCT		
F2-2	AGACCTAAATAAAATTAAT		
F3-1	ATTTTGATTTATTTGGTTG		
F3-2	TAGCAAACAATACAAAACG		
F4-1	CTTATTAGAGAATTTAAT		
F4-2	TCTGTAATCTTGCTTTGCTCTACC		
pBR322-7pieces-1	TTAAAAACAACAGGTAGAGCAAAAGCAAGAATTAC	Amplification of the pBR322 vector for 7-piece DNA assembly.	pBR322-amp-ccdB-rpsL
	AGAGTTTAAACACAAATGGCAAGGGCTAATG		
F5-1	TGTTGCTTGTGATGAGGCTG	Amplification of fragments for DNA assembly	MIT 9301 genomic DNA
F5-2	AGATTTAAGAAGGATTTTG		
F6-1	GAAATTTAATTGCCGATTT		
F6-2	CACATCCTGCACAATGAAAA		
pBR322-9pieces-1	AATGAGAAGAGAAAAGGGAATTTTCATTGTGCAG	Amplification of the pBR322 vector for 9-piece DNA assembly.	pBR322-amp-ccdB-rpsL
	GATGTGGTTTAAACACAAATGGCAAGGGCTAATG		
F7-1	ATAATGAGAAGAGAAAAGGGA	Amplification of fragments for DNA assembly	MIT 9301 genomic DNA
F7-2	ATTTTGAAGGCCTGGAATTA		

续表

Oligos	Sequence (5'→3')	Purpose	Template
F8-1	GCAAAACTACTTTGCTTAAT		
F8-2	TTTTGATTTAAGTAAAAGAT		
pBR322-11pieces-1	TAATTCCTTATCCAATGTTTATCTTTACTTAAATCA AAAGTTTAAACACAAATGGCAAGGCTAATG	Amplification of the pBR322 vector for 11-piece DNA assembly. Using together with pBR322-2.	pBR322-amp-ccdB-rpsL
F9-1	AGTAATCTTTATCCAATGT		
F9-2	TCAATCTTACTAAAAAAT		
F10-1	ATTCTTGAATAATTTACTCT		
F10-2	GAATGATCCAGTATTA AAAAG		
BAC-11pieces-1	TTTTAGTTATTTTTAAGACTTTTTAATACTGGATCATTC GTTTAAACCTCTAGAGTCGACCTGCAGG	Amplification of the BAC vector for 11-piece DNA assembly.	pBeloBAC11-hygccdB
BAC-11pieces-2	GGGAACTGTGGAATCTTAAATTAATACCTTGTCGAG GTGTTTAAACCGGTACCGAGCTCGAATTCG		
BAC-49kb-1	CCTTAAAAGATTGATTATTTTCAACCATTATTCAATTTT CCCATCATAGATGGCATTAGTTATTCTCCAGGTTAAACG CGGCCGCCTCTAGAGTCGACCTGCAGG	Amplification of the BAC vector to clone the 49 kb fragment	
BAC-49kb-2	GAATTCGAGCTCGGTACCCGGCGCCGCGGTTAAACA TAGATTGTGTATTGGCGTGTAAAAAATAATGGAAGA ATGAATTAATACAGATTTAGTAATACTA		
BAC-21kb-1	TCAATGTCCAGAACTTAAAGAGGTATTACGCCTCATG AAATAGGTCACTTTAGCAATGGGGATAAAGAATTGTTAA ACGCGGCCGCCTCTAGAGTCGACCTGCAGG	Amplification of the BAC vector to clone the 21 kb fragment	
BAC-21kb-2	GAATTCGAGCTCGGTACCCGGCGCCGCGGTTAAACAT TGCTAATCCCCTCCTGCATCTGCAAAGATACTTGTGTA GGAGTAGTGCAACTATAGCTCCAGAAATA		
BAC-82kb-1	TTTTACCTGGCATATCTGGCGGAATAAATACTGACATATT TTCTACAATTCCTAGTAAAGGTACTCCGAGTTGTTAAAC CTCTAGAGTCGACCTGCAGG	Amplification of the BAC vector to clone the 82 kb fragment	
BAC-82kb-2	GAATTCGAGCTCGGTACCCGGTTAAACAATCTTATTGA AGACATTAGTGAGGAAAAAATTACCTCCTAATATCGTT GAAGCAGCCTTGCGCGAAGCT		
BAC-65kb-1	ATTTACCAAATCCAAGGAATGAAGCAGTAGAAAATGAT CTAATTGTTGATAATAAATGCTTTATAGAATTAGGTTAAA CCTCTAGAGTCGACCTGCAGG	Amplification of the BAC vector to clone the 65 kb fragment	
BAC-65kb-2	GAATTCGAGCTCGGTACCCGGTTAAACCTTTATTAAT AGCTTGACCGGACATGATCCATTCTTTGTTATGGCGGACT TTGAAGACTACCTAAACAAAC		
BAC-22kb-1	TATGGGATCAAAAAGAAATCCCAAGAATTGCTCAATTG GTATATTTGGAGTTAATTACGACGGGACATGTTTCGTTAAA CCTCTAGAGTCGACCTGCAGG	Amplification of the BAC vector to clone the 22 kb fragment	
BAC-22kb-2	GAATTCGAGCTCGGTACCCGGTTAAACCCAACTACTC TTGATAATGGCCTATTAGAAGAAGTTGTTGAAGTTGCTAA AAAATACTCCAATAGATGTGAT		
BAC-18kb-A1	ATCAAATAAGGTATCTTTACTCAATTTGTAAAAAAGTTA CTTACTATCCGATAATAGGATTAATCTGAGGCTGTTAAAC CTCTAGAGTCGACCTGCAGG	Amplification of the BAC vector to clone the 18 kb-A fragment	
BAC-18kb-A2	GAATTCGAGCTCGGTACCCGGTTAAACAATGATGAAT GCCATAAGATTTATAAATATAAATGTGCCAGGTTTGCCAA ATAATGATTTGTTGATTCAG		
BAC-54kb-1	CAAGGTAACAACCTGCGAGAGTTAGAATAAGAACGTAA GCAAAAAGCTTCCATAAGATTAAGTAATTAAGTTAGTTTA AACCTCTAGAGTCGACCTGCAGG	Amplification of the BAC vector to clone the 54 kb fragment	

