



(12)发明专利

(10)授权公告号 CN 104024398 B

(45)授权公告日 2016.10.05

(21)申请号 201280050875.1

(22)申请日 2012.10.19

(65)同一申请的已公布的文献号
申请公布号 CN 104024398 A

(43)申请公布日 2014.09.03

(30)优先权数据
61/549375 2011.10.20 US

(85)PCT国际申请进入国家阶段日
2014.04.16

(86)PCT国际申请的申请数据
PCT/US2012/061027 2012.10.19

(87)PCT国际申请的公布数据
W02013/059595 EN 2013.04.25

(73)专利权人 斯卡拉布基因组有限责任公司
地址 美国威斯康星州

(72)发明人 F.R.布拉特纳 B.乔尔高
G.波斯菲

(74)专利代理机构 中国专利代理(香港)有限公
司 72001
代理人 杜艳玲 权陆军

(51)Int.Cl.
C12N 1/08(2006.01)
C12N 15/10(2006.01)

审查员 刘贺

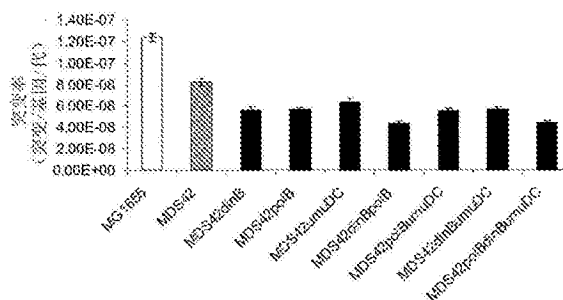
权利要求书1页 说明书12页 附图4页

(54)发明名称

具有提高的遗传稳定性的基因组简化的细菌

(57)摘要

本发明提供了具有提高的遗传稳定性的基因组简化的细菌。也提供了使用具有提高的遗传稳定性的基因组简化的细菌生产多肽的方法。



1. 一种基因组简化的大肠杆菌细菌,其基因组被遗传工程改造为比它的天然亲本菌株的基因组小了5%至30%,其中所述细菌缺少所有插入序列且缺少有功能的polB基因,和缺少有功能的dinB基因,并且任选地缺少有功能的umuDC基因,其中所述细菌已经从其中缺失了至少下述DNA区段:大肠杆菌K-12菌株MG1655的b0245-b0301、b0303-b0310、b1336-b1411、b4426-b4427、b2441-b2450、b2622-b2654、b2657-b2660、b4462、b1994-b2008、b4435、b3322-b3338、b2349-b2363、b1539-b1579、b4269-b4320、b2968-b2972、b2975-b2977、b2979-b2987、b4466-4468、b1137-b1172、b0537-b0565、b0016-b0022、b4412-b4413、b0577-b0582、b4415、b2389-b2390、b2392-b2395、b0358-b0368、b0370-b0380、b2856-b2863、b3042-b3048、b0656、b1325-b1333、b2030-b2062、b2190-b2192、b3215-b3219、b3504-b3505、b1070-b1083、b1878-b1894、b1917-b1950、b4324-b4342、b4345-b4358、b4486、b0497-b0502、b0700-b0706、b1456-b1462、b3481-b3484、b3592-b3596、b0981-b0988、b1021-b1029、b2080-b2096、b4438、b3440-b3445、b4451、b3556-b3558、b4455、b1786、b0150-b0153和b2945。

2. 根据权利要求1所述的细菌,其中所述细菌的天然亲本菌株是B菌株。

3. 根据权利要求2所述的细菌,其中所述细菌的天然亲本菌株是菌株BL21(DE3)。

4. 根据权利要求1所述的细菌,其中所述细菌的天然亲本菌株是K12菌株。

5. 根据权利要求4所述的细菌,其中所述细菌的天然亲本菌株是K12菌株MG1655。

6. 根据权利要求5所述的细菌,其中所述细菌通过使polB、dinB和任选地umuDC基因无功能而从MDS42产生。

7. 根据权利要求1所述的细菌,所述细菌具有有功能的umuDC基因。

8. 根据权利要求1所述的细菌,所述细菌缺少有功能的umuDC基因。

9. 根据任意前述权利要求所述的细菌,所述细菌包含异源核酸。

10. 根据权利要求9所述的细菌,其中所述异源核酸包含可操作地连接到表达控制序列的编码多肽的核酸。

11. 一种用于生产多肽的方法,所述方法包括:在适合表达所述多肽的条件下培养根据权利要求10所述的细菌,和收集所述多肽。

具有提高的遗传稳定性的基因组简化的细菌

[0001] 本专利申请要求2011年10月20日提交的美国临时申请号61/549,375在美国法典第35篇第119(e)条下的权益,其整个内容特此通过引用并入。

技术领域

[0002] 本发明涉及具有非常低突变率的基因组简化的细菌和使用所述基因组简化的细菌的方法。所述基因组简化的细菌对于核酸的高保真度维持和已经被证实难以在细菌宿主中克隆的基因的稳定表达而言是特别有用的。

背景技术

[0003] 产生多样性的固有机制对于细菌群体在自然界中存在的动态变化的环境条件中的存活而言是重要的。但是,在实验室的受控环境中,这些机制可以导致不希望的遗传型和表型改变,并且建立的生产菌株或克隆文库的自发性遗传修饰通常是非常不希望的。

[0004] 大肠杆菌(*Escherichia coli*, *E. coli*)是一种通用的克隆宿主,并且是在研究和工业中最常见的用于生产蛋白、代谢物和次级代谢物的生物。在这些场合中已经做出几种改进来改善大肠杆菌宿主的性能,所有这些改进遵循以下基本原理:将增加的给定生物材料生产的流线型代谢途径与不希望的副产物的减少相偶联。与此同时,已经从大肠杆菌背景除去多种非必需基因以形成可存活的基因组简化的大肠杆菌菌株,其几乎没有或根本没有显著的生长减弱。

[0005] 尽管已经证实这些基因组简化的细菌在许多方面是有益的,有些处于它们的功能形式的基因仍然难以或不可能克隆进细菌载体、甚至基因组简化的细菌中。因此,需要可以在其中克隆这样的基因的、具有非常低突变率的稳定细菌宿主。

发明内容

[0006] 本发明提供了一种基因组简化的细菌,其中编码3种易出错的DNA聚合酶Pol II、Pol IV和Pol V中的至少一种的基因是无功能的。在一个优选的实施方案中,编码Pol II和Pol IV的基因是无功能的,且编码Pol V的基因是有功能的或无功能的。在一个特别优选的实施方案中,这些基因在基因组简化的细菌中都是无功能的。通过从所述细菌的基因组中部分地或完全地缺失所述基因,可以使所述基因为无功能的,或者通过破坏所述基因而使其为无功能的。

[0007] 在一个实施方案中,所述基因组简化的细菌的基因组具有这样的基因组:其被遗传工程改造成比它的天然亲本菌株的基因组小约5%至约30%,且缺少所有插入序列。基因组简化的细菌可以通过从细菌的天然亲本菌株中缺失选定的基因而产生,或者可以例如完全合成为预选基因的集合。从本文中的讨论容易明白,与在进行比较并共有某些必需基因的天然亲本菌株中发现的基因的完整补体(full complement)相比,基因组简化的细菌更小。

[0008] 还提供了采用所述基因组简化的细菌生产多肽的方法。在一个实施方案中,如下

生产多肽:在适合表达所述多肽的条件下,培养基因组简化的细菌,所述细菌具有一种或多种无功能的基因,所述基因选自编码Pol II、PolIV和PolV的基因,且所述细菌进一步包含可操作地连接到表达控制序列的编码所述多肽的核酸。在一个有关的实施方案中,所述核酸编码“有毒”多肽,其难以或不可能克隆进具有功能性PolIII、PolIV和PolV基因的细菌中。

