



(12) 发明专利申请

(10) 申请公布号 CN 106701808 A

(43) 申请公布日 2017. 05. 24

(21) 申请号 201510456157. X

(22) 申请日 2015. 07. 29

(71) 申请人 深圳华大基因研究院

地址 518083 广东省深圳市盐田区北山工业  
区综合楼

(72) 发明人 郑越

(74) 专利代理机构 北京清亦华知识产权代理事

务所(普通合伙) 11201

代理人 李志东

(51) Int. Cl.

C12N 15/70(2006. 01)

C12N 15/54(2006. 01)

C12N 15/63(2006. 01)

C12N 1/21(2006. 01)

C12R 1/19(2006. 01)

权利要求书2页 说明书12页

序列表7页 附图7页

(54) 发明名称

DNA 聚合酶 I 缺陷型菌株及其构建方法

(57) 摘要

本发明公开一种构建 DNA 聚合酶 I 缺陷型菌株的方法,该方法基于 CRISPR 系统进行 DNA 聚合酶 I 缺陷型菌株构建,该方法以 klenow1、klenow2 和 klenow3 中的至少之一作为靶序列, klenow1、klenow2 和 klenow3 分别为 SEQ ID NO :1 所示 klenow 片段基因序列中的第 30-52 位、第 370-390 位和第 1048-1068 位中的任意连续的 20bp。本发明还公开一种 DNA 聚合酶 I 缺陷型菌株、一种 CRISPR 系统和一种用于构建 DNA 聚合酶 I 缺陷型菌株的试剂盒等。

1. 一种构建 DNA 聚合酶 I 缺陷型菌株的方法,其特征在于,所述方法基于 CRISPR 系统进行所述构建,所述方法以 klenow1、klenow2 和 klenow3 中的至少之一作为靶序列,

所述 klenow1 为 SEQ ID NO :1 所示 klenow 片段基因序列中的第 30-52 位中的任意连续的 20bp,

所述 klenow2 为 SEQ ID NO :1 所示 klenow 片段基因序列中的第 370-390 位中的任意连续的 20bp,

所述 klenow3 为 SEQ ID NO :1 所示 klenow 片段基因序列中的第 1048-1068 位中的任意连续的 20bp。

2. 权利要求 1 的方法,其特征在于,所述 klenow1 为 SEQ ID NO :1 所示 klenow 片段基因序列中的第 33-52 位,

所述 klenow2 为 SEQ ID NO :1 所示 klenow 片段基因序列中的第 371-390 位,

所述 klenow3 为 SEQ ID NO :1 所示 klenow 片段基因序列中的第 1049-1068 位。

3. 权利要求 1 或 2 的方法,其特征在于,所述方法包括以下步骤:

(1) 基于所述靶序列,获得能够与所述靶序列结合的双链 DNA 序列;

(2) 构建以获得所述 CRISPR 系统,所述 CRISPR 系统包括 sgRNA 表达质粒和 Cas9 表达质粒,其中,

构建所述 sgRNA 表达质粒包括,

将 (1) 中的双链 DNA 序列连接到 sgRNA- 载体 1 上,以获得所述 sgRNA 表达质粒,

构建所述 Cas9 表达质粒包括,

将 Cas9 连接到载体 2 上,以获得所述 Cas9 表达质粒;

(3) 利用 (2) 中的 CRISPR 系统转染所述菌株的感受态细胞,以获得所述 DNA 聚合酶 I 缺陷型菌株。

4. 权利要求 3 的方法,其特征在于,所述菌株为大肠杆菌。

5. 权利要求 3 的方法,其特征在于,能够与 klenow1 结合的双链 DNA 序列的两条链分别为 SEQ ID NO :2 和 3,

能够与 klenow2 结合的双链 DNA 序列的两条链分别为 SEQ ID NO :4 和 5,

能够与 klenow3 结合的双链 DNA 序列的两条链分别为 SEQ ID NO :6 和 7。

6. 权利要求 3 的方法,其特征在于,能够与 klenow1 结合的双链 DNA 序列的两条链分别为 SEQ ID NO :8 和 9,

能够与 klenow2 结合的双链 DNA 序列的两条链分别为 SEQ ID NO :10 和 11,

能够与 klenow3 结合的双链 DNA 序列的两条链分别为 SEQ ID NO :12 和 13。

7. 权利要求 3 的方法,其特征在于,所述 sgRNA- 载体 1 的载体 1 和所述载体 2 具有不同的复制起始子。

8. 一种 DNA 聚合酶 I 缺陷型菌株,其利用权利要求 1-7 任一方法构建获得。

9. 一种 sgRNA 表达质粒,其包含能够结合权利要求 1-7 任一方法中的靶序列的序列。

10. 一种 CRISPR 系统,其包括权利要求 9 的 sgRNA 表达质粒,任选的包括 Cas9 表达质粒。

11. 权利要求 10 的 CRISPR 系统在构建 DNA 聚合酶 I 缺陷型菌株中的用途。

12. 一种用于构建 DNA 聚合酶 I 缺陷型菌株的试剂盒,其包括权利要求 10 的 CRISPR 系

统,任选的包括所述菌株的感受态细胞。

## DNA 聚合酶 I 缺陷型菌株及其构建方法

### 技术领域

[0001] 本发明涉及基因工程领域,具体的,本发明涉及一种构建 DNA 聚合酶 I 缺陷型菌株的方法、一种 DNA 聚合酶 I 缺陷型菌株、一种 sgRNA 表达质粒、一种 CRISPR 系统、该 CRISPR 系统在构建 DNA 聚合酶 I 缺陷型菌株中的用途和一种试剂盒。

### 背景技术

[0002] CRISPR(Clustered regularly interspaced short palindromic repeats)最早发现于原核生物的免疫系统中,这些成簇规则间隔的短回文重复序列旁边经常伴随出现保守基因,而这些保守基因编码的蛋白统称为 CRISPR 相关蛋白(Cas, CRISPR-associated proteins)。CRISPR 系统和 TALEN(transcription activator-like effector nucleases)系统、ZFN(zinc finger nucleases)系统是现在常用的三种基因组编辑工具,可以用于复杂的基因组编辑。目前 CRISPR 系统成功已应用于细菌、酵母、斑马鱼、小鼠及人细胞的基因组精确修饰,以及基因转录翻译调控以及其他方面。由于其突变效率高、突变成本低、作用物种广泛,是一种非常有前景的基因组定点改造分子工具。

[0003] CRISPR 系统主要由两个部分组成:sgRNA(single guide RNA)和 Cas9 蛋白。sgRNA 序列也是由两部分组成:1)靶序列,长 20bp,和基因组上目标序列互补,位于 PAM(Protospacer-Adjacent Motif)序列之前,需要自行设计;2)crRNA-tracrRNA,由原核生物上的序列改造而成,形成可以被 Cas9 蛋白识别的茎环结构。Cas9 是一种核酸酶。当 sgRNA 和 Cas9 共转化受体细胞时,sgRNA 可以将 Cas9 带到基因组上靶序列的位置,在 PAM 序列上游 2-3 碱基处由 Cas9 将靶基因的 DNA 双链切断。如果新的 DNA 模板也一并导入的话,断裂的 DNA 双链会发生同源重组,如果没有 DNA 模板导入,则在非同源末端连接修复机制下,会产生 Indel 突变。

[0004] DNA 聚合酶 I 缺陷型菌株可以用来作为聚合酶突变体筛选的宿主细胞。市面上现有的 DNA 聚合酶 I 缺陷型菌株是 Addgene 的 JS200 温度敏感型 DNA 聚合酶 I 缺陷菌株。此菌株非常容易老化,Po1 I 缺陷型基因型很容易丢失,很难用来作为宿主细胞。传统缺陷型菌株的获得是通过从自然界中分离或者通过传统的同源重组的方法来获得,这些方法耗时久,成功率低。

### 发明内容

[0005] 本发明旨在解决上述问题至少之一或者至少提供一种商业选择。为此,本发明通过设计构建出一种适于构建 DNA 聚合酶缺陷型菌株的 CRISPR 系统,来获得 DNA 聚合酶缺陷型菌株。

[0006] 依据本发明的第一方面,本发明提供一种构建 DNA 聚合酶 I 缺陷型菌株的方法,该方法基于 CRISPR 系统进行 DNA 聚合酶 I 缺陷型菌株的构建,该方法以 klenow1、klenow2 和 klenow3 中的至少之一作为靶序列,所述 klenow1 为 SEQ ID NO:1 所示 klenow 片段中的第 30-52 位中的任意连续的 20bp,所述 klenow2 为 SEQ ID NO:1 所示 klenow 片段中的第

370-390 位中的任意连续的 20bp,所述 klenow3 为 SEQ ID NO:1 所示 klenow 片段中的第 1048-1068 位中的任意连续的 20bp。klenow1、klenow2 和 klenow3 是发明人经过多次模拟和试验确定的,目标基因序列上的这三个位置的至少之一都能使利用的 CRISPR 系统高成功率的获得目标缺陷型菌株。

[0007] 本发明的这一方面的构建 DNA 聚合酶 I 缺陷型菌株的方法,步骤简单、敲除效率高、特异性好,可以很好的对菌株的基因组进行编辑,例如对大肠杆菌中的 DNA 聚合酶 I 序列的至少一部分进行编辑,即利用对发明人多次筛选确定出的 Klenow 片段基因序列上的特定位置——klenow1、klenow2 和 klenow3 至少之一进行切割,使菌株突变成 DNA 聚合酶 I 缺陷。通过该方法的利用 CRISPR 系统来获得 DNA 聚合酶缺陷型菌株,耗时短,成功率高,且 CRISPR 系统构建好后,目标缺陷型菌株容易获得。

[0008] 根据本发明的一个实施例,所述 klenow1 为 SEQ ID NO:1 所示 klenow 片段基因序列中的第 33-52 位,所述 klenow2 为 SEQ ID NO:1 所示 klenow 片段基因序列中的第 371-390 位,所述 klenow3 为 SEQ ID NO:1 所示 klenow 片段基因序列中的第 1049-1068 位。klenow1、klenow2 和 klenow3 是发明人多次模拟和试验确定的,目标基因序列上的这三个位置的至少之一都能使其对应的 CRISPR 系统高成功率的获得目标缺陷型菌株。

[0009] 根据本发明的一个实施例,所述方法包括以下步骤:(1) 基于所述靶序列,获得能够与所述靶序列结合的双链 DNA 序列;(2) 构建以获得所述 CRISPR 系统,所述 CRISPR 系统包括 sgRNA 表达质粒和 Cas9 表达质粒,其中,构建所述 sgRNA 表达质粒包括,将(1)中的双链 DNA 序列连接到 sgRNA-载体 1 上,以获得所述 sgRNA 表达质粒,构建所述 Cas9 表达质粒包括,将 Cas9 连接到载体 2 上,以获得所述 Cas9 表达质粒;(3) 利用(2)中的 CRISPR 系统感染所述菌株的感受态细胞,以获得所述 DNA 聚合酶 I 缺陷型菌株。较佳的,所述 sgRNA-载体 1 所使用的载体 1 和所述载体 2 具有不同的复制起始子,这样,利于避开共转染两个质粒时两个质粒之间的竞争,避免感受态细胞中只存留一个质粒。

