论文

染色体上大片段 DNA 的直接克隆及移植

朱莹^{①②}, 杨艳^{①②}, 邓平平^{①③}, 黄勇^④, 倪孟祥^{②*}, 方宏清^{①③*}

① 中国人民解放军军事医学科学院生物工程研究所, 北京 100071;

② 中国药科大学生命科学与技术学院, 南京 210009;

③ 安徽大学健康科学研究院, 合肥 230039;

④ 中国人民解放军军事医学科学院微生物与流行病研究所, 北京 100071

* 联系人, E-mail: Nimx_2000@aliyun.com; fanghongqing@vip.sina.com; fanghq@nic.bmi.ac

收稿日期: 2015-06-15; 接受日期: 2015-09-07; 网络版发表日期: 2015-12-29

国家重点基础研究发展计划(批准号: 2011CBA00800)和国家自然科学基金(批准号: 81373286)资助

摘要 应用 Red 重组体内直接克隆染色体上大于 10 kb 的 DNA 序列非常困难. 利用一种抗性拆分-融合策略成功地从供体菌染色体上直接克隆了约 50 kb 的 DNA 序列, 并将其移植到受体菌的染色体上. 首先将抗性基因(如 *cat*)的互补片段 *cmb* 定向整合到染色体上拟克隆区域的一侧, 然后将含有抗性基因另一互补片段 *cma* 的线性质粒载体电转到细胞内进行定向捕获, 通过抗性筛选获得克隆. 使用该方法, 成功地将供体菌 DH1-Z 基因组上约 48 kb 的大片段 DNA 捕获至质粒载体上, 捕获阳性率达 80%. 然后, 将该质粒转化到受体菌 DH36, 利用双链断裂促进同源重组策略, 替换受体菌染色体上对应区域, 一次敲除基因组上 4 个非必需区, 获得重组菌 DH40. 该方法也可用于微生物天然产物生物合成途径的异源表达、基因组组装和大片段 DNA 的功能研究.

关键词 Red 同源重组, 抗性拆分-融合, 直接克隆, 移植

Red 重组广泛应用于基因敲除、基因突变和基因置换^[1-5]. 利用依赖同源序列的 Red 重组技术可以在细胞内直接克隆目的 DNA 到克隆载体上^[6-9]. 根据目的 DNA 来源是环状 DNA 还是线性 DNA, 可以将重组克隆方法分为两类^[7]: 线性加环状同源重组法(linear plus circular homologous recombination, LCHR)和线性加线性同源重组法(linear plus linear homologous recombination, LLHR). 利用 LCHR 方法可以轻松地从 BAC(Bacterial artificial chromosome)质粒上直接克隆约 30 kb 的目的 DNA 片段, 或从染色体上直

接克隆 3 kb 的目的 DNA 片段^[6]. 由于染色体上 DNA 序列极其复杂, 当直接克隆的序列长度超过 10 kb 时, 几乎难以获得正确的阳性克隆.

为从染色体上直接克隆几十 kb 的 DNA 序列, 采用了一种抗性互补策略(图 1). 该策略将线性质粒载体上的抗性标记 *cat* 基因分为两部分: *cma*(含 135 bp 与 *cmb* 同源的同源臂)和 *cmb*. 它们每一部分单独都没有抗性, 仅当融合在一起时才会恢复抗性. 首先, 将 *cmb* 和完整的 *kan* 抗性标记基因(作为这一步的筛选标记)一起定向整合到染色体上拟克隆区域的一侧.

引用格式: 朱莹, 杨艳, 邓平平, 等. 染色体上大片段 DNA 的直接克隆及移植. 中国科学: 生命科学, 2016, 46: 183-191

Zhu Y, Yang Y, Deng P P, et al. Direct cloning and transplanting of large DNA fragments from *Escherichia coli* chromosome. Sci Sin Vitae, 2016, 46: 183-191, doi: 10.1360/N052015-00182

图1 抗性互补策略直接克隆大片段 DNA 的原理

然后, 将一端含有 *cma* 的线性质粒载体通过电转方式导入细胞内, 该线性质粒载体另一端含有与染色体上拟克隆区域另一侧同源的 50 bp 序列. 拟克隆区域一侧的 *cmb* 与线性质粒载体的 *cma* 通过 Red 同源重组融合形成完整的 *cat* 基因, 借助 Cm^R 抗性筛选大大提高了筛选效率. 应用该策略成功克隆了大肠杆菌 DH1-Z 染色体上长度为 48.2 kb 的 Z 单元(该区域中 4 个非必需区已被无痕删除). 然后将重组质粒导入受体菌 DH36 细胞内, 利用双链断裂促进同源重组技术使 Z 单元定向置换受体菌染色体上与 Z 单元部分同源的对应区域, 获得了基因组上 4 个非必需区同时被删除的删减菌 DH40. 本方法可用于基因组组装及基因组功能研究.

