(19)国家知识产权局

(12)发明专利

(10)授权公告号 CN 109153999 B
(45)授权公告日 2022.06.24

(21)申请号 201680083261.1

(22)申请日 2016.12.30

(65)同一申请的已公布的文献号
申请公布号 CN 109153999 A

(43)申请公布日 2019.01.04

(30)优先权数据
1600417.8 2016.01.10 GB

(85)PCT国际申请进入国家阶段日
2018.09.06

(86)PCT国际申请的申请数据
PCT/EP2016/082942 2016.12.30

(87)PCT国际申请的公布数据
WO2017/118598 EN 2017.07.13

(73)专利权人 斯尼普技术有限公司
地址 英国伦敦

(72)发明人 J.克卢贝

(74)专利代理机构 中国专利代理(香港)有限公司 72001
专利代理师 梁谋 万雪松

(51)Int.Cl.
C12N 15/63(2006.01)
C12N 9/22(2006.01)
C12N 15/90(2006.01)

(56)对比文件
WO 2015095804 A1,2015.06.25
WO 2015095804 A1,2015.06.25
HISANO, Y.等.Precise in-frame integration of exogenous DNA mediated by CRISPR/Cas9 system in zebrafish..《SCIENTIFIC REPORTS》.2015,第5卷第8841篇第1-7页.
MOORE, R.等.CRISPR-based self-cleaving mechanism for controllable gene delivery in human cells..《Nucleic Acids Research》.2014,第43卷(第2期),第1297-1303页.
ZHANG, W.等.Multiple copies of a linear donor fragment released in situ from a vector improve the efficiency of zinc-finger nuclease-mediated genome editing..《Gene Therapy》.2014,第21卷第282-288页.

审查员 陈皓

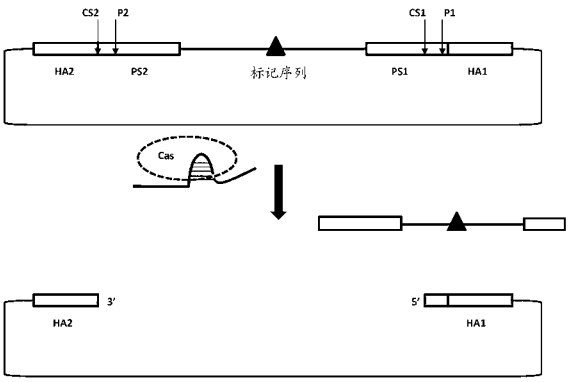
权利要求书3页 说明书61页 附图23页

(54)发明名称

原位重组产生的核酸酸链

(57)摘要

本发明涉及通过与已切割的载体的同源重组来回收或修饰靶核酸,诸如宿主细胞染色体DNA,以原位提供重组发生的核酸链。



CN 109153999 B

1.从供体核酸链回收核酸序列的体外方法,所述方法包括

(a) 提供包含第一核酸链的核酸载体,其中所述第一链包含:-

i. 用于链之间的同源重组的第一同源臂(HA1),其与所述供体链的第一核苷酸序列(NS1)同源;

ii. 用于链之间的同源重组的第二同源臂(HA2),其与所述供体链的第二核苷酸序列(NS2)同源;和

iii. 将HA1的3'末端连接至HA2的5'末端的间插核苷酸序列;

其中HA1在所述第一链中的HA2的5'侧;且NS1在所述供体链中的NS2的3'侧,其中所述供体链包含NS1和NS2之间的回收序列(RS);

(b) 将所述第一链与所述供体链组合;

(c) 核酸酶切割HA1的5'末端或侧接所述末端的5'的第一链的第一切割位点(CS1);和/或核酸酶切割HA2的3'末端或侧接所述末端的3'的第一链的第二切割位点(CS2);

其中步骤(b)和(c)以任何顺序或同时进行;

(d) 进行切割的第一链与所述供体链的同源重组,由此所述第一链的缺口修复产生其中在HA1和HA2之间回收RS的载体;和

(e) 任选地分离或测序所述载体或其RS序列。

2.权利要求1的方法,其中步骤(d)的产物是环状载体。

3.权利要求1或2的方法,其中步骤(b)至(d)在宿主细胞中进行,且任选地所述载体在步骤(a)中是环状的,且在步骤(b)中将环状载体引入细胞中。

4.任一前述权利要求的方法,其中所述载体是BAC、YAC、PAC、粘粒或质粒。

5.任一前述权利要求的方法,其中所述方法使用相应的切割位点基序来指导步骤(c)的每个所述切割,其中每个所述基序被同源重组破坏。

6.任一前述权利要求的方法,其中所述供体链是细胞或该细胞的染色体或附加体的链。

7.任一前述权利要求的方法,其中RS长度为至少10kb。

8.任一前述权利要求的方法,其中所述核酸酶切割是Cas、大范围核酸酶、锌指蛋白介导的切割或TALEN介导的切割。

9.任一前述权利要求的方法,其中所述切割通过Cas核酸酶切割,其中在步骤(a)中,所述载体包含(以5'至3'方向):

A. 包含CS1的第一前间隔区序列(PS1);

B. 与PS1的3'末端相邻的第一PAM (P1);和

C. HA1;

其中PS1的种子序列(S1)将CS1连接至P1;且

任选地,其中HA1包含S1和P1;和/或

其中在步骤(a)中,所述载体包含(以5'至3'方向):

D. HA2;

E. 包含CS2的第二前间隔区序列(PS2);

F. 与PS2的3'末端相邻的第二PAM (P2);

其中PS2的种子序列(S2)将CS2连接至P2;且

任选地,其中PS2的剩余部分在HA2的3'末端。

10.任一前述权利要求的方法,其中在步骤(c)中,相同的指导RNA (gRNA1)用于指导CS1和CS2的Cas切割。

11.任一前述权利要求的方法,其中在步骤(d)之前或期间,所述供体链在NS2的5'末端或侧接所述末端的5'处切割;和/或在NS1的3'末端或侧接所述末端的3'处切割。

12.任一前述权利要求的方法,其进一步包括将产物载体的所述RS的全部或部分在体外插入另外的核酸链中,任选地其中与所述另外的链的同源重组中,使用HA1和/或HA2将RS序列插入另外的链中。

13.权利要求12的方法,其中所述另外的链由细胞包含,且在所述同源重组之前将步骤(d)的载体产物插入细胞中。

14.权利要求12或13的方法,其中所述另外的链由胚胎干细胞或免疫细胞包含;任选地,其中所述方法包括将干细胞产物发育成组织或非人生物体。

15.权利要求12至14中任一项的方法,其用于体外的基因疗法,其中通过所述将RS插入另外的链中来除去另外的链的突变;所述方法任选地包括在所述产物中分离或测序另外的链产物或RS。

16.权利要求12至15中任一项的方法,其用于在靶人细胞中体外校正遗传突变,其中所述靶细胞包含所述另外的链,其中所述另外的链包含所述基因突变,所述方法包括

(i) 获得第二人细胞,并在第二细胞内实施权利要求1至11中任一项的方法,其中步骤(d)的产物载体中的RS包含所述遗传缺陷;

(ii) 使用基因工程改造来校正缺陷,由此产生包含校正的RS序列的产物载体;

(iii) 将校正的RS序列插入未校正的靶细胞中,其中校正的RS序列互补或替换未校正的RS序列,由此产生校正的靶细胞;和

(iv) 任选地培养校正的细胞或从校正的细胞或其后代发育组织。

17.权利要求16的方法,其中所述方法使用靶细胞中未校正的RS序列与侧接校正的RS序列的HA1和/或HA2的同源重组,由此替换未校正的RS序列;任选地,其中未校正和校正的RS序列仅通过校正的突变而不同。

18.权利要求16或17的方法,其中在步骤(iii)中,将闭合的环状校正产物载体插入靶细胞或后代中;任选地,其中校正的RS长度为至少10kb。

19.权利要求16至18中任一项的方法,其中所述细胞是相同人受试者的细胞和任选相同的细胞类型。

20.任一前述权利要求的方法,其用于产生嵌合抗原受体T细胞(CAR-T细胞)。

21.任一前述权利要求的方法,其中在步骤(b)和/或(d)中,所述载体是闭合的环状载体。

22.权利要求21的方法,其中在步骤(b)中,所述载体是超螺旋的。

23.任一前述权利要求的方法,其中从第一链缺失HA1和HA2之间的序列,和/或从供体链缺失NS1和NS2之间的序列。

24.权利要求23的方法,其中从供体链缺失RS。

25.任一前述权利要求的方法,其中所述切割通过Cas核酸酶切割,其中在HA1的5'末端或侧接所述末端的5'的前间隔区/PAM组合被步骤(d)破坏,由此使得PS1在第一链中无功

能;和/或其中在HA2的3'末端或侧接所述末端的3'的前间隔区/PAM组合被步骤(d)破坏,由此使得PS2在第一链中无功能。

26.任一前述权利要求的方法,其中将供体链序列插入第一链和/或从第一链缺失序列,以形成新的前间隔区/PAM组合或新的标记序列。

27.权利要求26的方法,其包括在步骤(d)之后Cas切割新组合的前间隔区,和任选地随后将另外的序列插入第一链中。

28.任一前述权利要求的方法,其中从供体链中缺失供体序列以形成新的前间隔区/PAM组合或新的标记序列。

29.权利要求28的方法,其包括在步骤(d)之后在供体链中Cas切割新组合的前间隔区,和任选地随后将另外的序列插入供体链或从供体链缺失。

30.任一前述权利要求的方法,所述方法包括分离或复制修饰的供体链或其中发生缺失的供体核酸序列的区域。

31.任一前述权利要求的方法,其中RS在步骤(d)中插入第一链中,并且插入的序列对于从修饰的第一链表达是有功能的,所述方法任选地包括检测或分离插入序列的表达产物。

32.任一前述权利要求的方法,其中步骤(b)至(d)在宿主细胞中进行,其中RS从供体链缺失并且(i)所述缺失对宿主细胞是致死的;或(ii)所述方法下调细胞活力、生长或增殖,或(iii)所述方法上调细胞活力、生长或增殖。

33.任一前述权利要求的方法,其中所述方法在细菌或古细菌细胞、任选地大肠杆菌细胞中进行。

34.任一前述权利要求的方法,其中所述方法在原核细胞群体中进行以杀死细胞,其中所述细胞是相同物种或菌株的。

35.任一前述权利要求的方法,其中步骤(b)至(d)在宿主细胞中进行,其中所述核酸酶是Cas,并且在步骤(c)中,Cas切割是在第一链中的前间隔区序列中,其中所述前间隔区包含种子序列,其中所述种子序列包含在细胞基因组中罕见或未发现的核苷酸序列。

36.不从属于权利要求5时的权利要求16至19中任一项所述的校正的细胞、校正的载体或校正的RS序列,其用于治疗人中的遗传缺陷,其中所述校正的细胞、校正的载体或校正的RS序列包含切割位点基序。

37.可通过权利要求1至4和6至35中任一项的方法获得的载体产物、核酸链产物、细胞产物或RS产物,其用于施用于人或动物受试者以治疗或预防受试者中的疾病或病况,其中所述载体产物、核酸链产物、细胞产物或RS产物包含切割位点基序。

38.用于产生多种载体的方法,所述多种载体包含回收序列(RS)的集合,其中所述集合包含彼此不同的第一和第二RS序列,所述方法包括进行权利要求1至35中任一项所述的方法以获得每种载体。

39.权利要求38的方法,其中载体是BAC或YAC载体,其包含从细胞样品回收的基因组片段的集合。

原位重组产生的核酸酸链

发明领域

[0001] 本发明涉及通过与已在细胞中切割的载体(例如,使用Cas核酸酶精确地切割)的同源重组来回收或修饰靶核酸,诸如宿主细胞染色体DNA,以原位提供重组产生的核酸链。

[0002] 发明背景

[0003] 转化相关的重组(TAR)克隆是一种用于从依赖于酵母酿酒酵母(*Saccharomyces cerevisiae*)的哺乳动物基因组分离大染色体区域的方法。在TAR克隆中,从细胞提取总基因组DNA,并与携带目标基因特异性的靶向序列的载体一起共转化至酵母细胞中。在共转化至酵母中后,在载体的钩和侧接目标基因的靶向基因组序列之间发生同源重组,以形成环状YAC(酵母人工染色体)。已通过TAR分离大小范围至250kb的染色体区域。如果需要序列在非YAC载体中,诸如将序列从YAC移动至BAC用于下游使用,则需要后续操作。

[0004] 参考Nat Commun. 2015年9月1日;6:8101. doi: 10.1038/ncomms9101; "Cas9-Assisted Targeting of CHromosome segments CATCH enables one-step targeted cloning of large gene clusters", Jiang W等人。作者描述了一种克隆方法,其中在两个指定基因座处通过RNA指导的Cas9核酸酶从细菌染色体体外切除靶基因组区段,并通过Gibson组装连接至克隆载体。

[0005] 工程改造的CRISPR/Cas系统已经用于精确修饰各种类型的原核和真核细胞(范围从细菌至动物和植物细胞)中的核酸(例如,参见Jiang W等人(2013))。已经提出切割靶核酸以在靶核酸中生成重组产生的游离末端。这些末端促进通过同源性介导修复(HDR)机制与伴侣核酸的同源重组,例如,用于将配偶体序列插入靶标中。然而,目前基于Cas的技术受限于需要冲洗宿主细胞基因组(或其他靶核酸)以找到匹配所需Cas的PAM位点,且然后仅选择与基因组中足够独特的前间隔区(protospacer)序列相邻的那些。理想地,前间隔区/PAM组合应该对于使减少脱靶切割的机会最小化是独特的,但这在实践中难以实现。这限制了在人类和其他精确度至关重要的设置中进行基因疗法的效用。此外,使用已经存在于基因组中的位点的需求降低了自由靶向基因组的所需区域用于同源重组的灵活性。

[0006] 发明声明

[0007] 在第一种配置中,本发明提供了:-

[0008] 从供体核酸链回收核酸序列的方法,所述方法包括

[0009] (a) 提供包含第一核酸链的核酸载体,其中所述第一链包含:-

[0010] i. 用于链之间的同源重组的第一同源臂(HA1),其与所述供体链的第一核苷酸序列(NS1)同源;

[0011] ii. 用于链之间的同源重组的第二同源臂(HA2),其与所述供体链的第二核苷酸序列(NS2)同源;和

[0012] iii. 将HA1的3'末端连接至HA2的5'末端的间插核苷酸序列;

[0013] 其中HA1在所述第一链中的HA2的5'侧;且NS1在所述供体链中的NS2的3'侧,其中所述供体链包含NS1和NS2之间的回收序列(RS);

[0014] (b) 将所述第一链与所述供体链组合;

[0015] (c) 核酸酶切割HA1的5’末端或侧接所述末端的5’的第一链的第一切割位点(CS1)；和/或核酸酶切割HA2的3’末端或侧接所述末端的3’的第一链的第二切割位点(CS2)；

[0016] 其中步骤(b)和(c)以任何顺序或同时进行；

[0017] (d) 进行切割的第一链与所述供体链的同源重组，由此所述第一链的缺口修复产生其中在HA1和HA2之间回收RS的载体；和

[0018] (e) 任选地分离或测序所述载体或其RS序列。

[0019] 在一个实例中，步骤(b)至(d)在宿主细胞中进行；和/或步骤(d)的产物是环状载体。

[0020] 在一个替代方案中，HA1在所述第一链中的HA2的5’侧；且NS1在所述供体链中的NS2的5’侧，其中所述供体链包含NS1和NS2之间的回收序列(RS)；其中步骤(c)反而包括核酸酶切割HA1的5’末端或侧接所述末端的5’的第一链的第一切割位点(CS1)；和/或核酸酶切割HA2的3’末端或侧接所述末端的3’的第一链的第二切割位点(CS2)。有利地，所述载体在步骤(c)之前是线性的。

[0021] 在第二种配置中，本发明提供了：-

[0022] 修饰核酸序列的方法，所述方法包括

[0023] (a) 提供包含第一核酸链的核酸载体，其中所述第一链包含：-

[0024] i. 用于链之间的同源重组的第一同源臂(HA1)，其与另外的核酸链的第一核苷酸序列(NS1)同源；

[0025] ii. 用于链之间的同源重组的第二同源臂(HA2)，其与所述供体链的第二核苷酸序列(NS2)同源；和

[0026] iii. 将HA1的3’末端连接至HA2的5’末端的间插序列(IS)；

[0027] 其中HA2在所述第一链中的HA1的5’侧；且NS2在所述另外的链中的NS1的5’侧，其中所述另外的链任选地包含NS1和NS2之间的序列(DS)；

[0028] (b) 将所述第一链与所述另外的链组合；

[0029] (c) 使用核酸酶(例如，Cas核酸酶)切割HA1的3’末端或侧接所述末端的3’的第一链的切割位点(CS3)；和/或Cas核酸酶切割HA2的5’末端或侧接所述末端的5’的第一链的切割位点(CS4)；

[0030] 其中步骤(b)和(c)以任何顺序或同时进行；

[0031] (d) 进行第一链与另外的链的同源重组或位点特异性重组，由此产生产物另外链，其中IS被插入NS1和NS2之间，并且任选地DS从另外的链缺失；和

[0032] (e) 在(d)之后任选地分离或测序所述载体、产物另外链或其IS。

[0033] 在第三种配置中，本发明提供了：-

[0034] 制备适于与第二核酸链同源重组的第一核酸链的方法，例如，在根据任何前述配置的方法中，所述制备方法包括：-

[0035] (a) 鉴定所述第二链的第一和第二核苷酸序列(NS1和NS2)，任选地，其中选择NS1使得其在第二链中的NS2的3’侧；

[0036] (b) 在另一核酸链中组合

[0037] i. 用于HA1和NS1之间的同源重组的第一同源臂(HA1)，其与所述第二链的第一预

定靶核苷酸序列NS1互补；

[0038] ii. 用于HA2和NS2之间的同源重组的第二同源臂(HA2)，其与所述第二链的第二预定靶核苷酸序列NS2互补；和

[0039] iii. 与PAM (P1)相邻的第一前间隔区序列(PS1)，其中P1与第一Cas核酸酶同源，所述第一Cas核酸酶用于在第一切割位点(CS1)处切割PS1，其中CS1在HA1的5’末端或侧接所述末端的5’侧；和/或 与PAM (P2)相邻的第二前间隔区序列(PS2)，其中P2与所述第一或不同的Cas核酸酶同源，所述第一或不同的Cas核酸酶用于在第二切割位点(CS2)处切割PS2，其中CS2在HA2的3’末端或侧接所述末端的3’侧；