续表

Oligos	Sequence (5'→3')	Purpose	Template
BAC-54kb-2	GAATTCGAGCTCGGTACCCGGTTAAACACACATTTAC ACTTAGTGTTTGAATTTCAAGTCTATTTCCGGTTGAATTA ATTGTTTTTTATAGTAGTTA		
BAC-25kb-1	CACTGTAATTAGTGAAATAAACCCCTTCTGTCTCCTTTT TGATTTCTTCAGTCCATCTTTCTGTGGAGCAAGTTTAAA CCTCTAGAGTCGACCTGCAGG	Amplification of the BAC vector to clone the 25 kb fragment	
BAC-25kb-2	GAATTCGAGCTCGGTACCCGGTTAAACGGTCAACGAA TTTCTAACGGTCCGATTTTCGACGAACAAACGGTAAATT AAAAGAAGGGAACAAATCTTTA		
BAC-20kb-1	TCCCATAACATCAATTAAGCGCTATTTCTATAAGAA ATCTATTATTGAAGCGTTGCCAGAAAGTTCTAAGTTTAAA CCTCTAGAGTCGACCTGCAGG	Amplification of the BAC vector to clone the 20 kb fragment	
BAC-20kb-2	GAATTCGAGCTCGGTACCCGGTTAAACAGAGGAGAA CGAAAAAAGGTAGTTCTCTGTTCACAGGATCTGAGGTG CAATCTCAGGCCAGTGGTCAAGC		
BAC-18kb-B1	AAAATTTATATTAGAAACCATAGTCTCTTTAGTTTACTT TTATTAGTTATTTTAAATCAATAATTAATTCGTTTAAACCT CTAGAGTCGACCTGCAGG	Amplification of the BAC vector to clone the 18kb-B fragment	
BAC-18kb-B2	GAATTCGAGCTCGGTACCCGGTTAAACAGGGTTCTAT AAAAGTTTAAATTAATCAATAAGTGTATTGAAAATCCA CCAAAAAATATAGAGAGCTTA		
BAC-10kb-1	TTGAGATGTAGATATTGTAGTTAACAAAAAAGAGAA GTGTTTAAACGGTTTAAATAAACATAGGGACTATAGTTTAA ACCTCTAGAGTCGACCTGCAGG	Amplification of the BAC vector to clone the 10 kb fragment	
BAC-10kb-2	TAAATATTAATGATATTTCTTTTATTCAATCCTTTACTAT TTGAGCGGTGCAATATGCCATCTTTGGTGTGTTTAAACC GGGTACCGAGCTCGAATTC		
BAC-11kb-1	ACGGAATGCAAGAGGATGGTGCTAGTAACGAAACATAT ACAGCATTAACGGTTTCTATTCAATTTGATAACGGTTTAAA CCTCTAGAGTCGACCTGCAGG	Amplification of the BAC vector to clone the 11 kb fragment	
BAC-11kb-2	GAATTCGAGCTCGGTACCCGGTTAAACCGTTATCAAAT GAATAATAACCATTAATGCCATGTATGTATCGTTGTCGTT ACCATCTTCGATGTTTGCCT		
BAC-7kb-1	TACAGGCATTCGGCAGTTTAAAATTACTGCCGAAAA TGTTTCTCTGATTTCTCAAGAGCTGGAACCTGGTTTAA ACCTCTAGAGTCGACCTGCAGG	Amplification of the BAC vector to clone the 7 kb fragment	
BAC-7kb-2	GAAGTACGCTAATGTTTCGTCCACATTTGCGTCAGCAA CATCGAGATCTCCGACTTCATAACCAGCACTGATGTTTAA ACCGGGTACCGAGCTCGAATTC		
BAC-12kb-1	TAAAAGAATTGCAAAATGAGTCATATATAATTAAGAAAA ATGATCAGGAAGGGGTTATAAAAAGTAATACTGGTTTAAA CCTCTAGAGTCGACCTGCAGG	Amplification of the BAC vector to clone the 12 kb fragment	
BAC-12kb-2	GAATTCGAGCTCGGTACCCGGTTAAACCTTTAATAGAC TCTTTGGCAATGAGTATCGCAGGAACCTCGATTCTGTTT TTTTATCTTACTTCTATGTT		
BAC-67kb-1	TCCTTGATTTAATGTCAGTTGGGATAGTTATTTACCTAA AATGAAAAAATATAAAAAGTTATAGGACATGGTTTAAAC CTCTAGAGTCGACCTGCAGG	Amplification of the BAC vector to clone the 67 kb fragment	
BAC-67kb-2	GAATTCGAGCTCGGTACCCGGTTAAACGGATATCACT CTAAAGATATAAGGAATTTGAAAATATCCCAAAAATCAA TAAGTTGATGTTTGGATAGCT		

续表

Oligos	Sequence (5'→3')	Purpose	Template
BAC-50kb-1	GAGTTCAGACCCAGCAGATATTCCTCGATATAAAGCAG CCTAAATTACAAAAAGCTTGAAAACCTATATAAATGTTTAA ACCTCTAGAGTCGACCTGCAGG	Amplification of the BAC vector to clone the 50 kb fragment	
BAC-50kb-2	GAATTCGAGCTCGGTACCCGGTTTAAACCTACCCAGGG AGAAGACATATCATGTAAATTAGTCTGAAATCCCATTCTT CTAATTCATAGCCTATTGGAT		

1.1.2 培养基

LB培养基 (1 L): 胰蛋白胨 10 g, 酵母提取物 5 g, 氯化钠 1 g。

Pro99培养基 (1 L): 贫营养海水 1 L (0.22 μm 混合纤维素微孔滤膜抽滤, 121 °C, 20 min 高压灭菌), Na₂HPO₄·H₂O [6.90 g/L, 0.22 μm Millipore Express (PES) 膜过滤器过滤除菌] 2 mL, NH₄Cl [42.80 g/L, 0.22 μm Millipore Express (PES) 膜过滤器过滤除菌] 1.6 mL, 微量元素溶液 100 μL。

微量元素溶液: Na₂EDTA·2H₂O 1.17×10⁻⁶ mol/L, FeCl₃·6H₂O 1.17×10⁻⁶ mol/L, ZnSO₄·7H₂O 8.00×10⁻⁹ mol/L, CoSO₄·7H₂O 5.00×10⁻⁹ mol/L, MnCl₂·4H₂O 9.00×10⁻⁸ mol/L, Na₂MoO₄·2H₂O 3.00×10⁻⁹ mol/L, Na₂SeO₃ 1.00×10⁻⁸ mol/L, NiCl₂·6H₂O 1.00×10⁻⁸ mol/L, 0.22 μm Millipore Express (PES) 膜过滤器过滤除菌。

1.1.3 主要试剂和仪器

试剂: PrimeSTAR Max DNA Polymerase、Tks Gflex™ DNA Polymerase 和 DNA Markers 购自 Takara 公司; 蛋白酶 K 购自 Roche 公司; 蛋白胨和酵母粉购自 OXOID 公司; 琼脂粉、琼脂糖、25×TAE 和 10×TBE 购自 Solarbio 公司; RNase A 购自 Thermo Scientific 公司; 限制性内切酶、Gibson 组装克隆试剂盒和 T4 DNA 聚合酶购自 New England Biolabs 公司; DNA 纯化回收试剂盒和溶菌酶购自 天根生化科技有限公司; DNA 提取液 (苯酚: 氯仿: 异戊醇=25: 24: 1) 购自北京鼎国昌盛生物技术有限责任公司。