[0009] 在下文中更详细地描述了本发明的这些和其它实施方案。

附图说明

[0010] 图1解释了与MDS42和MG1655对照相比,具有无功能的(在该情况下,缺失)dinB、polB和umuDC基因(单独地和以每种可能的组合)的基因组简化的细菌(MDS42)的自发突变率。通过对在cycA中发生的突变进行波动分析,确定突变率。野生型MG1655至MDS42的突变率下降是由于插入事件的缺失,而从MDS42至不同的易出错的聚合酶突变体的进一步下降是由于更低的点突变率。值为4个独立测量结果的平均值。

[0011] 图2解释了不同的应激对不同菌株的突变率的影响。显示了如下施加的应激的影响:过度产生绿色荧光蛋白(GFP),从pSG-ORF238过度产生有毒肽,和用丝裂霉素C处理。使用cycA波动测定进行所有测量,值为3个各自独立的测量结果的平均值。在有0.1μg/ml丝裂霉素C存在下,BL21(DE3)和MDS42recA没有生长。

[0012] 图3解释了不同菌株的突变谱的对比。条形图显示了通过聚合酶链式反应(PCR)分析检测出的cycA突变类型的分布。在MG1655、MDS42和MDS42polBdinBumuDC中的缺失的份额低至难以看到,并且在BL21(DE3)中没有检测到缺失。

[0013] 图4解释了通过利福平抗性测定测得的MDS42和MDS42polBdinBumuDC的突变率。值为3个各自独立的测量结果的平均值。

[0014] 图5解释了SinI酶的过度产生对MDS42-T7(mcrBC⁻)宿主的生长的毒性效应。从IPTG-可诱导的pSin32质粒过度产生SinI。两个测量结果是25个各自独立的菌落的O.D.540值的平均值,所述O.D.540值每5分钟使用Bioscreen C自动化仪器进行测量。

[0015] 图6解释了SinI甲基转移酶的过度产生对突变率的影响。从pSin32质粒过度产生SinI甲基转移酶。使用cycA波动测定来测量突变率。值为3个各自独立的测量结果的平均值。

[0016] 图7解释了含有突变的sinI的质粒在不同宿主中的积累。从pSin32表达SinI甲基转移酶。将质粒在不同的间隔分离,并针对导致该酶功能丧失的突变进行筛选(通过转化进McrBC⁺和McrBC⁻宿主中)。值为3个各自独立的测量结果的平均值。

[0017] 图8解释了表达SinI的不同菌株的培养物生长至OD=0.7所需的时间。测量了携带pSin32的菌株的生长曲线,其在OD=0.2(0min)时用IPTG诱导。选择OD=0.7作为截止值,以指示样品已经克服了过度产生的SinI甲基转移酶的毒性效应。显示的值为每种菌株的50个独立样品的测量结果的平均值。

[0018] 图9解释了能够转化MG1655的pSin32质粒上所携带的sinI的突变。对分离自McrBC⁺宿主的8个pSin32质粒样品测序。它们中的7个携带移码突变,所述移码突变会在sinI中产生新的终止密码子(位置182、183、194、195、196、208和862)。所述突变之一是导致Asn→Thr变化(位置880)的A→C转变。相对于sinI的第一个核苷酸显示了每个突变的核苷酸位置。

具体实施方式

[0019] 尽管本发明能够以不同的形式体现,但作出几个实施方案的以下描述,条件是本发明公开内容应认为是本发明的例证,且无意将本发明限于所阐述的具体实施方案。标题仅仅为了方便而提供,且不应解释为以任何方式限制本发明。在任何标题下阐述的实施方案可与在任何其它标题下阐述的实施方案组合。

[0020] 除非另外明确指出,否则在本申请中指出的不同范围内的数值的应用,象在所述范围内的最小值和最大值前面都加有词语“约”一样,作为近似值描述。以此方式,超过和低于所述范围的轻微变化,可以用于达到基本上与该范围内的值相同的结果。当提及数值时,本文中使用的术语“约”和“大约”应当具有相关技术领域的技术人员已知的它们的平常和普通含义。并且,范围的公开意图作为连续范围,包括在列举的最小值和最大值之间的每个值以及可以由这样的值形成的任意范围。这包括含有或不含有有限上限和/或有限下限的可形成的范围。这也包括通过将给定的公开数字除以另一个公开的数字而可衍生出的比率。因此,技术人员会明白,许多这样的比率、范围和比率范围可以从本文中提供的数据和数字明确地衍生出,并且都代表本发明的不同实施方案。

[0021] 术语“基因组简化的细菌”在本文中是指这样的细菌:它的基因组(例如编码蛋白的基因)缺失了约1%至约75%,例如缺失了约5%、约10%、约20%、约30%、约40%、约50%或约60%的基因组。在一个实施方案中,在本发明的实践中使用的基因组简化的细菌具有这样的基因组:其优选地遗传工程改造成比天然亲本菌株的基因组小了至少百分之二(2%)且至多百分之二十(20%)(包括其间的任意数字)。优选地,所述基因组比天然亲本菌株的基因组小了至少百分之五(5%)且至多百分之三十(30%)。更优选地,所述基因组比天然亲本菌株的基因组小了百分之八(8%)至百分之十四(14%)至百分之二十(20%)(包括其间的任意数字)或更多。可替换地,所述基因组可以工程改造成比天然亲本菌株的基因组小了小于20%、小于30%、小于40%或小于50%。术语“天然亲本菌株”是指在科学共同体通常理解的自然或天然环境中发现的细菌菌株,并且可以对其基因组做出一系列缺失以制备具有更小基因组的细菌菌株。天然亲本菌株也表示与经工程改造的菌株进行对比的菌株,并且其中所述经工程改造的菌株小于天然亲本菌株的完整补体。如下计算在一系列缺失以后基因组减小的百分比:将“所有缺失之后缺失的碱基对的总数”除以“所有缺失之前基因组中的碱基对总数”,然后乘以100。类似地,如下计算基因组比天然亲本菌株减小的百分比:将具有较小基因组(不论产生它的方法)的菌株中的核苷酸的总数除以天然亲本菌株中的核苷酸的总数,然后乘以100。

[0022] 在一个实施方案中,术语“基因组简化的细菌”表示这样的细菌:给其除去上述量的基因组不会不可接受地影响所述生物体的在基本培养基上生长的能力。在本发明的上下文中,2种或更多种基因的除去是否“不可接受地影响”所述生物体的在基本培养基上生长,取决于具体的应用。例如,增殖速率的30%下降对于一种应用而言可能是可接受的,但是对于另一种应用而言是不可接受的。另外,通过诸如改变培养条件等措施,可以减少从基因组缺失DNA序列的不良作用。这样的措施可能将否则是不可接受的不良作用转化为可接受的作用。在一个实施方案中,所述增殖速率与亲本菌株大约相同。但是,在比亲本菌株的增殖速率低约5%、10%、15%、20%、30%、40%至约50%的范围内的增殖速率是在本发明范围

内。更具体地,本发明的细菌的倍增时间可以在约5分钟至约3小时的范围内。合适的基因组简化的细菌的非限制性例子以及从细菌(诸如大肠杆菌)缺失DNA的方法公开在美国专利号6,989,265和7,303,906、美国专利公开号20060270043、2006/0199257和2007/0054358以及WIPO公开号W02003/070880中,它们中的每一篇特此通过引用并入本文。