[0040] 由此产生第一链。

[0041] 在第四种配置中，本发明提供了：-

[0042] 抑制原核宿主细胞的活力、生长或增殖的方法，所述方法包括

[0043] (a) 提供第一载体，其包含与PAM (P1)相邻的前间隔区序列(PS1)，所述PAM (P1)与第一Cas核酸酶同源；

[0044] (b) 提供毒素或毒素前体；或包含编码毒素或毒素前体的序列的载体，其中所述毒素对原核细胞有毒；

[0045] (c) 将所述细胞与所述载体组合，其中将所述载体引入所述细胞，并在所述细胞中提供所述毒素；和

[0046] (d) 使用Cas核酸酶切割第一载体的PS1，其使用第一Cas核酸酶，其中切割促进细胞中的毒素活性，由此杀死细胞或减少细胞生长或增殖。

[0047] 在第五种配置中，本发明提供了：-

[0048] 转化宿主细胞的方法，所述方法包括

[0049] (a) 提供第一核酸载体，所述载体处于第一状态，其包含

[0050] i. 与PAM (P1)相邻的前间隔区序列(PS1)，所述PAM (P1)与第一Cas核酸酶同源；和

[0051] ii. 第一核苷酸序列，其包含用于标记载体的第一状态的第一标记序列；

[0052] (b) 将细胞与载体组合，其中将载体引入细胞中；和

[0053] (c) 使用Cas核酸酶切割第一载体的PS1，其使用第一Cas核酸酶，其中切割将载体转化为第二状态，其中使得标记序列无功能；和

[0054] (d) 任选地通过检测标记物在细胞中无功能来检测第二状态。

[0055] 在第六种配置中，本发明提供了：-

[0056] 用于转化真核细胞的载体，其中所述载体包含第一核酸链，且所述细胞包含第二核酸链，所述第一链包含

[0057] (a) 侧接第一同源臂(HA1)的插入序列(IS)，其中HA1与第二链的预定核苷酸序列(NS1)互补，其中HA1能够在真核细胞中与NS1同源重组以修饰一条或两条链；

[0058] (b) 与PAM (P1)相邻的第一前间隔区序列(PS1)，其中P1与第一Cas核酸酶同源，所述第一Cas核酸酶用于在第一切割位点(CS1)处切割PS1；

[0059] (c) 其中CS1在HA1的末端或侧接HA1，其中CS1能够被第一Cas切割以产生切割的第一链，其中HA1能够与NS1同源重组。

[0060] 在第七种配置中，本发明提供了：-

[0061] 产生前体载体的方法,其中所述载体是第三种配置中所述载体的前体,所述前体载体包含

[0062] (a) 以5'至3'顺序,与PAM (P1)相邻的第一前间隔区序列(PS1),其中P1与第一Cas核酸酶同源,所述第一Cas核酸酶用于在第一切割位点(CS1)处切割PS1;和

[0063] (b) P1的紧邻3'的第一核苷酸序列,其中所述第一序列包含与P1的3'末端相邻或侧接所述末端的3'的第一克隆位点,用于将第二核酸序列克隆至所述前体载体中,其中所述第二序列包含HA1和HA2中的一种或两种;

[0064] 所述方法包括

[0065] (c) 获得前体核酸;

[0066] (d) 鉴定核酸中的克隆位点或在核酸中产生克隆位点核酸;和

[0067] (e) 鉴定核酸中的PS1/P1组合或在核酸中产生组合;

[0068] (f) 其中(d)和(e)以任何顺序或同时进行,且其中(d)和(e)中的至少一者包括所述产生步骤,由此产生并任选地分离所述前体载体。

[0069] 在第八种配置中,本发明提供了:-

[0070] 修饰真核细胞中的核酸的方法,所述方法包括

[0071] (a) 提供真核细胞,其含有

[0072] i. 闭合环状核酸载体,其中所述载体包含第一核酸链;

[0073] ii. 第二核酸链,其中所述第二链由细胞的染色体包含;和

[0074] iii. 编码Cas核酸酶的核苷酸序列和编码核定位信号(NLS)的核苷酸序列;

[0075] (b) 使用Cas核酸酶切割来切割细胞内载体的第一链以产生重组产生的载体核酸;和

[0076] (c) 在切割的第一链和第二链之间的细胞内使用同源重组,和在第一和第二链之间交换至少10kb的连续核苷酸的序列,由此形成至少一条修饰的链;和

[0077] (d) 任选地分离或测序修饰的链或细胞。

[0078] 本发明还提供了根据任何配置的载体、噬菌体、细胞、组合物、集合、群体、药物、抗生素、用途和方法。本发明提供了可通过本发明的方法获得的修饰的链、第一链、供体链、另外的链、载体、组合物、药物或试剂盒。

[0079] 在一个实例中,本文中的一种配置上的任何特征与本发明的不同配置组合,用于在本文的一个或多个权利要求中可能包括这样的组合。

[0080] 附图简述

[0081] 图1是显示在真核细胞中组合闭合的环状载体与原位Cas切割用于交换大核苷酸序列的示意图。

[0082] 图2A和B显示用于本发明中的单臂载体。

[0083] 图3A显示一种变型,其中环状载体包含两个Cas切割位点(例如,用作根据本发明的回收载体或插入载体)。

[0084] 图3B显示一种变型,其中已经在同源臂中鉴定PAM位点,将PAM位点工程改造至所述臂中,或已经从靶核酸或细胞的序列选择所述臂以含有适当放置的PAM。

[0085] 图4A显示标记片段的释放,以在回收载体中产生两个重组产生的末端。如图4B中所示,与供体链的HR回收含有所需RS的供体片段。如图4C中所示,闭合的环状载体可用于基

因疗法中或使用HR在体外或在任何原核或真核细胞中修饰靶核酸，其中含有RS的大或非常大的序列可凭借闭合的环状载体携带来管控。如图4D中所示，这允许RSCE的RMCE插入以及还有将改变特异性引入靶标中。

[0086] 图5A显示重组产生的核酸片段的释放，所述重组产生的核酸片段可用于随后的与靶标的HR过程(图5B)，用于插入插入序列(例如，调节元件或外显子序列)。在图5C中，载体同源臂与在靶链中间隔(由包括缺失的序列(矩形框)的序列间隔)的序列互补。没有插入IS。

[0087] 图6A和6B显示插入载体的应用，其中Cas切割在靠近第一同源臂的一个位点处。在本实施例中，发生靶序列的替代(图6B)。

[0088] 图7A和7B显示如何进行插入(即，不是替代，其涉及靶序列缺失)。

[0089] 图8A说明本发明的载体与Cas载体的共转化。图8B显示其中在本发明的载体中包括Cas9序列的变型。

[0090] 图9A显示本发明的前体载体。将克隆位点工程改造至载体中，随后用于落入一个或多个同源臂(HA1和/或HA2)中。这在图9B和9C中使用具有不相容的SSR(lox位点)的RMCE来说明。图9D和9E显示使用限制性内切核酸酶的变型。

[0091] 发明详述

[0092] 本发明以其各种配置解决了本领域的一个或多个问题，并提供了一种解决方案，其聚焦于载体核酸，而不是宿主细胞靶核酸，以实现特定的预定用于切割(例如，Cas切割)的位点的设计。以这种方式，本发明提供了一种灵活的系统，其中使用者自由选择和工程改造用于切割的特定位点，由此避免仅选择靶核酸中的有限位点的需求。可以使用标准重组工程改造或其他技术简单地在体外进行工程改造以制备其中切割(例如，Cas切割)位点已经仔细设计和定位的载体。这允许使用者将切割(例如，Cas切割)自由定位靠近同源臂和标记物，其可以用于随后的与靶核酸的同源重组方法中。

[0093] 尽管参考Cas切割的使用描述了本发明的各方面，但在替代方案中，可以使用其他切割方式，例如限制性内切核酸酶切割，TALEN介导的切割，锌指介导的切割和大范围核酸酶切割。在一个实例中，所述切割是酶促的，诸如核酸酶切割、指导的核酸酶切割或RNA指导的核酸酶切割。

[0094] 因此，本发明涉及通过与已使用Cas核酸酶或其他切割方式精确切割的载体的同源重组来修饰靶核酸，诸如宿主细胞染色体DNA，以提供重组产生的核酸链。在一个方面，本发明涉及精确切割以便将重组产生的末端定位靠近载体核酸中的同源臂，以促进同源重组。以这种方式，可以促进精度或效率。在一个实例中，该方面在回收或插入载体的背景下特别有用，并且避免在靶标中找到适当的位置用于切割以促进同源重组的需求。

[0095] 本发明的另一个优点是能够构建具有非常大的插入序列(例如，100kb或更多)的载体。在初始步骤中，可以将空载体例如作为闭合的环状载体引入宿主细胞中。用Cas或通过其他方式在细胞内进行随后的载体切割，原位产生重组产生的末端，用于将大插入序列同源重组插入宿主核酸中。这避免了切割外部细胞以产生难以处理和纯化且较少被靶细胞吸收的大序列的现有技术问题(参见例如，Jiang等人(2015)同上)。Jiang等人(2015)鉴定，由于在细胞之外操作期间长DNA片段的机械剪切以及与较长DNA的Gibson组装的连接效率较低，因此其技术效率降低。此外，将细胞嵌入琼脂糖凝胶中的要求严重限制了Gibson组装

的足够DNA的可用性。

[0096] 有利地，通过聚焦于进入的载体核酸的工程改造，本发明提供了用于引导切割的基序(诸如前间隔区)的精确序列的灵活性。在这方面，使用者能够选择具有在宿主细胞类型中未发现或在此类宿主细胞中罕见或在细胞基因组中的RS中未发现的序列基序的序列，诸如前间隔区序列。在一个实例中，所述载体前间隔区可以基于在原核生物中发现的序列(诸如细菌Pribnow盒的序列；细菌前导序列(例如，Pel B)；细菌-10序列或细菌-35元件序列(例如，TATAAT或TTGACA)；细菌上游启动子元件(UP元件)序列；细菌调节或操纵子序列；或细菌复制起点序列)。以这种方式，Cas切割由真核细胞(诸如人细胞)中没有或极少发现的序列引导，并且这大大有助于避免靶细胞内的脱靶Cas切割。在一个替代方案中，本发明允许使用完全合成的前间隔区序列，其可以针对宿主细胞基因组序列进行预先检查以确定它们的独特性。对于实施方案，这些益处可用于细胞和生物体的基因疗法中。本发明还适用于靶核酸的体外修饰，例如，使用体外重组工程改造。

[0097] 因此，在某些配置中，本发明的优点基于聚焦于在载体中的切割(例如，Cas切割)，预先设计和简单地设计切割位置的能力，预先检查和控制切割基序(例如，前间隔区)独特性的能力，选择切割位点相对于同源臂、插入序列和标记物的定位的能力，以及通过将切割限制在靶细胞内以产生重组产生的物质来处理和使用相容载体形式的非常大的序列的能力。有利地，避免了对PCR (其可以引入序列错误，例如，在回收的序列中)的依赖。此外，回收使用本发明的某些配置回收的序列以产生载体，所述载体即用于许多下游应用中进一步使用，例如在随后的重组工程改造中产生修饰的序列，在细胞的基因工程改造中，用于细胞或生物体(例如，人、小鼠、植物、酵母或细菌)的基因疗法等。

[0098] 此外，序列回收不限于特定细胞类型，诸如酵母细胞，如现有技术诸如转化相关的重组(TAR)所必需的。TAR依赖于细胞外的方法，用于提取和制备细胞DNA，用于在转化至酵母细胞中之前将片段引入线性TAR载体中。在细胞外生成基因组片段(通过剪切基因组DNA)也是Jiang等人2015(下文)的CATCH方法的特征。所述方法必须在保护性凝胶培养基中的细胞外的DNA剪切步骤进行，随后将片段体外Gibson组装至载体中。与这些现有方法不同，本发明在一些配置中有利地在细胞的保护环境内进行，用于生成重组产生的末端和序列回收至可以即用于进一步使用的载体中，以及额外剪切，控制这一点的需求不是这些发明配置的特征。在本发明的方面，可以通过在细胞中原位产生重组产生的核酸末端来有利地构建载体，所述细胞使得能够在细胞内捕获大基因组片段(例如，大约100-250kb或最高达1Mb，诸如YAC)直接进入可以容纳这种插入物大小的新生线性环状载体(诸如BAC和YAC)。有利地，在实施方案中，这使得能够直接产生具有大插入物的环状(例如，闭合环状，例如，超螺旋状)载体，所述插入物对于下游程序(诸如，引入靶细胞用于将一些或所有回收的序列插入靶细胞的基因组或附加体)是非常有用的。为此目的，有用的是本发明的载体中的回收序列可以侧接有同源臂，所述同源臂可以容易地用于将回收的序列同源重组插入靶细胞基因组中。

[0099] 同源重组用通过切割DNA产生的重组产生(游离)的末端开始。切割的末端增强效率，如技术人员所知。本发明有用地将切割(例如，Cas切割)聚焦至可以预先设计(即，工程改造)的第一链中的位点，而不是已经存在于宿主细胞基因组中的位点。在一个实例中，第一链可以在载体(诸如BAC或质粒或噬菌体)中。这允许预先工程改造第一链，使得在与第二

或供体链(例如,细胞的染色体或附加体的链)进行同源重组之前,可以精确地选择和预先确定前间隔区/PAM组合或其他切割基序的位置和性质。在一个实例中,用于后续同源重组的切割位置是已知的,并且可以相对于第一链或载体中的其他所需元件(例如,启动子、选择标记物、lox位点等)的位置进行控制(例如,图2A、3A、3B、5A、6A、8A和9A)。因此,这提供了相对于使用已经存在于宿主细胞基因组中的位点的优势,因为后者决定了切割基序或前间隔区的性质和位置以及相关(同源)PAM的性质。在该情况下,核酸酶的选择由可用切割基序的性质决定,例如,Cas由可接受的PAM的性质决定(反之亦然),而在本发明中,基序/核酸酶(例如,PAM/Cas组合)的选择可以自由地制备,因为在进行该方法之前,第一链的工程改造允许这样。此外,通过选择或设计预定的前间隔区PS1(和任选的PS2),本发明允许将同源重组靶向至宿主细胞基因组的任何所需区域(而不是由预先存在的宿主细胞位点决定)。在本发明中,通过设计和包括一个或多个同源臂(HA1、HA2)用于与载体中(而不是靶核酸中)的预定切割基序(例如,前间隔区/PAM组合)一起使用,存在靶核酸中同源序列的更自由选择,这意味着可以控制和更自由地选择靶中同源重组的位置。在现有方法中,仔细选择独特的前间隔区、适当的PAM和匹配的Cas的需要严格地限制并决定切割靶链和与靶链的同源重组的精确位置的选择。当想要在宿主细胞基因组中选择待修饰的精确位置(例如,用于生物体的基因疗法或用于治疗的人细胞的修饰) - 精确位置可能没有适当的前间隔区/PAM组合或其他切割基序用于使用时,这不是期望的。相反,本发明通过将这些选择转移至第一链的设计来避免这些问题。因此,通过第一链的有目的设计,本发明可以在选择宿主细胞链中待修饰的位置方面提供更大的灵活性,并且可以为选择待使用的Cas或其他核酸酶提供更大的灵活性。另外,可以实现和控制与同源臂序列相邻的切割的、重组产生的末端的紧密选址,并且可以预先确定载体标记物或插入序列相对于切割末端的定位,以允许在细胞内或体外的原位同源重组方法期间从载体可检测缺失和插入载体(参见例如图5A、5B、6B和8A)。本发明还提供了产生新的标记物的可能性,所述新的标记物指示原位同源重组的正确切割(参见例如图5A)。在一个实例中,这可以对重组工程改造方法有用。

[0100] 有用地,本发明还提供了从供体或靶链至第一链的序列回收的可能性,例如在载体中(参见例如图4B)。这可用于回收序列,用于在载体的背景中进一步修饰,例如,在重组工程改造和/或产生载体用于后续用于修饰另外的靶链或细胞的方法中。在一个实例中,这可用于随后的将回收的序列同源重组插入靶链的方法中(例如,图4C)或使用工程改造至产物第一链(诸如侧接回收的插入物)中的cre-lox位点用于随后将回收的序列插入靶链或细胞基因组中的方法中(例如,图4D)。在一个实例中,在后续的细胞修饰方法(例如,用于人、动物或植物基因疗法)中使用第一链中回收的序列。在一个替代方案中,本发明提供了使用第一链的切割(例如,Cas切割)来释放重组产生的核酸用于随后的与靶链同源重组或位点特异性重组(例如,使用lox位点,例如,使用不相容lox位点的RMCE)用于将序列插入靶链(参见例如图5A和5B)或从靶链缺失序列(参见例如图5C)的可能性。如图5A中所示,重组产生的链的释放可以产生新的第一链标记序列(例如,可通过PCR或表达可检测标记物来检测)。

[0101] 因此,在一个实施方案中,本发明提供了在同源重组或位点特异性重组(例如,使用cre-lox)的方法中使用插入第一链的序列用于修饰靶细胞(例如,原核、真核、植物、动物或人细胞)的靶核酸链或基因组的另外的步骤。在一个实例中,一些或所有插入序列用于替

代或插入靶标中的序列(伴随或不伴随靶序列的缺失)或缺失靶标中的序列。所述方法任选地包括分离或测序由靶细胞包含的产物靶链或修饰的核酸。在一个实例中,所述插入或替代改变靶核酸链中的1、2、3、4、5、6、7、8、9或10个核苷酸或密码子。在一个实例中,所述方法使靶链序列(例如,由靶链包含的基因)活化,失活,或增加活性或降低其活性(例如,其中所述活性是由该基因编码的蛋白的表达。在一个实例中,这校正靶标中的遗传缺陷(例如,用于人细胞基因疗法)。