仪器: 杭州柏恒科技有限公司 PCR 仪, 杭州奥盛仪器有限公司离心机和恒温混匀仪, Eppendorf 转化仪, 美国 Quawell 超微量分光光度计。

1.2 方法

1.2.1 原绿球藻 MIT 9301 基因组制备

将 50 mL 培养 7 d 的 MIT 9301 培养液 8000 r/min

离心 10 min, 弃上清。用 1 mL 灭菌的 ddH₂O 重悬菌体, 吹打混匀并转入 2 mL EP 管中, 10 000 r/min 离心 5 min, 弃上清。加入 400 μL SET 溶液 (75 mmol/L NaCl, 25 mmol/L EDTA, 20 mmol/L Tris, pH 8.0) 吹打混匀菌体。加入 30 μL 溶菌酶 (20 mg/mL), 颠倒混匀, 37 °C 水浴 2 h。加入 30 μL 蛋白酶 K (20 mg/mL), 颠倒混匀后加入 40 μL 10% SDS, 颠倒混匀, 50 °C 水浴 1 h, 直至溶液澄清。加入 100 μL 5 mol/L NaCl, 颠倒混匀。加入 600 μL DNA 提取液 (苯酚: 氯仿: 异戊醇=25: 24: 1), 颠倒混匀至溶液呈乳白色, 10 000 r/min 离心 30 min, 用剪掉枪尖的 200 μL 枪头取 500 μL 上清至新的 2 mL EP 管中。加入 35 μL 3 mol/L 醋酸钠 (pH 7.5) 颠倒混匀, 然后加入 1.2 mL 的无水乙醇颠倒混匀后, 管中会出现絮状 DNA。用 200 μL 枪头挑取絮状 DNA 并转入装有 1 mL 75% 乙醇的 1.5 mL EP 管中, 10 000 r/min 离心 2 min, 弃上清, 倒置于吸水纸上, 室温晾干至基因组呈透明状。加入 350 μL ddH₂O 溶解基因组 DNA, 测定浓度后 4 °C 保存备用。

MIT 9301 基因组酶切: 400 μL 酶切体系含有 10 μg 基因组 DNA, 40 μL 10 × NEB Cutsmart 缓冲液, 10 U *Pme* I, 2 μL RNaseA (10 mg/mL), 37 °C 酶切 4 h。酶切结束后用 DNA 提取液 (苯酚: 氯仿: 异戊醇=25: 24: 1) 对酶切产物进行抽提, 然后进行乙醇沉淀 (具体过程同基因组提取步骤)。

1.2.2 PCR 制备克隆载体和组装用 DNA 片段

PCR 扩增 pBR322 克隆载体, 以 pBR322-amp^r-ccdB-rpsL^[19] 为模板, PCR 产物长度为 2114 bp。PCR 扩增原绿球藻基因组片段, PCR 长度为 3000 bp。扩增 pBeloBAC11 克隆载体, 以 pBeloBAC11-hygccdB 为模板, PCR 产物长度为 7501 bp。PCR 反应所用的酶为 PrimeSTAR Max DNA Polymerase。PCR 所用引物见表 2。PCR 反应体系 (50 μL): 25 μL PrimeSTAR Max Premix (2×); Primer 1、Primer 2 各 2 μL; 模板 200 ng; ddH₂O 补齐至 50 μL。PCR 反

应条件：94 °C 2 min; 98 °C 10 s, T_m 5 s, 72 °C 5 s/kb, 30 cycles; 72 °C 5 min; 4 °C hold.

1.2.3 ExoCET 直接克隆和多片段组装

① T4 聚合酶 (T4pol) 反应体系 载体 200 ng, PCR 扩增基因组片段 200 ng 或 *Pme* I 酶切的 MIT 9301 基因组 DNA 10 μ g, 2 μ L 10 \times NEB Buffer 2.1 和 0.13 μ L T4 DNA 聚合酶, 用 ddH₂O 补齐至 20 μ L。混匀后放入 PCR 仪中进行体外反应, 具体程序为: 25 °C 1 h, 75 °C 20 min, 50 °C 1 h, 4 °C hold。反应产物在室温下 MF-Millipore 0.025 μ m MCE 混合纤维素酯亲水膜脱盐 40 min 备用。

② Gibson 组装反应体系 DNA 用量与 T4 聚合酶反应相同, DNA 混合液与 Gibson 组装预混液等比例混匀后放入 PCR 仪中进行体外反应, 具体程序为: 50 °C 1 h, 4 °C hold。反应产物在室温下脱盐 40 min 备用。

③ 电转化感受态细胞制备 转接 40 μ L 过夜培养的含有 pSC101-BAD-ETgA-tet 质粒的 *E. coli* GB05-dir 菌株培养液至 1.3 mL 含有 4 μ g/mL 四环素的 LB 培养基中, 30 °C, 950 r/min 培养 2 h, 之后加入 35 μ L 10% L-阿拉伯糖溶液, 37 °C, 950 r/min 培养 40 min。9500 r/min, 离心 1 min, 收集菌体, 加入 1 mL ddH₂O 重悬后, 9500 r/min, 离心 1 min, 收集菌体, 然后再重复洗一遍。之后将脱盐的 T4 聚合酶反应或 Gibson 组装反应产物与菌体混匀后转入

1 mm 电转杯中, 1350 V 电击。然后将 1 mL LB 培养基加入电转杯中重悬菌体, 再将菌液转入 1.5 mL EP 管中, 复苏 1 h 后, 9500 r/min, 离心 1 min 收集菌体, 涂布到含有相应抗生素的 LB 平板上, 37 °C 培养过夜。

2 结果和分析

2.1 ExoCET 组装高 AT 含量 DNA 片段时, 体外重组使用 Gibson 较 T4 聚合酶能获得更高效率

首先为了测试 ExoCET 组装高 AT 含量 DNA 的可行性, 对 PCR 扩增的 4 个 3 kb MIT 9301 基因组片段 (4 个片段的 AT 含量分别为 69%、69%、73%、70%) 和 1 个 2 kb pBR322 载体, 共计 5 个 DNA 片段, 先在体外分别利用 T4pol 体系和 Gibson 体系进行处理, 然后转化到表达 RecET 重组酶的大肠杆菌中进行 ExoCET 组装 [图 1(a)]。这 5 个 DNA 片段相互之间在末端带有 40 bp 同源臂。结果显示体外同源重组不论使用 T4pol 体系还是 Gibson 体系均获得了 87.5% 的正确率 [图 1(b)], 而且 Gibson+RecET 的 ExoCET 获得了更多的菌落数, 是 T4pol+RecET 的 22 倍 [图 1(c)]。这说明 ExoCET 能高效组装高 AT 含量 DNA 片段。在体外同源重组中, Gibson 反应体系中包含 T5 核酸外切