[0023] 在几个实施方案中,提供了基因组简化的细菌,其具有至少一种无功能的基因,所述基因选自编码DNA聚合酶II、DNA聚合酶IV和DNA聚合酶V的基因。与具有相同遗传背景细菌(但是其中这些基因是有功能的)相比,基因组简化的细菌(其中这些基因中的一种或多种是无功能的)表现出遗传稳定性的大幅提高。在一个方面,通过缺失,例如通过在美国专利号6,989,265的第8列第45行至第14列第41行处描述的“无痕”缺失方法,使得所述基因是无功能的。这些方法会造成靶基因的精确缺失,从所述缺失方法(即“无痕”缺失)不会产生插入的DNA,并且因此是优选的缺失方法。但是,应当理解,可以采用本领域完全地或部分地已知的缺失靶基因的任何方法来使得编码DNA聚合酶II、DNA聚合酶IV和DNA聚合酶V的基因中的一种或多种基因是无功能的。可替换地,通过使用本领域已知的任意技术破坏基因,可以使得编码DNA聚合酶II、DNA聚合酶IV和DNA聚合酶V的基因中的一种或多种基因是无功能的。例如,可以如下破坏靶基因:通过同源重组,用无功能的等位基因替代所述基因。可以联合使用破坏和缺失,以在细菌中产生无功能的编码DNA聚合酶II、DNA聚合酶IV和DNA聚合酶V的基因的任意组合。

[0024] 在一个实施方案中,可以使得编码DNA聚合酶II、IV和V的基因中的任一种基因是无功能的,并且剩余的2种基因可以是有功能的。在其它实施方案中,可以在基因组简化的细菌中使得编码DNA聚合酶II、IV和V的基因中的2种基因的任意组合是无功能的,并且剩余的基因可以是有功能的。例如,可以使得编码DNA聚合酶II和IV的基因是无功能的,并且编码DNA聚合酶V的基因可以是有功能的。可替换地,可以使得编码DNA聚合酶II和V的基因是无功能的,并且编码DNA聚合酶IV的基因可以是有功能的。可替换地,可以使得编码DNA聚合酶IV和V的基因是无功能的,并且编码DNA聚合酶II的基因可以是有功能的。在一个优选的实施方案中,在基因组简化的细菌中,编码DNA聚合酶II和DNA聚合酶IV的基因是无功能的,并且编码DNA聚合酶V的基因是有功能的或无功能的。在一个特别优选的实施方案中,在基因组简化的细菌中,编码DNA聚合酶II、DNA聚合酶IV和DNA聚合酶V的基因都是无功能的。

[0025] 在另一个优选的实施方案中,基因组简化的细菌(其具有一种或多种无功能的基因,所述基因选自编码DNA聚合酶II、DNA聚合酶IV和DNA聚合酶V的基因)具有这样的基因组:其被遗传工程改造为比天然亲本菌株的基因组小了至少百分之五(5%)且至多百分之三十(30%)(包括其间的任意数字)。在另一个优选的实施方案中,基因组简化的细菌具有在4.41Mb至3.71Mb之间、4.41Mb至3.25Mb之间、或4.41Mb至2.78Mb之间的基因组。

[0026] 本发明的基因组简化的细菌的亲本可以是任意细菌菌株。在一个优选的实施方案中,本发明的基因组简化的细菌的亲本是大肠杆菌菌株,诸如大肠杆菌K-12或B菌株。大肠杆菌K12菌株包括衍生菌株诸如MG1655、W3110、DH1、DH10B、DH5 α 、Inva、Top10、Top10F、JM103、JM105、JM109、MC1061、MC4100、XL1-Blue、EC100、BW2952或EC300。大肠杆菌B菌株包括REL606、BL/R和BL21(DE3)。

[0027] 亲本菌株的基因组的核苷酸序列可以是部分地或完全地已知的。几种大肠杆菌和其它常用实验室微生物的完整基因组序列是已知的(参见例如Blattner等人,Science,

277:1453-74(1997);GenBank登记号U00096;NCBI数据库,登记号AP009048,Perna等人,Nature,409,529-533(2001);Hayashi等人,DNA Res.,8,11-22(2001);Welch等人,Proc.Natl.Acad.Sci.,USA99:17020-17024(2002),GenBank登记号AE014075,EMBL登记号CP000948,EMBL登记号CP001637,EMBL登记号CP001396,EMBL登记号CP000819,和EMBL登记号CP001509,它们中的每一篇通过引用并入本文)。

[0028] 在一个优选的实施方案中,本发明的基因组简化的细菌的亲本是大肠杆菌菌株K12MG1655(注解版本m56)(NCBI登记号U000961),其具有4,639,674个碱基对的基因组。在另一个优选的实施方案中,基因组简化的细菌的亲本是大肠杆菌菌株BL21(DE3)(EMBL登记号CP001509),其具有4,557,508个碱基对的基因组。在表1中提供了编码DNA聚合酶II(polB)、DNA聚合酶IV(dinB)和DNA聚合酶V(umuDC)的基因在大肠杆菌K12MG1655基因组中的坐标。

[0029] 表1

[0030]

基因	坐标
polB(b0060)	63429-65780
dinB(b0231)	250898-251953
umuDC(b1183-b1184)	1229990-1231677

[0031] 在一个特别优选的实施方案中,提供了基因组简化的大肠杆菌细菌,其基因组比天然亲本菌株的基因组小了百分之五(5%)至百分之三十(30%),且缺少所有插入序列(IS)元件,并且具有至少一种无功能的基因,所述基因选自编码DNA聚合酶II、DNA聚合酶IV和DNA聚合酶V的基因。在美国专利公开号2003/138937(其内容通过引用并入本文)的图1和表2中显示了IS元件在大肠杆菌MG1655的基因组图谱(注解版本54)上的位置。在大肠杆菌中常见的且可以被除去的插入序列元件包括、但不限于,IS1、IS2、IS3、IS4、IS5、IS30、IS150、IS186、IS600、IS911和IS10。在一个特别优选的实施方案中,提供了基因组简化的大肠杆菌,其缺少所有插入序列且具有无功能的polB和dinB基因,且甚至更优选地,具有无功能的polB、dinB和umuDC基因。

[0032] 在一个有关的实施方案中,所述基因组简化的细菌是这样的大肠杆菌细菌,其至少缺少下述基因(基于在Blattner等人,Science,277:1453-74和GenBank登记号400096中阐述的命名法,用“b”编码来标识):b0245-b0301、b0303-b0310、b1336-b1411、b4426-b4427、b2441-b2450、b2622-b2654、b2657-b2660、b4462、b1994-b2008、b4435、b3322-b3338、b2349-b2363、b1539-b1579、b4269-b4320、b2968-b2972、b2975-b2977、b2979-b2987、b4466-4468、b1137-b1172、b0537-b0565、b0016-b0022、b4412-b4413、b0577-b0582、b4415、b2389-b2390、b2392-b2395、b0358-b0368、b0370-b0380、b2856-b2863、b3042-b3048、b0656、b1325-b1333、b2030-b2062、b2190-b2192、b3215-b3219、b3504-b3505、b1070-b1083、b1878-b1894、b1917-b1950、b4324-b4342、b4345-b4358、b4486、b0497-b0502、b0700-b0706、b1456-b1462、b3481-b3484、b3592-b3596、b0981-b0988、b1021-b1029、b2080-b2096、b4438、b3440-b3445、b4451、b3556-b3558、b4455、b1786、b0150-b0153和b2945,并且也具有一种或多种无功能的选自polB、dinB和umuDC的基因。在一个特别优选的实施方案中,polB和dinB是无功能的,且甚至更优选地polB、dinB和umuDC都是无功能的。