[0102] 在一个实例中,本发明还包括在重组工程改造的方法中使用产物第一链以产生包含在步骤(d)中插入第一链的全部或部分供体序列的另外的链。任选地,所述方法包括在靶细胞基因组的同源或位点特异性重组修饰的方法中使用另外的链产物,和任选地分离产物靶细胞和/或将靶细胞施用于人或动物受试者用于医学用途和/或培养此类靶细胞的群体并分离该群体或其样品。在一个实例中,从供体链(例如,来自人受试者的人细胞染色体DNA)回收序列,回收的序列在产物第一链中。然后通过基因工程改造(例如,重组工程改造或突变,例如,以校正回收的序列中的不期望的突变)来修饰该回收的序列。这产生另外的链。然后将该链的序列(其中所述序列包含修饰(例如,突变校正))插入靶细胞的链中(例如,插入人细胞(例如,不同人的细胞或来自人受试者,例如,其中所述细胞与第一细胞具有相同或不同的类型)的染色体的DNA链中)。在一个实例中,回收的序列从人受试者的第一细胞分离,所述方法在体外使用以校正回收序列中的遗传缺陷,并且将校正插入衍生自相同人受试者的iPS细胞中以产生包含遗传校正的iPS细胞(任选地,其中遗传缺陷已在iPS细胞的基因组中缺失)。任选地,培养iPS细胞以产生iPS细胞群体或分化细胞的群体用于人类疗法(例如,用于人类疗法的组织)。在一个实例中,所述疗法是所述人受试者的疗法。以这种方式,可以进行基因疗法,其中所述方法能够离体精确校正遗传缺陷(例如,单核苷酸多态性,SNP),而无需使用来自不同人的人DNA序列(例如,使用人BAC来源,其中所述序列可以与人受试者基因组相比具有未知或不需要的变化)。因此,本发明用于治疗人受试者而不引入另一人的DNA突变或序列。这对于人受试者的人基因疗法的精确和高度保守的遗传变化是有用的。本发明出于任何和所有这些目的提供本文的任何产物链、载体、群体或细胞。

[0103] 在一个实例中,所述校正是校正人受试者中的基因(例如,基因的外显子和/或调节序列)中的一个、两个或三个或更多个SNP而用于受试者的基因疗法。在一个替代方案中,所述受试者是非人动物(例如,啮齿动物、小鼠或大鼠)。在一个替代方案中,所述受试者是植物或酵母。在一个替代方案中,所述受试者是昆虫。在一个替代方案中,所述受试者是细菌或原生动物。

[0104] 本发明通过在同源重组步骤中修饰靶链来实现精确的小的和大的缺失,使所述同源重组步骤从靶链中缺失序列。在一个实例中,这对于失活靶链中的基因或调节序列是有用的(例如,用于修饰包含供体链的宿主细胞,诸如当供体链是染色体或附加体(例如,由质粒包含)时)。在一个实例中,可以从这种宿主细胞的基因组中缺失缺陷序列。在另一个实例中,可以通过修饰来杀死细胞或抑制其生长或增殖,例如,其中所述细胞是原核细胞,诸如人病原体细胞。一个实例是细菌细胞,例如大肠杆菌、金黄色葡萄球菌、铜绿假单胞菌、化脓脓链球菌、嗜热链球菌或本文公开的另一种细菌细胞。在一个实例中,所述细胞是植物、酵母、哺乳动物、人、啮齿动物、原核、真核、细菌或神经细胞,其中所述方法用于用来自第一链的序列替代宿主细胞序列。任选地,从第一链回收供体序列(将供体序列插入第一链)。所述

第一链可以由载体(诸如BAC、PAC、YAC或质粒)包含。

[0105] 在一个实例中,第一链由选自细菌人工染色体(BAC)、酵母人工染色体(BAC)、PAC和质粒的载体包含,并将回收的序列插入载体中。所述产物载体任选地包含在回收序列的5'末端或侧接所述末端的5'侧的第一Cas核酸酶切割位点,和/或在回收序列的3'末端或侧接所述末端的3'侧的Cas核酸酶切割位点。任选地,存在两个切割位点并且Cas位点与相同的Cas核酸酶(例如,Cas9)同源。本发明提供了通过本发明的方法获得或可获得的任何此类载体。本发明提供了本发明的用于产生所述载体的文库(例如,BAC文库)的方法,其中所述载体通过本发明的方法获得或可获得。在一个实例中,所述回收的序列或各回收的序列是人序列,例如来自相同或不同(例如2、3或4)个人基因组。在一个实例中,不同的人基因组是如1000个基因组项目所定义的相同种族群体的人的基因组(例如,在一个实例中,不同的人基因组是如由1000个基因组项目定义的不同种族群体的人的基因组)。在一个实例中,所述种族群体选自1-14,例如,数字3、4或9:

[0106] 1=ASW;2=CEU;3=CHB;4=CHS;5=CLM;6=FIN;7=GBR;8=IBS;9=JPT;10=LWK;11=MXL;12=PUR;13=TSI;14=YRI。

[0107] 在一个实例中,本发明提供了这种载体(例如,BAC)文库。包括Cas切割位点促进根据本发明产生具有一个或两个游离末端的重组产生的核酸,用于随后与如本文所述的另外的链的同源重组。在一个实例中,当所述文库包含人回收的序列时,通过将回收的人序列插入细胞的基因组中(任选地同时替代细胞基因组的序列),重组产生的核酸可用于修饰人或非人动物细胞。在一个实例中,人细胞是iPS细胞。在一个实例中,所述细胞是啮齿动物、小鼠或大鼠细胞。本发明提供了通过该方法获得或可获得的这种细胞,例如用于人或动物治疗。本发明提供了由通过该方法获得或可获得的这种细胞发育的组织,例如用于人或动物治疗,其中所述细胞是多能性细胞、多能细胞或全能细胞(例如,人或动物iPS细胞),其中所述组织不是人组织,或不是由人包含。

[0108] 在本发明的任何方法的实例中,将最多达300、200、100、75、50、25、15、10、5或1kb的链DNA(例如,作为连续序列)插入第一链和/或另外的链中。所述插入可以在多个步骤(例如,根据本发明的方法的每个步骤)中。在一个实例中,所述方法包括将DNA多次插入第一和/或另外的链中,其中将最多达3Mb、2Mb、1Mb或900、800、700、600、500、400、300、200、100、50、25、15、10、5或1kb的回收的DNA(例如,作为连续序列)插入该链或各链中。本发明提供这种产物第一或另外的链或包含其的细胞。另外或可替代地,插入不超过3Mb、2Mb、1Mb或900、800、700、600、500、400、300kb。

[0109] 在本发明的任何方法的实例中,将最多达300、200、100、75、50、25、15、10、5或1kb的第一链DNA(例如,作为连续序列)插入供体链中(有或没有序列回收至第一链中)。在一个实例中,另外或可替代地,将最多达300、200、100、50、75、25、15、10、5或1kb的供体链DNA(例如,作为连续序列)插入第一链中。所述插入可以在多个步骤(例如,根据本发明的方法的每个步骤)中。在一个实例中,所述方法包括将第一链DNA多次插入供体链中,其中将最多达3Mb、2Mb、1Mb或900、800、700、600、500、400、300、200、100、50、25、15、10、5或1kb的第一链DNA(例如,作为连续序列)插入供体链中。本发明提供这种产物供体链或包含其的宿主细胞。另外或可替代地,插入不超过3Mb、2Mb、1Mb或900、800、700、600、500、400、300kb。

[0110] 在本发明的任何方法的实例中,缺失最多达3Mb、2Mb、1Mb或900、800、700、600、

500、400、300、200、100、50、25、15、10、5或1kb的第一链DNA(例如,作为连续序列)。在一个实例中,另外或可替代地,缺失最多达3Mb、2Mb、1Mb或900、800、700、600、500、400、300、200、100、50、25、15、10、5或1kb的供体或另外的链DNA(例如,作为连续序列)。本发明提供这种产物供体链或包含其的宿主细胞。所述缺失可以在多个步骤(例如,根据本发明的方法的每个步骤)中。在一个实例中,从第一链缺失和插入第一链是同时的。在一个实例中,从供体链缺失和插入供体链是同时的。在一个实例中,从另外的链缺失和插入另外的链是同时的。另外或可替代地,缺失不超过3Mb、2Mb、1Mb或900、800、700、600、500、400、300kb。

[0111] 本发明提供了包括一种或多种用于原位产生指导RNA的序列的可能性。所述序列可以编码单一指导RNA(即,嵌合gRNA)或与tracrRNA组合以产生gRNA的cRNA。在一个实例中,gRNA或tracrRNA由第一链(例如,包含链的第一载体)包含或由不同核酸(例如,第二载体)包含的一个或多个序列编码(参见例如图8A和8B)。或者,tracrRNA由其中引入第一链用于在细胞内Cas切割的靶细胞编码。在这种情况下,所述靶细胞包含内源CRISPR/Cas系统(例如,是已经被工程改造以包含这种系统的真核或原核细胞或者是天然包含这种系统的细菌或古细胞)并且tracrRNA是内源性tracrRNA。

[0112] 可以将Cas核酸酶(例如,Cas9或Cas3)引入靶细胞(或在体外与第一核酸混合)或可以由第一链(或第一载体)或另一种核酸(例如,第二载体)包含的序列编码,所述第一链(或第一载体)或另一种核酸(例如,第二载体)被引入靶细胞或体外与第一链和指导RNA或编码这些的序列组合(参见例如图8A和8B)。在一个替代方案中,所述靶细胞包含内源CRISPR/Cas系统,并且Cas是内源Cas核酸酶(例如Cas9或Cas3)。在一个替代方案中,代替Cas切割,本发明使用cpf1切割和同源位点,并且本文的公开内容经必要的变更适用于cpf1,而不是Cas。

[0113] 本发明的另一个优点是构建可用于适于产生预定载体的通用前体载体的可能性。所述通用载体可以被设计为包括一个或多个克隆位点用于插入同源臂,所述同源臂被选择以靶向靶细胞基因组或靶核酸的期望和预定区域。提供具有一个或多个切割位点(例如,Cas核酸酶切割位点)的通用载体,每个切割位点侧接或紧邻相应位点(“克隆位点”)用于插入同源臂。以这种方式,可以相对于同源臂控制重组产生的末端的定位,用于随后的同源重组。可以控制载体中相对于SSR位点(例如,lox位点)的定位,随后用于将序列cre-lox插入如本文所述的另外的或第二链中。如图9A-E中所示,在一个实例中,每个克隆位点可以是限制性内切核酸酶位点(例如,EcoR1、BamH1或其他位点)或SSR位点(例如,loxP、lox211、lox2272、frt或其他位点)。如所示,本发明的方法包括提供这种通用前体载体并将同源臂组合至载体中与Cas核酸酶位点相邻处。在一个实施方案中,所述臂由已被限制性内切核酸酶切割的核酸包含,其中所述内切核酸酶也已用于切割所述前体载体的克隆位点,其中将所述切割位点重组以将所述臂插入所述前体载体中,由此产生包含第一链的载体,用于本发明的方法中。在一个替代方案中,前体载体的克隆位点中和包含该臂的核酸中的相容性SSR位点用于位点特异性重组,其将所述臂插入载体中与Cas核酸酶位点相邻或侧接Cas核酸酶位点,由此产生包含第一链的载体,用于本发明的方法中。

[0114] 在本发明的通用前体载体概念的另外的应用中,提供了产生第一载体集合的方法,其中每种载体包含用于本发明的方法中的第一链。在该产生方法中,将多种前体载体与多种核酸组合,其各自包含相应的同源臂,其中所述方法包括将同源臂克隆至载体中,其中

各产物载体包含至少一个与切割(例如,Cas核酸酶)位点相邻或侧接切割(例如,Cas核酸酶)位点的所述同源臂,由此产生载体集合,其各自包含用于本发明的方法中的第一链。在一个实例中,所述核酸各自用相同的限制性内切核酸酶切割,其中前体载体的克隆位点用相同的内切核酸酶切割,其中通过重组前体载体中的切割位点与切割核酸来将同源臂插入前体载体中。在一个替代方案中,SSR重组用于将所述臂插入前体载体中的相容性SSR位点。在一个实例中,所述核酸通过切割第一细胞的基因组DNA(例如,全基因组DNA,或细胞的一个或多个染色体或附加体的DNA)获得,其中所述切割用所述限制性内切核酸酶,由此产生切割核酸群体,其与多种前体载体组合,所述前体载体具有用相同内切核酸酶切割的克隆位点,其中将多种切割的核酸序列插入多种前体载体中以产生产物载体的集合,其各自包含用于本发明的方法中的所述第一链。这对于回收细胞的许多不同基因组序列或靶核酸的不同序列是有用的,其中回收的序列可以充当产物载体中的同源臂,用于本发明中,例如,用于从细胞(例如,与第一细胞相同类型的细胞或来自不同物种或来自不同物种的不同成员的细胞,例如,第一细胞是第一人受试者的细胞,并且回收的序列来自不同人受试者或相同人受试者)回收另外的序列。额外地或可替代地,所述产物载体包含与回收序列相邻或侧接回收序列的插入序列,其中回收的序列用作同源臂,用于将插入序列插入另外的核酸链或靶细胞基因组或附加体中。在一个实例中,这使得能够在一个细胞或细胞群体中的多个位点插入,以确定插入序列在群体中的一个或多个细胞中的各种影响。这可以用于例如找到在靶细胞基因组中对于插入序列(例如,外源基因序列)的表达所需的一个或多个位点。在一个实例中,所述靶细胞是原核(例如,细菌)或酵母细胞,并且所述方法可用于生成用于工业或实验室制造插入序列的产物的细胞群体。在另一个实例中,所述靶细胞是真核细胞(例如,人细胞,例如,CHO或HEK293或Cos细胞),且所述方法可用于生成用于工业或实验室制造插入序列的产物的细胞群体。本发明包括产生这种原核或真核细胞用于表达插入序列的产物的方法,其中所述方法包括表达插入序列的产物和任选地分离产物。本发明包括用于本文公开的任何工业中、例如用于医疗用途或用于医药、化妆品、食品、环境、饮料、水或石化产品或处理的产品。

[0115] 根据本发明引入细胞的每种核酸(例如,第一链或载体)可以包含一种或多种核定位信号(NLS),如常规已知,以定位至细胞核。

[0116] 在一个实施方案中,将一个或多个核苷酸序列插入细胞的基因组中(例如,插入染色体或附加体中)。然后插入的序列可以充当诱饵位点,即与本发明的方法中各自的载体(第一链)同源臂同源且杂交的序列。例如,N1和/或N2可以是已插入细胞基因组中的序列,并且这些分别与HA1和HA2同源。诱饵位点的插入可以使用任何可用的方法(诸如使用侧接每个诱饵序列的一个或多个位点特异性重组位点(例如,lox或frt位点)的SSR插入)实现。可以随机插入一个或多个诱饵位点(其然后能够从供体链回收随机(非预定)RS序列或将IS或RS插入受体细胞基因组中的随机位点中,例如,以评价插入的序列在多个受体细胞中的不同位置处的效果。在一个实施方案中,每个诱饵核苷酸序列由转座子的长末端重复序列(LTR)侧接(例如,由piggyBac或Sleeping Beauty LTR侧接),并且使用表达的相应转座酶将所得核酸引入细胞中,用于将诱饵序列的拷贝随机且任选地多次插入细胞基因组中。在一个实例中,

[0117] 将以5'至3'顺序包含相同转座子类型的转座子LTR、NS1、任选的间隔区序列、NS2

和LTR的核苷酸序列引入细胞中，并且将NS1/NS2组合插入细胞基因组中，其即用于随后与本发明的链或载体的同源臂进行HR。以这种方式，可以使用NS1/NS2诱饵序列将RS或由适当同源臂侧接的任何其他序列插入细胞基因组中。在一个实例中，仅将单一类型的诱饵序列(例如NS1或NS2)插入细胞基因组中，并且使用包含同源HA的载体(第一链)(例如，单臂链或载体)将其用于本发明的RS回收方法中。使用预定序列的诱饵序列是有利的，因为可以选择或设计在原始细胞基因组中未发现的序列，这降低了在非预期的基因组位点处与本发明的链或载体的HR事件的机会。例如，所述细胞是真核细胞(例如，人、啮齿动物、小鼠或大鼠细胞)，并且所述诱饵序列是细菌或病毒序列。

[0118] 因此，本发明提供了以下实施方案：-

[0119] 1．从供体核酸链回收核酸序列的方法，所述方法包括

[0120] (a) 提供包含第一核酸链的核酸载体，其中所述第一链包含：-

[0121] i．用于链之间的同源重组的第一同源臂(HA1)，其与所述供体链的第一核苷酸序列(NS1)同源；

[0122] ii．用于链之间的同源重组的第二同源臂(HA2)，其与所述供体链的第二核苷酸序列(NS2)同源；和

[0123] iii．将HA1的3’末端连接至HA2的5’末端的间插核苷酸序列；

[0124] 其中HA1在所述第一链中的HA2的5’侧；且NS1在所述供体链中的NS2的3’侧，其中所述供体链包含NS1和NS2之间的回收序列(RS)；

[0125] (b) 将所述第一链与所述供体链组合；

[0126] (c) 切割(例如，核酸酶或Cas核酸酶切割)HA1的5’末端或侧接所述末端的5’的第一链的第一切割位点(CS1)；和/或Cas核酸酶切割HA2的3’末端或侧接所述末端的3’的第一链的第二切割位点(CS2)；

[0127] 其中步骤(b)和(c)以任何顺序或同时进行；

[0128] (d) 进行切割的第一链与所述供体链的同源重组，由此所述第一链的缺口修复产生其中在HA1和HA2之间回收RS的载体；和

[0129] (e) 任选地分离或测序所述载体或其RS序列。

[0130] 在一个替代方案中，步骤(d)使用SSR (而不是HR)。

[0131] 在一个实例中，使用Cas(例如Cas9)进行切割。

[0132] 在一个替代方案中，实施方案1提供了：-

[0133] 1a．修饰核酸序列的方法，所述方法包括

[0134] (a) 提供包含第一核酸链的核酸载体，其中所述第一链包含：-

[0135] i．用于链之间的同源重组的第一同源臂(HA1)，其与另外的核酸链的第一核苷酸序列(NS1)同源；

[0136] ii．用于链之间的同源重组的第二同源臂(HA2)，其与所述供体链的第二核苷酸序列(NS2)同源；和

[0137] iii．将HA1的3’末端连接至HA2的5’末端的间插序列(IS)；

[0138] 其中HA2在所述第一链中的HA1的5’侧；且NS2在所述另外的链中的NS1的5’侧，其中所述另外的链任选地包含NS1和NS2之间的序列(DS)；

[0139] (b) 将所述第一链与所述另外的链组合；

[0140] (c) 切割(例如,使用Cas核酸酶切割)HA1的3'末端或侧接所述末端的3'的第一链的切割位点(CS3);和/或Cas核酸酶切割HA2的5'末端或侧接所述末端的5'的第一链的切割位点(CS4);

[0141] 其中步骤(b)和(c)以任何顺序或同时进行;