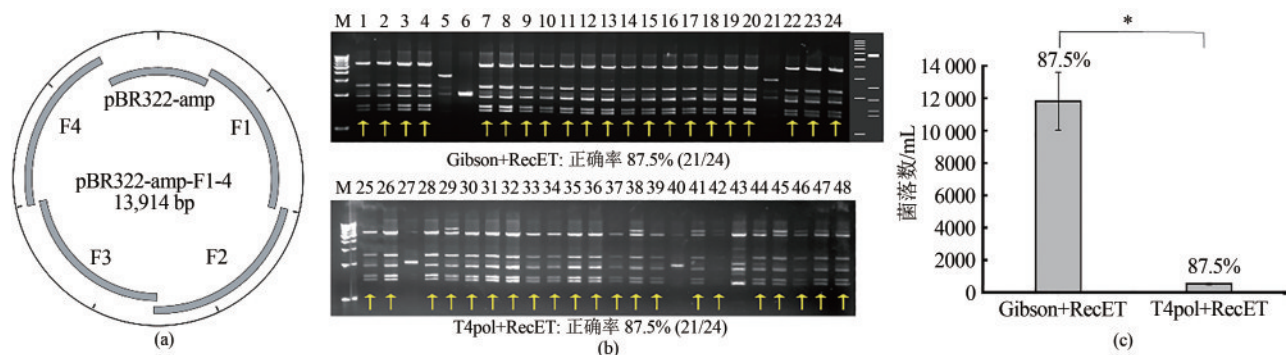


图1 ExoCET 技术组装 5 个 DNA 片段

(a) 5 个片段组装示意图; (b) 组装获得重组质粒的 *Eco*R I 酶切图谱, 胶图右侧为酶切理论图谱, 1~24 为 Gibson+RecET; 25~48 为 T4pol+RecET; (c) Gibson+RecET 和 T4pol+RecET 五片段组装效率比较。采用单因素方差分析差异。 $P < 0.05$ 认为具有显著性差异 (** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$)

Fig. 1 Assembly of 5 DNA fragments using ExoCET

(a) Schematic diagram for the assembly. (b) *Eco*R I digestion map for assembled plasmids. The right side of the gel images is the theoretical map for the restriction digestion. 1-24: Gibson+RecET, and 25-48: T4pol+RecET assembly. (c) Comparison of efficiencies for the assembly. Significance was analyzed by one-way ANOVA, and $P < 0.05$ was considered statistically significant (** $P < 0.001$, ** $P < 0.01$, and * $P < 0.05$)

酶、Phusion DNA 聚合酶和 Taq 连接酶，在这些酶的作用下形成的 DNA 拼接产物是不含缺口的共价键连接的 DNA 分子，而 T4pol 反应获得 DNA 拼接产物是带有缺口的 DNA 分子。高 AT 含量的 DNA 片段在体外同源重组中通过同源臂区碱基互补配对形成的共价键连接的双链 DNA 分子与带缺口的双链 DNA 分子相比更稳定，因此体外重组使用 Gibson 较 T4 聚合酶能获得更高效率。

2.2 单拷贝 BAC 载体更适合高 AT 含量基因的克隆

为了测试 ExoCET 组装更多高 AT 含量 DNA 片段的效率，分别将 PCR 扩增的 6 个、8 个和 10 个 3 kb MIT 9301 基因组片段与 1 个 2 kb pBR322 载体，先在体外用 Gibson 体系进行处理，然后转化到表达 RecET 重组酶的大肠杆菌中进行 ExoCET 组装 [图 2(a)]。结果显示 7 个片段和 9 个片段均成功组

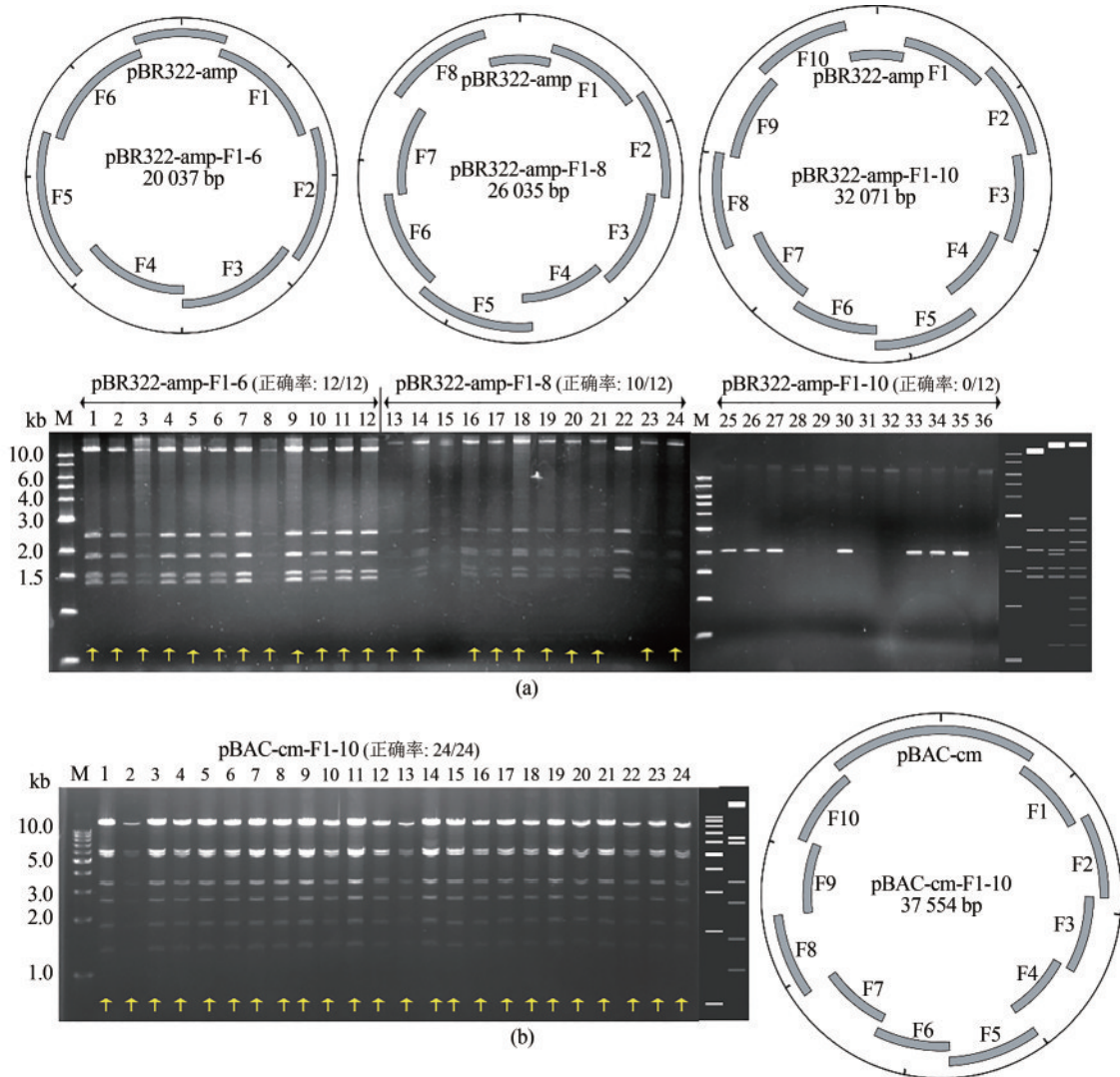


图2 ExoCET 技术组装 7~11 个 DNA 片段

- (a) 利用 pBR322 载体组装的示意图和 *EcoRI* 酶切鉴定图：1~12 为 7 个片段组装；13~24 为 9 个片段组装；25~36 为 11 个片段组装；胶图右侧从左到右依次为 pBR322-amp-F1-6、pBR322-amp-F1-8、pBR322-amp-F1-10 的酶切理论图谱
(b) 利用 BAC 载体组装 11 个片段的示意图和 *EcoRV* 酶切鉴定图，胶图右侧为酶切理论图谱