所述基因组简化的大肠杆菌细菌可以是菌株MDS42,其基因组缺少所有插入序列,其具有一种或多种无功能的polB、dinB和umuDC基因,优选地具有无功能的polB和dinB基因,且甚至更优选地所有3种基因都是无功能的。所述基因组简化的菌株也可以是菌株MDS43或MDS66(或任意衍生菌株),其具有一种或多种无功能的polB、dinB和umuDC基因,优选地所有3种基因都是无功能的。

[0033] 可以缺失不同的编码蛋白的基因以形成基因组简化的细菌。在大肠杆菌和其它细菌中,可以缺失的DNA序列的类型包括通常不利地影响生物体或该生物体的基因产物的稳定性的那些类型。这样的造成不稳定性的元件包括、但不限于可能在基因组不稳定性中起作用的转座元件、插入序列和其它“自在DNA”元件。例如,插入序列(IS)元件和它们的有关转座常见于细菌基因组中,且因而是缺失的靶标。IS序列常见于大肠杆菌中,并且它们都可以被缺失。在该文件中为了清楚起见,我们使用术语IS元件和转座元件泛指可以在基因组中从一个点移至另一个点的完整的或有缺陷的DNA元件。IS元件在科学和技术中的有害作用的一个例子是以下事实:它们可以在繁殖过程中从宿主大肠杆菌的基因组跳进BAC质粒中进行排序。通过从宿主细胞缺失所有IS元件,可以防止该伪像。对于具体应用而言,也可以缺失与基因组不稳定性有关的其它具体基因,诸如有活性的和无活性的原噬菌体。

[0034] 本发明的基因组简化的细菌也可以经工程改造成缺少,例如,但不限于,细菌的生长和代谢不必要的某些基因、假基因、原噬菌体、不希望的内源限制修饰基因、致病性基因、毒素基因、菌毛基因、周质蛋白基因、透明质酸酶基因、脂多糖基因、III类分泌系统、噬菌体毒力决定簇、噬菌体受体、致病岛、RHS元件、未知功能的序列和没有发现在细菌的相同天然亲本物种的2个菌株之间共有的序列。还可以缺失或省略细胞存活不需要的其它DNA序列。

[0035] 本发明的基因组简化的细菌可以包含编码多肽的异源核酸。所述多肽可以是治疗性蛋白诸如胰岛素、白介素、细胞因子、生长激素、生长因子、促红细胞生成素、集落刺激因子、干扰素或抗体。可以将所述异源核酸置于载体(诸如质粒)内,且可操作地连接到启动子和任选的额外调节序列。

[0036] 具有一种或多种无功能的polB、dinB和umuDC基因、优选地至少具有无功能的polB和dinB基因、且进一步缺少所有插入序列的基因组简化的细菌会表现出惊人的遗传稳定性,所述稳定性会实现甚至在缺乏所有插入序列的基因组简化的细菌中难以或不可能分离或维持的有毒核酸的克隆。“有毒”核酸可以是这样的核酸:当其在宿主菌株中繁殖时,会导致升高的突变率。有毒核酸也可能导致升高的IS元件转座率。通过与繁殖对照核酸的宿主菌株进行对比,可以确定有毒核酸的升高的突变率。

[0037] 包含一种或多种无功能的polB、dinB和umuDC基因的基因组简化的细菌可以用于生产多肽。简而言之,可以在足以允许多肽产物表达的条件下温育本发明的细菌,其包含如上所述可操作地连接到表达控制序列的编码多肽的异源核酸。

[0038] 即使良好耐受的目标蛋白的过表达也可能导致升高的IS转座率,并活化细胞的应激反应,从而导致显著增加的突变率。编码目标蛋白的质粒在这些条件下快速地获得功能缺失突变,并且携带这些突变质粒的细菌会快速地变成在培养物中占优势,这至少部分地归因于编码过表达的蛋白的完整质粒的生长抑制作用。本发明的表现出惊人的基因组稳定性和保真度的细菌会延迟这样的突变质粒的出现,并且所述细胞可以在延长的时间段内生产有功能的毒蛋白。

[0039] 可以在周质或细胞质中表达重组蛋白。在周质中的蛋白表达常规地用于工业用途,并且已经综述在Hanahan,J.Mol.Biol.,166:557-580(1983);Hockney,Trends Biotechnol.,12:456-632(1994);和Hannig等人,Trends Biotechnol.,16:54-60(1998),它们中的每一篇通过引用并入本文。通过表达融合蛋白,可以在周质中生产重组蛋白,在所述融合蛋白中,所述重组蛋白连接至造成分泌进周质间隙中的信号肽。在周质中,所述信号肽可以被特定信号肽酶切掉。运输进周质间隙中的蛋白可以是生物学上有活性的。

[0040] 可以与分子伴侣/二硫键形成酶一起共表达重组蛋白,它们可以提供重组蛋白的适当折叠。这样的可用于重组蛋白的周质表达的蛋白的核酸序列包括、但不限于,在美国专利号5,747,662、5,578,464和6,022,952(它们中的每一篇通过引用并入本文)中描述的那些。

[0041] 实施例1

[0042] 基因组简化的大肠杆菌的生产

[0043] 如在国际专利公开号W02003/070880(其通过引用并入本文)中所述,生产基因组简化的菌株MDS39。简而言之,通过从亲本菌株大肠杆菌MG1655产生一系列39个累积的核酸序列缺失(基因组的大约14.1%),生产一系列基因组简化的菌株(MDS01-MDS39)。

[0044] 与含有K-12序列和IS数据库中的所有序列的基因组扫描芯片(NimbleGen Systems, Madison, WI)的杂交揭示,MDS39(被设计成缺少所有IS元件的第一菌株)意外地含有IS元件的额外拷贝,其在它的生产过程中跳至新位置。缺失这些IS元件以产生MDS40。从MDS40缺失fhuACDB(tonA基因座)以产生MDS41。为了产生MDS01-MDS41而做出的每个累积缺失的位置和功能可以参见美国申请公开号2007/0054358(其整个内容通过引用并入本文)的表2。然后从MDS41缺失endA基因以产生MDS42。

[0045] 使用如在美国专利号6,989,265和Feher等人,Methods Mol.Biol.,416:251-259(2008)中描述的基于自杀质粒的方法,用质粒pST76-A和pSTKST,以无痕方式从MDS42的基因组中缺失编码DNA聚合酶II(polB)、DNA聚合酶IV(dinB)和DNA聚合酶V(umuDC)的基因。基因缺失个别地进行,并且也以所有可能的组合进行连接,以产生下述菌株:MDS42polB、MDS42dinB、MDS42umuDC、MDS42polBdinB、MDS42polBumuDC、MDS42dinBumuDC和MDS42polBdinBumuDC。如下组合各个缺失:对缺失构建体的标记的(含有整合的自杀质粒)中间体进行P1噬菌体转导,随后进行内切核酸酶切割刺激的质粒的脱离重组(Out-recombination)和损失。通过聚合酶链式反应(PCR)来验证所有缺失,并使用侧接引物测序。在研究中使用的引物序列列出在表2中:

[0046] 表2

[0047]