[0142] (d) 进行第一链与另外的链的同源重组或位点特异性重组,由此产生产物另外链,其中IS被插入NS1和NS2之间,并且任选地DS从另外的链缺失;和

[0143] (e) 在(d)之后任选地分离或测序所述载体、产物另外链或其IS。

[0144] 任选地,在方法1a中,第一链不在CS1和CS2两者处切割。

[0145] 对本文实施方案1的参考可替代地被读作对实施方案1a的参考。

[0146] 在一个实例中,本文使用的Cas是Cas9,任选地化脓脓链球菌或金黄色葡萄球菌Cas9。在一个实例中,HA1与NS1的序列相同(或至少90%或95%相同)和/或HA2与NS2的序列相同(或至少90%或95%相同)。

[0147] 在本文中,在本发明的任何配置中,本文的第一链、载体或任何其他链或载体可以任选地包含一个或多个核定位信号序列NLS(例如,当该方法在真核细胞中进行时)。例如,所述Cas核苷酸序列、gRNA编码序列或IS序列可以包含一个或多个NLS。

[0148] 当CS1在HA1的5'末端或侧接所述末端的5'侧时,其意味着CS1紧邻HA1的5'末端(即,切割(例如,Cas切割)紧邻HA1的第一个(最5')核苷酸);或切割(例如,Cas切割)在HA1的第一个核苷酸的5'侧(例如,与末端分开末端的5'的不超过1kb或500、250、200、150、50、25、20、10或5个核苷酸)。当CS2在HA2的3'末端或侧接所述末端的3'侧时,其意味着CS2紧邻HA2的3'末端(即,切割(例如,Cas切割)紧邻HA2的最后一个(最3')核苷酸);或切割(例如,Cas切割)在HA2的最后一个核苷酸的3'侧(例如,与末端分开末端的3'的不超过1kb或500、250、200、150、50、25、20、10或5个核苷酸)。

[0149] 当CS3在HA1的3'末端或侧接所述末端的3'侧时,其意味着CS3紧邻HA1的3'末端(即,切割(例如,Cas切割)紧邻HA1的第一个(最3')核苷酸);或切割(例如,Cas切割)在HA1的最后一个核苷酸的3'侧(例如,与末端分开末端的3'的不超过1kb或500、250、200、150、50、25、20、10或5个核苷酸)。当CS4在HA2的5'末端或侧接所述末端的5'侧时,其意味着CS4紧邻HA2的5'末端(即,切割(例如,Cas切割)紧邻HA2的第一个(最5')核苷酸);或切割(例如,Cas切割)在HA2的第一个核苷酸的5'侧(例如,与末端分开末端的5'的不超过1kb或500、250、200、150、50、25、20、10或5个核苷酸)。

[0150] 本文中"在末端"和"侧接末端"的概念应被类似地解释。

[0151] 所述方法任选地包括使用切割位点基序在所述同源重组(HR)之前引导供体或另外的链的切割(例如,Cas、大范围核酸酶、锌指蛋白介导的或TALEN介导的切割),其中HR用切割的第一和供体/另外的链进行。任选地,所述基序通过同源重组破坏(例如,破坏用于切割供体链的供体或另外的链中的前间隔区和/或PAM)。

[0152] 所述方法在所述RS的同源重组(HR)回收之前任选地包括使用相应的切割位点基序来引导第一链或回收载体的每次切割(例如,Cas、大范围核酸酶-、锌指蛋白介导的或TALEN介导的切割或限制性内切核酸酶切割),其中所述基序通过同源重组破坏(例如,破坏用于切割第一链或载体的供体或另外的链中的前间隔区和/或PAM)。

[0153] 在一个实例中,第一链产物中插入的核酸序列在一侧或两侧侧接(例如,在实施方

案1的HR步骤之后添加)位点特异性重组(SSR)位点(例如,loxP或不相容性lox位点,例如,loxP和lox511或loxP和lox2272 – 这使得重组酶介导的盒交换(RMCE)能够与不相容性lox位点一起使用,使得插入序列可以随后通过与靶标中的匹配SSR位点进行RMCE而插入另外的核酸链(例如,靶细胞的染色体)中(参见例如图4D)。因此,在一个实例中,本发明还包括在与靶核酸链(例如,由靶细胞(例如人、细菌、植物、酵母或啮齿动物细胞,例如,在体外)的染色体或附加体包含)的位点特异性重组的方法中使用该第一链产物,其中将位点特异性位点之间的插入序列通过与靶核酸中的相容性SSR位点的重组而插入靶核酸链中。

[0154] 在一个实例中,包含供体或另外的链的细胞和/或靶细胞是人细胞(例如,在体外)。在一个实例中,所述细胞是淋巴细胞、B细胞或T细胞(例如,人细胞)。在一个实施方案中,本发明的方法用于修饰人T细胞,例如通过将DNA插入T细胞的基因组中和/或从T细胞的基因组缺失,例如通过插入一个或多个DNA序列用于在产物T细胞中产生抗原受体。本发明提供了通过该方法获得或可获得的这种T细胞产物。任选地,所述方法包括分离T细胞或其后代,并任选地将细胞或后代(或其群体)施用于人受试者用于治疗或预防疾病或病况(例如,癌症或自身免疫疾病或病况)。在一个实例中,本发明提供了通过本发明的方法获得或可获得的细胞(例如,T细胞或B细胞)产物,其用于治疗或预防疾病或病况(例如,癌症或自身免疫疾病或病况)。在一个实例中,本发明提供了这种细胞或其后代或其群体,此时其由用于治疗或预防疾病或病况(例如,癌症或自身免疫疾病或病况)的药物包含。

[0155] 在一个实例中,包含供体或另外的链的细胞和/或靶细胞是细菌物种或菌株的细胞(例如,其为人或动物病原体或由食品、饮料、石化物质或在环境(例如,水路)或工业水库中包含)。在一个实例中,所述细胞是人或动物微生物组物种或菌株(例如,人或动物肠道、口腔、阴道、皮肤或腋窝微生物组物种或菌株)的细菌细胞。在一个实例中,本发明提供了用于治疗或预防由所述细菌菌株或物种介导的人或动物中的疾病或病况的方法的产物细胞(或其后代)。在一个实例中,本发明提供了用作为人或动物病原体的物种或菌株的细菌或人或动物微生物组物种或菌株(例如,人或动物肠道、口腔、阴道、皮肤或腋窝微生物组物种或菌株)的细菌来治疗或预防人或动物的感染的方法。在一个实例中,本发明提供了用于杀死或抑制食品、饮料、石化物质或环境(例如水路)或工业水库中的细菌的生长或增殖的方法。

[0156] 2. 实施方案1的方法,其中步骤(b)至(d)在宿主细胞中进行;任选地,其中所述供体链是细胞的染色体或附加体的链。

[0157] 在一个实例中,所述细胞是真核细胞(例如,人、非人动物、酵母或植物细胞)。

[0158] 在一个实例中,所述方法在人(例如,其中所述方法是化妆方法)、胚胎(例如,非人胚胎)、受精卵(例如,非人受精卵)、生殖细胞(例如,人或非人生殖细胞)或非人动物中进行。

[0159] 在一个实例中,所述方法在体外进行。在一个实例中,所述方法在体内(例如,但不在人或人胚胎中)进行。

[0160] 为了用于本发明中,可以将载体引入细胞中,例如通过注射或电穿孔,或通过细胞融合。

[0161] 3. 实施方案2的方法,其中所述载体在步骤(a)中是环状的,且在步骤(b)中将环状载体引入细胞中。

[0162] 4．实施方案1、2或3的方法，其中第一、供体或另外(例如，第一链)链由BAC包含。

[0163] 例如，所述供体链由BAC包含，且其中RS是所述BAC的人或啮齿动物(例如小鼠)序列。

[0164] 5．任一前述实施方案的方法，其中所述载体在步骤(a)中是环状的，和/或其中步骤(d)的产物是环状载体。

[0165] 该方面有利于在载体中容纳大(> 10kb)或非常大(> 100kb)的IS或RS。这避免了处理和纯化用于HR或SSR的大线性核酸片段的领域中已知的问题，因为本发明的该实施方案原位产生用于HR的重组产生的线性底物；然而，较早或较晚的处理步骤(诸如将链组合或将第一链引入靶细胞中)是用环状底物进行，这有利于使用较大的回收片段或插入较大的片段。此外，环状载体可用于本发明，因为它们容易被细胞摄取，用于在细胞中实施本发明的方法。在一个实例中，所述载体是闭合环状载体(例如，超螺旋闭合环状载体)。发明人意识到这非常有效地被细胞摄取。超螺旋DNA转化大肠杆菌比松弛的开环DNA好至少一百倍。

[0166] 6．任一前述实施方案的方法，其中所述载体是BAC、YAC、PAC、质粒或粘粒；任选地，其中所述供体链由BAC、YAC、PAC、粘粒或质粒包含。

[0167] 关于YAC作为环状YAC/BAC的回收率和潜在效用的细节，参考Nucleic Acids Res.2000 Sep 1; 28(17):e81; PMCID:PMC110718; Massimo Cocchia等人。在一个实例中，本发明的载体是环状YAC，其包含最多达250、500或670kb的IS或RS。

[0168] 在一个实例中，所述载体是环状BAC，其包含最多达250或300kb的IS或RS。

[0169] 7．任一前述实施方案的方法，其中RS或IS长度为至少1kb。

[0170] 8．实施方案7的方法，其中RS或IS长度为至少10kb。

[0171] 9．实施方案7的方法，其中RS或IS长度为至少100kb。

[0172] 例如，RS长度为至少1、10、100、200、300、400、500、600、700、800或900 kb或1Mb，且步骤(d)的产物是包含RS的闭合环状载体。例如，IS长度为至少1、10、100、200、300、400、500、600、700、800或900 kb或1Mb，且步骤(a)的载体是包含IS的闭合环状载体。

[0173] 10．任一前述实施方案的方法，其中CS1或CS3距HA1不超过250nt(即250个连续核苷酸)。

[0174] 术语“nt”是指连续核苷酸。

[0175] 11．任一前述实施方案的方法，其中CS2或CS4距HA2不超过250nt。

[0176] 12．任一前述实施方案的方法，其中HA1和HA2各自长度为至少15nt。

[0177] 在一个实例中，HA1长度为至少50、100、150、200、300、400、500、600、700、800或900nt，或长度为至少1、2、3、4、5、6、7、8、9、10、20、30、50或100kb。在一个实例中，HA2长度为至少50、100、150、200、300、400、500、600、700、800或900nt，或长度为至少1、2、3、4、5、6、7、8、9、10、20、30、50或100kb。

[0178] 13．任一前述实施方案的方法，其中HA1和HA2包含至少0.2kb的总长度。

[0179] 在一个实例中，HA1和HA2包含长度为至少0.2、0.3、0.4、0.5、0.6、0.7、0.8、0.9、1、2、3、4、5、6、7、8、9、10、20、30、50或100kb的总长度或由其组成。

[0180] 14．任一前述实施方案的方法，其中HA1和HA2分别与NS1和NS2同基因(即，100％相同的核苷酸序列)。

[0181] 在一个实施方案中，HA1和HA2各自与NS1和NS2的核苷酸序列分别至少80、85、90或

95％相同。

[0182] 15．任一前述实施方案的方法，其中Cas切割，将RS回收至第一链中，或将IS插入另外的链中，将来自第一链或载体的标记序列(例如，用于表达第一可检测标记物的序列)缺失或失活，任选地，其中所述方法包括检测标记序列的缺失(例如，由标记序列编码的可检测标记物的不存在)。

[0183] 标记物的实例(例如，当在细胞中进行该方法时)是抗生素抗性标记物，例如neo，潮霉素、卡那霉素、氯霉素、嘌呤霉素(例如，puro-deltaTK)抗性等，如本领域技术人员已知。在一个实例中，所述标记物是通过本发明的方法产生的序列，其中标记序列在该方法的产物中，但不存在于该方法的起始链中。因此，标记序列可通过PCR (例如连接PCR)检测。

[0184] 在一个实例中，所述标记序列是HPRT序列，且该方法在细胞内进行。通过对6-硫鸟嘌呤(6-TG)的抗性来选择HPRT缺陷的细胞。因此，在该实例中，可以选择通过RS插入第一链或从第一链缺失IS(和插入另外的链)来缺失或失活HPRT序列，因为所得细胞对6-TG具有抗性。另一个实例是用于表达胸苷激酶的tk序列，其在下面进一步解释。

[0185] 16．任一前述实施方案的方法，其中第一链或载体包含未通过该方法缺失的标记序列(例如，用于表达第二可检测标记物)；并且任选地，该方法包括在步骤(d)之后检测标记序列(或其表达产物)的存在，而检测或不检测通过实施方案15的方法缺失的标记序列(或表达产物)(例如，HPRT序列)的不存在。

[0186] NeoR、HygR和PuroR广泛用作哺乳动物和其他细胞中的标记物。携带抗性序列的细胞可以通过施用G418(NeoR序列赋予的抗性)、潮霉素(HygR序列赋予的抗性)或嘌呤霉素(PuroR序列赋予的抗性)来确定。因此，在一个实例中，该实施方案编号16中的标记物选自NeoR、HygR和PuroR。

[0187] 因此，用于本发明中的适当选择剂是：-

[0188] G418：与庆大霉素类似的氨基糖苷类抗生素。除了原生动物和蠕虫以外，还对细菌、酵母、高等植物和哺乳动物细胞有毒性。

[0189] 潮霉素：抑制细菌、真菌和高等真核生物中的蛋白合成的氨基糖苷类抗生素。

[0190] 嘌呤霉素：抑制原核生物和真核生物中的蛋白合成的广谱抗生素。

[0191] 在一个实例中，该方法包括使用阳性-阴性选择标记物检测来确定RS已插入第一链或IS已从第一链中缺失。在该实例中，该方法在细胞中进行，并且已经历该方法的细胞在阴性选择剂中存活并在阳性剂(例如，抗生素抗性，诸如上述上面提及的抗生素抗性)存在的情况下生长。更昔洛韦将杀死任何含有胸苷激酶(tk)序列的细胞。因此，本发明的方法可以在细胞中进行，并且第一链包含用于表达胸苷激酶的tk作为根据上述实施方案15的标记物。在本发明的方法中经历成功同源重组或SSR的细胞在抗生素(以选择重组)和更昔洛韦(以杀死没有成功地经历根据该方法的重组的任何细胞(即，保留tk序列))中生长。

[0192] 17．任一前述实施方案的方法，其中在步骤(a)中，所述载体包含(以5’至3’方向)：

[0193] A．包含CS1的第一前间隔区序列(PS1)；

[0194] B．与PS1的3’末端相邻的第一PAM (P1)；和

[0195] C．HA1；

[0196] 其中PS1的种子序列(S1)将CS1连接至P1；且

[0197] 任选地，其中HA1包含S1和P1。

[0198] 18．任一前述实施方案的方法，其中在步骤(a)中，所述载体包含(以5’至3’方向)：

[0199] D．HA2；

[0200] E．包含CS2的第二前间隔区序列(PS2)；

[0201] F．与PS2的3’末端相邻的第二PAM (P2)；

[0202] 其中PS2的种子序列(S2)将CS2连接至P2；且

[0203] 任选地，其中PS2的剩余部分在HA2的3’末端。

[0204] 在本文发明的任何配置中，每个前间隔区(例如，PS1或PS2)由15至45个连续核苷酸(例如，18、19、20、21、22、23、24或25个连续核苷酸)组成。例如，PS1由18-21个核苷酸(例如20个核苷酸)组成。例如，PS2由18-21个核苷酸(例如20个核苷酸)组成。在本文发明的任何配置中，每个种子序列(例如，S1或S2)由2至15个连续核苷酸(例如，12、11、10、9、8、7、6、5、4、3或2个连续核苷酸)组成。例如，S1由5-12或8-12个核苷酸(例如12个核苷酸)组成。例如，S2由5-12或8-12个核苷酸(例如12个核苷酸)组成。

[0205] 在本文发明的任何配置中，每个PAM(例如，P1或P2)是2、3、4或5个连续核苷酸。PAM的选择与待用于切割的Cas匹配，正如技术人员所知。在一个实例中，PAM是本文公开的任何PAM。

[0206] 在本文发明的任何配置中，每个Cas切割位点例如是同源PAM的紧邻5’的2-5个，例如2、3、4或5个核苷酸。例如，CS1是P1的5’侧的3或4个核苷酸。例如，CS2是P2的5’侧的3或4个核苷酸。

[0207] 19．任一前述实施方案的方法，其中相同的核酸酶(例如，Cas核酸酶(例如，Cas9))用于切割CS1和CS2。

[0208] 20．任一前述实施方案的方法，其中相同的指导RNA (gRNA1)用于指导CS1和CS2的Cas切割。

[0209] 在一个实例中，P1和P2由相同的PAM序列组成，CS1是P1的5’侧的核苷酸，CS2是P2的5’侧的核苷酸，前者的核苷酸数目与后者相同。额外地或可替代地，PS1由与PS2相同的序列组成。额外地或可替代地，S1由与S2相同的序列组成。在一个实例中，PS1至(且包括)P1的序列与PS2至(且包括)P2的序列相同，且任选地，P1和P2与相同Cas9同源。

[0210] 本领域技术人员将容易地理解术语“与...同源(cognate to)”或“与...同源(cognate with)”或类似术语。例如，PAM NGG与化酿脓链球菌(*S pyogenes*) Cas9一起发挥功能，因此该PAM与化酿脓链球菌Cas9同源。

[0211] 21．从属于实施方案18的实施方案20的方法，其中所述PS2的剩余部分由NS2包含(例如，在NS2的3’末端)或在供体链中与NS2相邻(例如，与NS2的3’末端相邻或侧接所述末端的3’)。

[0212] 22．任一前述实施方案的方法，其中NS1和NS2不在步骤(c)中使用的Cas核酸酶识别的前间隔区/PAM位点处或与其相邻。

[0213] 因此，在该方法中，供体或另外的链不被Cas核酸酶切割。在一个替代实施方案中，供体或另外的链被Cas核酸酶切割。这对HR是有用的。

[0214] 23．实施方案1至21中任一项的方法，其中在步骤(d)之前或期间，所述供体或另

外的链在NS2的5’末端或侧接所述末端的5’处切割;和/或在NS1的3’末端或侧接所述末端的3’处切割。

[0215] 24．任一前述实施方案的方法,其进一步包括将产物载体的所述RS的全部或部分插入另外的核酸链中;任选地,其中所述产物载体是闭合环状载体并且所述载体在RS之外切割(例如,在HA1和HA2之外)(例如,使用Cas)切割,和/或在RS插入另外的链之前切割另外的链(例如,使用Cas)。