[0010] 根据本发明的一个实施例,载体 1 和载体 2 中的其中的一个可以为 PSB1A2 和 PSB1C3 中的一个,另一个为 PSB2K3,因为 PSB1A2 和 PSB1C3 的复制起始子都是 pMB1,而 PSB2K3 的复制起始子与它们不同。

[0011] 根据本发明的一个实施例,所述菌株为大肠杆菌,包括但不限于 XL-10 和 DH5a。在筛选聚合酶突变体时,一般需把菌株本身的聚合酶给突变掉,而大肠杆菌的聚合酶 I 和要筛选的聚合酶 klenow 片段结构功能相似,并且实验证明失活聚合酶 I 后大肠杆菌细胞不会死亡,只会生长缓慢,因此可以选用大肠杆菌。

[0012] 根据本发明的一个实施例,能够与 klenow1 结合的双链 DNA 序列的互补的两条链分别为 SEQ ID NO:2 和 3,能够与 klenow2 结合的双链 DNA 序列的两条链分别为 SEQ ID NO:4 和 5,能够与 klenow3 结合的双链 DNA 序列的两条链分别为 SEQ ID NO:6 和 7。根据本发明的另一个实施例,能够与 klenow1 结合的双链 DNA 序列的两条链分别为 SEQ ID NO:8 和 9,能够与 klenow2 结合的双链 DNA 序列的两条链分别为 SEQ ID NO:10 和 11,能够与 klenow3 结合的双链 DNA 序列的两条链分别为 SEQ ID NO:12 和 13。其中,SEQ ID NO:8 和 9 分别对应于 SEQ ID NO:2 和 3——在 SEQ ID NO:2 的左侧加上 TAGC、在其右侧加上 G 碱基即为 SEQ ID NO:8,在 SEQ ID NO:3 的左侧加上 AAAC 四个碱基,在其右侧加上一个 G 碱基即为 SEQ ID NO:9,SEQ ID NO:8 和 9 退火形成的双链 DNA 序列相较于 SEQ ID NO:2 和

3 带有 BspQI 酶切位点 ;类似的, SEQ ID NO :10 和 11 分别对应于 SEQ ID NO :4 和 5——在 SEQ ID NO :4 的左侧加上 TAGC、在其右侧加上 G 碱基即为 SEQ ID NO :8,在 SEQ ID NO :5 的左侧加上 AAAC 四个碱基,在其右侧加上一个 G 碱基即为 SEQ ID NO :11, SEQ ID NO :10 和 11 退火形成的双链 DNA 序列相较于 SEQ ID NO :4 和 5 带有 BspQI 酶切位点 ;类似的, SEQ ID NO :12 和 13 分别对应于 SEQ ID NO :6 和 7——在 SEQ ID NO :6 的左侧加上 TAGC、在其右侧加上 G 碱基即为 SEQ ID NO :12,在 SEQ ID NO :7 的左侧加上 AAAC 四个碱基,在其右侧加上一个 G 碱基即为 SEQ ID NO :13, SEQ ID NO :12 和 13 退火形成的双链 DNA 序列相较于 SEQ ID NO :6 和 7 带有 BspQI 酶切位点。

[0013] 根据本发明的一个实施例,在进行步骤 (1) 时,分别基于靶序列 klenow1、klenow2 和 klenow3,设计合成能够分别与 klenow1、klenow2 和 klenow3 结合的双链 DNA 序列 SEQ ID NO :2 和 3、SEQ ID NO :4 和 5 以及 SEQ ID NO :6 和 7 ;而考虑到后续需要通过酶切连接的方式将双链 DNA 序列插入到 sgRNA 质粒骨架 (sgRNA-载体 1) 上去,因此分别在 SEQ ID NO :2、4 和 6 的左侧加上 TAGC,右侧加上 G 碱基,分别在 SEQ ID NO :2、4 和 6 各自的反向互补序列左侧需加上 AAAC 四个碱基,右侧加上一个 G 碱基,即相应获得 SEQ ID NO :8 和 9、SEQ ID NO :10 和 11 以及 SEQ ID NO :12 和 13,磷酸化各链,以期退火形成的双链 DNA 序列带有 BspQI 酶切位点,帮助将 SEQ ID NO :2、4 和 6 分别插入到 sgRNA 质粒骨架上。

[0014] 依据本发明的第二方面,本发明提供一种 DNA 聚合酶 I 缺陷型菌株,其利用上述本发明一方面或者任一实施例的方法构建获得。DNA 聚合酶 I 缺陷型菌株可以用来作为聚合酶突变体筛选的宿主细胞。相较于市面上现有的 DNA 聚合酶 I 缺陷型菌株,例如 Addgene 公司的 JS200 温度敏感型 DNA 聚合酶 I 缺陷菌株,本发明这一方面的菌株不易老化,Pol I 缺陷型基因型不易丢失,能很好的作为宿主细胞。

[0015] 依据本发明的第三方面,本发明提供一种 sgRNA 表达质粒,该 sgRNA 表达质粒包含能够结合上述本发明一方面或者任一实施例的方法中的靶序列的序列,即包含能结合 klenow1、klenow2 和 klenow3 至少之一的序列。在本发明的一些实施例中,构建该 sgRNA 表达质粒,包括 :基于所述靶序列,获得能够与所述靶序列结合的双链 DNA 序列 ;将双链 DNA 序列连接到 sgRNA-载体 1 上,以获得所述 sgRNA 表达质粒。sgRNA-载体 1 由 sgRNA 质粒骨架或者说 sgRNA 序列和载体连接构成,可以预先连接获得保存,可以在构建该表达质粒时同时获得,也可以通过市售获得。该 sgRNA 表达质粒可用于编辑宿主细胞靶序列区域,使宿主细胞的 DNA 聚合酶 I 突变。在本发明的一个实施例中,载体 1 为 PSB1C3。

[0016] 依据本发明的第四方面,本发明提供一种 CRISPR 系统,其包括上述本发明一方面的 sgRNA 表达质粒,任选的还包括 Cas9 表达质粒。在本发明的一个实施例中,Cas 表达质粒是由 Cas9 组蛋白连接在 PSB1A2 载体或者 PSB2K3 载体上构成的。CRISPR 系统能够用于编辑宿主细胞中的 DNA 聚合酶 I,使宿主突变,获得 DNA 聚合酶 I 缺陷型菌株。因为该 CRISPR 系统以质粒的形式存在,后续可以很方便获得缺陷型菌株。

[0017] 依据本发明的第五方面,本发明提供上述本发明一方面的 CRISPR 系统在构建 DNA 聚合酶 I 缺陷型菌株中的用途。

[0018] 依据本发明的第六方面,本发明提供一种用于构建 DNA 聚合酶 I 缺陷型菌株的试剂盒,其包括上述本发明一方面的 CRISPR 系统,任选的还包括宿主细胞。利用本发明这一方面的试剂盒构建得的缺陷型菌株,不易老化,Pol I 缺陷型基因型不易丢失,能很好的作

为宿主细胞。

### 附图说明

[0019] 本发明的上述和 / 或附加的方面和优点从结合下面附图对实施方式的描述中将变得明显和容易理解, 其中 :

[0020] 图 1 是本发明一个实施例中的 CRISPR 系统的 sgRNA 质粒图谱 ;

[0021] 图 2 是本发明一个实施例中的 CRISPR 系统的 Cas9 质粒图谱 ;

[0022] 图 3 是本发明一个实施例中的对照组的 CRISPR 系统转染效率结果图 ; 图 3A 显示阴性对照结果, 其左边图片为 PSB1C3 质粒阴性对照, 其右边图片为 PSB2K3-J04450 质粒阴性对照 ; 图 3B 显示阳性对照结果 ;

[0023] 图 4 是本发明一个实施例中的实验组的 CRISPR 系统转染效率结果图 ; 图 4A、4B 和 4C 分别显示实验组 1、2 和 3 结果, 图 4A、4B 和 4C 各自的右图分别为各自的左图的局部示意图 ;

[0024] 图 5 是本发明一个实施例中的实验组 1 获得的菌株的 klenow1 区域的序列与原始序列的比对结果示意图 ;

[0025] 图 6 是本发明一个实施例中的实验组 2 获得的菌株的 klenow2 区域的序列与原始序列的比对结果示意图 ;

[0026] 图 7 是本发明一个实施例中的实验组 3 获得的菌株的 klenow3 区域的序列与原始序列的比对结果示意图。

### 具体实施方式

[0027] 根据本发明的一个实施例提供的构建 DNA 聚合酶 I 缺陷型菌株的方法, 该方法基于 CRISPR 系统进行 DNA 聚合酶 I 缺陷型菌株的构建, 该方法以 klenow1、klenow2 和 klenow3 中的至少之一作为靶序列, 所述 klenow1 为 SEQ ID NO : 1 所示 klenow 片段中的第 30-52 位中的任意连续的 20bp, 所述 klenow2 为 SEQ ID NO : 1 所示 klenow 片段中的第 370-390 位中的任意连续的 20bp, 所述 klenow3 为 SEQ ID NO : 1 所示 klenow 片段中的第 1048-1068 位中的任意连续的 20bp。所称 klenow1、klenow2 和 klenow3 都是 klenow 片段 (Klenow fragment) 基因序列中的一段。Klenow 片段, 又名 DNA 聚合酶 I 大片段、克列诺片段 (Klenow fragment) 或克列诺酶 (Klenow enzyme), 是 E. coli DNA 聚合酶 I 经胰蛋白酶或枯草杆菌蛋白酶部分水解生成的 C 末端 605 个氨基酸残基片段。该片段保留了 DNA 聚合酶 I 的 5' -3' 聚合酶和 3' -5' 外切酶活性, 但缺少完整酶的 5' -3' 外切酶活性。DNA 聚合酶 I (DNA-pol I) 断开后的存在另一个 323 个氨基酸残基片段, 保留 5' -3' 外切酶活性。