引物名称	序列(5'-3')	应用
polB-A	ccgaattcagtatccaggcgagt	缺失polB
polB-BR	caggcaggtgtggcggagggaatact	缺失polB
polB-BF	tccgccacacctgcctgcgccacgct	缺失polB
polB-C	ccggatccattggcggcattgt	缺失polB
polB-D	tgctgaacaccagtttgct	缺失polB
polB-E	aaccggtgaagtgggtga	缺失polB

dinB-A	ccggtaccgggcataccgatgcga	缺失dinB
dinB-BR	cagaatatacattgtctacctctcaacact	缺失dinB
dinB-BF	gaggtgagcaatgtatatcttctggtgtgca	缺失dinB
dinB-C	ccggatccgccggttaacgcataca	缺失dinB
dinB-D	gtgttcgactcgctcgat	缺失dinB
dinB-E	gagtcgctcgtagagtgcata	缺失dinB
umuDC-A	ggaattcggatgagcgtcgtcgcca	缺失umuDC
umuDC-BR	ttgagcgcaacaacagcagcgtgacaa	缺失umuDC
umuDC-BF	gctgctgttgttgcgctcaatgaacctt	缺失umuDC
umuDC-C	gctgcagatcgcttacctgattgtc	缺失umuDC
umuDC-D	aatgctccatctgcggtt	缺失umuDC
umuDC-E	gctctatccttcgcggtt	缺失umuDC
lexA-A	gttatggctcgcatcttggata	修饰lexA
lexA-BR	gatatctttcatcgCcatcccgtgacgcga	修饰lexA
lexA-BF	ggatgGcgatgaaagatatcgga	修饰lexA
lexA-C	ccggatcccagcaacggaacggt	修饰lexA
lexA-D	cgggtgctgattgccatta	修饰lexA
lexA-E	gggctatcaagatgacca	修饰lexA
recA-D	cggctagcgacgggatgttgattc	缺失recA
recA-E	gtgctgattatgccgtgt	缺失recA
BMD30-A	ccgaattcagtcgcgacgcaactt	缺失mcrBC
BMD30-BR	ctcgcccttaattacatacttttgggtgc	缺失mcrBC
BMD30-BF	tatgtaaattaaggcgagattattaaa	缺失mcrBC
BMD30-C	ccggatccacatggcgcgttacaa	缺失mcrBC
BMD30-D	tgataccgcccgcacaaca	缺失mcrBC
BMD30-E	actgggtgtgtctcgcaag	缺失mcrBC
cycA-D	ctgatgccggtaggttct	分析cycA突变
cycA-E	gcgccatccagcatgata	分析cycA突变
AK54-D	atgataatgaatgacatca	对sinI测序
AK55-E	ctcgagttagaccaactctccaaa	对sinI测序
Sce2	attaccctgttatcccta	pST76测序引物
T7	taatacgactcactataggg	pST76测序引物

[0048] 使用以A、C、BF和BR标记的引物通过重组PCR建立同源区域,用于自杀质粒的基因组整合。以D或E标记的引物与侧接基因组区域是同源的,并用于通过PCR和测序来检查缺失/等位基因替换。在lexA引物中的大写字母指示在该基因中引入的点突变。

[0049] 实施例2

[0050] 基因组简化的大肠杆菌中的自发突变率

[0051] 然后使用D-环丝氨酸抗性测定来确定每种菌株的自发突变率,所述测定检测cycA基因中的所有类型的突变,如在Feher等人,Mutat.Res.595(1-2):184-190(2006)中所述。

简而言之,在波动测定中,给20试管的1ml补充了0.2%葡萄糖的MS培养基(描述在Hall, Mol.Biol.Evol.,15(1):1-5(1998))各自接种大约 10^4 个细胞,并将培养物培养至稳定期早期。然后将取自每个试管的50 μ l等分试样涂布在含有D-环丝氨酸(0.04mM)的MS平板上。使用Ma-Sandri-Sarkar最大可能性方法(Sarkar等人,Genetica,85(2):173-179(1992)),从菌落数计算每个试管的估测突变数(m)。使用得自Stewart等人,Genetics,124(1):175-185(1990)的方程式41,将得到的对50 μ l有效的m值外推至1ml。仅当总细胞数目的差异是可忽略的时($<3\%$, $P \leq 0.6$,使用双尾未配对t检验),做出m值的统计对比。通过将得自3个随机试管的稀释物涂布在非选择性的平板上,计算试管中的细胞总数。将每个试管的突变数除以试管中的细胞的平均总数,得到突变率(突变/细胞/代)。

[0052] 每个基因的缺失本身会导致通过该方法测得的突变率的至少20%下降(所有值是显著的, $P < 0.05$,双尾未配对t检验)。结果图示在图1中。将不同的缺失进行组合,会进一步降低突变率,最低突变率为MDS42polBdinB和三重缺失菌株MDS42polBdinBumuDC的突变率。组合polB和dinB缺失的效应是倍增的,从而指示这些聚合酶的独立作用模式。当其它2种易错聚合酶中的任一种缺失时,umuDC的缺失没有造成突变率的额外降低,这可能指示所述基因或它们的产物之间的相互作用。与亲本MDS42菌株相比,菌株MDS42polBdinB和MDS42polBdinBumuDC表现出自发突变率的接近50%下降(8.2×10^{-8} 个突变/细胞/代分别下降至 4.34×10^{-8} 和 4.45×10^{-8})。

[0053] 为了验证编码DNA聚合酶II、IV和V的基因对适应性没有不良作用,在MOPS基本培养基中测量了不同菌株的生长速率。挑选10个平行培养物(源自每种菌株的10个单独菌落),并在Bioscreen C仪器中培养。通过跟踪每种培养物在540nm的光密度(O.D.),测量生长曲线。所述缺失都对在MOPS基本培养基中的适应性没有显著影响,即使当在三重缺失菌株MDS42polBdinBumuDC中组合时。

[0054] 为了确定经由调节剂突变体对整个SOS应答的上游灭活是否对自发突变率具有与编码DNA聚合酶II、IV和V的基因的消除相同的效果,制备了MDS42recA和MDS42lexA。MDS42recA包含recA的无痕缺失(MG1655的坐标2820783-2821861),其产物是LexA的自体溶解的诱导所必需的。MDS42lexA包含无功能的等位基因对lexA基因的替换,在所述等位基因中,在位置119处的丝氨酸被丙氨酸替换(S119A)。为了去抑制SOS调节子基因,需要这些基因(recA和lexA)中的每一个基因。因此,MDS42recA或MDS42lexA都不能诱导SOS途径。这些修饰都对菌株的总体适应性没有不良作用,正如如上所述通过菌株在MOPS基本培养基中的生长速率所测得的。令人惊奇地,任何菌株都没有表现出相对于MDS42的显著自发突变率下降。在MDS42lexA的情况下,实际上观察到轻微增加(2.07×10^{-7} 相对于MDS42的 8.2×10^{-8})。结果图示在图2中(空白条(未应激))。

[0055] 实施例3

[0056] 应激诱导的基因组简化的大肠杆菌中的突变率

[0057] 然后测量并对比了MDS42recA、MDS42lexA和MDS42polBdinBumuDC在应激条件下的突变率。

[0058] 丝裂霉素-C(一种造成双链DNA中的损伤的DNA交联剂)会直接地活化SOS应答,从而导致DNA聚合酶II、IV和V的增量调节。使用亚抑制浓度($0.1 \mu\text{g/ml}$)的丝裂霉素-C使细胞应激,并分析对突变率的影响。结果图示在图2中(带阴影线的条(丝裂霉素))。