1 材料与方法

1.1 菌株、质粒、引物、酶和试剂

实验中所用菌株及质粒见表 1. 所有引物(表 2)由上海生工生物工程公司合成.

内切酶 *Dpn* I 购自 NEB 公司(美国); 其余酶及 DNA Marker 均购自宝生物工程有限公司(大连). 离心柱型质粒小提试剂盒和离心柱型细菌基因组 DNA 提取试剂盒购自天根生化科技有限公司(北京). DNA 片段快速胶回收试剂盒购自北京康为世纪生物科技有限公司. 所有抗生素购自 Amresco 公司(美国).

LB 培养基: 蛋白胨(Oxoid, 英国)10 g/L, 酵母粉(Oxoid) 5 g/L, NaCl 10 g/L, pH 7.0. LB 固体培养基需另加琼脂粉 20 g/L. 蔗糖 LB 固体培养基还需另加蔗糖 60 g/L, 但不加 NaCl. 需要时, 培养基中添加抗生素的浓度为: 氨苄青霉素(Amp) 100 mg/L, 氯霉素

(Cm) 50 mg/L, 卡那霉素(Kan) 50 mg/L, 四环素(Tet) 10 mg/L.

1.2 构建 *kan* 抗性基因和互补片段 *cmb* 的融合基因 *KCb*

以质粒 pKD4 为模板, K-F 与 K-R 为引物 PCR 扩增约 1.5 kb 的 *kan* 抗性基因片段; 以质粒 p15AD2IC-SacB 为模板, ICB-F 与 ICB-R 为引物 PCR 扩增约 0.6 kb 的 *cmb* 片段. 引物设计时需引入 *Eco*R I, *Bam*H I 酶切位点和部分重叠序列. 再以 *kan* 片段和 *cmb* 片段为模板, 以 K-F 与 ICB-R 为引物重叠 PCR 扩增约 2.1 kb 的融合基因 *KCb*. 后者再以 *Eco*R I 与 *Bam*H I 双酶切后, 与同样经双酶切处理的 pUC19 载体连接, 连接产物转化 *E. coli* DH5 α , 获得质粒 pUC19-*KCb*.

1.3 利用抗性互补策略直接克隆 *E. coli* DH1-Z 染色体上 Z 单元序列

Z 单元为 *E. coli* DH1 染色体上(2344629...2491415)之间的部分序列. 删除 *E. coli* DH1 染色体上 4 个非必需区域: 35 区(2344852...2350586)、36 区(2396224...2478151)、37 区(2478889...2480087)和 38 区(2481055...2490788), 构建 *E. coli* DH1-Z(图 2). *E. coli* DH1-Z 染色体上的 Z 单元长度为 48.2 kb, 相比 DH1 基因组上 Z 单元长度减少了约 100 kb. 利用 Red 重组直接克隆 Z 单元的过程分两个步骤.

(1) 利用类似 Gene-Doctoring 方法^[12]将融合基因 *KCb* 插入 *E. coli* DH1-Z 基因组中 Z 单元左侧(2344628...2344629)位置上, 构建 *E. coli* DH1-Z-ZL*KCb*. 以 *E. coli* DH1 基因组为模板, 以 ZL-L5 与 ZL-L3 为引物, PCR 扩增约 0.5 kb 的左同源臂 ZLL;