[0028] 所称靶序列与 CRISPR 系统中的 sgRNA 靶序列是互补的, 即如果靶序列在基因序列上, sgRNA 靶序列则选择其反义链。sgRNA 靶序列一般长度为 20bp, 在 PAM 序列上游。在确定靶序列的过程中, 发明人经过多次设计和试验摸索发现, 使靶序列上的种子序列, 即靠近 3' 端的 12bp, 不与目标基因上其他位置的序列完全一致, 能够获得较好的工作效率 ; 进一步的, 靶序列选择目标基因序列上面 -35 到 1/10 基因序列全长区域内的位置能够获得较好的工作效率。在本发明的一个实施例中, 根据 DNA 聚合酶 I 不同区域的 DNA 序列设计 sgRNA 靶序列, 每个区域设计一条 sgRNA 靶序列。klenow1、klenow2 和 klenow3 是发明人经过多次

模拟和试验确定的,目标基因序列上的这三个位置的至少之一都能使其对应的 CRISPR 系统高成功率的获得目标缺陷型菌株。

[0029] 上述构建 DNA 聚合酶 I 缺陷型菌株的方法,步骤简单、敲除效率高、特异性好,可以很好的对菌株的基因组进行编辑,通过利用 CRISPR 系统对宿主菌株,例如大肠杆菌中的 DNA 聚合酶 I 序列的至少一部分进行编辑,即利用对多次筛选确定出的 Klenow 片段基因序列上的特定位置——klenow1、klenow2 和 klenow3 中的至少之一进行切割,使菌株突变成 DNA 聚合酶 I 缺陷。通过该方法的利用 CRISPR 系统来获得 DNA 聚合酶缺陷型菌株,耗时短,成功率高,且 CRISPR 系统构建好后,目标缺陷型菌株容易获得。

[0030] 根据本发明的一个实施例,所述 klenow1 为 SEQ ID NO :1 所示 klenow 片段基因序列中的第 33-52 位,所述 klenow2 为 SEQ ID NO :1 所示 klenow 片段基因序列中的第 371-390 位,所述 klenow3 为 SEQ ID NO :1 所示 klenow 片段基因序列中的第 1049-1068 位。klenow1、klenow2 和 klenow3 是发明人多次模拟和试验确定的,目标基因序列上的这三个位置的至少之一都能使利用的 CRISPR 系统高成功率的获得目标缺陷型菌株。

[0031] 根据本发明的一个实施例,所述方法包括以下步骤:(1) 基于所述靶序列,获得能够与所述靶序列结合的双链 DNA 序列;(2) 构建以获得所述 CRISPR 系统,所述 CRISPR 系统包括 sgRNA 表达质粒和 Cas9 表达质粒,其中,构建所述 sgRNA 表达质粒包括,将(1)中的双链 DNA 序列连接到 sgRNA-载体 1 上,以获得所述 sgRNA 表达质粒,构建所述 Cas9 表达质粒包括,将 Cas9 连接到载体 2 上,以获得所述 Cas9 表达质粒;(3) 利用(2)中的 CRISPR 系统感染所述菌株的感受态细胞,以获得所述 DNA 聚合酶 I 缺陷型菌株。较佳的,所述 sgRNA-载体 1 所使用的载体 1 和所述载体 2 具有不同的复制起始子,这样,利于避开共转染两个质粒时两个质粒之间的竞争,避免感受态细胞中只存留一个质粒。

[0032] 较佳的,载体 1 和载体 2 是不同的。在本发明的一个实施例中,其中的一个可以为 PSB1A2 和 PSB1C3 中的一个,另一个为 PSB2K3,因为 PSB1A2 和 PSB1C3 的复制起始子都是 pMB1,而 PSB2K3 的复制起始子与它们不同。

[0033] 根据本发明的一个实施例,所述菌株为大肠杆菌,包括但不限于 XL-10 和 DH5a。在筛选聚合酶突变体时,一般需把菌株本身的聚合酶给突变掉,而大肠杆菌的聚合酶 I 和要筛选的聚合酶 klenow 片段结构功能相似,并且实验证明失活聚合酶 I 后大肠杆菌细胞不会死亡,只会生长缓慢,因此可以选用大肠杆菌。

[0034] 根据本发明的一个实施例,能够与 klenow1 结合的双链 DNA 序列的互补的两条链分别为 SEQ ID NO :2 和 3,能够与 klenow2 结合的双链 DNA 序列的两条链分别为 SEQ ID NO :4 和 5,能够与 klenow3 结合的双链 DNA 序列的两条链分别为 SEQ ID NO :6 和 7。根据本发明的另一个实施例,能够与 klenow1 结合的双链 DNA 序列的两条链分别为 SEQ ID NO :8 和 9,能够与 klenow2 结合的双链 DNA 序列的两条链分别为 SEQ ID NO :10 和 11,能够与 klenow3 结合的双链 DNA 序列的两条链分别为 SEQ ID NO :12 和 13。其中,SEQ ID NO :8 和 9 分别对应于 SEQ ID NO :2 和 3——在 SEQ ID NO :2 的左侧加上 TAGC、在其右侧加上 G 碱基即为 SEQ ID NO :8,在 SEQ ID NO :3 的左侧加上 AAAC 四个碱基,在其右侧加上一个 G 碱基即为 SEQ ID NO :9,SEQ ID NO :8 和 9 退火形成的双链 DNA 序列相较于 SEQ ID NO :2 和 3 带有 BspQI 酶切位点;类似的,SEQ ID NO :10 和 11 分别对应于 SEQ ID NO :4 和 5——在 SEQ ID NO :4 的左侧加上 TAGC、在其右侧加上 G 碱基即为 SEQ ID NO :8,在 SEQ ID NO :5 的



左侧加上 AAAC 四个碱基,在其右侧加上一个 G 碱基即为 SEQ ID NO :11, SEQ ID NO :10 和 11 退火形成的双链 DNA 序列相较于 SEQ ID NO :4 和 5 带有 BspQI 酶切位点;类似的, SEQ ID NO :12 和 13 分别对应于 SEQ ID NO :6 和 7——在 SEQ ID NO :6 的左侧加上 TAGC、在其右侧加上 G 碱基即为 SEQ ID NO :12,在 SEQ ID NO :7 的左侧加上 AAAC 四个碱基,在其右侧加上一个 G 碱基即为 SEQ ID NO :13, SEQ ID NO :12 和 13 退火形成的双链 DNA 序列相较于 SEQ ID NO :6 和 7 带有 BspQI 酶切位点。

[0035] 根据本发明的一个实施例,在进行步骤 (1) 时,分别基于靶序列 klenow1、klenow2 和 klenow3,设计合成能够分别与 klenow1、klenow2 和 klenow3 结合的双链 DNA 序列 SEQ ID NO :2 和 3、SEQ ID NO :4 和 5 以及 SEQ ID NO :6 和 7;而考虑到后续需要通过酶切连接的方式将双链 DNA 序列插入到 sgRNA 质粒骨架或者 sgRNA-载体 1 上去,因此分别在 SEQ ID NO :2、4 和 6 的左侧加上 TAGC,右侧加上 G 碱基,分别在 SEQ ID NO :2、4 和 6 各自的反向互补序列左侧需加上 AAAC 四个碱基,右侧加上一个 G 碱基,即相应获得 SEQ ID NO :8 和 9、SEQ ID NO :10 和 11 以及 SEQ ID NO :12 和 13,磷酸化各链,以期退火形成的双链 DNA 序列带有 BspQI 酶切位点,帮助将 SEQ ID NO :2、4 和 6 分别插入到 sgRNA 质粒骨架上。本领域技术人员能够理解,也可以在序列的左右两端加上其它碱基,使得退火形成的序列具有相对应的酶切位点,较佳的,所利用的酶的酶切位点和识别位点相同。

[0036] 根据本发明的一个实施例,提供一种 DNA 聚合酶 I 缺陷型菌株,其利用上述本发明任一实施例的方法构建获得。DNA 聚合酶 I 缺陷型菌株可以用来作为聚合酶突变体筛选的宿主细胞。相较于市面上现有的 DNA 聚合酶 I 缺陷型菌株,例如 Addgene 公司的 JS200 温度敏感型 DNA 聚合酶 I 缺陷菌株,本发明这一方面的菌株不易老化,Pol I 缺陷型基因型不易丢失,能很好的作为宿主细胞。

[0037] 根据本发明的的一个实施例,提供一种 sgRNA 表达质粒,该 sgRNA 表达质粒包含能够结合上述本发明任一实施例的方法中的靶序列的序列,即其 sgRNA 靶序列为能结合 klenow1、klenow2 和 klenow3 中的至少之一。在本发明的一些实施例中,构建该 sgRNA 表达质粒,包括:基于所述靶序列,获得能够与所述靶序列结合的双链 DNA 序列;将双链 DNA 序列连接到 sgRNA-载体 1 上,以获得所述 sgRNA 表达质粒。sgRNA-载体 1 由 sgRNA 质粒骨架或者说 sgRNA 序列和载体连接构成,可以预先连接获得保存,可以在构建该表达质粒时同时获得,也可以通过市售获得。该 sgRNA 表达质粒可用于编辑宿主细胞靶序列区域,使宿主细胞的 DNA 聚合酶 I 突变。在本发明的一个实施例中,载体 1 为 PSB1C3。

[0038] 根据本发明的一个实施例,提供一种 CRISPR 系统,其包括上述本发明的实施例中的 sgRNA 表达质粒,任选的还包括 Cas9 表达质粒。sgRNA 表达质粒可参照上述实施例获得。在本发明的一个实施例中,Cas 表达质粒是由 Cas9 组蛋白连接在 PSB1A2 载体或者 PSB2K3 载体上构成的。CRISPR 系统能够用于编辑宿主细胞中的 DNA 聚合酶 I,使宿主突变,获得 DNA 聚合酶 I 缺陷型菌株。因为该 CRISPR 系统以质粒的形式存在,后续可以很方便获得缺陷型菌株。

[0039] 根据本发明的一个实施例,提供上述本发明实施例的 CRISPR 系统在构建 DNA 聚合酶 I 缺陷型菌株中的用途。

[0040] 根据本发明的一个实施例,提供一种用于构建 DNA 聚合酶 I 缺陷型菌株的试剂盒,其包括上述本发明上述实施例中的 CRISPR 系统,任选的还包括宿主细胞,例如大肠杆菌。

利用本发明这一方面的试剂盒构建得的缺陷型菌株,不易老化,Pol I 缺陷型基因型不易丢失,能很好的作为聚合酶突变体筛选的宿主细胞。

[0041] 以下结合附图和具体实施例对本发明的构建 DNA 聚合酶 I 缺陷型菌株的方法、CRISPR 系统、sgRNA 表达质粒、试剂盒等进行详细的描述。下面示例,仅用于解释本发明,而不能理解为对本发明的限制。在本发明的描述中,除非另有说明,“多个”的含义是两个或两个以上。

[0042] 除另有交待,以下实施例中涉及的未特别交待技术手段为本领域技术人员所熟知的常规手段,可以参照《分子克隆实验指南》第三版或者相关产品说明书进行,未特别交待的试剂、序列、载体和细胞等也均为可商业获得的。未详细描述的各种过程和方法是本领域中公知的常规方法,所用试剂的来源、商品名以及有必要列出其组成成分者,均在首次出现时标明,其后所用相同试剂如无特殊说明,均以首次标明的内容相同。