[0216] 例如,如实施方案23中所定义的那样切割另外的链。

[0217] 25．实施方案24的方法,其中与所述另外的链的同源重组中,使用HA1和/或HA2将RS序列插入另外的链中;任选地,其中所述另外的链包含NS1和NS2或各自与NS1和NS2至少80、90或95％相同(或100％相同)的序列。

[0218] 26．实施方案25的方法,其中所述另外的链由细胞包含,并且在所述同源重组之前将步骤(d)的载体产物(例如,闭合环状载体或超螺旋载体)插入细胞中。

[0219] 27．实施方案24、25或26的方法,其中所述另外的链由干细胞、例如诱导的专能干细胞(iPS细胞)(例如,体外的人iPS细胞);免疫细胞(例如,T细胞或B细胞);胚胎干细胞(例如,小鼠或大鼠ES细胞)包含;任选地,其中所述方法包括将干细胞产物发育成组织或非人生物体(例如,小鼠或大鼠)。

[0220] 例如,所述干细胞是专能干细胞、多能干细胞或全能干细胞。在一个实例中,所述干细胞是非人和体外的。

[0221] 28．用于基因疗法的实施方案24至27中任一项的方法,其中通过所述将RS同源重组插入另外的链中来除去另外的链的突变;所述方法任选地包括在所述产物中分离或测序另外的链产物或RS。

[0222] 在一个实例中,所述基因疗法是人、动物、植物或酵母细胞的疗法。在另一个实例中,所述疗法是人、动物或植物的疗法。

[0223] 29．用于校正靶人细胞或其后代中的体外遗传突变的实施方案24至28中任一项的方法(例如,当从属于实施方案1时,其使用步骤(d)中的HR),其中所述细胞或后代包含所述另外的链,其中所述另外的链包含所述基因突变,所述方法包括

[0224] (i) 获得第二人细胞(例如,靶细胞的后代或祖先,例如来自相同培养物的培养细胞),并在第二细胞内实施实施方案1至24中任一项的方法,其中步骤(d)的产物载体中的RS包含所述遗传缺陷;

[0225] (ii) 使用基因工程改造(例如重组工程改造)来校正缺陷,由此产生包含校正的RS序列的产物载体;

[0226] (iii) 将校正的RS序列插入(例如,通过将校正的产物载体,注入闭合的环状或超螺旋载体插入)未校正的靶细胞或其未校正的后代(例如,iPS细胞)中,其中校正的RS序列互补或替换未校正的RS序列,由此产生校正的靶细胞或其后代;和

[0227] (iv) 任选地培养校正的细胞或从校正的细胞发育组织。

[0228] 30．实施方案29的方法,其中所述方法使用靶细胞或后代中未校正的RS序列与侧接校正的RS序列的HA1和/或HA2的同源重组,由此替换未校正的RS序列;任选地,其中未校正和校正的RS序列仅通过校正的突变而不同。

[0229] 31．实施方案29或30的方法,其中在步骤(iii)中,将闭合的环状校正产物载体插

入靶细胞或后代中;任选地,其中校正的RS长度为至少10或100kb。

[0230] 任选地,校正的RS最高达300kb和/或载体是BAC。

[0231] 32. 实施方案29至31中任一项的方法,其中所述细胞是相同人受试者的细胞和任选相同的细胞类型(例如,来自相同人受试者的成体干细胞或脐带血干细胞)。

[0232] 33. 如实施方案29至32中任一项所述的校正的细胞、校正的载体或校正的RS序列,其用于治疗人中的遗传缺陷。

[0233] 在一个实例中,所述方法提供了医学治疗人受试者的方法,所述方法包括向受试者施用校正的细胞或后代(例如,iPS细胞)或由其发育的组织,由此获得受试者中的医学治疗。

[0234] 34. 实施方案24至32中任一项的方法,其中所述另外的链由免疫细胞,例如T细胞或B细胞(例如,人细胞)包含;任选地,其中所述另外的链中的RS编码抗原受体或其链或结构域(例如,T细胞受体或B细胞受体,例如嵌合抗原受体或其组分)。

[0235] 35. 实施方案1至32和34中任一项的方法,其用于产生嵌合抗原受体T细胞(CAR-T细胞)(例如,人CAR-T细胞)。

[0236] 36. 实施方案1至35中任一项的方法,其中第一载体是腺病毒载体。

[0237] 37. 实施方案24、25或26的方法,其中所述另外的链由腺病毒载体包含。

[0238] 38. 实施方案1至26中任一项的方法,其中所述方法是重组工程改造方法。

[0239] 在一个实例中,所述方法在细菌细胞(例如,大肠杆菌细胞)中在体外进行。

[0240] 39. 任一前述实施方案的方法,其包括将载体产物、链产物或细胞产物或RS产物施用于人或动物受试者以治疗或预防受试者中的疾病或病况;或可通过任一前述实施方案的方法获得的载体产物、核酸链产物或细胞产物或RS产物,其用于施用于人或动物受试者以治疗或预防受试者中的疾病或病况。

[0241] 40. 任一前述实施方案的方法,其中在步骤(b)和/或(d)中,所述载体是闭合的环状载体。

[0242] 例如,所述载体是共价闭合的环状载体,其任选地是超螺旋的。在一个实例中,所述方法在细胞中进行,并且在步骤(b)中,将所述细胞与多种所述载体组合。在一个实例中,所述多种包含多种所述闭合的环状载体,其中将一种或多种所述环状载体引入细胞中。在一个实例中,通过注射、电穿孔或细胞融合将一种或多种载体引入细胞中。

[0243] 41. 实施方案40的方法,其中在步骤(b)中,所述载体是超螺旋的。

[0244] 这对于本发明中细胞的有效转化是特别有用的,其中在所述方法中回收或插入大(> 10kb)或非常大(> 100kb)的RS或IS。

[0245] 42. 任一前述实施方案的方法,其中从第一链缺失HA1和HA2之间(例如,在HA2的3'末端和HA1的5'末端之间,例如,当载体在步骤(a)中是环状时)的序列,和/或从供体或另外的链缺失NS1和NS2之间的序列。

[0246] 43. 从属于实施方案1的实施方案42的方法,其中从供体链缺失RS。

[0247] 44. 从属于实施方案1的任一前述实施方案的方法,其中在HA1的5'末端或侧接所述末端的5'的前间隔区/PAM组合被步骤(d)破坏,例如,PS1序列被缺失,由此使得PS1在第一链中无功能。

[0248] 这对于锁定所述方法产生的改变是有用的,下一实施方案也是如此。

[0249] 45．从属于实施方案1的任一前述实施方案的方法，其中在HA2的3’末端或侧接所述末端的3’的前间隔区/PAM组合被步骤(d)破坏，例如，PS2序列和/或P2被缺失，由此使得PS2在第一链中无功能。

[0250] 46．任一前述实施方案的方法，其中将供体或另外的链序列插入第一链和/或从第一链缺失序列(例如，IS)，以形成新的前间隔区/PAM组合(例如，当所述载体在实施方案1的步骤(d)中为环状时，在HA2的3’末端和HA1的5’末端之间)或新的标记序列。

[0251] 47．从属于实施方案1的实施方案46的方法，其中将供体链序列紧邻 CS1的5’插入，并且供体序列和PS1序列一起形成与P1相邻的新的前间隔区(或形成用于检测所述同源重组已发生的新的标记序列)；和/或将供体序列紧邻 CS2的3’插入，并且供体序列和PS2序列一起形成新的前间隔区(或形成用于检测所述同源重组已发生的新的标记序列)。

[0252] 新的标记物可以是，例如，如本文公开的抗生素抗性标记序列。任选地，所述方法包括检测新的标记序列或其表达产物的存在。

[0253] 48．实施方案46或47的方法，其包括在步骤(d)之后Cas切割新组合的前间隔区，例如，用于随后将另外的序列插入第一链中或用于实施实施方案24至32和34至37中任一项的方法；所述方法任选地包括使用与步骤(c)中使用的Cas不同的Cas核酸酶切割新组合的前间隔区；或所述方法包括检测所述新的标记物。

[0254] 49．任一前述实施方案的方法，其中从供体或另外的链中缺失供体序列以形成新的前间隔区/PAM组合(例如，在NS2的3’末端和NS1的5’末端之间)或新的标记序列。

[0255] 50．实施方案49的方法，其中在步骤(d)之前，新的前间隔区/PAM序列或标记的序列由在供体或另外的链中发现的核苷酸序列(NS3和NS4)包含，并且在步骤(d)中缺失供体序列(例如RS)或另外的链序列并将NS3与NS4连接后形成新的前间隔区/PAM组合或标记序列。

[0256] 51．实施方案49或50的方法，其包括在步骤(d)之后在供体或另外的链中Cas切割新组合的前间隔区，例如，用于随后将另外的序列插入供体或另外的链或从供体或另外的链缺失；所述方法任选地包括使用与步骤(c)中使用的Cas不同的Cas核酸酶切割供体或另外的链中的新组合的前间隔区；或所述方法包括检测供体链中的所述新的标记物。

[0257] 52．实施方案51的方法，其中在步骤(d)之前，NS3侧接待缺失的供体或另外的链序列的5’，并且其中NS4侧接待缺失的供体或另外的序列的3’，其中在步骤(d)中，所述序列缺失，并且NS3的3’末端与NS4的5’末端连接，由此在供体或另外的链中形成新的前间隔区/PAM组合或标记序列。

[0258] 在一个实例中，NS3紧邻待缺失的序列的5’末端，或在所述末端的5’侧的10个核苷酸内。在一个实例中，NS4紧邻待缺失的序列的3’末端，或在所述末端的3’侧的10个核苷酸内。

[0259] 53．任一前述实施方案的方法，所述方法包括分离或复制修饰的供体或另外的链或其中发生缺失的供体或另外的核酸序列的区域。

[0260] 例如，所述分离或复制使用PCR。

[0261] 54．任一前述实施方案的方法，其中RS或IS包含调节元件、蛋白编码序列、外显子或标记序列或由其组成。

[0262] 55．从属于实施方案1的任一前述实施方案的方法，其中RS在步骤(d)中插入第一

链中，并且插入的序列对于从修饰的第一链表达是有功能的，所述方法任选地包括检测插入序列的表达产物。

[0263] 56．实施方案55的方法，其中插入的序列是蛋白编码序列、外显子或标记序列，其与第一链调节元件以功能性排列插入，由此所述调节元件能够控制来自修饰的第一链的供体序列的表达。

[0264] 57．任一前述实施方案的方法，其中插入的序列是标记序列，并且所述方法包括检测修饰的第一或另外的链中的标记序列；或者检测表达的标记物，其中在步骤(d)之前不能检测到所述标记物。

[0265] 58．实施方案2或从属于实施方案2的任一前述实施方案的方法，其中RS从供体链缺失并且所述缺失对宿主细胞是致死的。

[0266] 59．实施方案2或从属于实施方案2的任一前述实施方案的方法，其中所述方法下调细胞活力、生长或增殖(例如，杀死细胞)。

[0267] 60．实施方案2或从属于实施方案2的实施方案3至58中任一项的方法，其中所述方法上调细胞活力、生长或增殖(例如，杀死细胞)。

[0268] 61．实施方案2或从属于实施方案2的任一前述实施方案的方法，其中所述方法在细菌或古细菌细胞(例如大肠杆菌细胞)中进行。

[0269] 62．实施方案61的方法，其中所述细胞是人病原体物种的。

[0270] 63．实施方案61或62的方法，其中所述细胞是人微生物组物种(例如人肠道微生物组物种)的细胞。

[0271] 64．实施方案61至63中任一项的方法，其中所述第一链是噬菌体核酸链，其中通过用噬菌体感染宿主细胞将噬菌体链引入宿主细胞。

[0272] 65．实施方案2或从属于实施方案2的任一前述实施方案的方法，其中所述细胞包含内源CRISPR/Cas系统，并且其中内源宿主Cas用于在步骤(c)中切割。

[0273] 66．实施方案2或从属于实施方案2的任一前述实施方案的方法，其中所述方法在不将外源Cas或编码外源Cas的序列引入细胞的情况下进行。

[0274] 67．实施方案2或从属于实施方案2的任一前述实施方案的方法，其中所述方法在细胞中进行以杀死原核细胞群体，其中所述细胞是相同物种或菌株的。

[0275] 这可用作抗生素方法，例如用于治疗、去除或预防人、动物、食品、饮料、化妆品、石化或环境物质(例如，土壤或水或含水流体)中的细菌群体。

[0276] 68．实施方案67的方法，其中所述群体由原核(例如，细菌)细胞的混合群体构成，并且所述方法选择性杀死混合群体中的一些、但不是所有细胞，其中杀死的细胞是其中引入第一链并通过Cas切割CS1和/或CS2的细胞。

[0277] 69．实施方案67或68的方法，其中PS1和P1与被杀死的细胞的Cas核酸酶(例如，内源性Cas)同源，但不与未被杀死的群体中的细胞同源，其中CS1在杀死的细胞中被选择性切割；和/或其中PS2和P2与被杀死的细胞的Cas核酸酶(例如，内源性Cas)同源，但不与未被杀死的群体中的细胞同源，其中CS2在杀死的细胞中被选择性切割。

[0278] 70．实施方案67至69中任一项的方法，其包括分离产物群体或所述产物的样品。

[0279] 71．任一前述实施方案的方法，其中所述供体链、细胞或群体由以下包含：食品、食品成分或前体成分；饮料；水(例如，意欲用于人食用)；工业或环境物质(例如，油、石油产

品、土壤或水路或水库;或用于回收或处理油、石油产品、土壤、水、食品、食品成分或前体、或饮料或前体的饮料成分的设备)。

[0280] 72. 任一前述实施方案的方法,其中步骤(c)包括将第一链暴露于指导RNA(gRNA1,例如单一gRNA),所述指导RNA (gRNA1,例如单一gRNA)与PS1杂交以指导Cas切割PS1;任选地,其中步骤(c)包括从在步骤(b)中使用(例如,作为第一链的一部分)的核苷酸序列表达其gRNA1或crRNA序列;和/或其中步骤(c)包括将第一链暴露于gRNA1或第二指导RNA(gRNA2,例如单一gRNA),所述第二指导RNA(gRNA2,例如单一gRNA)与PS2杂交以指导Cas切割PS2;任选地,其中步骤(c)包括从在步骤(b)中使用(例如,作为第一链的一部分)的核苷酸序列表达其gRNA1或gRNA2或crRNA序列;

[0281] 73. 任一前述实施方案的方法,其中第一链或每条链是dsDNA。

[0282] 例如,所有所述链(例如,第一和供体,或第一和另外的链)都是dsDNA链。

[0283] 74. 任一前述实施方案的方法,其中第一链或每条链是ssDNA。

[0284] 例如,所有所述链(例如,第一和供体,或第一和另外的链)都是ssDNA链。

[0285] 75. 任一前述实施方案的方法(例如,当在真核细胞、例如人细胞中进行时),其中在步骤(c)中,Cas切割是在第一链中的前间隔区序列(例如,PS1和/或PS2)中,其中前间隔区包含种子序列(例如,S1或S2),其中所述种子序列包含在细胞基因组中罕见或未发现的核苷酸序列。

[0286] 这对于使除了在所述方法中使用的链的期望位置之外的异地切割的机会最小化是有用的。例如,在人细胞基因疗法中应当避免异地切割。因此,本发明实现了使用原核序列引导真核(例如人或动物或植物或酵母)靶细胞中的Cas切割的效用。因此,本发明提供了如下的该概念的另外的实施方案。

[0287] 76. 实施方案75的方法,其中所述种子序列包含原核起始密码子,例如GTG和/或TTG密码子,并且任选地,所述种子序列不包含ATG密码子。

[0288] 77. 实施方案75或76的方法,其中所述种子序列包含Shine Dalgarno序列,例如所述种子序列包含GAGG或AGGAGG。

[0289] 78. 任一前述实施方案的方法,其中(例如,当在真核细胞、例如人细胞中进行时),其中在步骤(c)中,Cas切割在针对原核(例如,大肠杆菌)、古细菌或病毒使用优化或在细胞基因组中未发现的前间隔区序列。

[0290] 该实施方案基于以下观察:原核生物等中的序列被优化(密码子优化)用于那些生物体中,因此序列优化不同于用于真核用途的序列优化。当在真核(例如人,例如iPS)细胞中进行该方法时,这增加切割的特异性的机会。

[0291] 79. 实施方案78的方法,其中所述前间隔区序列(例如,其最高达25、21或20个连续核苷酸)包含细菌Pribnow盒的序列;细菌前导序列(例如,Pel B);细菌-10序列或细菌-35元件序列(例如,TATAAT或TTGACA);细菌上游启动子元件(UP元件)序列;细菌调节或操纵子序列;或细菌复制起点序列。

[0292] 80. 制备适于与第二核酸链同源重组的第一核酸链的方法,例如,在根据从属于实施方案1的任一前述实施方案的方法中,所述制备方法包括:-

[0293] (a) 鉴定所述第二链的第一和第二核苷酸序列(NS1和NS2),任选地,其中选择NS1使得其在第二链中的NS2的3'侧;

[0294] (b) 在另一核酸链中组合

[0295] i. 用于HA1和NS1之间的同源重组的第一同源臂(HA1),其与所述第二链的第一预定靶核苷酸序列NS1互补;

[0296] ii. 用于HA2和NS2之间的同源重组的第二同源臂(HA2),其与所述第二链的第二预定靶核苷酸序列NS2互补;和

[0297] iii. 与PAM (P1)相邻的第一前间隔区序列(PS1),其中P1与第一Cas核酸酶同源,所述第一Cas核酸酶用于在第一切割位点(CS1)处切割PS1,其中CS1在HA1的5'末端或侧接所述末端的5'侧;和/或 与PAM (P2)相邻的第二前间隔区序列(PS2),其中P2与所述第一或不同的Cas核酸酶同源,所述第一或不同的Cas核酸酶用于在第二切割位点(CS2)处切割PS2,其中CS2在HA2的3'末端或侧接所述末端的3'侧;

[0298] 由此产生第一链。

[0299] 81. 实施方案80的方法,其中第一链产物由闭合的环状载体包含,任选地,其中所述载体是超螺旋的。

[0300] 82. 实施方案80或81的方法,其中所述第一链产物载体包含,所述载体是BAC、YAC、PAC、粘粒或质粒。

[0301] 83. 实施方案80至82中任一项的方法,其中CS1距HA1的所述末端不超过250nt。

[0302] 在一个实例中,CS1距HA1的所述末端不超过3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、150、200或250nt。

[0303] 84. 实施方案80至83中任一项的方法,其中CS2距HA2的所述末端不超过250nt。

[0304] 在一个实例中,CS1距HA2的所述末端不超过3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、150、200或250nt。