表 1 本研究使用菌株和质粒

菌株或质粒	具体描述	来源或出处
大肠杆菌(<i>Escherichia coli</i>)菌株		
BL21(AI)	F ⁻ <i>ompT hsdS_B</i> (r _B ⁻ m _B ⁻) <i>gal dcm araB::T7RNAP-tetA</i> . 提供 <i>Tet^r</i> 基因	Invitrogen
DH5α	<i>supE44 lacZΔM15 hsdR17 recA1 endA1 gyrA96 thi-1 relA1</i>	Transgen
DH1	基因组序列号: CP001637	ATCC 33849
DH36	<i>E. coli</i> DH1 无痕敲除 36 区得到的基因组删减菌	本研究
DH1-Z	<i>E. coli</i> DH1 无痕敲除 Z 单元得到的基因组删减菌(图 2)	本研究
DH40	<i>E. coli</i> DH36 无痕敲除 Z 单元得到的基因组删减菌	本研究
质粒		
pKD4	提供卡那霉素抗性基因(<i>kan</i>)	[4]
pAAGBE	以氨苄青霉素抗性基因(<i>amp</i>)替代 pKOBEG 上的氯霉素抗性基因(<i>cat</i>)得到的 ^[10] 在 <i>L</i> -阿拉伯糖诱导下表达 Red 重组酶体系的温敏辅助质粒	本研究
pAAISGBE	在 <i>L</i> -阿拉伯糖诱导下表达 I - <i>Sce</i> I 归巢内切酶的温敏辅助质粒 <i>amp</i>	本研究
pAAICDGBE	在 <i>L</i> -阿拉伯糖诱导下表达 I - <i>Cre</i> I 归巢内切酶的温敏辅助质粒 <i>amp</i>	本研究
p15AD2IC-SacB	供体质粒包含 p15A ori, 两侧含 I - <i>Cre</i> I 归巢内切酶位点的 <i>cat</i> 基因(简称为 IC)以及负筛选标记 <i>sacB</i> (含 <i>sacB</i> 的质粒不能在含蔗糖的培养基中生长) ^[11]	本研究

表 2 本研究使用引物

引物名称	5'→3'
K-F	GGAATTC GTGTAGGCTGGAGCTGCT
K-R	GTGGTATTCACCTCCAGAGCATGGGAATTAGCCATGGTCC
ICB-F	GCTCTGGAGTGAATACCAC
ICB-R	CGGGATCC TACGACTCACTATAGGGAGA
35-55	GCCAAGCTTAACAGTAACCGCCACGAC
35-53	GCCGAATTCATCATCACATAACCATGACATTTTCTGATATTCCTCGGTAACGCCAGATAACGATGCCAAATTTACCGT
35-35	GCCGGATCCATCTGGCGTTACCGGGGA
35-33	GCCCTCGAGGTGGTGGTGCAGTGCTTT
35-0	GGCGGATTGATAAACAATAC
35-1	GAGAGCTACACATTATTGGA
AD3	GGATTGGCTGAGACGAAAA
AD5-Z	AGCATAATTTAATATGCCTGATGGTGTGTCACCATCAGGCATATTCGCGCCTGCGCTGAGCAATAA
M13F	TGTAACGACGCGCCAGT
M13R	CAGGAAACAGCTATGACC
AD-1	GACGCTAGTTATTGCTCAG
36-0	GCATCGCCCATTCAAACAT
36-1	GTGAACAAGCTGGCACAAAT
37-0	TGTGGGTGATTTGCGATGA
37-1	GACGGGACAGGTTGTAGAT
38-0	CCTGGTGGGATTTGACGC
38-1	GCCAGAGCTGCTTCGTAAT
ZL-L5	GCCAAGCTTCCTCT TCCAGTTCCA CCA
ZL-L3	GCCGAATTCATGGTATCGCCTCGCGAA
ZL-R5	GCCGGATCCGAGCCGATTTCATCCCCATA
ZL-R3	GCCCTCGAGAACGATGCCAAATTTACCGT
ZL-0	CCAATCATCA CCTCCTG
ZL-1	CCCATCCGTTGCGTTTGT
35-L5	AACA GTAACCGCCACGAC
35-L3	AGAGACCTCGTGTGCTACAACGATGCCAAATTTACCGT
35-R5	TCTCCTATAGTGAGTCGTAATCTGGCGTTACCG GGGA
35-R3	GTGGTGGTGCAGTGCTTT
38-L5	TGA TGCGTCTGAT GAGCG
38-L3	AGAGACCTCGTGTGCTACACAGGGTCTCCTTACAACAA
38-R5	TCTCCTATAGTGAGTCGTAAATAGCGTTTGC TGTGTTATTG
38-R3	TTTCAGTGGGCCATCAGC
BIT-5	GTAGCACACGAGGTCTCTTAGGGATAACAGGGTAATTAAGCAGCTCTAATGCGCA
BIT-3	TACGACTCACTATAGGGAGAATTACCCTGTTATCCCTA GCGAATTGGGCCCTCTAG
BIK-5	GTAGCACACGAGGTCTCTTAGGGATAACAGGGTAAT GTGAGGCTGGAGCTGCT
BIK-3	TACGACTCACTATAGGGAGAATTACCCTGTTATCCCTAATGGGAATTAGCCATGGTCC
ZR-R3	GCCAAGCTTCTGTTCTGTCGCCCCGCTTT