[0043] 实施例一

[0044] 构建 DNA 聚合酶 I 缺陷型菌株的方法,一般包括:

[0045] 1. sgRNA 靶序列设计及双链 DNA 序列合成

[0046] 根据 DNA 聚合酶 I 不同区域的 DNA 序列设计 sgRNA 靶序列,每个区域设计一条 sgRNA 靶序列。sgRNA 靶序列要求:长度为 20bp,在 PAM 序列上游;如果靶序列在基因序列上,sgRNA 靶序列需要选择反义链;靶序列上的种子序列(靠近 3' 端的 12bp)必须不能和基因组上其他位置的序列完全一致;靶序列最好选择基因上面 -35 到 1/10 基因全长区域内的位置来获得最好的工作效率。

[0047] 2. CRISPR 系统构建

[0048] 2.1 sgRNA 组件构建

[0049] sgRNA 组件的质粒上靶序列区域可以通过酶切连接置换掉。因此在合成靶序列时,左右两端要加上合适的酶切位点。

[0050] 在收到合成的靶序列后,合成的单链序列要进行磷酸化处理,然后进行退火处理。然后通过酶切连接,将靶序列插入到 sgRNA 质粒骨架上。

[0051] 2.2 Cas9 组件构建

[0052] 将 Cas9 组件通过酶切连接换到合适的载体上,以便能完成和 sgRNA 组件的质粒共转化。

[0053] 3. CRISPR 系统转化 XL-10 感受态细胞,检测 CRISPR 系统的工作效率

[0054] sgRNA 组件和 Cas9 组件共转化 DH5a 和 XL-10 感受态细胞,通过对 Cas9 表达量的调控来检测 CRISPR 系统的工作效率。

[0055] 4. 测序验证缺陷型菌株是否成功构建

[0056] 利用上述 CRISPR 系统能够快速、有效、高精度的获得 DNA 聚合酶 I 缺陷型大肠杆菌菌株。该方法步骤简单、敲除效率高、特异性好,可以很好的对大肠杆菌的基因组进行编辑,且因为 CRISPR 系统以质粒的形式存在,后续可以很方便获得缺陷型菌株。

[0057] 实施例二

[0058] 1. sgRNA 靶序列设计及合成

[0059] 因为想要用缺陷型大肠杆菌作为宿主细胞来筛选聚合酶突变体,因此需要把大肠杆菌本身的聚合酶给突变掉。大肠杆菌的聚合酶 I 和要筛选的聚合酶结构功能相似,并且

实验证明失活聚合酶 I 后细胞不会死亡,只会生长缓慢,因此考虑突变大肠杆菌聚合酶 I。

[0060] 从 klenow fragment 的前中部分别选取了三个位置作为靶序列,标记为 klenow1、klenow2、klenow3。

[0061] klenow 片段全长 1818bp, klenow 片段 DNA 序列如下所示 :5'-GTGATTTCTTATGACAACTACGTCACCATCCTTGATGAAGAAACACTGAAAGCGTGGATTGCGAAGCTGGAAAAAGCGCCGGTATTTGCATTTGATACCGAAACCGACAGCCTTGATAACATCTCTGCTAACCTGGTCGGGCTTTCTTTTGTATCGAGCCAGGCGTAGCGCATATATTCCGGTTGCTCATGATTATCTTGATGCGCCCGATCAAATCTCTCGCGAGCGTGCCTCGAGTTGCTAAAACCGCTGCTGGAAGATGAAAAGGCGCTGAAGGTCGGGCAAAACCTGAAATACGATCGCGGTATTCTGGCGAACTACGGCATTGAACTGCGTGGGATTGCGTTTGATAACCATGCTGGAGTCTTACATTCTCAATAGCGTTGCCGGGCGTCACGATATGGACAGCCTCGCGAAGCTTGGTTGAAGCACAAAACCATCACTTTTGAAGAGATTGCTGGTAAAGGCAAAAATCAACTGACCTTTAACCAGATTGCCCTCGAAGAAGCCGGACGTTACGCCGCCGAAGATGCAGATGTCACCTTGCGATTGCGATCTGAAAATGTGGCCGGATCTGCAAAAACACAAAAGGCGGTTGAACGTCTTCGAGAATATCGAAATGCCGCTGGTCCGGTGCTTTACGCATTGAACGTAACGGTGTGAAGATCGATCCGAAAGTGTGCACAATCATTCTGAAGAGCTCACCTTCGTCTGGCTGAGCTGGAAAAAGAAAGCGCATGAAATTGCAGGTGAGGAATTTAACCTTTCTCCACCAAGCAGTTACAAAACCATCTCTTTGAAAAACAGGGCATTAAACCGCTGAAGAAAACGCCGGGTGGCGCGCCGTCACCGTCGGAAAGGTAAGTACTGGAAGAACTGGCGCTGGACTATCCGTTGCCAAAAGTGATTCTGGAGTATCGTGGTCTGGCGAAGCTGAAATCGACCTACACCGACAAGCTGCCGCTGATGATCAACCCGAAAACCGGGCGTGTGCATACCTCTTATCACCAGGCAGTAACCTGCAACCGGACGTTTATCGTCAACCGATCCTAACCTGCAAAACATTCCGGTGCCTAACGAAGAAGTGCCTGATCCGCCAGGCGTTTATTGCGCCAGAGGATTATGTGATTGTCTCAGCGGACTACTCGCAGATTGAACTGCGCATTATGGCGCATCTTTGCGGTGACAAAAGGCTTGCTGACCGCATTCGCGGAAGGAAAAGATATCCACCGGGCAACGGCGGCAGAAAGTGTGGTTTGGTGGCAACTGGAAAACCGTCACCAGCGAGCAACGCCGTAGCGCGAAAGCGATCAACTTTGGTCTGATTATGGCATGAGTGCTTTGCGTCTGGCGCGCAATTGAACATTCCACGTAAAGAAGCGCAGAAGTACATGGACCTTTACTTCGAACGCTACCCTGGCGTGTGGAGTATATGGAACGCACCCGTGCTCAGGCGAAAGAGCAGGGCTACGTTGAAACGCTGGACGGACGCCGCTGTGATCTGCCGGATATCAAATCCAGCAATGGTGTCTCGTGTGCAGCGGCTGAACGTGCCATTAACCGCCAATGCAGGGAACCGCCGCCGACATTATCAAACGGCGATGATTGCCGTTGATGCGTGGTTACAGGCTGAGCAACCGCGTGTACGTATGATCATGCAGGTACACGATGAACTGGTATTTGAAGTTCATAAAGATGATGTTGATGCCGTCGCGAAGCAGATTCATCAACTGATGGA AAACTGTACCCGTCTGGATGTGCCGTTGCTGGTGGAAAGTGGGAGTGGCGAAAACCTGGGATCAGGCGCACTAA-3' (SEQ ID NO :1)。

[0062] klenow1 片段距离其起始密码子约 30bp, sgRNA 靶序列为 5'-CTTTCAGTGTTTCTTCATCA-3' (SEQ ID NO :2), 反向互补序列为 5'-TGATGAAGAAACACTGAAAG-3' (SEQ ID NO :3); klenow2 片段距离起始密码子约 370bp, 靶序列为 5'-GGCAACGCTATTGAGAATGT-3' (SEQ ID NO :4), 反向互补序列为 5'-ACATTCTCAATAGCGTTGCC-3' (SEQ ID NO :5); klenow3 片段距离起始密码子约 1048bp, 靶序列为 5'-AATGTTTTGCAGGTTAGGAT-3' (SEQ ID NO :6), 反向互补序列为 5'-ATCCTAACCTGCAAAACATT-3' (SEQ ID NO :7)。因为后续需要通过酶切连接的方式将靶序列 oligo 插入到 sgRNA 质粒骨架上去,因此 sgRNA 靶序列左侧需要加上 TAGC,右侧加上 G 碱基,反向互补序列左侧需加上 AAAC 四个碱基,右侧加上一个 G 碱基,即对应获得 5'-TAGCCTTTCAGTGTTCATCAG-3' (SEQ ID NO :8),反向互补序列为 5'-AAACTGATGAAGAAACACTGAAAGG-3' (SEQ ID NO :9); klenow2 片段距离起始密码子约 370bp, 靶序列为 5'-TAGCGCAACGCTA

TTGAGAATGTG-3' (SEQ ID NO:10), 反向互补序列为 5'-AAACACATTCTCAATAGCGTTGCCG-3' (SEQ ID NO:11); klenow3 片段距离起始密码子约 1048bp, 靶序列为 5'-TAGCAATGTTTTGCAGTTAGGATG-3' (SEQ ID NO:12), 反向互补序列为 5'-AAACATCCTAACCTGCAAAACATTG-3' (SEQ ID NO:13)。这样后期经退火后可以组成 BspQI 酶切位点, 帮助将 sgRNA 靶序列 oligo 插入到 sgRNA 质粒骨架上。

[0063] 2. CRISPR 系统构建

[0064] 通过以下构建得的 sgRNA 表达质粒和 Cas9 表达质粒各自的质粒图谱分别如图 1 和图 2 所示。

[0065] 2. 1sgRNA 组件构建

[0066] 2. 1. 1 合成的靶序列 oligo 和反向互补序列磷酸化处理、退火处理

[0067] 反应体系为:

[0068]

T4 PNK	1 $\mu$ l
T4 Ligase Buf.	2 $\mu$ l
ddH <sub>2</sub> O	13 $\mu$ l
sgRNA-Primer-F	2 $\mu$ l
sgRNA-Primer-R	2 $\mu$ l

[0069] 55°C 反应 30min 后, 将 PCR 管放于 80°C 水浴中 10min, 终止磷酸化反应, 然后放置于室温进行退火, 使 PCR 管温度恢复至室温, 完成退火处理。其中的 sgRNA-Primer-F 为 SEQ ID NO:8、10 或 12, 相应的 sgRNA-Primer-R 为 SEQ ID NO:9、11 或 13。

[0070] 2. 1. 2sgRNA 质粒骨架 BspQI 酶切

[0071] sgRNA 质粒骨架连在 PSB1C3 载体 (来自 iGEM, 氯霉素抗性) 上, 所有来自 iGEM 的载体都有 EcoRI、PstI、SpeI、XbaI 四个通用酶切位点, 以便 DNA 插入片段在不同的载体上进行切换。而 sgRNA 质粒骨架上靶序列的位置有 BspQI 酶切位点, 方便将不同的上步退火得的靶序列连入该质粒骨架上。

[0072] 反应体系为:

[0073] BspQI 2 $\mu$ l

[0074] 10xCutsmart Buffer 2 $\mu$ l

[0075] pSB1C3-sgRNA cassette 1 $\mu$ g

[0076] ddH<sub>2</sub>O 补充至 20 $\mu$ l

[0077] 50°C 反应 2h。1% 琼脂糖胶跑电泳 (150V, 30min), EB 染色 10min, 切胶回收。

[0078] 2. 1. 3 靶序列 oligo 连接入上步的酶切开的 sgRNA 质粒骨架

[0079] 通过 T4 DNA ligase 的作用将靶序列 oligo 连接到 PSB1C3-sgRNA 质粒骨架上。

[0080] 反应体系为:

[0081] 2x Rapid ligation buffer 10 $\mu$ l

[0082] T4 DNA ligase 1 $\mu$ l

[0083] PSB1C3-sgRNA 质粒骨架 10ng

[0084] 靶序列 oligo 50ng

- [0085] ddH<sub>2</sub>O 补足至 20ul
- [0086] 20℃反应 2h。
- [0087] 2. 1. 4 连接产物转化 DH5a 感受态细胞
- [0088] 从 -80℃冰箱拿出 DH5a 感受态细胞 (100ul), 置于冰上 10min。然后向感受态细胞中加入 10ul 连接产物, 轻轻混匀, 置于冰上静置 30min。然后 42℃热击 90s, 再置于冰上 5min, 后在超净台内加入 800ul SOC 无抗培养基。37℃, 200rpm 培养 45min, 使细胞恢复。然后取 100ul, 涂氯霉素平板。平板置于 37℃培养箱过夜培养。
- [0089] 2. 1. 5 提质粒, 送样测序
- [0090] 从平板上挑单菌落, 置于 5ml 氯霉素液体培养基中, 37℃, 200rpm 过夜培养, sgRNA 表达质粒小提, 送样测序。
- [0091] 2. 2Cas9 组件构建
- [0092] Cas9 组件连接在 PSB1A2 载体 (来自 iGEM, 氨苄抗性) 上, 而 PSB1A2 和 PSB1C3 的复制起始子都是 pMB1, 如果将两个质粒共转化的话, 因为它们使用相同的复制机制, 可能会引起质粒之间的竞争, 而导致感受态细胞中只存留一个质粒。因此考虑将 Cas9 组件连接到 PSB2K3 载体上 (来自 iGEM, 卡那抗性, P1 及 F' 复制起始子)。
- [0093] 2. 2. 1PSB1A2-Cas9 骨架及 PSB2K3 载体酶切
- [0094] 载体骨架反应体系:
- [0095] PSB2K3 1ug
- [0096] SpeI 3ul
- [0097] PstI 2ul
- [0098] NEB 2. 1 2ul
- [0099] ddH<sub>2</sub>O 补足至 20ul
- [0100] PSB1A2-Cas9 酶切体系:
- [0101] PSB1A2-Cas9 1ug
- [0102] PstI 2ul
- [0103] XbaI 1. 5ul
- [0104] NEB2. 1 2ul
- [0105] ddH<sub>2</sub>O 补足至 20ul
- [0106] 37℃反应 2h。1%琼脂糖胶跑电泳 (150V, 30min), EB 染色 10min, 切胶回收。
- [0107] 2. 2. 2Cas9 组件连接 PSB2K3 载体
- [0108] 反应体系为:
- [0109] 2xRapid ligation buffer 10ul
- [0110] T4 DNA ligase 1ul
- [0111] PSB2K3 10ng
- [0112] Cas9 30ng
- [0113] ddH<sub>2</sub>O 补足至 20ul
- [0114] 20℃, 反应 2h。
- [0115] 2. 2. 3 连接产物转化 DH5a 感受态细胞
- [0116] 从 -80° 冰箱拿出 DH5a 感受态细胞 (100ul), 置于冰上 10min。然后向感受态细

胞中加入 10u1 连接产物,轻轻混匀,置于冰上静置 30min。然后 42℃热击 90s,再置于冰上 5min,后在超净台内加入 800u1 SOC 无抗培养基。37℃,200rpm 培养 45min,使细胞恢复。然后取 100u1,涂卡那抗性平板。平板置于 37℃培养箱过夜培养。

[0117] 2.2.4 提质粒,送样测序

[0118] 从平板上挑单菌落,置于 5ml 卡那霉素液体培养基中,37℃,200rpm 过夜培养,质粒小提,送样测序。测序结果如预期。

[0119] 3. CRISPR 系统转化 XL-10 感受态细胞,检测 CRISPR 系统的工作效率

[0120] 配制 Kan/Chl 双抗板,Kan 终浓度为 50ug/ml,Chl 终浓度为 17ug/ml。

[0121] 共转化体系:

[0122] 阴性对照:PSB1C3 质粒,PSB2K3-J04450 质粒

[0123] 阳性对照:PSB1C3&PSB2K3-J04450 质粒共转化

[0124] 实验组 1:PSB2K3-Cas9&PSB1C3-klenow1 共转化

[0125] 实验组 2:PSB2K3-Cas9&PSB1C3-klenow2 共转化

[0126] 实验组 3:PSB2K3-Cas9&PSB1C3-klenow3 共转化

[0127] 转化时质粒各转 50ng,从 -80° 冰箱拿出 XL-10 感受态细胞 (100u1),置于冰上 10min。然后向感受态细胞中加入 50ng 质粒,轻轻混匀,置于冰上静置 30min。然后 42℃热击 90s,再置于冰上 5min,后在超净台内加入 800u1 SOC 无抗培养基。37℃,200rpm 培养 45min,使细胞恢复。然后取 100u1,涂卡那/氯霉素双抗性平板,涂板时,实验组的平板上加上终浓度为 0.1mM 的 IPTG,以及 aTc 来启动 PSB2K3 载体的高拷贝复制 (PSB2K3 载体上有两个复制起始子,在一般情况下,受 F' 起始子调控,产生低拷贝质粒,而在 IPTG 的诱导下,P1 起始子被启动,产生高拷贝质粒) 以及解除对 Cas9 组件的抑制 (Cas9 组件本身受 PLtet0-1 启动子调控,此启动子受 TetR 抑制子的抑制,在 aTc 存在的情况下,TetR 不能再抑制 PLtet0-1 启动子,Cas9 组件被表达)。平板置于 37℃培养箱过夜培养。

[0128] 实验结果如图 3 和图 4 所示,图 3 显示对照组的菌落生长情况。图 3A 显示阴性对照无菌落长出,图 3A 的左图为 PSB1C3 质粒阴性对照,图 3A 的右图为 PSB2K3-J04450 质粒阴性对照;图 3B 显示阳性对照菌落生长情况,由图 3B 可看出阳性对照有大量菌落长出,说明抗性板正常,质粒共转化效率正常。图 4 显示实验组菌落生长情况,图示显示实验组也有菌落长出,只是相较阳性对照,菌落极小,说明 CRISPR 系统发挥作用,菌落生长受到限制,生长缓慢。图 4A-C 分别对应实验组 1-3,图 4A-C 的右图分别为左图的局部放大,对比图 4A、4B 和 4C,可看出和 klenow1、klenow2 相比,klenow3 的菌落虽然不及阳性对照菌落大小,但是比 klenow1、klenow2 菌落都大,说明它对 DNA 聚合酶 I 的作用小于 klenow1/klenow2。

[0129] 4. 测序验证缺陷型菌株是否成功构建

[0130] 对各实验组的菌株的目标区域进行测序。

[0131] 测序结果证明了各实验组获得的菌株均为 DNA 聚合酶 I 缺陷型菌株。

[0132] 图 5-7 分别显示各测序结果与参考序列 (原始序列) 的比对情况。图 5 显示实验组 1 获得的菌株的 klenow1 区域的测序序列与原始序列 (DNA Pol I 区域) 的比对情况。图 6 显示实验组 2 获得的菌株的 klenow2 区域的测序序列与原始序列 (DNA Pol I 区域) 的比对情况。图 7 显示实验组 3 获得的菌株的 klenow3 区域的测序序列与原始序列 (DNA Pol I 区域) 的比对情况。图 5-7 中显示的每两行序列比对,上边一行是测序结果,下边一行是

DNA Pol I 原始序列,带下划线标记的位点表示突变位点,这些单核苷酸个位点的突变编码终止密码子,使菌株失去全部或部分 DNA Pol I 酶功能。

[0133] 实施例三

[0134] sgRNA 表达质粒。可参照实施例二中的 2.1 构建获得该 sgRNA 表达质粒。

[0135] 实施例四

[0136] 可用于构建 DNA 聚合酶 I 缺陷型菌株的试剂盒,该试剂盒包括 CRISPR 系统。CRISPR 系统参照实施例二构建。

[0137] 在本说明书的描述中,参考术语“一个实施例”、“一些实施例”、“示例”、“具体示例”、或“一些示例”等的描述意指结合该实施例或示例描述的具体特征、结构、材料或者特点包含于本发明的至少一个实施例或示例中。在本说明书中,对上述术语的示意性表述不一定指的是相同的实施例或示例。而且,描述的具体特征、结构、材料或者特点可以在任何的一个或多个实施例或示例中以合适的方式结合。

[0138] 尽管已经示出和描述了本发明的实施例,本领域的普通技术人员可以理解:在不脱离本发明的原理和宗旨的情况下可以对这些实施例进行多种变化、修改、替换和变型,本发明的范围由权利要求及其等同物限定。

[0001]

## SEQUENCE LISTING

<110> 深圳华大基因研究院

<120> DNA 聚合酶 I 缺陷型菌株及其构建方法

<130> PIDC3152552

<160> 13

<170> PatentIn version 3.5

<210> 1

<211> 1818

<212> DNA

<213> Escherichia coli

<220>

<221> gene

<222> (1)..(1818)

<220>

<221> primer\_bind

<222> (30)..(52)

<223> 30-52 位中的任意 20nt

<220>

<221> primer\_bind

<222> (370)..(390)

<223> 370-390 位中的任意 20nt

<220>

<221> primer\_bind

<222> (1048)..(1068)