Fig. 2 Assembly of 7-11 fragments using ExoCET

- (a) Schematic diagrams and *EcoRI* restriction analysis for the assembly using pBR322 vectors. 1-12: 7-piece assembly, 13-24: 9-piece assembly, and 25-36: 11-piece assembly. The right side of the gel image is the theoretical map for the restriction digestion of pBR322-amp-F1-6, pBR322-amp-F1-8 and pBR322-amp-F1-10 from left to right. (b) Schematic diagrams and *EcoRV* restriction analysis for the 11-piece assembly using the BAC vector. The right side of the gel image is the theoretical map for the restriction digestion

装,而且菌落数均大于27 000个/mL。在进行限制性酶切分析时,对相同培养时间并且相同体积的培养液进行质粒提取,然后溶解于相同体积的ddH₂O后,取相同体积溶解好的质粒进行酶切。从胶图中可以看出,9片段组装质粒的浓度明显比7片段组装质粒的浓度低[图2(a)]。推测F7和/或F8片段中的DNA对大肠杆菌有轻微毒性,影响了质粒拷贝数。11个片段的组装失败,可能是由于F9和/或F10片段对大肠杆菌也有毒性,导致正确重组的质粒无法在大肠杆菌中复制。接下来,尝试利用BAC载体与F1-F10片段进行组装,结果获得了大于2200个/mL的菌落数,正确率为100%[图2(b)]。

使用pBR322质粒组装11个片段失败,而使用BAC载体可以成功,推测这是由质粒拷贝数导致的。可能由于F9和/或F10片段的细胞毒性,当克隆载体选用拷贝数约为20的pBR322载体时,其表达量较高,毒性较强,超过细胞耐受水平,导致克隆失败。而选用单拷贝的BAC载体时,其表达量控制在细胞可耐受的较低水平,所以能够克隆成功。此外,由于高AT含量DNA更容易发生呼吸作用(DNA breathing)形成单链DNA构成的泡状结构,使基因表达相关的蛋白更容易结合在这些区域从而促进转录^[14],导致毒性基因更容易表达而影响克隆。另一方面,如果pBR322质粒携带了易发生呼吸作用的DNA片段,其较高的拷贝数会占用宿主细胞大量的基因表达相关蛋白,进而影响宿主基因的转录导致细胞毒性,而单拷贝的BAC载体则可以避免此问题。

2.3 ExoCET成功从原绿球藻MIT 9301基因组上靶向抓取21~82 kb的大片段

为了测试ExoCET抓取高AT含量基因组大片的可行性,利用Pme I将MIT 9301基因组切割成1.4~206 kb的30个DNA片段[图3(a)],然后随机选取其中21 kb、25 kb、49 kb、50 kb、54 kb、65 kb、67 kb、82 kb的8个片段进行靶向抓取实验。首先在体外利用Gibson体系对Pme I酶切的MIT 9301基因组和PCR扩增的BAC载体进行处理,然后转化到表达RecET重组酶的大肠杆菌中进行ExoCET靶向抓取[图3(b)]。BAC载体两端带有目的基因组片段的两端的80 bp同源臂。结果显示ExoCET均能成功靶向抓取这8个21~82 kb

的基因组大片段[图3(c)~(j)],其中50 kb片段的正确率最高,为87.5%;82 kb片段的正确率最低,为4.2%。其余片段的正确率在11%~75%之间。克隆效率的总体趋势是片段越长克隆效率越低,这是由于基因组DNA在制备过程中容易断裂,致使越大的目的DNA片段在制备的基因组中含量越低,因此克隆效率也就越低。此外,长度接近片段的克隆效率也会存在较大差异,比如50 kb片段的正确率为87.5%,而49 kb片段的克隆正确率仅为23.5%,这可能是由于用于克隆49 kb片段的同源臂经过核酸外切酶处理后产生的单链DNA较容易产生二级结构,影响了载体和片段的退火。

2.4 ExoCET从原绿球藻MIT 9301基因组上一步同时抓取4个7~20 kb的大片段

系统开发基因组中的生物合成资源需要能够一步同时抓取多个片段的克隆技术。为了测试ExoCET同时靶向抓取2个片段的效率,随机选取了Pme I酶切MIT 9301基因组产生的2组大小接近的片段,分别为10 kb、11 kb和18 kb、22 kb片段。首先将Pme I酶切的MIT 9301基因组和PCR扩增的2个BAC载体混合(其中2个BAC载体两端分别含有2个对应目的片段两端的80 bp同源臂),并利用Gibson体系在体外进行处理,然后转化到表达RecET重组酶的大肠杆菌中进行ExoCET靶向抓取[图4(a)]。结果显示ExoCET均能实现上述两组片段的靶向抓取。从10 kb、11 kb组筛选平板上随机挑取24个菌落进行限制性酶切分析,发现其中3个含有10 kb片段,10个含有11 kb片段[图4(b)]。从18 kb、22 kb组筛选平板上随机挑取12个菌落进行限制性酶切分析,发现其中1个含有18 kb片段,6个含有22 kb片段[图4(c)]。每组中两个大小接近片段的克隆效率差异较大,具体原因有待进一步解析。

接下来,进一步测试了ExoCET同时靶向抓取4个片段的效率,此时随机选取了Pme I酶切MIT 9301基因组产生的大小差异较大的4个片段,分别为7 kb、12 kb、18 kb和20 kb。首先将Pme I酶切的MIT 9301基因组和PCR扩增的4个BAC载体混合(其中4个BAC载体两端分别含有4个对应目的片段两端的80 bp同源臂),并利用Gibson体系在