[0059] 蛋白过度产生会对宿主细胞造成应激。试验了良性蛋白绿色荧光蛋白(GFP)的过度产生对突变率的影响。将编码GFP的基因克隆到质粒上,作为受T7启动子控制的可诱导的构建体。为了表达GFP,通过用IPTG-可诱导的lac操纵子/T7聚合酶盒替换yahA-yaiL基因组区域,构建菌株MDS42polBdinBumuDC(MDS42polBdinBumuDC-T7)、MDS42recA(MDS42recA-T7)和MDS42lexA(MDS42lexA-T7)的编码T7RNA聚合酶的变体。还构建了MDS42(MDS42-T7)、MG1655(MG1655-T7)和广泛使用的蛋白生产菌株BL21(DE3)的编码T7RNA聚合酶的变体,并测量和对比了GFP的过表达对每种菌株的突变率的影响。结果图示在图2中(实心黑色条(pET-GFP))。

[0060] 接着,通过用质粒pSG-ORF238(一种IPTG-可诱导的、基于pSG1144的构建体,其能够过度产生ORF238蛋白)转化菌株,试验了毒蛋白(ORF238,一种小型的富含亮氨酸的疏水蛋白)的过度产生对细菌的突变率的影响。结果图示在图2中(浅色条(pSG-ORF238))。ORF238的过度产生显著增加了MDS42的突变率。MDS42polBdinBumuDC的值在相同条件下保持稳定。

[0061] 结果证实,除了MDS42recA和MDS42polBdinBumuDC以外,不同的应激会增加包括MDS42在内的所有菌株的突变率。有毒的ORF238蛋白的过度产生具有最大效应:测得突变率的大于5倍增加。亚抑制浓度的丝裂霉素-C造成突变率的大于2-3倍增加,BL21(DE3)和MDS42recA不能在这些条件下生长。GFP的过度产生具有相对较小的效应,造成突变率的1.5-2倍增加。

[0062] 相比而言,在有任何应激源存在下,在MDS42recA或MDS42polBdinBumuDC中都没有看到突变率的显著增加。令人感兴趣的是,MDS42lexA没有遵循该行为——该菌株响应于所有应激表现出突变率的增加。可以将MDS42polBdinBumuDC表征为遗传上最稳定的菌株,从而显示出最低自发突变率和表现出对应激条件的可忽略应答。

[0063] 当过度产生有毒的ORF238蛋白时,最常用的蛋白生产菌株BL21(DE3)表现出比MDS42polBdinBumuDC高接近2个数量级的突变率。为了分析该差异,通过环丝氨酸抗性的突变体中的cycA的PCR分析,研究了BL21(DE3)、MG1655、MDS42和MDS42polBdinBumuDC的突变谱。简而言之,使用引物对cycA1-D/cycA2-E从突变细胞扩增包括整个基因的1,877-bp基因组区段。通过分析得自每个平行平板的5个菌落,得到代表性的样品,从而产生每个实验共计96个样品。将扩增的片段在琼脂糖凝胶上分离,并与从野生型模板产生的片段进行对比。相同的大小指示仅影响一个或几个核苷酸的突变,尺寸的减小或扩增的失败指示缺失,并且可检测的尺寸增加指示IS插入。结果图示在图3中。在MG1655中,74%的突变被证实为点突变,24%为IS插入,2%为缺失。相比而言,在BL21(DE3)中,77%的cycA突变为IS插入。尽管在BL21(DE3)中的点突变的比例远远更小(在MG1655中的74%相对于在BL21(DE3)中的23%),在BL21(DE3)中的实际点突变率明显更高(2.28×10^{-7} 相对于在MG1655中的 9.2×10^{-8})。在BL21(DE3)的cycA等位基因中没有发现缺失。

[0064] 为了证实使用cycA波动测定得到的数据,还使用利福平抗性测定测量了MDS42和MDS42polBdinBumuDC在每种不同的应激条件下的突变率。该测定检测必需的rpoB基因中的点突变,如在Jin和Gross, J. Mol. Biol., 202(1):45-58(1988)中所述。简而言之,给20试管的1ml LB各接种 10^4 个细胞,并使培养物生长至稳定期早期。将适当的稀释物涂布在非选择性的LB琼脂平板和含有利福平(100 μ g/ml)的LB琼脂平板上。分别在24或48小时以后进行菌

落计数。将突变频率报告为利福平抗性菌落数相对于总活菌数的比例。结果对应于在每种菌株和条件的3个独立实验中得到的平均值。在必要时,以与在cycA测定中相同的方式提供不同的应激条件。使用利福平抗性测定得到的数据与cycA波动数据相一致,如在图4中所示。MDS42polBdinBumuDC具有与MDS42相比明显更低的自发突变频率。响应于有毒的ORF238蛋白的过度产生,以及在有丝裂霉素-C存在下,MDS42的突变率变得显著升高,而MDS42polBdinBumuDC的应答不太明显。

[0065] 实施例4

[0066] MDS42polBdinBumuDC会给表达毒蛋白的质粒提供提高的稳定性

[0067] 为了证实包含无功能的polB、dinB和/或umuDC基因的基因组简化的细菌的惊人优点,设计了基于质粒的突变筛选。质粒pSin32携带sinI的可诱导拷贝,后者编码肠沙门氏菌(*Salmonella enterica*)婴儿血清变型的SinI甲基转移酶,其被克隆进pET3-His质粒的XhoI位点。SinI将DNA中在GG(A/T)CC位点处的内胞嘧啶甲基化,从而产生5-甲基胞嘧啶,由此产生McrBC内切核酸酶(其切割含有甲基胞嘧啶的DNA)的靶标。因此,携带甲基化的SinI位点的质粒(例如pSin32,在它的8个SinI位点处自甲基化)不能在mcrBC⁺宿主中自我复制。当引入mcrBC⁻宿主中时,所述质粒在诱导sinI的表达时被甲基化,但是可以得到维持。

[0068] 在MDS42的生产过程中缺失了mcrBC基因,因此所有MDS42菌株是mcrBC⁻。从BL21(DE3)缺失mcrBC基因以建立菌株BL21(DE3)mcrBC⁻。将质粒pSin32电穿孔进MDS42-T7、MDS42polBdinBumuDC-T7和BL21(DE3)mcrBC⁻中。在37℃在1ml LB中恢复温育1小时以后,将100μl转化的培养物置于100ml补充了氨苄西林(Ap)的LB中,并在37℃温育。从剩余的900μl中,根据标准方案分离质粒DNA。温育7小时以后,培养物达到O.D.₅₄₀ = ~0.2,在此时,用IPTG(1mM终浓度)诱导样品。在此时也取用于质粒制备的样品(8-小时样品),随后每2小时取额外样品直到18小时,然后转化后生长24和36小时。然后将纯化的pSin32质粒样品(9个得自每种菌株)转化进MDS42(McrBC⁻)和MG1655(McrBC⁺)。通过计数每个质粒样品的转化的MG1655和MDS42菌落,可以计算突变质粒的相对数目。为了得到突变质粒的数目的绝对值,用携带Ap抗性盒的对照(pST76-A)质粒转化每批电感受态的MDS42和MG1655指示菌株。然后使用MG1655和MDS42转化体的比率作为校正因子,以计算每个样品的突变pSin32质粒数目的绝对值。

[0069] 在用pSin32转化BL21(DE3)mcrBC⁻、MDS42-T7和MDS42polBdinBumuDC-T7以后发现,在IPTG诱导后,SinI酶的过度产生具有中等的生长抑制效应,即使在McrBC⁻菌株中也是如此(图5)。尽管该中等毒性会导致MDS42-T7的突变率的升高,该效应在MDS42polBdinBumuDC-T7中弱得多(图6),从而支持了上面讨论的发现。