<223> 1048-1068 位中的任意 20nt

<400> 1

gtgatttctt atgacaacta cgtcaccatc cttgatgaag aaacactgaa agcgtggatt 60

gcgaagctgg aaaaagcggc ggtatttgca ttgataccg aaaccgacag ccttgataac 120

atctctgcta acctgtcgg gctttcttt gctatcgagc caggcgtagc ggcatatatt 180

[0002]



ccggttgctc atgattatct tgatgcgccc gatcaaatct ctgcgagcg tgcactcgag	240
ttgctaaaac cgctgctgga agatgaaaag gcgctgaagg tcgggcaaaa cctgaaatac	300
gatcgcggta ttctggcgaa ctacggcatt gaactgcgtg ggattgcgtt tgataccatg	360
ctggagtctt acatttcaa tagcgttgcc gggcgtcacg atatggacag cctcgcggaa	420
cgttggttga agcacaaaac catcactttt gaagagattg ctggtaaagg caaaaatcaa	480
ctgaccttta accagattgc cctcgaagaa gccggacgtt acgcccccga agatgcagat	540
gtcaccttgc agttgcatct gaaaatgtgg ccggatctgc aaaaacacaa agggccgttg	600
aacgtcttcg agaatatcga aatgccgctg gtgccgtgc ttccacgat tgaacgtaac	660
gggtggaaga tcgatccgaa agtgcgtcac aatcattctg aagagctcac cttcgtctg	720
gctgagctgg aaaagaaagc gcatgaaatt gcaggtgagg aatttaacct ttctccacc	780
aagcagttac aaaccattct cttgaaaaa cagggcatta aaccgctgaa gaaaacgccg	840
ggtgccgctc cgtcaacgtc ggaagagga ctggaagaac tggcctgga ctatccgttg	900
ccaaaagtga ttctggagta tcgtggtctg gcgaagctga aatcgacct caccgacaag	960
ctgccgctga tgatcaacc gaaaaccggg cgtgtgcata cctcttatca ccaggcagta	1020
actgcaacgg gacgtttatc gtcaaccgat cctaacctgc aaaacattcc ggtgcgtaac	1080
gaagaaggtc gtcgtatccg ccaggcgttt attgcgccag aggattatgt gattgtctca	1140
gcggactact cgcagattga actgcgcatt atggcgcac ttccgctga caaaggcttg	1200
ctgaccgcat tcgcggaagg aaaagatata caccgggcaa cggcggcaga agtgtttggt	1260
ttgccactgg aaaccgtcac cagcgagcaa gcccgtagcg cgaaagcgat caactttggt	1320
ctgatttatg gcatgagtgc ttccggtctg gcgcggcaat tgaacattcc acgtaaagaa	1380
gcgcagaagt acatggacct ttacttcgaa cgctaccctg gcgtgctgga gtatatgaa	1440
cgcacccgtg ctcaggcgaa agagcagggc tacgttgaaa cgctggacgg acgccgtctg	1500

[0003]

tatctgccgg atatcaaadc cagcaatggt gctcgtcgtg cagcggctga acgtgcagcc	1560
attaacgcgc caatgcaggg aaccgccgcc gacattatca aacggcgat gattgccgtt	1620
gatgcgtggt tacaggctga gcaaccgctg gtacgtatga tcatgcaggt acacgatgaa	1680
ctggtatttg aagttcataa agatgatggt gatgccgtcg cgaagcagat tcatcaactg	1740
atggaaaact gtaccctctt ggatgtgccg ttgctggtgg aagtggggag tggcgaanaac	1800
tgggatcagg cgcaactaa	1818
<210> 2	
<211> 20	
<212> DNA	
<213> 人工序列	
<220>	
<221> misc_feature	
<222> (1)..(20)	
<223> 能结合到 SEQ ID NO: 1 的 30-52 位中的 20nt 的序列	
<400> 2	
ctttcagtgt ttctcatca	20
<210> 3	
<211> 20	
<212> DNA	
<213> 人工序列	
<220>	
<221> misc_feature	
<222> (1)..(20)	
<223> SEQ ID NO: 2 的反向互补序列	
<400> 3	
tgatgaagaa acaactgaaag	20
<210> 4	

[0004]

<211>	20	
<212>	DNA	
<213>	人工序列	
<220>		
<221>	misc_feature	
<222>	(1)..(20)	
<223>	能结合到 SEQ ID NO: 1 的 370-390 位中的 20nt 的序列	
<400>	4	
	ggcaacgcta ttgagaatgt	20
<210>	5	
<211>	20	
<212>	DNA	
<213>	人工序列	
<220>		
<221>	misc_feature	
<222>	(1)..(20)	
<223>	SEQ ID NO: 2 的反向互补序列	
<400>	5	
	acattctcaa tagcgttgcc	20
<210>	6	
<211>	20	
<212>	DNA	
<213>	人工序列	
<220>		
<221>	misc_feature	
<222>	(1)..(20)	
<223>	能结合到 SEQ ID NO: 1 的 1048-1068 位中的 20nt 的序列	
<400>	6	
	aatgttttgc aggttaggat	20

[0005]

<210> 7  
 <211> 20  
 <212> DNA  
 <213> 人工序列

<220>  
 <221> misc\_feature  
 <222> (1)..(20)  
 <223> SEQ ID NO: 6 的反向互补序列

<400> 7  
 atcctaacct gcaaaacatt 20

<210> 8  
 <211> 25  
 <212> DNA  
 <213> 人工序列

<220>  
 <221> misc\_feature  
 <222> (5)..(24)  
 <223> SEQ ID NO: 2

<400> 8  
 tagcctttca gtgtttcttc atcag 25

<210> 9  
 <211> 25  
 <212> DNA  
 <213> 人工序列

<220>  
 <221> misc\_feature  
 <222> (5)..(24)  
 <223> SEQ ID NO: 3

<400> 9  
 aaactgatga agaaacactg aaagg 25

[0006]

<210> 10  
 <211> 25  
 <212> DNA  
 <213> 人工序列

<220>  
 <221> misc\_feature  
 <222> (5)..(24)  
 <223> SEQ ID NO: 4

<400> 10  
 tagcggcaac gctattgaga atgtg 25

<210> 11  
 <211> 25  
 <212> DNA  
 <213> 人工序列

<220>  
 <221> misc\_feature  
 <222> (5)..(24)  
 <223> SEQ ID NO: 5

<400> 11  
 aaacacattc tcaatagcgt tgccg 25

<210> 12  
 <211> 25  
 <212> DNA  
 <213> 人工序列

<220>  
 <221> misc\_feature  
 <222> (5)..(24)  
 <223> SEQ ID NO: 6

<400> 12  
 tagcaatggt ttgcagggtta ggatg 25

[0007]

<210> 13

<211> 25

<212> DNA

<213> 人工序列

<220>

<221> misc\_feature

<222> (5)..(24)