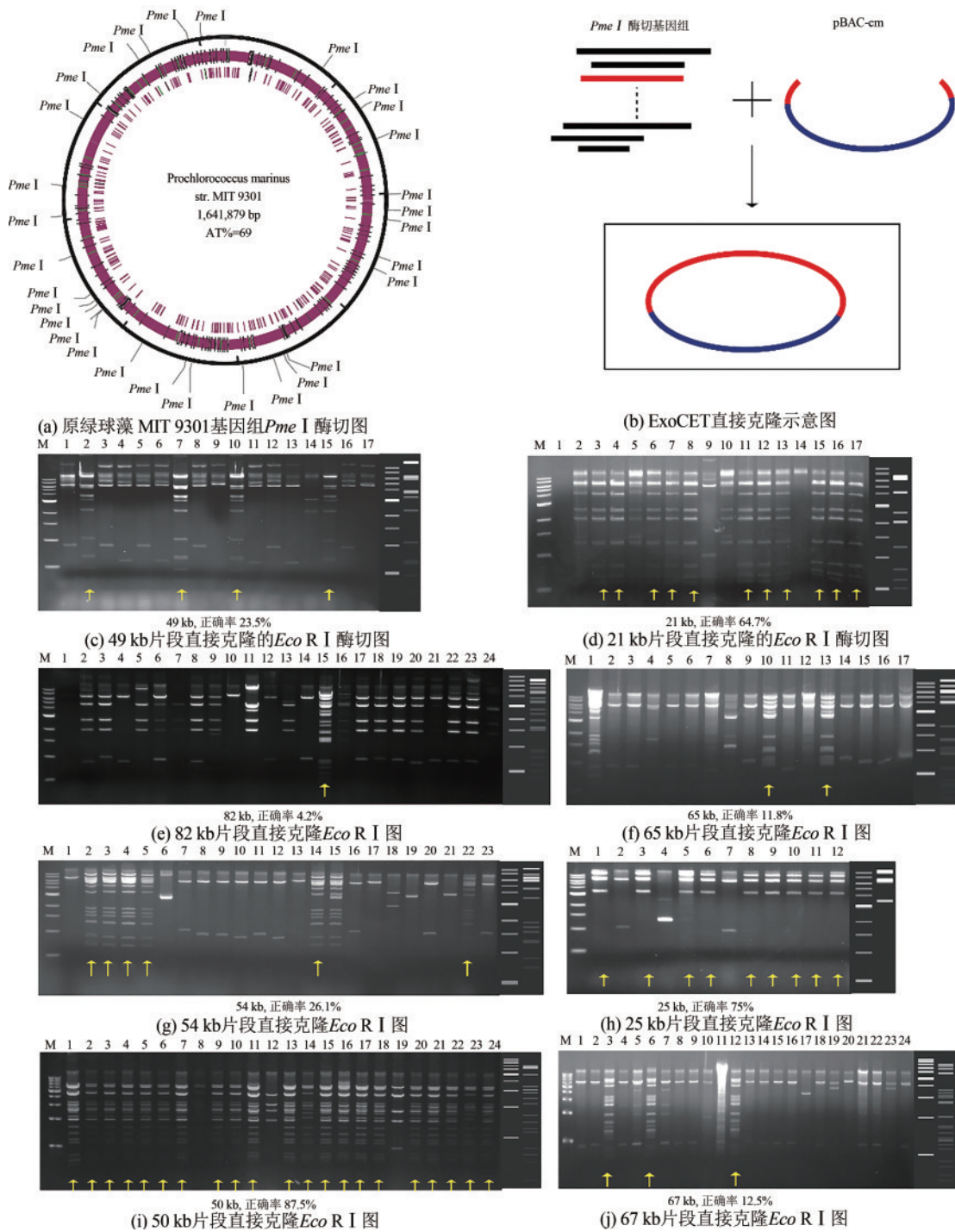


图3 利用 ExoCET-BAC 策略进行直接基因克隆

Fig. 3 ExoCET-BAC strategy for direct gene cloning

(a) *Pme I* restriction map of *Prochlorococcus* MIT 9301 genome, (b) Strategy for ExoCET direct cloning, and *Eco R I* restriction map of recombinants from the 49 kb cloning (c), 21 kb cloning (d), 82 kb cloning (e), 65 kb cloning (f), 54 kb cloning (g), 25 kb cloning (h), 50 kb cloning (i), and 67 kb cloning (j), respectively. Yellow arrows indicate correct clones. The right side of the (c)-(j) gel images is the theoretical map for the restriction digestion

体外进行处理，然后转化到表达 RecET 重组酶的大肠杆菌中进行 ExoCET 靶向抓取 [图 5(a)]。结

果显示 ExoCET 能实现上述 4 个片段的同时靶向抓取。从筛选平板上随机挑取 12 个菌落进行限制性

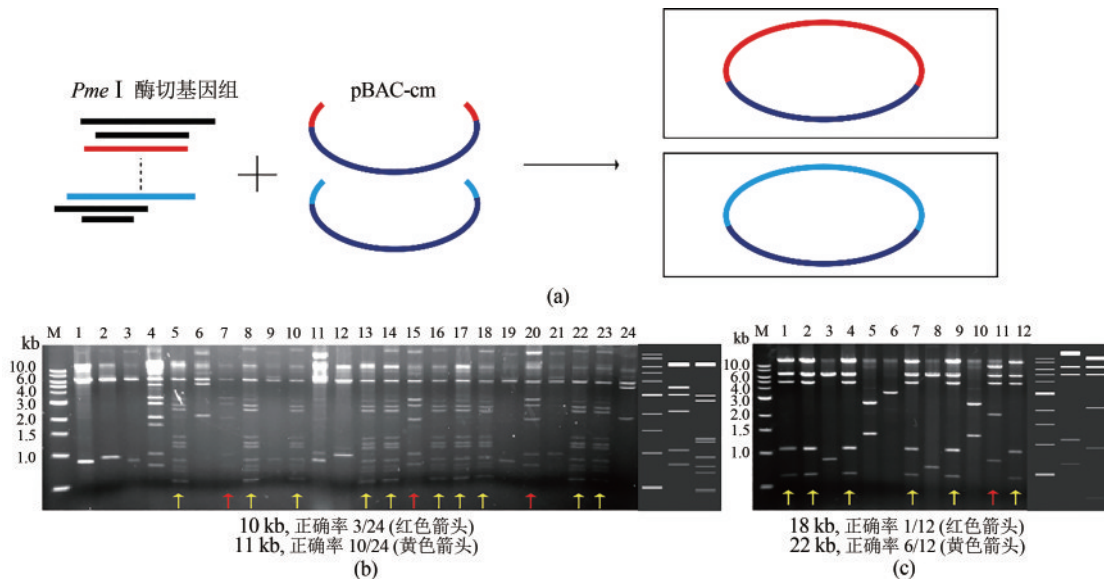


图4 利用 ExoCET-BAC 策略进行两个片段一步抓取

(a) 一步抓取两个片段示意图；(b) 直接克隆 10 kb、11 kb 片段 *EcoR* I 酶切图（胶图右侧分别为直接克隆 10 kb 和 11 kb 基因组大片的酶切理论图谱）；(c) 直接克隆 18 kb、22 kb 片段 *EcoR* I 酶切图（胶图右侧从左到右依次为直接克隆 22 kb 和 18 kb 基因组大片的酶切理论图谱）

Fig. 4 ExoCET-BAC strategy for capturing two DNA fragments simultaneously

(a) Schematic diagram for the strategy. (b) *EcoR* I restriction map for the direct cloning of 10 kb and 11 kb fragments. The right side of the gel image is the theoretical restriction map for the direct cloning of 10 kb and 11 kb fragments. (c) *EcoR* I restriction map for the direct cloning of 18 kb and 22 kb fragments. The right side of the gel image is the theoretical restriction map for the direct cloning of 18 kb and 22 kb fragments from left to right