以 ZL-R5 与 ZL-R3 为引物, PCR 扩增约 0.5 kb 的右同源臂 ZLR. 将酶切片段 ZLL(*EcoR* I 与 *Hind*III), ZLR(*Bam*H I 与 *Xho* I), KCb(*EcoR* I 与 *Bam*H I)和酶切的 p15AD2IC-SacB 载体(*Hind*III与 *Xho* I)混合连接, 构建 p15AD2-ZLLR-KCb-SacB. 将质粒 p15AD2-ZLLR-KCb-SacB 和质粒 pAAICDGBE 共转化至 *E. coli* DH1-Z 中, 经 Amp 与 Kan 双抗平板筛选后, 挑单菌落于含有 Amp 的 LB 培养基中 0.2%L-阿拉伯糖诱导培养 4 h, 菌液 10 倍稀释后涂布于含 Kan 的蔗糖 LB 平板过夜培养. 单菌落用引物 ZL-0 与 ZL-1 进行 PCR 鉴定, 阳性菌落条带约 5 kb, 命名为 *E. coli* DH1-Z-ZLRKCb.

(2) 将辅助质粒 pAAGBE 转化到 *E. coli* DH1-Z-ZLKCb 中, 制备 Red 重组电感受态. 以 *Bam*H I 单酶切后的质粒 p15AD2IC-SacB 为模板, AD5-Z 与 AD3 为引物, PCR 扩增约 2 kb 的线性质粒载体 ADZ, ADZ 两端分别带有与 *cmb* 同源的 135 bp 同源臂以及与基因组拟克隆区域同源的 50 bp 同源臂. 将 ADZ 电转至制备好的电感受态细胞中, 活化后涂布 Cm 抗性蔗糖 LB 平板进行初步筛选. 以两对引物: M13F 与 35-1(阳性条带约 2.5 kb)和 38-0 与 AD-1(阳性条带约 1.5 kb)进行 PCR 筛选. 将阳性重组质粒转化到 *E. coli* DH1 细胞内, 重新提取质粒后进行 PCR 及酶切鉴定, 获得质粒 p15AD2IC-Z.

1.4 将 Z 单元的移植到 *E. coli* DH36 的染色体上

直接克隆至质粒载体的 Z 单元与 *E. coli* DH36 染色体上对应区域高度同源, 实现完整置换目标区域有一定的难度. 在 *E. coli* DH36 染色体被置换区域两侧引入 I -*Sce* I 识别位点, 通过双链断裂促进同源重组的策略, 用 48.2 kb 的 Z 单元置换 *E. coli* DH36 染色体上对应约 146.8 kb 的区域. 具体设计方案如图 5, 在 *E. coli* DH36 的染色体上被置换区域的两侧即 35 区和 38 区整合带有 I -*Sce* I 酶切位点的抗性基因 *IS-kan* 和 *IS-tet*, 诱导表达 I -*Sce* I 归巢内切酶切割染色体上 I -*Sce* I 位点产生双链断链, 中间部分被切除, 两端各有与 Z 单元两侧 500 bp 以上的同源序列. 在 Red 重组酶作用下, Z 单元与断裂位点区域发生同源重组修复双链断裂.