<223> SEQ ID NO: 7

<400> 13

aaacatccta acctgcaaaa cattg

25

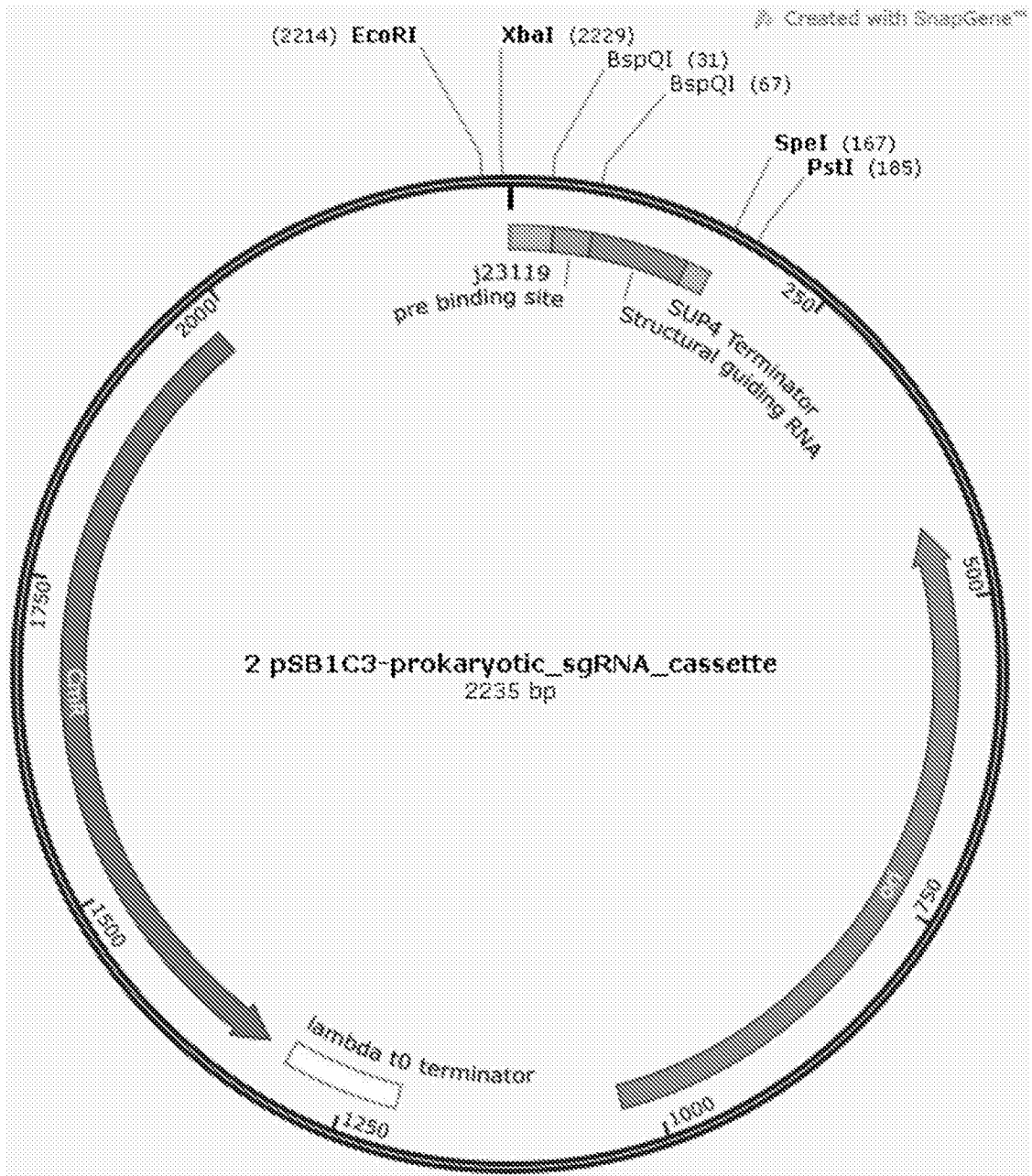


图 1

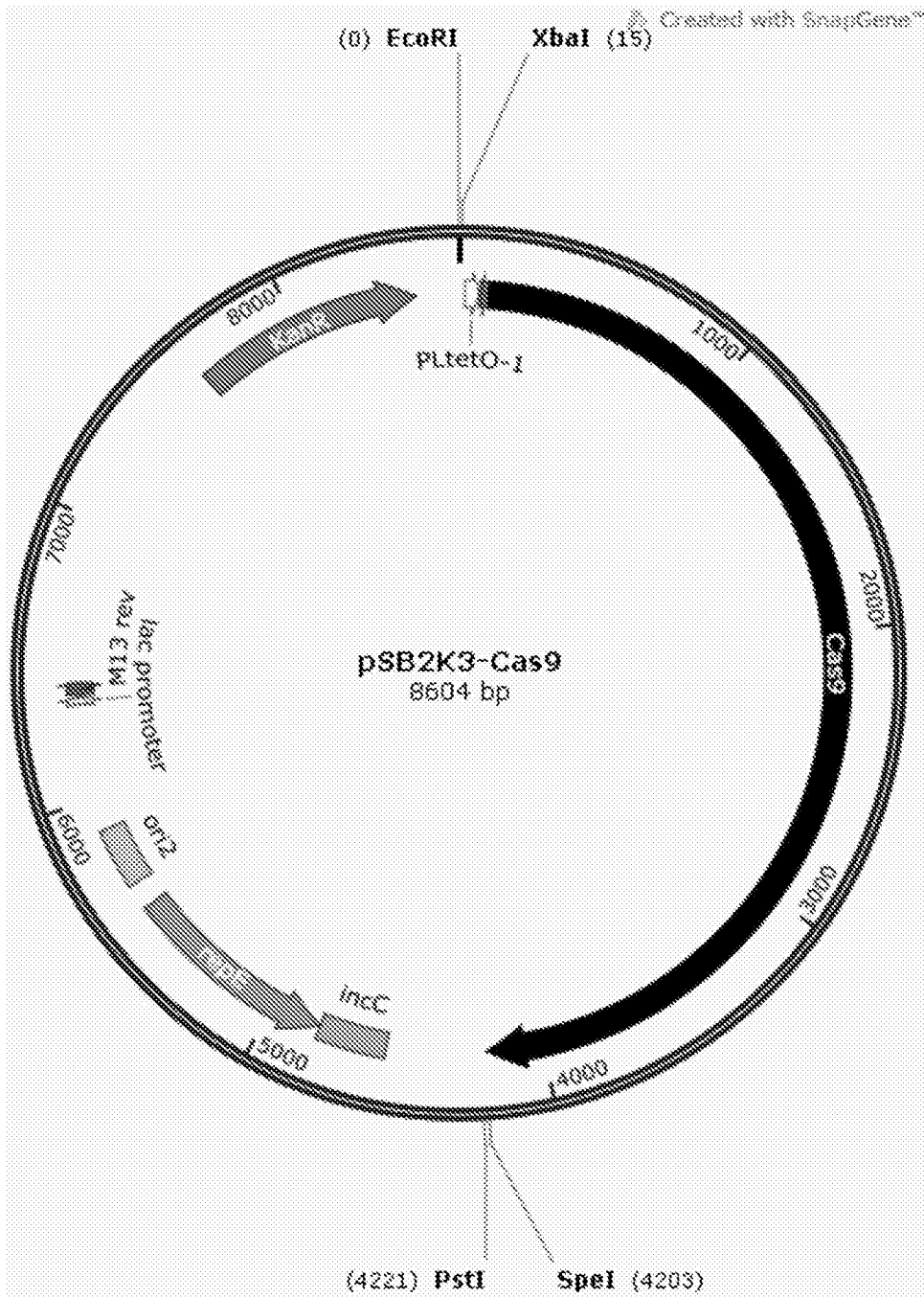


图 2



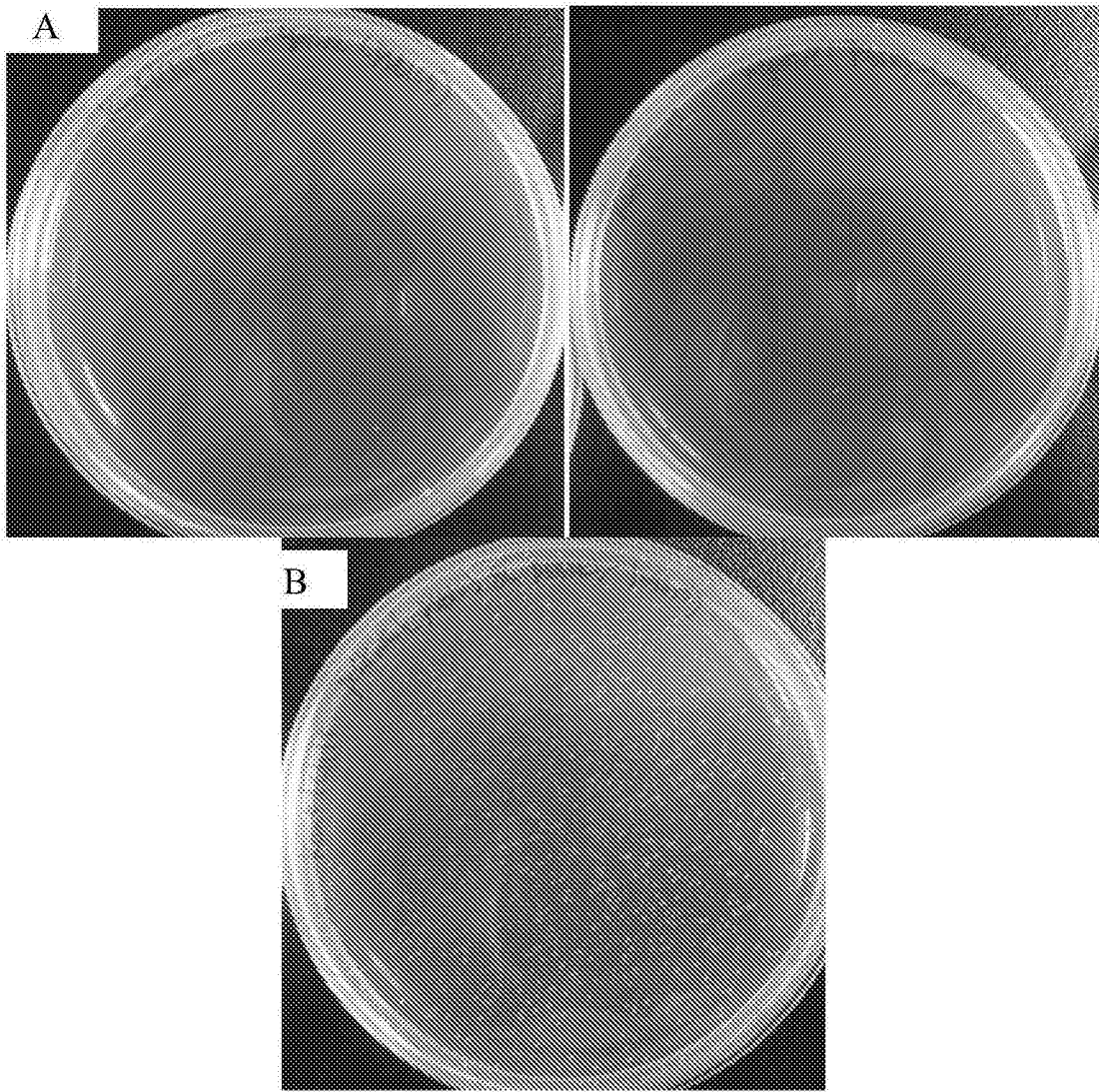


图 3

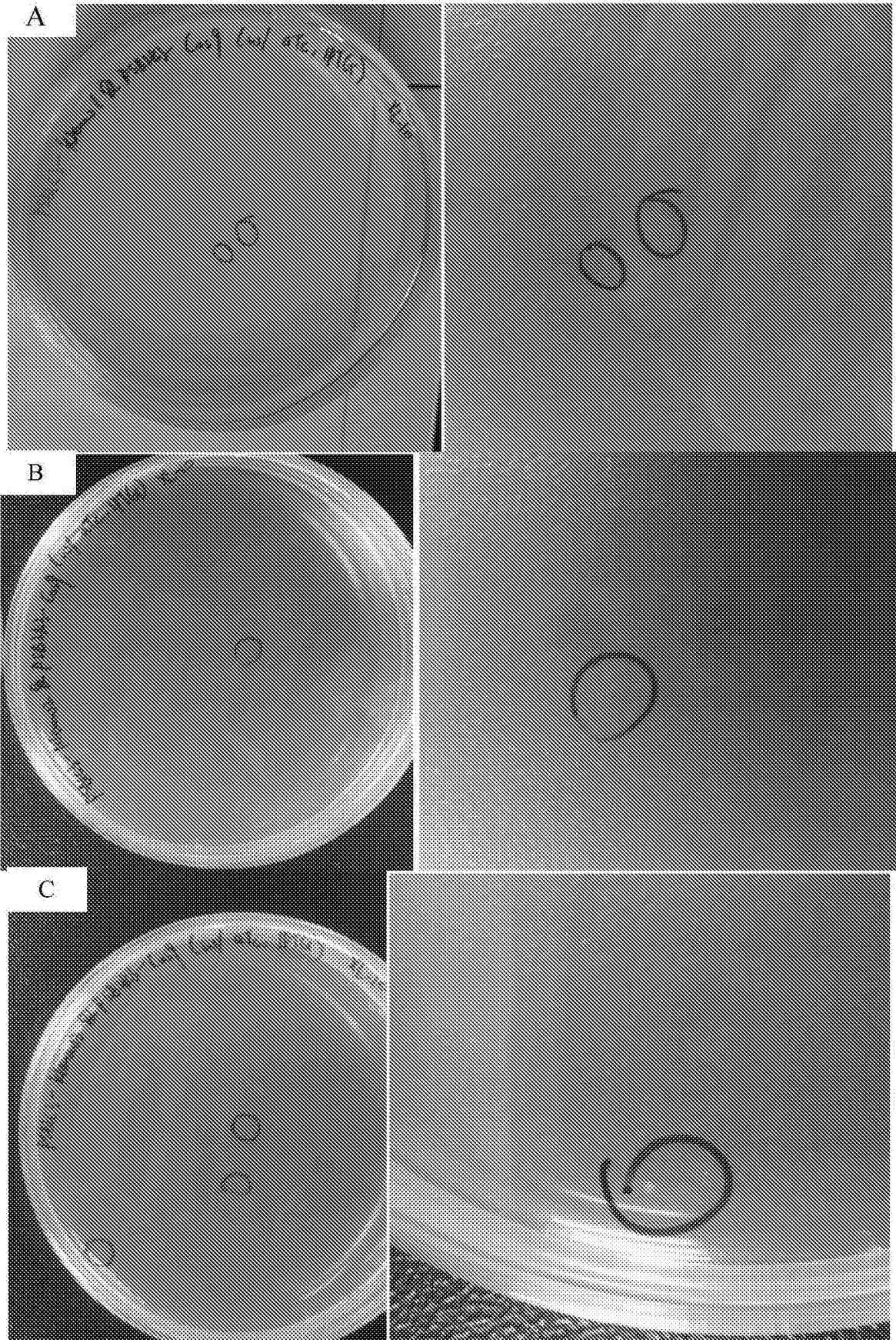


图 4

**klenow1**

```

878 AGCGGGCRAATGGTTACAGGCCAAAGGGGCAAAACCAGCCGCGAAGCCACAGGAAACCAG
  || |||
1  GAGCGGGCAATGGTTACAGGCCAAAGGGGCAAAACCAGCCGCGAAGCCGCRGGAACCAG
936 TGTTCACACGAGGCACCCAGAAGTGCAGCCCAACGGTGAATTCCTTATGACAACCTACGT CAC
|||
61  TGTTCACACGAGGCAGCCAGAGTGCAGCCCAACGGTGAATTCCTTATGACAACCTACGT CAC
996 CATCTTGTATGAAGAAA CACTGAAGCGTGGATTGCGAAGCTGGAAAAGCCGCGGTATT
|||
121  CATCTTGTATGAAGAAA CACTGAAGCGTGGATTGCGAAGCTGGAAAAGCCGCGGTATT
1856 TGCATTTGATACCGAAACCGACAGCCCTTGATAACATCTCTGCTAACCTGGTCGGGCTTTC
|||
181  TGCATTTGATAACCGAAACCGACAGCCCTTGATAACATCTCTGCTAACCTGGTCGGGCTTTC
1116 TTTTGTCTATCGAGCCAGGCGTAGCGGCATATATTCGGTTGCTCATGATATCTTGTATGC
|||
241  TTTTGTCTATCGAGCCAGGCGTAGCGGCATATATTCGGTTGCTCATGATATCTTGTATGC
1176 GCCCGATCAAATCTCTCGCGAGCGTGCACCTCGAGTTGCTAAAACCGCTGCTGGAAGATGA
|||
301  GCCCGATCAAATCTCTCGCGAGCGTGCACCTCGAGTTGCTAAAACCGCTGCTGGAAGATGA
1236 AAAGGCGCTGAAGGTGGGCAAAACCTGAAATACGATCGCGTATTCTGCGGAACCTACGG
|||
361  AAAGGCGCTGAAGGTGGGCAAAACCTGAAATACGATCGCGTATTCTGCGGAACCTACGG
1296 CATTGAACCTGCGTGGGATTGCGTTTGNIAACCATGCTGGAGTCTACATCTCAATAGCGT
|||
421  CATTGAACCTGCGTGGGATTGCGTTTGNIAACCATGCTGGAGTCTACATCTCAATAGCGT
1356 TGCCGGGCGTCAAGATATGGA CAGCCTCGCGGAACGTTGGTTEAAGCACAAAACCATCAC
|||
481  TGCCGGGCGTCAAGATATGGA CAGCCTCGCGGAACGTTGGTTEAAGCACAAAACCATCAC
1416 TTTTGAAGAGATTGCTGGTAAAGGCAAAAATCAACTGACCTTTAACCAGAT
|||
541  TTTTGAAGAGATTGCTGGTAAAGGCAAAAATCAACTGACCTTTAACCAGAT

```

图 5

```

klenow2
61  GTGTTGCAGACGAAGCCG CAGAAGTGACGGCAACGGTGATTTCCTATGACAACTACGTCA
   |||
935  GTGTTGCAGACGAAGSCACCAGAAGTGACGGCAACGGTGATTTCCTATGACAACTACSTCA
   |||

121  CCATCCTTGATGAAGAAACACTGAAAGCGTGGATTCGGAAAGCTGGAAAAAGCGCCGGTAT
   |||
995  CCATCCTTGATGAAGAAACACTGAAAGCGTGGATTCGGAAAGCTGGAAAAAGCGCCGGTAT
   |||

181  TTGCATTTGATACCGAAACCGCACAGCCTTGATAACATCTCTGCTAACCTGGTCCGGCTTT
   |||
1055  TTGCATTTGATACCGAAACCGCACAGCCTTGATAACATCTCTGCTAACCTGGTCCGGCTTT
   |||

241  CTTTTGCTATCGAGCCAGGGCGTAGCGGCAATATATCCGGTTGCTCATGATTATCTTGAAG
   |||
3115  CTTTTGCTATCGAGCCAGGGCGTAGCGGCAATATATCCGGTTGCTCATGATTATCTTGAAG
   |||

361  CGCCCGATCAAATCTCTCGGAGCGTGCACTCGAGTTGCTAAAACCGCTGCTGGAAGATG
   |||
1175  CGCCCGATCAAATCTCTCGGAGCGTGCACTCGAGTTGCTAAAACCGCTGCTGGAAGATG
   |||

361  AAAAGGCGCTGAAAGTCCGGGCAAAACCTGAAATACGATCGCGGTATTCCTGGCGAACTACG
   |||
1235  AAAAGGCGCTGAAAGTCCGGGCAAAACCTGAAATACGATCGCGGTATTCCTGGCGAACTACG
   |||

421  GCATTGAAGTGGCGTGGGATTGCGTTTGTATACCATGCTGGAGTCCCTACATTCTCAATAGCG
   |||
1295  GCATTGAAGTGGCGTGGGATTGCGTTTGTATACCATGCTGGAGTCCCTACATTCTCAATAGCG
   |||