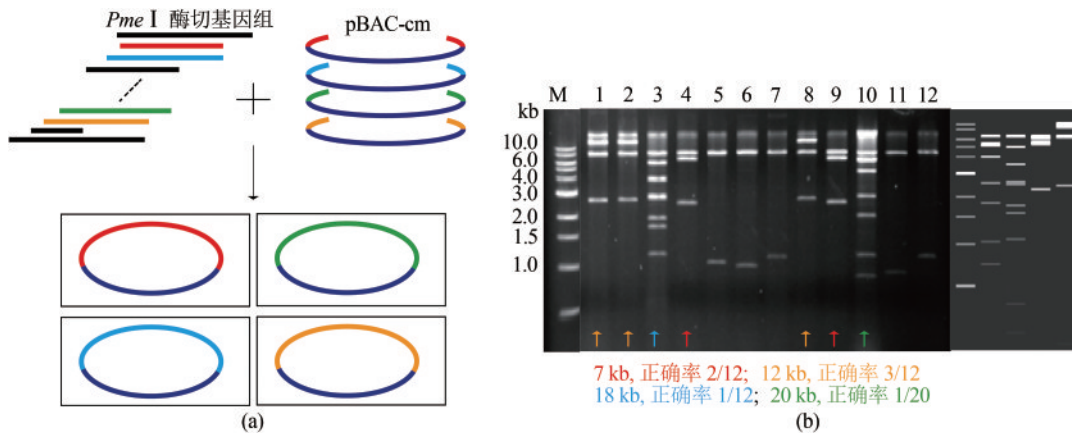


图5 利用 ExoCET-BAC 策略一步抓取 4 个片段

(a) 一步抓取 4 个片段示意图；(b) 直接克隆 7 kb、12 kb、18 kb 和 20 kb 的 *EcoR* I 酶切结果图（胶图右侧从左到右依次为直接克隆 20 kb、18 kb、7 kb 和 12 kb 基因组大片的酶切理论图谱）

Fig. 5 ExoCET-BAC strategy for capturing four DNA fragments simultaneously

(a) Schematic diagram for the strategy. (b) *EcoR* I restriction map for the direct cloning of 7 kb, 12 kb, 18 kb and 20 kb fragments. The right side of the gel image is the theoretical restriction map for the direct cloning of 20 kb, 18 kb, 7 kb and 12 kb fragments from left to right

酶切分析，发现其中 2 个含有 7 kb 片段，3 个含有 12 kb 片段，1 个含有 18 kb 片段，1 个含有 20 kb 片段 [图 5(b)]。

还测试了 >22 kb 片段的多片段同时靶向抓取，但是均失败了，因此以目前的方法 ExoCET 只能实现 4 个小于 20 kb 片段的的同时靶向抓取。

3 结论

根据 NCBI 2021 年 6 月公布的基因组数据，已测序的 327 130 个原核生物基因组中，有 99 340 个 (30.36%) 微生物的基因组 AT 含量 ≥ 60%。已测序的 16 037 个真核生物基因组中，有 5274 个 (32.89%)

真核生物基因组 AT 含量 $\geq 60\%$ 。开发这些生物基因组中的生物合成资源需要更通用和便利的克隆技术。虽然已经有很多的直接克隆和多片段组装方法可以对 DNA 大片段进行多片段组装和直接克隆,但是仍然有很多限制,导致高 AT 含量的 DNA 大片段克隆效率低下,或者无法进行克隆。本研究以 AT 含量为 69% 原绿球藻基因组为对象,探讨了利用 ExoCET 技术进行高 AT 含量基因组大片段克隆的可行性,通过优化体外同源重组方式和载体拷贝数建立了 ExoCET-BAC 策略。随机选取的 16 个 7~82 kb 的基因组大片段全部靶向抓取成功,其中对 82 kb 的 DNA 大片段进行靶向抓取的效率为 4.2%。而且以 100% 的正确率组装了 11 个 3 kb 的 DNA 片段。ExoCET-BAC 技术还能从原绿球藻基因组上一步同时抓取 4 个 7~20 kb 的基因组大片段,进一步证明了该技术的高效率。

最近, David J. Newman 和 Gordon M. Cragg 对 1981~2019 年所有 FDA 批准的小分子药物的分析显示,这些化合物中有 34% 是天然产物或天然产物衍生物,其中相当一部分是由微生物产生的^[27]。目前次级代谢产物研究最深入的微生物包括放线菌、黏细菌等^[28]的基因组 AT 含量均 $\leq 40\%$ 。链球菌、葡萄球菌和梭菌等高 AT 含量微生物也含有丰富的有待开发的次级代谢产物生物合成基因簇^[16, 29]。ExoCET-BAC 技术将为这些生物合成资源的开发提供新手段。

随着合成生物学的不断发展,对生物基因组进行设计合成越来越受到人们关注^[30-35],该研究建立的 ExoCET-BAC 技术不仅有利于高 AT 含量生物基因组“自上而下”的功能解析,还有利于高 AT 含量生物基因组“自下而上”的从头合成、组装和替换,丰富了合成生物学的使能技术工具箱。