利用电转线性片段的方法在 *E. coli* DH36 染色体上非必需的 35 区与 38 区分别插入带有 I -*Sce* I 识别位点的卡那霉素抗性筛选标记 *IS-kan* 和四环素抗性筛选标记 *IS-tet*. 以 *E. coli* DH1 基因组为模板, 分别

以 35L-5 与 35L-3 为引物 PCR 扩增约 0.5 kb 的左同源臂 35L, 以 35R-5 与 35R-3 为引物 PCR 扩增约 0.5 kb 的右同源臂 35R. 以 pKD4 为模板, BIK-5 与 BIK-3 为引物 PCR 扩增约 1.5 kb 的 *IS-kan* 片段. 以 *E. coli* BL21(AI)基因组为模板, BIT-5 与 BIT-3 为引物 PCR 扩增约 1.5 kb 的 *IS-tet* 片段. 然后, 以 35L, 35R, *IS-kan* 为模板, 以 35L-5 与 35R-3 为引物, 融合 PCR 扩增约 2.5 kb 的片段 35LR-IK; 同法获得约 2.5 kb 的片段 38LR-IT. 应用 Red 重组方法通过电转将抗性标签 *IS-kan* 和 *IS-tet* 分别插入到 *E. coli* DH36 基因组中的 35 区和 38 区位置, 通过抗性筛选及 PCR 鉴定获得 *E. coli* DH36-35LRIK-38LRIT.

供体质粒 p15AD2IC-Z 和辅助质粒 pAAISGBE 共同转化到 *E. coli* DH36-35LRIK-38LRIT 中, 经 Cm 和 Amp 抗性平板筛选. 挑单菌落于含有 Cm 和 Amp 的 LB 培养基中过夜培养, 再转接到仅含有 Amp 的 LB 培养基中 0.2%L-阿拉伯糖诱导培养 4 h, 菌液稀释 10 倍后涂布无抗 LB 平板过夜培养. 于 96 孔板中筛选在无抗 LB 培养基中生长、在 Kan 抗性培养基和 Tet 抗性培养基中都不长的单菌落. 再用 4 对引物: ZL-R5 与 35-1, 36-0 与 36-1, 37-0 与 37-1 和 38-0 与 ZR-R3, 进行 PCR 鉴定, 获得 *E. coli* DH40.

2 结果

2.1 Z 单元的直接克隆及鉴定

为了用已删除 4 个非必需区的 Z 单元替换 *E. coli* DH36 染色体上对应区域, 进一步减小基因组, 首先需要定向捕获 Z 单元. 起初采用常规 Red 重组方法^[6]未能捕获到 Z 单元. 而利用本文的抗性互补策略(图 2)从 *E. coli* DH1-Z 染色体上成功地克隆了长达 48.2 kb 的 Z 单元, 采用两对引物 38-0 与 AD-1, 35-1 与 M13F 进行 PCR 筛选, 阳性克隆高达 89%(17/19)(图 3A 和 B). 为避免残余 *E. coli* DH1-Z 基因组对 PCR 鉴定结果的影响, 将质粒 p15AD2IC-Z 转至 *E. coli* DH1 后重新提取. 以 4 对引物 ZL-R5 与 35-1, 36-0 与 36-1, 37-0 与 37-1, 38-0 与 38-1 对提取质粒进行 PCR 鉴定, 证明所克隆的 Z 单元中 4 个非必需区均已敲除(图略). 根据酶切位点分析, 选择了 *Apa* I, *Sac* I 和 *Xba* I 3 个较合适的内切酶对质粒进行酶切分析. 根据序列预测 p15AD2IC-Z 在三酶切后产生如下条带: 约 17.2, 14.7, 14.4, 3.3, 1.4 kb(图 4A). 从三酶切

图2 抗性互补策略直接克隆Z单元的原理图

IS: I-Sce I 识别位点

电泳结果看出, 因前三条带大小接近难以分开, 在 15 kb 附近有很亮且宽的条带, 3.3 和 1.4 kb 的条带清晰可见(图 4B)。酶切质粒大小鉴定结果结合 PCR 鉴定结果, 说明此克隆得到的质粒 p15AD2IC-Z 为阳性重组质粒。获得质粒 p15AD2IC-Z。

2.3 通过置换Z单元同时删除受体菌染色体上的4个非必需区

将供体质粒 p15AD2IC-Z 和辅助质粒 pAAISGBE

转化入 *E. coli* DH36-35LRIK-38LRIT 中(图 5)。添加阿拉伯糖诱导辅助质粒表达 Red 重组酶体系和 I-Sce I 归巢内切酶, 将诱导后的菌液涂无抗 LB 平板, 37°C 培养过夜。通过抗性筛选获得在含有 Kan 液体 LB 培养基和含有 Tet 液体 LB 培养基中都不长的菌落, 将其菌液从无抗液体 LB 转接至 Kan, Tet, Cm, Amp 4 种单抗液体 LB 中, 再次验证所有抗性都已去除。