[0305] 85. 实施方案80至84中任一项的方法,其中HA1和HA2各自长度为至少15nt。

[0306] 在一个实例中,HA1长度为至少20、30、40、50、60、70、80、90、100、150、200、250、300、400、500、600、700、800或900nt,长度为至少1、2、3、4、5,6、7、8、9、10、50、100、200、300、400或500kb。在一个实例中,HA2长度为至少20、30、40、50、60、70、80、90、100、150、200、250、300、400、500、600、700、800或900nt,长度为至少1、2、3、4、5,6、7、8、9、10、50、100、200、300、400或500kb。

[0307] 86. 实施方案80至85中任一项的方法,其中HA1和HA2包含至少0.2kb的总长度。

[0308] 在一个实例中,HA1和HA2包含至少0.2、0.3、0.4、0.5、0.6、0.7、0.8、0.9、1、2、3、4、5、6、7、8、9、10、20、30、50或100kb的总长度或由其组成。

[0309] 87. 实施方案80至86中任一项的方法,其中HA1和HA2分别与NS1和NS2同基因或至少80%、90%或95%相同。

[0310] 88. 实施方案80至87中任一项的方法,其中所述第一链产物包含(以5'至3'方向):-

[0311] G. 包含CS1的第一前间隔区序列(PS1);

[0312] H. 与PS1的3'末端相邻的第一PAM (P1);和

[0313] I. HA1;

[0314] 其中PS1的种子序列(S1)将CS1连接至P1;且

[0315] 任选地,其中HA1包含S1和P1。

[0316] 和/或所述第一链产物包含:-

[0317] J. HA2;

[0318] K. 包含CS2的第二前间隔区序列(PS2);

[0319] L. 与PS2的3'末端相邻的第二PAM (P2);

[0320] 其中PS2的种子序列(S2)将CS2连接至P2;且

[0321] 任选地,其中PS2的剩余部分在HA2的3'末端。

[0322] 89. 实施方案80至88中任一项的方法,其包括在第一链中提供编码所述第一Cas(例如,Cas9)的可表达的核苷酸序列;任选地,其中所述第一链包含(或第一链与包含以下的另外的链组合)可表达的核苷酸序列,其编码与PS1互补的gRNA (gRNA1,例如单一gRNA);或所述序列编码在宿主细胞中形成gRNA1的可表达的crRNA,其任选地具有由宿主细胞序列编码的tracrRNA。

[0323] 在一个实例中,所述tracrRNA由载体序列(例如,由第一载体或与第一载体组合的另一载体)编码(例如,共转化至其中进行该方法的细胞中)。

[0324] 90. 实施方案80至89中任一项的方法,所述方法包括在步骤(b)之后Cas切割CS1和/或CS2,任选地用相同的Cas核酸酶。

[0325] 91. 实施方案80至90中任一项的方法,其中所述方法在细胞(例如原核细胞)中进行。

[0326] 92. 实施方案91的方法,其中所述方法包括重组工程改造。

[0327] 93. 实施方案80至92中任一项的方法,其中NS1和NS2侧接必需基因、抗生素抗性基因或毒力基因的序列,例如原核细胞的序列。

[0328] 在一个实例中,所述原核细胞是本文所述的任一原核细胞(例如,人病原体细胞)。

[0329] 94. 实施方案93的方法,其进一步产生抗生素组合物或试剂盒,其包含所述方法的产物与抗生素的组合,其中NS1和NS2侧接抗生素抗性基因(例如,所述原核细胞)的序列,其中所述抗生素是与核酸组合的抗生素。

[0330] 95. 实施方案80至94中任一项的方法,其进一步在第一链中(或在除了第一链以外的另外的核酸链上)包括编码指导RNA (gRNA1,例如单一gRNA)的核苷酸序列,其与PS1和/或PS2互补,或所述序列编码在细胞中形成gRNA1的crRNA(其任选地具有由宿主细胞序列编码的tracrRNA);和/或其进一步在第一链(或在除了第一链以外的另外的核酸链上)包括编码第二指导RNA (gRNA2,例如单一gRNA)的核苷酸序列,其与PS2互补,或所述序列编码在细胞中形成gRNA2的crRNA(其任选地具有由宿主细胞序列编码的tracrRNA)。

[0331] 在一个实例中,所述tracrRNA由载体序列(例如,由第一载体或与第一载体组合的另一载体)编码(例如,共转化至其中进行该方法的细胞中)。

[0332] 96. 实施方案80至95中任一项的方法,其进一步包括产生包含实施方案80至95中任一项的第一链、组合物或试剂盒产品的组合物,其中所述组合物是农业试剂(例如除草剂或肥料)、食品、食品成分或前体成分;饮料;水组合物(例如,意欲用于人食用);工业或环境物质(例如,油、石油产品、土壤或水路或水库;或用于回收或处理油、石油产品、土壤、水、食品、食品成分或前体、或饮料或前体的饮料成分的设备)。

[0333] 97. 实施方案80至95中任一项的方法,其进一步包括产生包含实施方案80至95中任一项的第一链、组合物或试剂盒产品的药物用于治疗和/或预防人或非人动物中的疾病

或病况。

[0334] 98. 实施方案80至97中任一项的方法,步骤(b)包括

[0335] (i) 获得包含克隆的细胞基因组序列的核酸链(例如,由BAC包含),所述序列包含(以5'至3'方向)HA2和HA1以及任选地其之间的连接序列(JS1);

[0336] (ii) 获得包含(以5'至3'方向)第一同源盒(HB2)和第二同源盒(HB1)的靶向序列;其中HB2的序列与HA2的序列或JS1的5'末端处的序列或跨越HA2和JS1的5'末端的序列相同;且其中HB1的序列与HA1的序列或JS1的3'末端处的序列或跨越HA1和JS1的3'末端的序列相同;其中所述靶向序列包含将HB1的5'末端连接至HB2的3'末端的连接序列(JS2);其中JS2包含所述PS1/P1/CS1组合和/或所述PS2/P2/CS2组合,其中CS1距HA1的5'末端不超过250nt;且其中CS2距HA2的3'末端不超过250nt;

[0337] (iii) 进行步骤(i)的核酸链与靶向序列的同源盒之间的同源重组(例如,使用重组工程改造),以产生核酸链(第一核酸链),其包含(以5'至3'方向)

[0338] a. HA2、PS2(包含CS2)、P2、PS1(包含CS1)、P1和HA1,其任选地包含CS1和CS2之间的选择标记序列;

[0339] b. HA2、PS2(包含CS2)、P2和HA1;或

[0340] c. HA2 PS1 (包含CS1)、P1和HA1。

[0341] 99. 实施方案98的方法,用于产生用于人细胞的基因疗法的第一链或载体或产生CAR-T细胞,其中所述细胞是步骤(i)的细胞或其后代或祖先;任选地,其中所述方法包括进行实施方案28至32、34和35中任一项的方法以进行所述人细胞基因疗法或产生所述CAR-T细胞。

[0342] 100. 可通过实施方案80至99中任一项的方法获得的第一链、载体、组合物、药物或试剂盒,其任选地包含HA1和HA2之间的回收序列(RS)或IS,其中RS如实施方案1至79中任一项中所定义,例如,用于人或动物的基因疗法或用于治疗或预防人或动物中的细菌感染。

[0343] 101. 实施方案100的第一链或载体,其用于人细胞的基因疗法中,所述链或载体包含第一链中的前间隔区序列(例如,PS1和/或PS2),其中所述前间隔区包含种子序列(例如,S1或S2),其中所述种子序列包含在人细胞基因组中罕见或未发现的核苷酸序列。

[0344] 102. 实施方案101的第一链或载体,其中所述种子序列包含原核起始密码子,例如GTG和/或TTG密码子,且任选地,所述种子序列不包含ATG密码子。

[0345] 103. 实施方案101或102的第一链或载体,其中所述种子序列包含Shine Dalgarno序列,例如,所述种子序列包含GAGG或AGGAGG。

[0346] 104. 实施方案101至103中任一项的第一链或载体,其中所述链包含针对原核(例如,大肠杆菌)、古细菌或病毒使用优化或在人细胞基因组中未发现的前间隔区序列。

[0347] 105. 实施方案104的第一链或载体,其中所述前间隔区序列(例如,其最高达25或20个连续核苷酸)包含细菌Pribnow盒的序列;细菌前导序列(例如,Pel B);细菌-10序列或细菌-35元件序列(例如,TATAAT或TTGACA);细菌上游启动子元件(UP元件)序列;细菌调节或操纵子序列;或细菌复制起点序列。

[0348] 106. 实施方案101至105中任一项的第一链或载体,其用于人或人细胞的基因疗法(例如,在根据实施方案28至32中任一项的方法中)或用于产生嵌合抗原受体T细胞(CAR-T细胞)(例如,人CAR-T细胞)用于施用于人以治疗或预防人中的疾病或病况(例如,癌症或

自身免疫疾病或病况)。

[0349] 107. 抑制原核宿主细胞的活力、生长或增殖的方法,所述方法包括

[0350] (e) 提供第一载体,其包含与PAM (P1)相邻的前间隔区序列(PS1),所述PAM (P1)与第一Cas核酸酶同源;

[0351] (f) 提供毒素或毒素前体;或包含编码毒素或毒素前体的序列的载体,其中所述毒素对原核细胞有毒;

[0352] (g) 将细胞与载体组合,其中将载体引入细胞,并在细胞中提供毒素;和

[0353] (h) 使用Cas核酸酶切割第一载体的PS1,其使用第一Cas核酸酶,其中切割促进细胞中的毒素活性,由此杀死细胞或减少细胞生长或增殖。

[0354] 在一个实例中,本发明的任一配置的方法不是在人或动物体或胚胎上实施的治疗、预防或诊断方法。在一个实例中,所述方法是用于生产或处理化妆品、药物、食品、饮料、水或石化产品的美容方法或方法。

[0355] 108. 实施方案107的方法,其中第一Cas由细胞中的与PS1杂交的第一指导RNA(gRNA1,例如单一指导RNA)指导,由此Cas切割PS1。

[0356] 任选地,编码gRNA1的序列由引入细胞的载体(例如,第一载体)包含。

[0357] 109. 实施方案108的方法,其中所述方法包括在步骤(d)之前在细胞中提供gRNA1或在细胞中产生gRNA1的序列,其中所述序列由在步骤(c)中引入细胞的第一或另一载体包含。

[0358] 在一个实例中,所述序列编码crRNA或包含可表达的CRISPR阵列。

[0359] 110. 实施方案107至109中任一项的方法,其中第一Cas核酸酶对细胞是内源的。

[0360] 111. 转化宿主细胞的方法,所述方法包括

[0361] (a) 提供第一核酸载体,所述载体处于第一状态,其包含

[0362] i. 与PAM (P1)相邻的前间隔区序列(PS1),所述PAM (P1)与第一Cas核酸酶同源;和

[0363] ii. 第一核苷酸序列,其包含用于标记载体的第一状态的第一标记序列;

[0364] (b) 将细胞与载体组合,其中将载体引入细胞;和

[0365] (c) 使用Cas核酸酶切割第一载体的PS1,其使用第一Cas核酸酶,其中切割将载体转化为第二状态,其中使得标记序列无功能或缺失;和

[0366] (d) 任选地通过检测标记物在细胞中无功能或从所述载体缺失来检测第二状态。

[0367] 所述标记序列可以是本文公开的任何标记序列,例如tk或HPRT序列。

[0368] 112. 实施方案111的方法,其中通过允许切割载体经历与细胞的染色体或附加体核酸的同源重组来使标记序列无功能或缺失。

[0369] 113. 实施方案112的方法,其中将载体序列插入宿主核酸中。

[0370] 114. 实施方案111或112的方法,其中将宿主核酸的序列插入载体中。

[0371] 115. 实施方案114的方法,其中通过插入宿主序列在载体中破坏PS1和/或P1。

[0372] 116. 实施方案111至115中任一项的方法,其中所述标记序列从载体Cas切除并且所述载体经历通过宿主核酸的序列同源重组插入载体的缺口修复。

[0373] 117. 实施方案107至116中任一项的方法,其中通过切割在载体中破坏PS1和/或P1。

[0374] 118．实施方案107至116中任一项的方法，其中所述方法在不将外源Cas或编码外源Cas的序列引入细胞的情况下在一种或多种细胞中进行。

[0375] 例如，所述细胞是如本文所述的原核细胞，例如细菌或古细菌或人病原体或微生物组细胞。

[0376] 119．实施方案107至118中任一项的方法，其中所述方法在细胞中进行，并且PS1和P1与用于进行步骤的切割的细胞内源的Cas同源。

[0377] 120．实施方案107至119中任一项的方法，其中所述方法在细胞中进行以杀死原核细胞群体，其中所述细胞是相同物种或菌株的。

[0378] 121．实施方案120的方法，其中所述群体由以下包含：食品、食品成分或前体成分；饮料；水(例如，意欲用于人食用)；工业或环境物质(例如，油、石油产品、土壤或水路或水库；或用于回收或处理油、石油产品、土壤、水、食品、食品成分或前体、或饮料或前体的饮料成分的设备)。

[0379] 122．实施方案120或121的方法，其中所述群体由原核(例如，细菌)细胞的混合群体构成，并且所述方法选择性杀死混合群体中的一些、但不是所有细胞，其中杀死的细胞是其中引入第一载体或第一链并通过Cas切割PS1的细胞。

[0380] 123．实施方案122的方法，其中所述第一载体或第一链包含编码与细胞中的PS1杂交的gRNA (gRNA1，例如单一gRNA)的序列或所述载体或链编码在细胞中形成gRNA1的crRNA，其任选地具有由宿主细胞序列编码的tracrRNA。

[0381] 任选地，所述tracrRNA由载体序列、例如由第一载体编码。

[0382] 124．实施方案122或123的方法，其中PS1和P1与被杀死的细胞的Cas核酸酶(例如，内源性Cas)同源，但不与未被杀死的群体中的细胞同源，其中PS1在杀死的细胞中被选择性切割。

[0383] 125．实施方案122至124中任一项的方法，其中混合群体是人或非人动物微生物组群体，例如肠道、阴道、腋窝或口腔微生物组群体。

[0384] 126．实施方案125的方法，其包括分离实施方案125的产物群体或所述产物的样品。

[0385] 127．用于转化细胞(例如，真核细胞)的载体，其中所述载体包含第一核酸链，且所述细胞包含第二核酸链，所述第一链包含

[0386] (a) 侧接第一同源臂(HA1)的插入序列(IS)，其中HA1与第二链的预定核苷酸序列(NS1)互补，其中HA1能够在细胞中与NS1同源重组以修饰一条或两条链；

[0387] (b) 与PAM (P1)相邻的第一前间隔区序列(PS1)，其中P1与第一Cas核酸酶同源，所述第一Cas核酸酶用于在第一切割位点(CS1)处切割PS1；

[0388] (c) 其中CS1在HA1的末端或侧接HA1，其中CS1能够被第一Cas切割以产生切割的第一链，其中HA1能够与NS1同源重组。

[0389] 在一个实例中，IS紧邻HA1，例如，邻近HA1的5’或3’末端。在一个实例中，P1紧邻PS1的(3’)。在一个实例中，CS1不由IS包含。例如，CS1在最远离IS的HA1的末端，或者所述末端在CS1和IS之间。

[0390] 128．实施方案127的载体，其中所述载体是能够转化细胞的闭合的环状载体并且IS另外包含至少10kb的核苷酸序列(例如，人序列)，其中所述环状载体能够转化细胞用于

在HA1和NS1之间进行同源重组。

[0391] 例如,IS长度为至少50、100、200、300、400、500、600、700、800或900 kb或1Mb。

[0392] IS序列能够通过在HA1和NS1之间的同源重组插入第二链。

[0393] 129. 实施方案128的载体,其中所述载体是超螺旋的。

[0394] 130. 实施方案127、128或129的载体,其中所述第一链与指导RNA (gRNA1,例如单一指导RNA)组合;或与用于产生gRNA1的可表达的核苷酸序列组合或包含用于产生gRNA1的可表达的核苷酸序列,其中gRNA1包含与PS1互补的核苷酸序列,用于指导Cas以切割PS1。

[0395] gRNA1与PS1、P1和第一Cas同源,其用于指导Cas至PS1用于切割PS1。

[0396] 131. 实施方案127至130中任一项的载体,其中CS1远离HA1不超过250nt。

[0397] 在一个实例中,CS1距HA1的所述末端不超过3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、150、200或250nt。

[0398] 132. 实施方案127至131中任一项的载体,其中

[0399] (a) 所述第一链包含第二同源臂(HA2),其与第二链的第二核苷酸序列(NS2)同源,用于细胞中HA2和NS2之间的同源重组和所述一条或两条链的修饰;且

[0400] (b) IS在HA1和HA2之间;

[0401] (a) 所述第一链包含与PAM (P2)相邻的第二前间隔区序列(PS2),其中P2与所述第一Cas或不同的Cas核酸酶同源,用于在切割位点(CS2)处切割PS2;

[0402] (c) 其中CS2在HA2的末端或侧接HA2,其中CS2能够被Cas切割以产生切割的第一链,其中HA2能够与NS2同源重组。

[0403] (d) 在HA1的5'末端或侧接所述末端的5'的第一链的CS1;和/或Cas核酸酶(例如,Cas9,任选化脓脓链球菌或金黄色葡萄球菌Cas9)切割HA2的3'末端或侧接所述末端的3'的第一链的第二切割位点(CS2)。

[0404] 例如,P1紧邻PS1的(3')。在一个实例中,CS1不由IS包含。在一个实例中,CS1在最远离IS的HA1的末端,或者所述末端在CS1和IS之间。

[0405] 在一个实例中,IS将HA1直接连接至HA2。在一个实例中,所述载体(例如,环状载体)是回收载体,HA1在所述第一链中的HA2的5'侧;且NS1在所述第二链中的NS2的3'侧,其中所述第二链包含NS1和NS2之间的回收序列(RS);且其中IS任选地包含当HA1和HA2分别与NS1和NS2重组并且RS被回收至第一链中时缺失的标记序列。

[0406] 在另一个实例中,所述载体是插入载体,且HA1在所述第一链中的HA2的3'侧;且NS1在所述第二链中的NS2的3'侧。当HA1和HA2分别与NS1和NS2重组时,将IS的序列(例如至少10kb的序列)插入第二链中。任选地,IS在HA1和HA2之间并且包含标记序列,当HA1和HA2分别与NS1和NS2重组时,所述标记序列也插入第二链中和/或从第一链缺失。这对于确定已经发生成功的同源重组是有用的。所述第二链任选地包含将NS1连接至NS2的核苷酸序列(或者NS1与NS2直接相邻)。