[0070] 在IPTG-诱导以后,定期取质粒样品。如下检测携带sinI-废除突变的质粒样品(未甲基化质粒)的比例:将所述质粒样品转化回MG1655(mcrBC⁺)。通过将样品同时转化进MDS42中,确定每个样品的总质粒数目。在用得自每组电感受态细胞的对照质粒的转化体数目校正每个值以后,计算编码有功能的/无功能的sinI的质粒的比例。结果图示在图7中。

[0071] 令人惊奇地,96.7%的源自MDS42的起始(0小时)质粒样品不能在MG1655中建立。这指示,即使在缺乏T7聚合酶的宿主中,sinI的假转录已经导致SinI表达并从而将sinI位点甲基化。通过SinI对它们的不可切割性,证实了原始质粒样品中的SinI位点的甲基化状态。

[0072] 在IPTG-诱导SinI表达以后,关于不同菌株中的克隆稳定性的差异变得明显。转化后36小时(IPTG-诱导后28小时),51.7%的在BL21(DE3)mcrBC细胞中含有的pSin32携带阻止活性SinI产生的突变。该值在MDS42-T7中明显更低(25.8%)。在MDS42polBdinBumuDC-T7中,突变的pSin32质粒的比例甚至更低(8.2%)。通过SinI对它们的可切割性,证实了携带突变sinI基因的质粒上的SinI位点的非甲基化状态。

[0073] 突变质粒在BL21(DE3)mcrBC和MDS42-T7中的积累,是由于应激诱导的诱变和表达SinI的质粒引起的生长抑制的组合效应。所述酶的过度产生会升高突变率和降低生长。在这些缓慢生长的培养物中,SinI-灭活突变随时间出现,所述培养物然后在已经恢复它们的正常生长速率以后,在培养的剩余时间中快速地过度生长。在低突变率MDS42polBdinBumuDC-T7中,形成SinI-灭活突变,平均而言,经历更长的时间段。MDS42-T7和MDS42polBdinBumuDC-T7(都携带pSin32质粒)的50个独立菌落的生长曲线测量结果支持了该观点(图8)。使用0.7的O.D.₅₄₀值作为截止值来指示培养物已经克服了诱导的质粒的生长阻碍效应。MDS42polBdinBumuDC-T7达到该密度水平所需的平均时间明显长于MDS42-T7(分别是727.8和571.8分钟; $P < 0.005$,双尾未配对t检验)。

[0074] 为了验证突变已经在质粒中实际发生(其允许McrBC⁺细胞的生长),对8种不同质粒样品(取自存活的pSin32转化的MG1655菌落)的sinI区域测序(图9)。在8个事件中的7个中,移码突变已经发生在sinI中,从而在该基因内产生新的终止密码子。第八个事件显示出A至C颠换,从而导致蛋白的N880T突变。由移码造成的7个新终止密码子中的6个位于该基因的前125个碱基对内。

[0075] 这些结果证实了具有无功能的编码DNA聚合酶II、IV和/或V的基因的基因组简化的细菌的明显且意外的实用优点,特别是在不含IS元件的遗传背景下。当过度产生SinI时,在质粒上携带的sinI基因获得功能缺失突变,在MDS42polBdinBumuDC中的频率为在MDS42中的大约1/3,并且低于在BL21(DE3)mcrBC中的频率的1/5。值得注意的是,在BL21(DE3)mcrBC中过度产生仅16小时以后,在质粒上编码的所有sinI基因中的几乎一半已经发生废除突变。

[0076] 过表达SinI的培养物中的突变克隆的意外高的比例不可单独通过应激诱导的诱变来解释,其总突变率的绝对值过低(在 10^{-6} 个突变/基因/代的量级)不足以造成这样的显著效应。相反,该现象在很大程度上是由于携带毒基因的质粒的生长抑制作用。事件的关联如下:在毒基因表达后,降低细胞的生长速率。与此同时,应激会增加突变率。一旦在质粒群体中产生不再表达有毒功能的突变体,携带它的细胞可以恢复正常生长并变得在培养物中占优势。在具有一种或多种无功能的polB、dinB和/或umuDC基因的基因组简化的细菌中,如MDS42polBdinBumuDC所例征的,这样的突变体的出现被延迟,并且所述细胞可以在延长的时间段内生产有功能的有毒产物。诸如MDS42polBdinBumuDC等菌株胜过亲本菌株MDS42和常用生产菌株BL21(DE3)的优点是显著的,并且随着过度生产产物的应激严重程度的增加而增加。如本文中所述的具有高基因组稳定性的细菌菌株在以核酸和/或蛋白产物的保真度为首要的治疗用途中是特别有价值的。在需要长期连续培养条件的情况下,本发明的细菌也是惊人地有用的。