对该菌株进一步以 ZL-L5 与 35-1, 36-0 与 36-1, 37-0 与 37-1, 38-0 与 ZR-R3 为鉴定引物, 进行菌液 PCR 鉴定。PCR 结果显示, 阳性菌液样品 4 个区域都是 1.2 kb 左右的阳性条带, 实际结果与理论结果相同(图 6)。接着, 以 35-0 与 35-1, 36-0 与 36-1, 37-0 与 37-1, 38-0 与 38-1 为引物, 进行高保真 PCR 扩增 35LR, 36LR, 37LR, 38LR, 回收后进行测序, 测序结果与预先设计的删除结果完全一致, 进一步确证这 4 个区域被无痕删除。此外, 还用其他 36 个已删除区域引物进行了验证, 也完全正确。另外, 为证明敲除片段内部不存在和预期不符的重组行为, 对 *E. coli* DH36, *E. coli* DH40 菌株进行基因组 DNA 重测序, 使用软件 CLC Genomics Workbench 比对分析的部分结果见图 7。从图中可以看出, *E. coli* DH40 与 *E. coli* DH36 基因组重测序比对出现的空白部分为已敲除区域 35, 36, 37, 38, 其他部分没有发生不符预期的重组。保存此阳性菌, 命名为 *E. coli* DH40, 该菌株中已有 40 个非必需区域被无痕删除, 删除总长度达到 945 kb, 基因组减小了 20.4%。

3 讨论

应用 Red 重组技术可以从 BAC 载体上直接克隆

图3 直接捕获Z单元的重组质粒初步鉴定电泳图

A: 引物 38-0 与 AD-1(阳性条带约 1.5 kb); B: 引物 M13F 和 35-1(阳性条带约 2.5 b). 1~19: 菌液样本; 20: 阴性对照 p15AD2IC-SacB; M: 250 bp DNA Ladder Marker

图 4 质粒 p15AD2IC-Z 酶切位点分析及酶切鉴定图

A: 质粒 p15AD2IC-Z 酶切位点分析图(红色部分为捕获骨架部分); B: 质粒 p15AD2IC-Z 经 *Apa* I, *Sac* I, *Xba* I 三酶切鉴定电泳图. M: DL 15000 DNA Marker; 1: p15AD2IC-Z

图 5 Z 单元转移示意图

几十 kb 的 DNA 片段, 但从染色体上直接克隆长片段 DNA 时, 成功率极低, 这是因为染色体的序列及结构比 BAC 载体的复杂得多^[6,7]. 直接克隆大片段 DNA 的方法可分为 LCHR 和 LLHR 两类^[7]. 基因组和 BAC 质粒上的 DNA 可以直接用 LCHR 法特异性直接克隆目的大片段, 也可以先使用限制性内切酶将基因组或 BAC 降解成包含目的 DNA 的碎片, 与骨架共电转至 *E. coli* Red/ET⁺ 细胞中通过 LLHR 法克隆目的 DNA.

应用 LCHR 法还可以对重组后的质粒进行改造, 如将野生型启动子替换为合成启动子、添加修饰基因等, 进一步改变重组质粒的拷贝数、表达方向和异源表达的宿主范围等^[13-15]. 但上述两种方法重组后都会产生大量骨架自连的空载体以及骨架 PCR 片段与其他非特异性区的重组产物, 也含有抗性标记, 因此假阳性多, PCR 筛选工作量大.

本文利用抗性互补策略对常规 Red 重组克隆方

图 6 Z 单元转移的 PCR 鉴定电泳图

1,2: 引物 ZL-L5 与 35-1; 3,4: 引物 36-0 与 36-1; 5,6: 引物 37-0 与 37-1; 7,8: 引物 38-0 与 ZR-R3. 1,3,5,7: *E. coli* DH40; 2,4,6,8: *E. coli* DH36-35LRK-38LRT/p15AD2IC-Z/pAAISGBE. M: 250 bp DNA Ladder Marker

法进行改进, 结合 LCHR 法成功地从染色体上直接克隆长 48.2 kb 的 DNA 序列, 克隆阳性率达到 80%, 而直接应用 Red 重组技术克隆该片段没有获得成功.