[0407] 133. 实施方案132的载体,其中所述第一链以5'至3'方向包含,(i) CS1、P1、HA1、IS、HA2、CS2和P2;或(ii) CS2、P2、HA2、IS、HA1、CS1和P1。

[0408] 134. 实施方案133的载体,其中选项(i)适用,且P1由HA1包含,例如侧接HA1的5'末端的3'或在所述末端;或选项(ii)适用,且其中P2由HA2包含,例如,侧接HA2的5'末端的3'或在所述末端。

[0409] 135. 实施方案132、133或134的载体，其中CS2距HA2不超过250nt。

[0410] 在一个实例中，CS2距HA2的所述末端不超过3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、150、200或250nt。

[0411] 136. 实施方案127至135中任一项的载体，其中HA1和HA2各自长度为至少100nt。

[0412] 在一个实例中，HA1长度为至少50或至少100、150、200、300、400、500、600、700、800或900nt，或长度为至少1、2、3、4、5、6、7、8、9、10、20、30、50或100kb。在一个实例中，HA2长度为至少50或至少100、150、200、300、400、500、600、700、800或900nt，或长度为至少1、2、3、4、5、6、7、8、9、10、20、30、50或100kb。

[0413] 137. 实施方案127至136中任一项的载体，其中HA1和HA2包含至少0.2kb的总长度。

[0414] 在一个实例中，HA1和HA2包含长度为至少0.2、0.3、0.4、0.5、0.6、0.7、0.8、0.9、1、2、3、4、5、6、7、8、9、10、20、30、50或100kb的总长度或由其组成。

[0415] 138. 实施方案127至137中任一项的载体，其中HA1和HA2分别与NS1和NS2同基因(即，100％相同的核苷酸序列)。

[0416] 在一个实例中，HA1和HA2各自与NS1和NS2的核苷酸序列分别至少80、85、90或95％相同。

[0417] 139. 实施方案127至138中任一项的载体，其中第一链或每条链是dsDNA。

[0418] 140. 实施方案127至138中任一项的载体，其中第一链或每条链是ssDNA。

[0419] 141. 实施方案127至140中任一项的载体，其中PS1和/或PS2各自包含种子序列，所述种子序列包含在细胞基因组(例如，人细胞基因组)中罕见或未发现的核苷酸序列。

[0420] 142. 实施方案141的载体，其中各种子序列包含原核起始密码子，例如GTG和/或TTG密码子，且任选地，所述种子序列不包含ATG密码子。

[0421] 143. 实施方案141或142的载体，其中各种子序列包含Shine Dalgarno序列，例如，所述种子序列包含GAGG或AGGAGG。

[0422] 144. 实施方案141至143中任一项的载体，其中PS1和/或PS2序列各自针对原核(例如，大肠杆菌)、古细菌或病毒使用优化，或在细胞基因组(例如，人细胞基因组)中未发现。

[0423] 145. 实施方案144的载体，其中各所述前间隔区序列(例如，其最高达25或20个连续核苷酸)包含细菌Pribnow盒的序列；细菌前导序列(例如，Pel B)；细菌-10序列或细菌-35元件序列(例如，TATAAT或TTGACA)；细菌上游启动子元件(UP元件)序列；细菌调节或操纵子序列；或细菌复制起点序列。

[0424] 146. 实施方案127至145中任一项的载体，其中所述载体与编码第一Cas(例如，Cas9)的可表达的核酸组合或包含编码第一Cas(例如，Cas9)的可表达的核酸；或其中所述载体与所述Cas组合。

[0425] 147. 实施方案127至146中任一项的载体，其中所述载体在所述真核细胞(例如，人细胞，例如，体外的人干细胞或iPS细胞)内。

[0426] 148. 包含一群闭合环状载体(例如，超螺旋载体)的载体集合，其中各所述环状载体根据实施方案127至147中任一项，例如，其中各载体中的IS包含长度为至少10kb的序列(例如，人序列)。

[0427] 149. 抗生素(例如,抗细菌或抗古细菌)组合物,其包含实施方案127至148中任一项的载体、组合或集合。

[0428] 150. 用于治疗或预防人中的疾病或病况(例如细菌感染或肥胖症)的药物,所述药物包含实施方案127至149中任一项的载体、组合或集合。

[0429] 151. 根据实施方案127至150中任一项所述的载体、组合、集合、抗生素或药物,其中所述载体或所述集合的各载体进一步根据实施方案100至106中任一项。

[0430] 152. 产生前体载体的方法,例如,其中所述载体是实施方案80中所述载体的前体,所述前体载体包含

[0431] (a) 以5'至3'顺序,与PAM (P1)相邻的第一前间隔区序列(PS1),其中P1与第一Cas核酸酶同源,所述第一Cas核酸酶用于在第一切割位点(CS1)处切割PS1;和

[0432] (b) P1的紧邻3'的第一核苷酸序列,其中所述第一序列包含与P1的3'末端相邻或侧接所述末端的3'的第一克隆位点,用于将第二核酸序列克隆至所述前体载体中,其中所述第二序列包含HA1和HA2中的一种或两种;

[0433] 所述方法包括

[0434] (c) 获得前体核酸;

[0435] (d) 鉴定核酸中的克隆位点或在核酸中产生克隆位点核酸;和

[0436] (e) 鉴定核酸中的PS1/P1组合或在核酸中产生组合;

[0437] 其中(d)和(e)以任何顺序或同时进行,且其中(d)和(e)中的至少一者包括所述产生步骤,由此产生并任选地分离所述前体载体。

[0438] 在一个替代方案中,提供了:-

[0439] 152a. 制备适于与第二核酸链同源重组的第一核酸链的方法,例如,在根据任一前述权利要求的方法中,所述制备方法包括:-

[0440] (a) 鉴定所述第二链的第一和第二核苷酸序列(NS1和NS2),任选地,其中选择NS1使得其在第二链中的NS2的3'侧;

[0441] (b) 在另一核酸链中组合

[0442] i. 用于HA1和NS1之间的同源重组的第一同源臂(HA1),其与所述第二链的第一预定靶核苷酸序列NS1互补;

[0443] ii. 用于HA2和NS2之间的同源重组的第二同源臂(HA2),其与所述第二链的第二预定靶核苷酸序列NS2互补;和

[0444] iii. 与PAM (P1)相邻的第一前间隔区序列(PS1),其中P1与第一Cas核酸酶同源,所述第一Cas核酸酶用于在第一切割位点(CS1)处切割PS1,其中CS1在HA1的5'末端或侧接所述末端的5'侧;和/或与PAM (P2)相邻的第二前间隔区序列(PS2),其中P2与所述第一或不同的Cas核酸酶同源,所述第一或不同的Cas核酸酶用于在第二切割位点(CS2)处切割PS2,其中CS2在HA2的3'末端或侧接所述末端的3'侧;

[0445] 由此产生第一链。

[0446] 153. 实施方案152或152a的方法,其中所述产物载体是环状载体,例如闭合的环状载体。

[0447] 154. 实施方案152或153的方法,其中克隆位点远离CS1不超过1、2、3、4、5、6、7、8、9、10、15、20、25、30、40、50、100、200或250nt。

[0448] 155．实施方案152至154中任一项的方法，其中所述前体载体包含第一克隆位点的3’侧的第二克隆位点，所述方法包括鉴定所述核酸中的第二克隆位点或在所述核酸中产生第二克隆位点。

[0449] 156．实施方案155的方法，其中所述前体载体以5’至3’顺序包含与PAM (P2)相邻的第二前间隔区序列(PS2)，其中P2与所述第一Cas或不同Cas核酸酶同源，所述第一Cas或不同Cas核酸酶用于在第二切割位点(CS2)处切割PS2；其中所述第二克隆位点与PS2的5’末端相邻，或侧接所述末端的5’。

[0450] 157．实施方案156的方法，其中所述方法包括鉴定或产生PS2/P2组合，其中产生所述第二克隆位点和PS2/P2组合中的至少一个。

[0451] 158．实施方案155、156或157的方法，其中所述前体载体包含克隆位点之间的标记序列(例如，用于表达可检测标记物)，其中当第二序列插入克隆位点时，所述标记物从载体中丢失。

[0452] 所述标记序列可以是本文公开的任何标记序列，例如tk或HPRT序列。

[0453] 159．实施方案152至158中任一项的方法，其中第一或每个克隆位点是限制性内切核酸酶位点(例如，其中所述位点被相同的内切核酸酶识别)。

[0454] 160．实施方案152至158中任一项的方法，其中所述第一或每个克隆位点是位点特异性重组(SSR)位点(例如，其中相同的SSR位点)；任选地，其中第一或每个位点是lox或frt位点。

[0455] 在一个实例中，使用不相容的lox位点(例如，loxP和lox511；或loxP和lox2272)。

[0456] 161．实施方案155至160中任一项的方法，其进一步包括获得所述第二核酸序列并将第二序列克隆至克隆位点中，其中将HA1和HA2插入P1的前体载体3’，且所述标记序列从载体缺失，并且任选地检测以下中的至少一种：(i)从载体丢失标记序列或丢失标记表达产物，(i)载体中存在HA1序列和(iii)载体中存在HA2序列；并且任选地分离所述产物载体。

[0457] 162．实施方案156至160的方法，其包括在产生产物前体载体之后，使用Cas切割CS1和/或CS2，任选地用相同的Cas核酸酶。

[0458] 163．实施方案152至162中任一项的方法，其进一步包括获得所述第二核酸序列并将第二序列克隆至第一克隆位点中，其中将HA1和HA2插入P1的前体载体3’；且任选地分离所述产物载体。

[0459] 164．实施方案161、162或163的方法，其后包括在根据实施方案1至79中任一项所述的另外的方法中使用产物载体，其中实施方案161、162或163的产物载体在另外的方法中提供所述第一链。

[0460] 165．实施方案80至99中任一项的方法，其中步骤(b)使用实施方案152至164中任一项的方法。

[0461] 166．可通过实施方案152至163中任一项的方法获得的前体载体。

[0462] 167．实施方案166的载体，其中所述载体是dsDNA载体并且任选地是环状的。

[0463] 168．实施方案166的载体，其中所述载体是ssDNA载体并且任选地是环状的。

[0464] 169．实施方案166至168中任一项的载体，其中PS1和/或PS2各自包含种子序列，所述种子序列包含在细胞基因组(例如，人细胞基因组)中罕见或未发现的核苷酸序列。

[0465] 170．实施方案169的载体，其中各种子序列包含原核起始密码子，例如GTG和/或

TTG密码子,且任选地,所述种子序列不包含ATG密码子。

[0466] 171. 实施方案169或170的载体,其中各种子序列包含Shine Dalgarno序列,例如,所述种子序列包含GAGG或AGGAGG。

[0467] 172. 实施方案169至171中任一项的载体,其中PS1和/或PS2序列各自针对原核(例如,大肠杆菌)、古细菌或病毒使用优化,或在细胞基因组(例如,人细胞基因组)中未发现。

[0468] 173. 实施方案172的载体,其中各所述前间隔区序列(例如,其最高达25或20个连续核苷酸)包含细菌Pribnow盒的序列;细菌前导序列(例如,Pel B);细菌-10序列或细菌-35元件序列(例如,TATAAT或TTGACA);细菌上游启动子元件(UP元件)序列;细菌调节或操纵子序列;或细菌复制起点序列。

[0469] 174. 实施方案166至173中任一项的载体,其中所述载体与编码第一Cas(例如,Cas9)的可表达的核酸组合或包含编码第一Cas(例如,Cas9)的可表达的核酸;或其中所述载体与所述Cas组合。

[0470] 175. 实施方案166至174中任一项的载体或组合,其中所述前体载体与指导RNA(gRNA1,例如单一指导RNA)组合;或与用于产生gRNA1的可表达的核苷酸序列组合或包含用于产生gRNA1的可表达的核苷酸序列,其中gRNA1包含与PS1互补的核苷酸序列,用于指导Cas以切割PS1。

[0471] gRNA1与PS1、P1和第一Cas同源,其用于指导Cas至PS1用于切割PS1。

[0472] 176. 实施方案166至175中任一项的载体或组合,其中所述前体载体可通过实施方案156的方法获得,并且与指导RNA (gRNA2,例如单一指导RNA)组合;或与用于产生gRNA2的可表达的核苷酸序列组合或包含用于产生gRNA2的可表达的核苷酸序列,其中gRNA2包含与PS2互补的核苷酸序列,用于指导Cas以切割PS2。

[0473] gRNA2与PS2、P2和第一Cas同源,其用于指导Cas至PS2用于切割PS2。

[0474] 本发明提供了以下项:-

[0475] 1. 在原核或真核细胞(例如,人细胞,例如,iPS细胞)中回收供体核酸序列的方法,所述方法包括

[0476] (a) 提供第一核酸链(例如,由工程改造的载体包含),所述链包含与PAM (P1)相邻的前间隔区序列(PS1),所述PAM (P1)与第一Cas核酸酶同源;

[0477] (b) 将第一链引入宿主细胞中,以将细胞内的供体核酸链与第一链组合;和

[0478] (c) 使用用第一Cas核酸酶的Cas核酸酶切割第一链的PS1;

[0479] (d) 允许切割的第一链经历与细胞内的供体链的同源重组,其中将供体链的序列插入第一链中以形成修饰的第一链,由此PS1和/或P1不通过插入供体序列在修饰的第一链中再形成。

[0480] 在一个实例中,所述供体链由细胞的染色体或附加体核酸包含。

[0481] 2. 在真核或原核细胞中缺失供体核酸序列的方法,所述方法包括

[0482] (a) 提供第一核酸链(例如,由工程改造的载体包含),所述链包含与PAM (P1)相邻的前间隔区序列(PS1),所述PAM (P1)与第一Cas核酸酶同源;

[0483] (b) 将第一链引入宿主细胞中,以将细胞内的供体核酸链与第一链组合;和

[0484] (c) 使用用第一Cas核酸酶的Cas核酸酶切割第一链的PS1;

[0485] (d) 允许切割的第一链经历与细胞内的供体链的同源重组,其中缺失供体链的序列以形成修饰的供体链,其中PS1和/或P1不在修饰的供体链中形成或侧接修饰的供体链中的缺失。

[0486] 3. 项1或2的方法,其中在步骤(c)或(d)中从第一链和/或供体链缺失序列。

[0487] 4.任一前述项的方法,其中缺失PS1序列,由此使PS1在第一链中无功能。

[0488] 5.任一前述项的方法,其中缺失P1序列,由此使P1在第一链中无功能。

[0489] 6.任一前述项的方法,其中将供体序列插入第一链中,或从第一链缺失序列,以形成与P1相邻的新的前间隔区;所述方法任选地包括在步骤(d)之后切割新的前间隔区。

[0490] 7.项6的方法,其中切割所述新的前间隔区以产生另外的修饰的第一链,其用于与另外的核酸链同源重组,用于序列插入另外的修饰的第一链或从其中缺失。

[0491] 8. 项1至5中任一项的方法,其中将供体序列插入第一链中,或从第一链缺失序列,以形成与PS1相邻的新的PAM;所述方法任选地包括在步骤(d)之后使用不同于第一Cas的Cas核酸酶切割PS1。

[0492] 9.项8的方法,其中用所述不同的Cas切割PS1以产生另外的修饰的第一链,其用于与另外的核酸链同源重组,用于序列插入另外的修饰的第一链或从其中缺失。

[0493] 10.项1至4中任一项的方法,其中步骤(c)或(d)缺失第一链的标记序列(例如,抗生素抗性基因序列),其中所述方法任选地包括检测步骤(d)之后的标记物的不存在或减少。

[0494] 11.任一前述项的方法,其中将第一链序列插入供体链中,或从供体链缺失序列,以形成(i)与PAM相邻的新的前间隔区;所述方法任选地包括在步骤(d)之后切割新的前间隔区和/或(ii)供体链中的新的标记序列,例如,其中所述标记序列对于表达可检测标记物有功能。

[0495] 12.项11的方法,其中(i)切割所述新的前间隔区以产生另外的修饰的供体链,其用于与另外的核酸链同源重组,用于序列插入另外的修饰的供体链或从其中缺失和/或(ii)检测标记物以确定步骤(d)中已发生同源重组。

[0496] 13.项11或12的方法,其中所述新的前间隔区和PAM序列或标记序列在步骤(d)之前由供体链中发现的核苷酸序列(NS3和NS4)包含,并且在步骤(d)中缺失供体序列并将NS3与NS4连接后形成新的前间隔区/PAM组合或标记序列。

[0497] 14.项13的方法,其中步骤(a)中的第一链包含第一同源臂(HA1),其中HA1与供体核酸链的预定第一核苷酸序列(NS1)同源;在步骤(d)中,在HA1和NS1之间进行同源重组以产生修饰的链;在步骤(d)之前,NS1侧接待缺失的供体序列的3'侧,且NS3在NS1的5'末端(在最多达10个核苷酸处或间隔最多达10个核苷酸),并且其中NS4侧接待缺失的供体序列的5'侧,其中在步骤(d)中,缺失所述供体序列,并且NS3的5'末端与NS4的3'末端连接,由此在供体链中形成新的前间隔区/PAM组合或标记序列。

[0498] 15.项13的方法,其中步骤(a)中的第一链包含第一同源臂(HA1),其中HA1与供体核酸链的预定第一核苷酸序列(NS1)同源;在步骤(d)中,在HA1和NS1之间进行同源重组以产生修饰的链;在步骤(d)之前,NS1侧接待缺失的供体序列的5'侧,且NS3在NS1的3'末端(在最多达10个核苷酸处或间隔最多达10个核苷酸),并且其中NS4侧接待缺失的供体序列的3'侧,其中在步骤(d)中,缺失所述供体序列,并且NS3的3'末端与NS4的5'末端连接,由此

在供体链中形成新的前间隔区/PAM组合或标记序列。

[0499] 16．项14或15的方法，其中步骤(a)中的第一链包含第一和第二同源臂(分别为HA1和HA2)，其中HA1与供体核酸链的NS1同源，并且HA2与供体核酸链的预定第二核苷酸序列(NS2)同源；其中间插序列在步骤(d)之前将NS1连接至供体链中的NS2，并且包含待缺失的序列，并且其中NS2包含NS4；任选地，其中所述缺失序列包含用于指示HA1和NS1之间以及HA2和NS2之间的同源重组的序列(标记序列)；其中在步骤(d)中通过HA1和NS1之间以及HA2和NS2之间的同源重组从供体链缺失所述缺失序列。