致谢:感谢山东省泰山学者项目 tsqn201812008 和山东大学齐鲁青年学者项目对本工作的支持。

参 考 文 献

- [1] YAMANAKA K, REYNOLDS K A, KERSTEN R D, et al. 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.
- [2] LEE N C O, 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.
- [3] FU Jun, BIAN Xiaoying, HU Shenbiao, et al. Full-length RecE enhances linear-linear homologous recombination and facilitates direct cloning for bioprospecting[J]. Nature Biotechnology, 2012, 30(5): 440-446.
- [4] JIANG Wenjun, ZHAO Xuejin, GABRIELI T, et al. Cas9-Assisted Targeting of Chromosome segments CATCH enables one-step targeted cloning of large gene clusters[J]. Nature Communications, 2015, 6: 8101.
- [5] JIANG Wenjun, ZHU Ting F. Targeted isolation and cloning of 100-kb microbial genomic sequences by Cas9-assisted targeting of chromosome segments[J]. Nature Protocols, 2016, 11(5): 960-975.
- [6] WANG Hailong, LI Zhen, JIA Ruonan, et al. ExoCET: exonuclease in vitro assembly combined with RecET recombination for highly efficient direct DNA cloning from complex genomes[J]. Nucleic Acids Research, 2018, 46(5): e28.
- [7] TAO Weixin, CHEN Li, ZHAO Chunhua, et al. *In vitro* packaging mediated one-step targeted cloning of natural product pathway[J]. ACS Synthetic Biology, 2019, 8(9): 1991-1997.
- [8] ENGHAD B, HUANG Chunshuai, GUO Fang, et al. Cas12a-assisted precise targeted cloning using *in vivo* Cre-lox recombination[J]. Nature Communications, 2021, 12(1): 1171.
- [9] GREUNKE C, DUELL E R, D'AGOSTINO P M, et al. Direct Pathway Cloning (DiPaC) to unlock natural product biosynthetic potential[J]. Metabolic Engineering, 2018, 47: 334-345.
- [10] D'AGOSTINO P M, GULDER T A M. Direct pathway cloning combined with sequence- and ligation-independent cloning for fast biosynthetic gene cluster refactoring and heterologous expression[J]. ACS Synthetic Biology, 2018, 7(7): 1702-1708.
- [11] QUAN Jiayuan, TIAN Jingdong. Circular polymerase extension cloning for high-throughput cloning of complex and combinatorial DNA libraries[J]. Nature Protocols, 2011, 6(2): 242-251.
- [12] SHAO Zengyi, ZHAO Hua, Zhao Huimin. DNA assembler, an *in vivo* genetic method for rapid construction of biochemical pathways[J]. Nucleic Acids Research, 2009, 37(2): e16.
- [13] SHAO Zengyi, ZHAO Huimin. Manipulating natural product biosynthetic pathways *via* DNA assembler[J]. Current Protocols in Chemical Biology, 2014, 6(2): 65-100.
- [14] VON HIPPEL P H, JOHNSON N P, MARCUS A H. Fifty years of DNA “breathing”: reflections on old and new approaches[J]. Biopolymers, 2013, 99(12): 923-954.
- [15] GODISKA R, MEAD D, DHODDA V, et al. Linear plasmid vector for cloning of repetitive or unstable sequences in *Escherichia coli*[J]. Nucleic Acids Research, 2009, 38(6): e88.

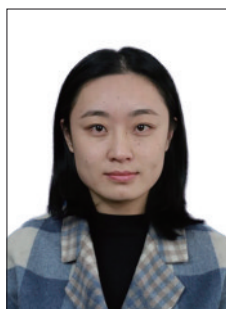
- [16] HAO Tingting, XIE Zhoujie, WANG Min, et al. An anaerobic bacterium host system for heterologous expression of natural product biosynthetic gene clusters[J]. *Nature Communications*, 2019, 10: 3665.
- [17] ZHANG Youming, BUCHHOLZ F, MUYRERS J P P, et al. A new logic for DNA engineering using recombination in *Escherichia coli*[J]. *Nature Genetics*, 1998, 20(2): 123-128.
- [18] ZHANG Youming, MUYRERS J P, TESTA G, et al. DNA cloning by homologous recombination in *Escherichia coli*[J]. *Nature Biotechnology*, 2000, 18(12): 1314-1317.
- [19] WANG Hailong, LI Zhen, JIA Ruonan, et al. RecET direct cloning and Red α recombineering of biosynthetic gene clusters, large operons or single genes for heterologous expression[J]. *Nature Protocols*, 2016, 11(7): 1175-1190.
- [20] CHAI Yi, SHAN Shiping, WEISSMAN K J, et al. Heterologous expression and genetic engineering of the tubulysin biosynthetic gene cluster using red/ET recombineering and inactivation mutagenesis[J]. *Chemistry & Biology*, 2012, 19(3): 361-371.
- [21] SONG Chaoyi, LUAN Ji, CUI Qingwen, et al. Enhanced heterologous spinosad production from a 79-kb synthetic multi-operon assembly[J]. *ACS Synthetic Biology*, 2019, 8(1): 137-147.
- [22] FU Jun, WENZEL S C, PERLOVA O, et al. Efficient transfer of two large secondary metabolite pathway gene clusters into heterologous hosts by transposition[J]. *Nucleic Acids Research*, 2008, 36(17): e113.
- [23] ABBASI M N, FU Jun, BIAN Xiaoying, et al. Recombineering for genetic engineering of natural product biosynthetic pathways[J]. *Trends in Biotechnology*, 2020, 38(7): 715-728.
- [24] SONG Chaoyi, LUAN Ji, LI Ruijuan, et al. RedEx: a method for seamless DNA insertion and deletion in large multimodular polyketide synthase gene clusters[J]. *Nucleic Acids Research*, 2020, 48(22): e130.
- [25] WANG Hailong, BIAN Xiaoying, XIA Lihui, et al. Improved seamless mutagenesis by recombineering using ccdB for counterselection[J]. *Nucleic Acids Research*, 2013, 42(5): e37.
- [26] BIAN Xiaoying, PLAZA A, YAN Fu, et al. Rational and efficient site-directed mutagenesis of adenylation domain alters relative yields of luminide derivatives in vivo[J]. *Biotechnology and Bioengineering*, 2015, 112(7): 1343-1353.
- [27] NEWMAN D J, CRAGG G M. Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019[J]. *Journal of Natural Products*, 2020, 83(3): 770-803.
- [28] HUO Liujie, HUG J J, FU Chengzhang, et al. Heterologous expression of bacterial natural product biosynthetic pathways[J]. *Natural Product Reports*, 2019, 36(10): 1412-1436.
- [29] LI J S, DU Yongle, GU Di, et al. Discovery and biosynthesis of clostyrylpyrones from the obligate anaerobe *Clostridium roseum*[J]. *Organic Letters*, 2020, 22(21): 8204-8209.
- [30] ZHANG Weimin, ZHAO Guanghou, LUO Zhouqing, et al. Engineering the ribosomal DNA in a megabase synthetic chromosome[J]. *Science*, 2017, 355(6329): eaaf3981.
- [31] XIE Zexiong, LI Bingzhi, MITCHELL L A, et al. "Perfect" designer chromosome V and behavior of a ring derivative[J]. *Science*, 2017, 355(6329): eaaf4704.
- [32] WU Yi, LI Bingzhi, ZHAO Meng, et al. Bug mapping and fitness testing of chemically synthesized chromosome X[J]. *Science*, 2017, 355(6329): eaaf4706.
- [33] SHEN Yue, WANG Yun, CHEN Tai, et al. Deep functional analysis of synII, a 770-kilobase synthetic yeast chromosome[J]. *Science*, 2017, 355(6329): eaaf4791.
- [34] GIBSON D G, GLASS J I, LARTIGUE C, et al. Creation of a bacterial cell controlled by a chemically synthesized genome[J]. *Science*, 2010, 329(5987): 52-56.
- [35] FREDENS J, WANG Kaihang, DE LA TORRE D, et al. Total synthesis of *Escherichia coli* with a recoded genome[J]. *Nature*, 2019, 569(7757): 514-518.



通讯作者: 张友明(1964—),男,教授,博士生导师。研究方向为研究方向为基因组编辑与合成生物学等。
E-mail: zhangyouming@sdu.edu.cn



通讯作者: 王海龙(1984—),男,教授,博士生导师。研究方向为微生物基因编辑与药物合成生物学等。
E-mail: wanghailong@sdu.edu.cn



第一作者: 姜娟娟(1990—),女,博士研究生。研究方向为微生物基因编辑与药物合成生物学等。
E-mail: chanjuanjiang@163.com