在实际应用中, 根据拟克隆的目的 DNA 片段的大小不同应该选择不同的载体骨架. 研究发现, 以高拷贝载体(如 pUC, pBluescript, 拷贝数 500~700)和中拷贝载体(如 p15A, pBR322, 拷贝数 15~20)为模板构建的载体骨架克隆的大片段 DNA 上限约为 25~80 kb^[15]. 如果要直接克隆长达上百 kb 的序列, 就需要使用单拷贝的 BAC 质粒. BAC 质粒能够承载 200~250 kb

的外源 DNA 序列, 而且单拷贝质粒在宿主体内更稳定^[16].

本研究以 Red 重组体系介导的 LCHR 法为基础, 设计抗性互补拼接方案: 将中拷贝的 p15A 骨架载体上完整的氯霉素抗性基因分为 *cma* 和 *cmb* 两部分, *cmb* 部分插入到基因组, *cma* 部分保留在骨架上, 这两部分都无法独立表达抗性蛋白, 只有两部分发生重组拼接形成完整质粒时才能具有氯霉素抗性, 因此避免了骨架自连产生的假阳性. 应用该策略, 本课题组成功克隆了 *E. coli* DH1-Z 基因组中约 48.2 kb 的 Z 单元序列. 原始 *E. coli* DH1 基因组中该区域中的 4 个非必需区被无痕删除, 致使该 Z 单元区域 DNA 大小由 146.8 kb 降为 48.2 kb. 直接克隆法得到的阳性重组质粒 p15AD2IC-Z 大小约 50 kb, 能够在大肠杆菌体内稳定复制. 接着应用双链断裂诱导同源重组方法, 将 Z 单元转移到基因组减小菌 *E. coli* DH36 中, 并置换了后者对应的同源区域, 获得基因组减小菌 *E. coli* DH40 菌株, 成功地一次删除了 4 个非必需区, 基因组减小了~100 kb.

本文建立的方法可以用于微生物天然产物生物合成途径的异源表达、基因组组装和大片段 DNA 的功能研究. 将未经基因组测序的基因簇直接克隆至表达载体或整合至异源基因组进行异源表达, 从而得知该基因簇的功能^[17]. 在此基础上, 通过同源重组

图 7 *E. coli* DH36, DH40 基因组 DNA 重测序比对

2344629...2491415 为对应于 *E. coli* DH1 染色体上序列区域; 红色部分为已删除区域

对基因簇调节区域进行改造, 提高天然前体量进而提高天然产物表达量. 传统的从头合成组装方法不利于合成几十甚至上百 kb 的混合来源的生物合成途径, 而直接克隆的方法不仅能获得目的基因簇, 将多个大片段拼接成超大基因簇家族^[16], 还能对其进行定点突变、插入、无痕敲除等基因改造, 合成天然产

物的衍生物. 对于实验室难以培养微生物, 可以直接克隆已知天然产物的相应合成路径异源表达, 产量高且安全有效, 节约生产成本^[18]; 该方法也可用于发现潜在的未知的大基因簇, 发现有价值的天然产物; 也可在异源表达天然产物的同时, 通过同源重组改造合成路径, 合成新颖的天然产物衍生物.