481  TTGCCGGCGTCAAGATATGGACAGCCTCGCGGACSTTGGTTGAAACACAAAACCTATCA
   |||
1355  TTGCCGGCGTCAAGATATGGACAGCCTCGCGGACSTTGGTTGAAACACAAAACCTATCA
   |||

541  CTTTTGAAGAGATTGCTGATAAAAGCAAAATCAACTGACCTTTAAACGATTAACCTCG
   |||
1415  CTTTTGAAGAGATTGCTGATAAAAGCAAAATCAACTGACCTTTAAACGATTAACCTCG
   |||

601  AAGAAAGCCGGACGTTACGCGCGCGAAGATGCAGATGTCACCTTGCAATTGCACTGAAAA
   |||
1475  AAGAAAGCCGGACGTTACGCGCGCGAAGATGCAGATGTCACCTTGCAATTGCACTGAAAA
   |||

661  TGTGGCCGGATCTGCAAAAACACAAAGGGCCGTTGAAAGTCTTCGAGAAATATCGAAATGC
   |||
1535  TGTGGCCGGATCTGCAAAAACACAAAGGGCCGTTGAAAGTCTTCGAGAAATATCGAAATGC
   |||

721  CGCTGGTGCCAGTGCTTTCACGCATTGAAAGTAAAGGTTGAAAGTGGATCGATCCGAAAGTGC
   |||
1595  CGCTGGTGCCAGTGCTTTCACGCATTGAAAGTAAAGGTTGAAAGTGGATCGATCCGAAAGTGC
   |||

781  TGCACATCATCTGAAAGAGCTCACCCCTTGGTCTGGCTGAGCTGGAAAAAGAAAGGCGATG
   |||
1655  TGCACATCATCTGAAAGAGCTCACCCCTTGGTCTGGCTGAGCTGGAAAAAGAAAGGCGATG
   |||

```

图 6

**klenow3:**

```

938 TTTCAGACGAGCACCACAGACTGACGGCAACGGCTGATTTCTTATGACCAACTACGTCACCA
61  TTTCAGACGAGCACCACAGACTGACGGCAACGGCTGATTTCTTATGACCAACTACGTCACCA
998 TCCCTTGATGACGAAACA CTGAAAGCGTGGATTCGGAACCTGGAAAAAGCCCGGTATTTG
121 TCCCTTGATGACGAAACA CTGAAAGCGTGGATTCGGAACCTGGAAAAAGCCCGGTATTTG
1058 CATTTGATACCGAAACCGACAGCCCTTGATTAACATCTCTGCTAACCTGGTCGGGCTTTCTT
181 CATTTGATACCGAAACCGACAGCCCTTGATTAACATCTCTGCTAACCTGGTCGGGCTTTCTT
1118 TTGCTATGAGCCAGGCGTAGCCGCATATATTCGGGTTGCTCATGATTAATCTTGAATGCGC
241 TTGCTATGAGCCAGGCGTAGCCGCATATATTCGGGTTGCTCATGATTAATCTTGAATGCGC
1178 CCGATCAAACTCTCTGCGGAGCGTGCCTGCACTGAGTTGCTAAACCGCTGCTGGAGATGAAA
301 CCGATCAAACTCTCTGCGGAGCGTGCCTGCACTGAGTTGCTAAACCGCTGCTGGAGATGAAA
1238 AGGCGCTGAAAGCTGGGGCAAACTGAAAATCGATCGCGGTATCTTGGCCAACTACGGCA
361 AGGCGCTGAAAGCTGGGGCAAACTGAAAATCGATCGCGGTATCTTGGCCAACTACGGCA
1298 TTGAACTGGCTGGATTGGTTTGAATACCATGCTGGAGTCTTACATTCCTCAAATAGCGTTG
421 TTGAACTGGCTGGATTGGTTTGAATACCATGCTGGAGTCTTACATTCCTCAAATAGCGTTG
1358 CCGGGGCTCAGATAATGACAGCGCTGGCGGAGCTTGGTTGAAAGCACAAGCCATCACTT
481 CCGGGGCTCAGATAATGACAGCGCTGGCGGAGCTTGGTTGAAAGCACAAGCCATCACTT
1418 TTGAGAGAGATTCCTGCTAAAGGCAAAATCAACTGACCTTTAACCAGATTCGCGCTGAGG
541 TTGAGAGAGATTCCTGCTAAAGGCAAAATCAACTGACCTTTAACCAGATTCGCGCTGAGG
1478 AAGCCGACCTTACCGCCGGAAGATGAGATGTCATCACTTGCAGTTGCTATCTGAAAATGT
601 AAGCCGACCTTACCGCCGGAAGATGAGATGTCATCACTTGCAGTTGCTATCTGAAAATGT
1538 GCGCGGATCTGCAAAAACAACAAAGGCGGCTTGAACCTCTTGGAGATAATGAAATGCGG
661 GCGCGGATCTGCAAAAACAACAAAGGCGGCTTGAACCTCTTGGAGATAATGAAATGCGG
1598 TGGTGCGGCTGCTTTCAACGATTAAGCTAAAGCTGTGAGATGATCCAAAGTCTGTC
721 TGGTGCGGCTGCTTTCAACGATTAAGCTAAAGCTGTGAGATGATCCAAAGTCTGTC
1658 ACAATCATCTTAAGAGCTCACCTTGGCTGGCTGAGCTGAAAAGAAAGCCCAATGAAA
781 ACAATCATCTTAAGAGCTCACCTTGGCTGGCTGAGCTGAAAAGAAAGCCCAATGAAA
1718 TTGCAGGTGAGGAATTAACCTTCTTCCACCAAGCAGTTACAAACAATCTCTTTGAAA
841 TTGCAGGTGAGGAATTAACCTTCTTCCACCAAGCAGTTACAAACAATCTCTTTGAAA

```

图 7