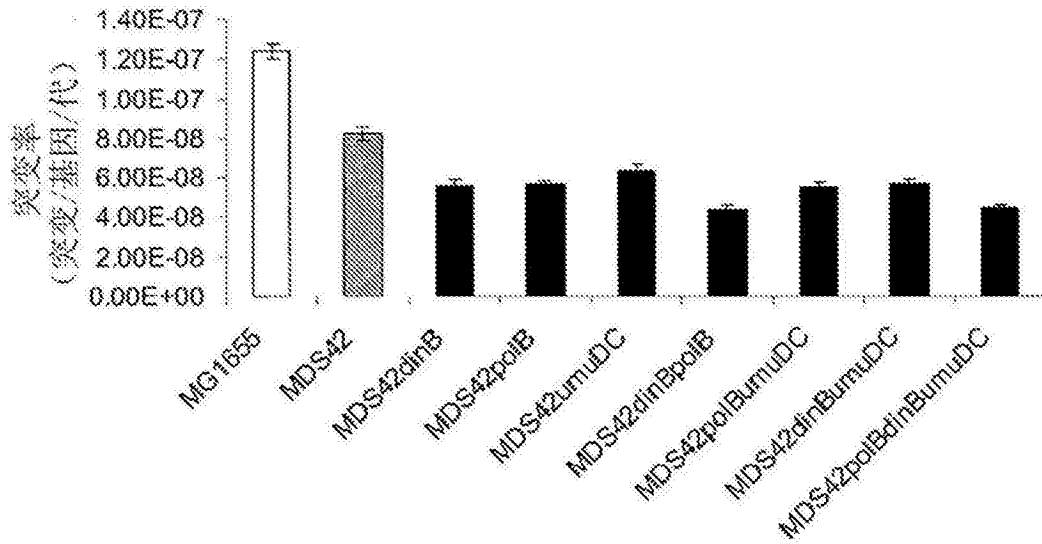


图1

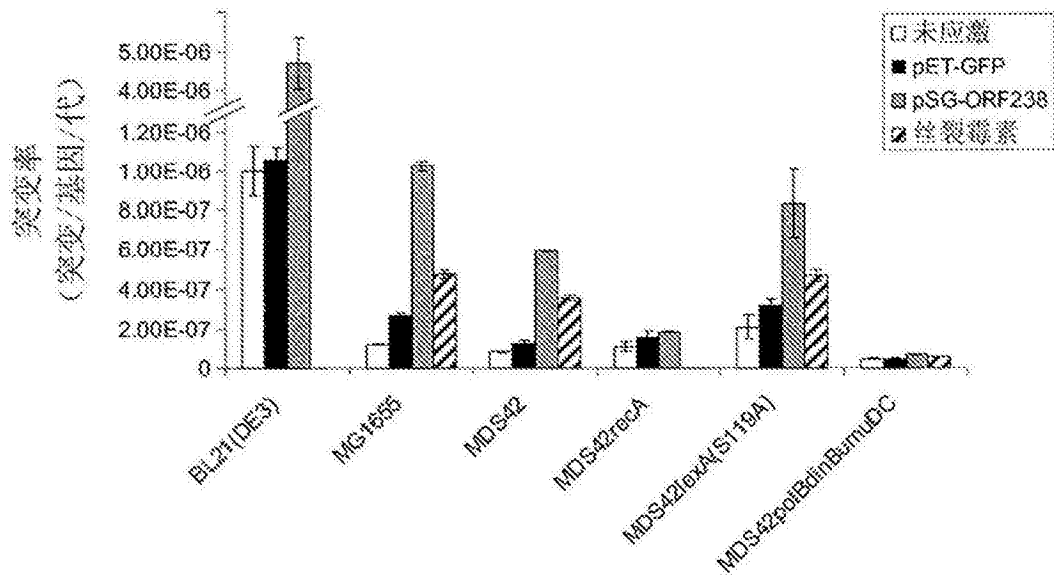


图2

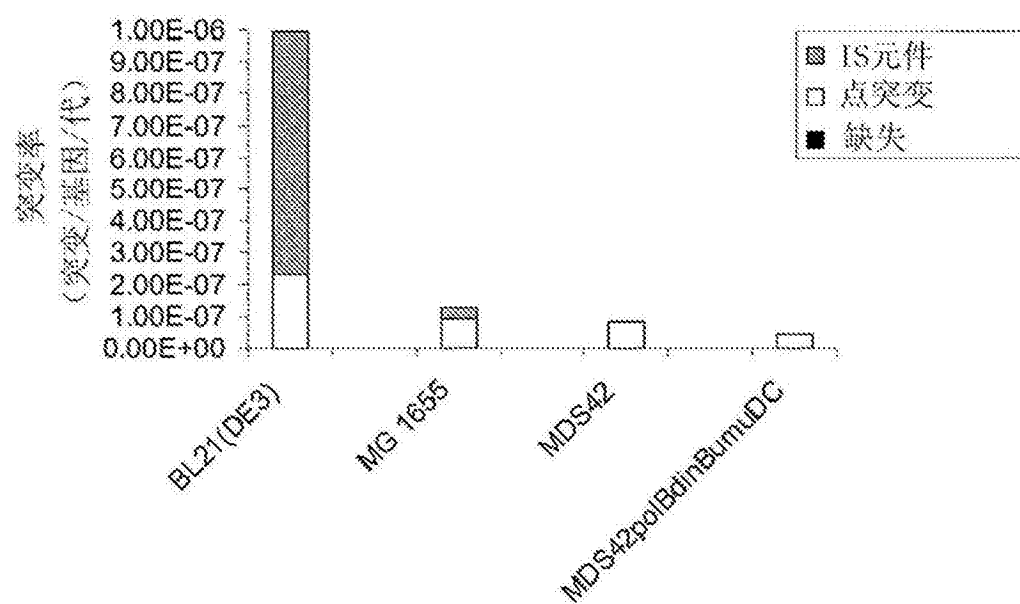


图3

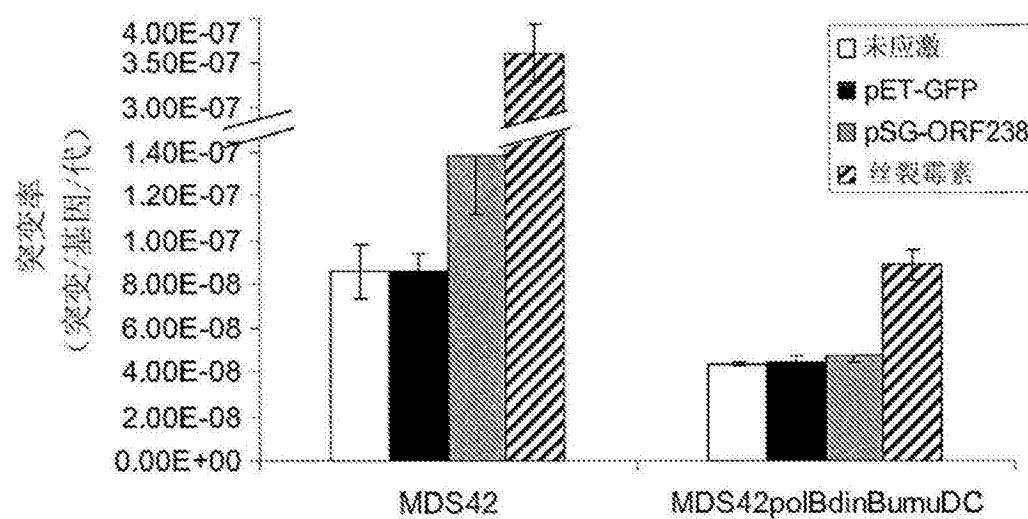


图4

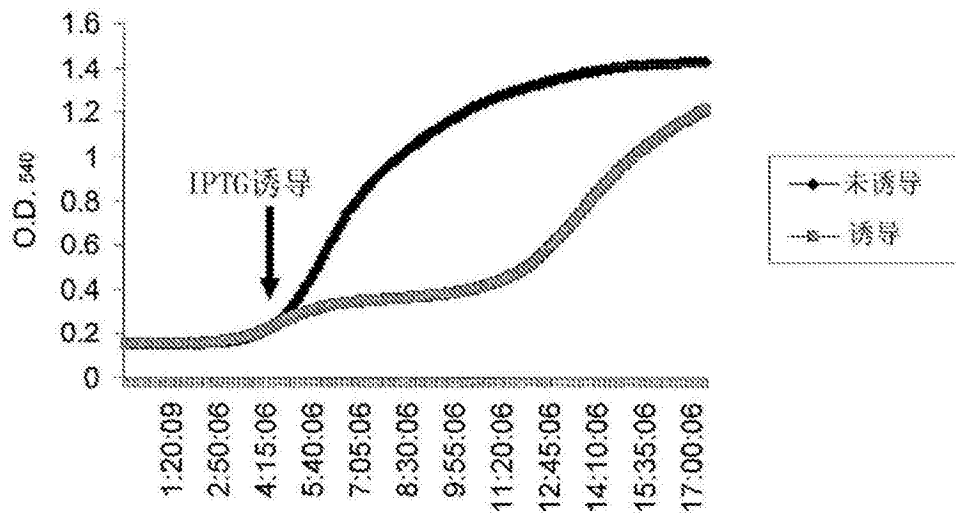


图5

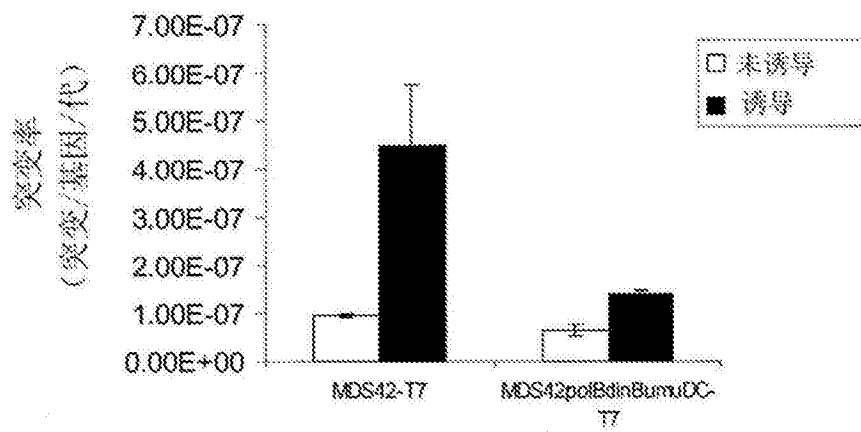


图6