参考文献

- Murphy K C. Use of bacteriophage λ recombination functions to promote gene replacement in *Escherichia coli*. *J Bacteriol*, 1998, 180: 2063–2071
- Muyrers J P, Zhang Y, Buchholz F, et al. RecE/RecT and Red α /Red β initiate double-stranded break repair by specifically interacting with their respective partners. *Genes Dev*, 2000, 14: 1971–1982
- Yu D, Ellis H M, Lee E C, et al. An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proc Natl Acad Sci USA*, 2000, 97: 5978–5983
- Datsenko K A, Wanner B L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci USA*, 2000, 97: 6640–6645
- Ellis H M, Yu D, DiTizio T, et al. High efficiency mutagenesis, repair, and engineering of chromosomal DNA using single-stranded oligonucleotides. *Proc Natl Acad Sci USA*, 2001, 98: 6742–6746
- Zhang Y, Muyrers J P, Testa G, et al. DNA cloning by homologous recombination in *Escherichia coli*. *Nat Biotechnol*, 2000, 18: 1314–1317
- Fu J, Bian X, Hu S, et al. Full-length RecE enhances linear-linear homologous recombination and facilitates direct cloning for bioprospecting. *Nat Biotechnol*, 2012, 30: 440–446
- Bian X, Huang F, Stewart F A, et al. Direct cloning, genetic engineering, and heterologous expression of the syringolin biosynthetic gene cluster in *E. coli* through Red/ET recombineering. *Chembiochem*, 2012, 13: 1946–1952
- Cobb R E, Zhao H. Direct cloning of large genomic sequences. *Nat Biotechnol*, 2012, 30: 405–406
- Chaveroche M K, Ghigo J M, d'Enfert C. A rapid method for efficient gene replacement in the filamentous fungus *Aspergillus nidulans*. *Nucleic Acids Res*, 2000, 28: E97
- Li X T, Thomason L C, Sawitzke J A, et al. Positive and negative selection using the tetA-SacB cassette: recombineering and P1 transduction in *Escherichia coli*. *Nucleic Acids Res*, 2013, 41: e204
- Lee D J, Bingle L E, Heurlier K, et al. Gene doctoring: a method for recombineering in laboratory and pathogenic *Escherichia coli* strains. *BMC Microbiol*, 2009, 9: 252
- Wenzel S C, Gross F, Zhang Y, et al. Heterologous expression of a myxobacterial natural products assembly line in pseudomonas via Red/ET recombineering. *Chem Biol*, 2005, 12: 349–356
- Sharan S K, Thomason L C, Kuznetsov S G, et al. Recombineering: a homologous recombination-based method of genetic engineering. *Nat Protoc*, 2009, 4: 206–223
- Venken K J, He Y, Hoskins R A, et al. P[acman]: a BAC transgenic platform for targeted insertion of large DNA fragments in *D. melanogaster*. *Science*, 2006, 314: 1747–1751
- Kotzamanis G, Huxley C. Recombining overlapping BACs into a single larger BAC. *BMC Biotechnol*, 2004, 4: 1
- Ongley S E, Bian X, Neilan B A, et al. Recent advances in the heterologous expression of microbial natural product biosynthetic pathways. *Nat Prod Rep*, 2013, 30: 1121–1138
- Ongley S E, Bian X, Zhang Y, et al. High-titer heterologous production in *E. coli* of lyngbyatoxin, a protein kinase C activator from an uncultured marine cyanobacterium. *ACS Chem Biol*, 2013, 8: 1888–1893

Direct Cloning and Transplanting of Large DNA Fragments from *Escherichia coli* Chromosome

ZHU Ying^{1,2}, YANG Yan^{1,2}, DENG PingPing^{1,3}, HUANG Yong⁴, NI MengXiang²
& FANG HongQing^{1,3}

1 Institute of Biotechnology, Academy of Military Medical Sciences, Beijing 100071, China;

2 Institute of Life Science and Technology, China Pharmaceutical University, Nanjing 210009, China;

3 Institute of Health Sciences, Anhui University, Hefei 230039, China;

4 Institute of Microbiology and Epidemiology, Academy of Military Medical Sciences, Beijing 100071, China

We applied a resistance split-fusion strategy to increase the *in vivo* direct cloning efficiency mediated by Red recombination. The *cat* cassette was divided into two parts: *cma* and *cmb* (has a homologous sequence with *cma*). Each of them has no resistance separately unless the two parts are fused together. The *cmb* fragment was first integrated into a side of the target cloning region, then a linear *cma*-containing plasmid vector was electroporated into the cells to directly capture the target region. Based on this strategy, we successfully cloned a DNA fragment about 48 kb from *E. coli* DH1-Z chromosome with positive frequency of about 80%. Then, combined with double-strand breakage-stimulated homologous recombination, we applied this strategy to replace the corresponding region of *E. coli* DH36 chromosome and knock out four nonessential genomic regions at one time. This strategy provides a powerful tool for the heterologous expression of microbial natural product biosynthetic pathways, for genome assembly and for the function study of large DNA sequence.

Red homologous recombination, resistance split-fusion, target cloning, planting

doi: 10.1360/N052015-00182