[0500] 17．任一前述项的方法，其包括在步骤(d)之前使用切割位点基序来引导供体链的切割(例如，Cas-、大范围核酸酶-、锌指蛋白- 或TALEN介导的切割)，其中步骤(d)用切割的第一和供体链进行。

[0501] 18．项17的方法，其中所述基序通过步骤(d)中的同源重组破坏(例如，破坏用于根据项17所述的切割的供体链中的前间隔区和/或PAM)。

[0502] 19.任一前述项的方法，其中步骤(a)中的第一链包含：-

[0503] (a) 第一同源臂(HA1)，其中HA1与供体核酸链的预定第一核苷酸序列(NS1)同源；和

[0504] (b) 由PS1包含的第一切割位点(CS1)，其中在步骤(c)中由第一Cas切割CS1；

[0505] (c) 其中(i)CS1在HA1的5’末端或侧接所述末端的5’侧；或(ii)其中CS1在HA1的3’末端或侧接3’末端的3’侧；且其中步骤(d)包括使切割的第一链的HA1经历与供体链的NS1的同源重组，其中将供体链的序列插入第一链和/或缺失供体序列。

[0506] 20．项19的方法，其中步骤(a)中的第一链包含

[0507] (c) 第二同源臂(HA2)，其中HA2与供体核酸链的预定第二核苷酸序列(NS2)同源；和

[0508] (d) 连接所述同源臂的间插序列，且

[0509] 其中

[0510] (e) 所述间插序列包含缺失序列，任选地，其中所述缺失序列包含用于指示HA1和NS1之间以及HA2和NS2之间的同源重组的序列(标记序列)；其中在步骤(c)或(d)中通过HA1和NS1之间以及HA2和NS2之间的同源重组从第一链缺失所述缺失序列；或

[0511] (f) 所述间插序列包含CS1(例如，其中CS1远离HA1和/或HA2间隔不超过250nt)且任选地NS1和NS2侧接从供体链缺失的序列的任一侧。

[0512] 20．项19或20的方法，其中步骤(a)中的第一链包含：-

[0513] (e) 第二同源臂(HA2)，其中HA2与供体核酸链的第二预定核苷酸序列(NS2)同源；和

[0514] (f) 第二前间隔区(PS2)和同源PAM(P2)，其在步骤(c)中由第一或不同(第二)Cas核酸酶在第二切割位点(CS2)处切割；其中(i)CS1在HA1的5’末端或侧接所述末端的5’侧，且CS2在HA2的3’末端或侧接HA2的所述末端的3’侧；或(ii)其中CS1在HA1的3’末端或侧接HA1的3’末端的3’侧，且CS2在HA2的5’末端或侧接HA2的5’末端的5’末端；且其中步骤(d)包括允许切割的第一链的HA1经历与供体链的NS1的同源重组且允许NS2经历与NS2的同源重组，其中将供体链的序列插入第一链。

[0515] 22．任一前述项的方法，其中步骤(a)的第一链包含第二前间隔区(PS2)和同源

PAM (P2)，其中在步骤(c)中通过第一或不同(第二)Cas核酸酶切割PS2，其中PS1和PS2的切割切除PS1和PS2之间的第一链序列。

[0516] 23．项22的方法，其中所述切除的序列使得第一链标记序列在所述细胞中无功能，其中所述方法任选地包括检测步骤(d)之后的标记物的不存在或减少。

[0517] 24．项21至23中任一项的方法，其中PS1和PS2的切割产生切割的第一链末端，其与供体核酸重组，其中将供体核酸的序列插入末端之间的第一链中。

[0518] 25．项20至24中任一项的方法，其中PS2和/或P2不通过插入供体序列在修饰的第一链中再形成。

[0519] 26.任一前述项的方法，所述方法包括分离或复制(例如，使用PCR)修饰的第一链或修饰的第一链中的供体核酸序列。

[0520] 27.任一前述项的方法，所述方法包括分离或复制(例如，使用PCR)修饰的供体链或其中发生缺失的供体核酸序列的区域。

[0521] 28．任一前述项的方法，其中将供体序列在步骤(d)中插入第一链中，并且插入的序列包含调节元件、蛋白编码序列、外显子或标记序列或由调节元件、蛋白编码序列、外显子或标记序列组成。

[0522] 29．任一前述项的方法，其中将供体序列在步骤(d)中插入第一链中，并且插入的序列对于从修饰的第一链表达是有功能的。

[0523] 30．项29的方法，其中插入的序列是蛋白编码序列、外显子或标记序列，其与第一链调节元件以功能性排列插入，由此所述调节元件能够控制来自修饰的第一链的供体序列的表达。

[0524] 31．任一前述项的方法，其中将供体序列在步骤(d)中插入第一链中，并且插入的序列是标记序列，并且所述方法包括检测修饰的第一链中的标记序列；或者检测表达的标记物，其中在步骤(d)之前不能检测到所述标记物。

[0525] 32.根据前述项的方法，所述方法包括在步骤(d)之前使用在待插入第一链的所述序列的一个或两个末端或侧接所述末端的供体核酸链的Cas切割，其任选地使用第一或第二Cas核酸酶。

[0526] 例如，切割在两个末端，侧接两个末端，在5’末端且侧接3’末端，侧接5’末端且在3’末端，在5’末端，在3’末端。

[0527] 33．任一前述项的方法，其中所述供体链由宿主细胞染色体或附加体核酸包含。

[0528] 34．根据项33所述的方法，其中在步骤(d)中从宿主染色体或附加体(例如质粒)缺失所述供体核酸序列。

[0529] 35.项34的方法，其中从染色体或附加体的缺失对宿主细胞是致命的。

[0530] 36.项33至35中任一项的方法，其中所述方法下调细胞活力、生长或增殖(例如，杀死细胞)。

[0531] 37.项33至35中任一项的方法，其中所述方法上调细胞活力、生长或增殖。

[0532] 38．任一前述项的方法，所述方法包括在步骤(d)之后分离修饰的细胞、修饰的第一链、修饰的供体链和/或修饰的宿主染色体或附加体核酸，其中所述宿主核酸序列已通过所述方法缺失。

[0533] 39.任一前述项的方法，其中所述方法在细菌或古细菌细胞(例如大肠杆菌细胞)

中进行。

[0534] 40. 项39的方法,其中所述细胞是人病原体物种的。

[0535] 41. 项39或40的方法,其中所述细胞是人微生物组物种(例如人肠道微生物组物种)的细胞。

[0536] 42. 任一前述项的方法,其中所述方法是重组工程改造方法。

[0537] 43.任一前述项的方法,其中所述细胞包含内源CRISPR/Cas系统,并且其中内源宿主Cas用于在步骤(c)中切割。

[0538] 44. 任一前述项的方法,其中所述方法在不将外源Cas或编码外源Cas的序列引入细胞的情况下进行。

[0539] 45. 任一前述项的方法,其中所述方法在细胞中进行以杀死原核细胞群体,其中所述细胞是相同物种或菌株的。

[0540] 46.项45的方法,其中所述群体由原核(例如,细菌)细胞的混合群体构成,并且所述方法选择性杀死混合群体中的一些、但不是所有细胞,其中杀死的细胞是其中引入第一链并通过Cas切割PS1的细胞。

[0541] 47. 项45或46的方法,其中PS1和P1与被杀死的细胞的Cas核酸酶(例如,内源性Cas)同源,但不与未被杀死的群体中的细胞同源,其中PS1在杀死的细胞中被选择性切割。

[0542] 48.项45至47中任一项的方法,其包括分离产物群体或所述产物的样品。

[0543] 49.项45至48中任一项的方法,其中所述细胞或群体由以下包含:食品、食品成分或前体成分;饮料;水(例如,意欲用于人食用);工业或环境物质(例如,油、石油产品、土壤或水路或水库;或用于回收或处理油、石油产品、土壤、水、食品、食品成分或前体、或饮料或前体的饮料成分的设备)。

[0544] 50.任一前述项的方法,其中所述切割包括将第一链暴露于指导RNA (gRNA1,例如单一gRNA),所述指导RNA (gRNA1,例如单一gRNA)与细胞中的PS1杂交以指导Cas切割PS1;任选地,其中所述步骤包括从在步骤(b)中引入(例如,作为第一链的一部分)的核苷酸序列表达其gRNA1或crRNA序列。

[0545] 51.任一前述项的方法,其中第一链或每条链是dsDNA。

[0546] 52. 任一前述项的方法,其中第一链或每条链是ssDNA。

[0547] 53.任一前述项的方法,其中所述第一链和/或供体链由BAC、PAC、YAC、粘粒或质粒包含。

[0548] 54. 任一前述项的方法,其中所述第一链在切割之前是环状的。

[0549] 在本文的任何配置中,所述第一链或载体(例如,第一链或供体或另外的链载体)是线性DNA。

[0550] 55.项1至52和54中任一项的方法,其中所述细胞是细菌细胞,且所述第一链是噬菌体核酸链,其中通过用噬菌体感染宿主细胞将噬菌体链引入宿主细胞。

[0551] 56.项19或从属于项19的任一前述项的方法,其中HA1远离CS1间隔不超过250nt(或本文公开的任何其他间隔)。

[0552] 57.项14至16和19中任一项或从属于项14至16和19中任一项的任一前述项的方法,或权利要求56的方法,其中HA1包含至少15个连续核苷酸(或本文公开的任何其他长度)。

[0553] 58. 项14至16和19中任一项或从属于项14至16和19中任一项的任一前述项的方法，或项56或57的方法，其中HA1的序列与NS1同源(或本文公开的任何其他％同一性)。

[0554] 59.项21或从属于项16或20的任一前述项的方法，或项56、57或58的方法，其中HA2远离CS2间隔不超过250nt(或本文公开的任何其他距离)。

[0555] 60.项16和20至22中任一项或从属于项16和20至22中任一项的任一前述项的方法，或项56至59中任一项的方法，其中HA2包含至少15个连续核苷酸(或本文公开的任何其他长度)。

[0556] 61.项16和20至22中任一项或从属于项16和20至22中任一项的任一前述项的方法，或项56至60中任一项的方法，其中HA2的序列与NS2同源(或本文公开的任何其他％同一性)。

[0557] 62. 制备适于与第二核酸链同源重组的第一核酸链(例如，由回收载体包含)的方法，例如，在根据任一前述项的方法中，所述制备方法包括在核酸链中组合

[0558] (a) 插入序列(IS)，其包含第一同源臂(HA1)(或由第一同源臂(HA1)组成)，所述第一同源臂(HA1)与第二链的第一预定靶核苷酸序列(NS1)互补，用于通过同源重组将IS插入第二链中；

[0559] (b) 与PAM (P1)相邻的第一前间隔区序列(PS1)，其中P1与第一Cas核酸酶同源，所述第一Cas核酸酶用于在第一切割位点(CS1)处切割PS1；其中CS1在HA1的末端或侧接HA1，其中当CS1被Cas切割时，HA1能够与NS1同源重组用于所述IS的插入。

[0560] 在本文的任何配置中，例如分离一条或每条链或载体。

[0561] 63.项62的方法，所述方法包括在核酸链中进一步组合

[0562] (c) 与第二链的第二预定序列(NS2)互补的第二同源臂(HA2)；和

[0563] (d) 任选地，与第二PAM (P2)相邻的第二前间隔区序列(PS2)，其中P2与Cas核酸酶同源，所述Cas核酸酶用于在第二切割位点(CS2)处切割PS2，其中当CS2被Cas切割时，HA2能够与NS2同源重组用于IS的插入。

[0564] 64.项63的方法，其包括用与P2同源的Cas核酸酶切割CS2。

[0565] 65.项64的方法，所述方法包括用相同的Cas核酸酶切割CS1和CS2。

[0566] 66.项62至65中任一项的方法，其中NS1和NS2侧接细胞的必需基因、抗生素抗性基因或毒力基因的序列。

[0567] 67. 项66的方法，其进一步产生抗生素组合物或试剂盒，其包含所述方法的产物与抗生素的组合，其中NS1和NS2侧接抗生素抗性基因的序列，其中所述抗生素是与核酸组合的抗生素。

[0568] 68.项62至67中任一项的方法，其中CS1远离HA1和/或HA2间隔不超过250nt。

[0569] 69.项62至68中任一项的方法，其中所述链如项51至61中任一项中所定义。

[0570] 70. 项62至69中任一项的方法，所述方法包括用第一Cas(例如，Cas9，任选地化脓链球菌或金黄色葡萄球菌Cas9)切割CS1。

[0571] 71.项52至70中任一项的方法，其进一步在第一链中(或在除了第一链以外的另外的核酸链上)包括编码gRNA (gRNA1，例如单一gRNA)的核苷酸序列，其与PS1互补，或所述序列编码在细胞中形成gRNA1的crRNA，其任选地具有由宿主细胞序列编码的tracrRNA。

[0572] 在一个实例中，所述tracrRNA由载体序列、例如由第一载体编码。

[0573] 72.项1至71中任一项的方法,其中所述方法的第一链产物在载体(例如,任选地超螺旋的闭合的环状载体)中。

[0574] 73.可通过项62至72中任一项的方法获得的工程改造的第一核酸链,所述链包含

[0575] (a) 第一同源臂(HA1),其中HA1与原核细胞的供体核酸链的第一预定核苷酸序列(NS1)同源;

[0576] (b) 与PAM (P1)相邻的第一前间隔区序列(PS1),其中P1与第一Cas核酸酶同源,所述第一Cas核酸酶用于在第一切割位点(CS1)处切割PS1;

[0577] (g) 第二同源臂(HA2),其中HA2与供体核酸链的第二预定核苷酸序列(NS2)同源;和

[0578] (c) 连接所述同源臂的间插序列,其中所述间插序列包含CS1;或其中CS1在所述间插序列之外且在HA1的末端处或侧接HA1的末端。

[0579] 74.项73的核酸链,其用于实施项16、20或21的方法,其中NS1和NS2侧接细胞的必需基因、抗生素抗性基因或毒力基因的序列。

[0580] 75.项74的核酸链,其与抗生素组合,其中NS1和NS2侧接抗生素抗性基因的序列,其中所述抗生素是与核酸组合的抗生素。

[0581] 76.项73、74或75的核酸链,其中CS1远离HA1和/或HA2间隔不超过250nt。

[0582] 77. 项73至76中任一项的第一核酸链,其任选地与另外的核酸链组合,其中所述第一链或另外的链包含编码gRNA(gRNA1,例如单一gRNA)的可表达的核苷酸序列,其与PS1互补;或所述序列编码在细胞中形成gRNA1的可表达的crRNA,其任选地具有由宿主细胞序列、第一链或另外的链编码的tracrRNA。

[0583] 在一个实例中,所述tracrRNA由载体序列、例如由第一载体编码。

[0584] 78.项73至77中任一项的第一核酸链,其中CS1在所述间插序列之外且在HA1的末端处或侧接HA1的末端;所述链包含第二前间隔区(PS2)和同源PAM (P2),其用于在步骤(c)中由第一或不同(第二)Cas核酸酶在第二切割位点(CS2)处切割PS2;其中CS2在所述间插序列之外且在HA2的末端处或侧接HA2的末端。

[0585] 79. 项78的第一核酸链,其任选地与所述或另一另外的核酸链组合,其中所述第一链或另外的链包含编码gRNA(gRNA2,例如单一gRNA)的可表达的核苷酸序列,其与PS2互补;或所述序列编码在细胞中形成gRNA2的可表达的crRNA,其任选地具有由宿主细胞序列、第一链或另外的链编码的tracrRNA。

[0586] 在一个实例中,所述tracrRNA由载体序列、例如由第一载体编码。

[0587] 80. 从属于项77的项78的第一核酸链,其中gRNA1-和gRNA2-编码序列(或相应的crRNA-编码序列)由相同的核酸链、任选地第一链包含。

[0588] 81.项77至80中任一项的第一核酸链,其与编码所述第一Cas(例如,Cas9)的可表达的核苷酸序列组合,其中第一Cas核苷酸序列由第一链、所述另外的核酸链或另一另外的核酸链包含;任选地,其中P2与第一Cas同源。

[0589] 82.项73至81中任一项的核酸链,其中所述链如项51至61中任一项中所定义。

[0590] 83.抗微生物组合物,其包含项73至82中任一项的核酸链。

[0591] 84. 包含项83的组合物的食品、饮料或工业流体(例如,储存的水)。

[0592] 85.包含一种或多种核酸载体的组合物,第一所述载体包含如项73至84中任一项

中所定义的第一核酸链,其中所述另外的链由所述组合物包含,并且每条另外的链由与第一载体不同的载体包含。

[0593] 86. 项85的组合物,其包含含有第一链的第一载体、含有编码所述第一Cas的可表达的核苷酸序列的第二载体和任选第三载体,其中所述第一、第二或第三载体包含项77中所述的所述编码gRNA1的核苷酸序列;任选地,所述第一、第二或第三载体包含项79中所述的所述编码gRNA2的核苷酸序列。

[0594] 在本文的任何配置中,任选地所述或每个Cas是切口酶,或Cas核酸酶之一(例如,第一Cas)是切口酶。

[0595] 87.项85或86的组合物,其中第一载体或各载体是dsDNA载体。

[0596] 88. 项85或86的组合物,其中第一载体或各载体是ssDNA载体。

[0597] 89.项85至88中任一项的组合物,其中第一载体或各载体是BAC、PAC、YAC或质粒,例如,所有载体都是质粒或第一载体是BAC,且其他载体是质粒。

[0598] 90. 项85至89中任一项的组合物,其中第一载体或各载体在切割之前是环状的。

[0599] 91. 第一载体或载体组合,其中

[0600] (a) 所述第一载体包含工程改造的第一核酸链,所述链包含与PAM (P1)相邻的第一前间隔区序列(PS1),其中P1与第一Cas核酸酶同源,所述第一Cas核酸酶用于在第一切割位点(CS1)处切割PS1;和

[0601] (b) 第一载体或所述组合的另一载体包含编码gRNA(gRNA1,例如单一gRNA)的可表达的核苷酸序列,其与PS1互补,用于指导Cas以切割PS1;或所述可表达的序列编码与tracrRNA形成gRNA1的crRNA。

[0602] 92.项91的载体或组合,其与编码所述第一Cas(例如,Cas9)的可表达的核苷酸序列组合,其中第一Cas核苷酸序列由第一载体包含,且所述(b)的可表达的序列由第一载体或所述组合的另一载体包含。

[0603] 93.项92的载体或组合,其中CS1由编码所述第一Cas的可表达的核苷酸序列或其调节元件的序列包含,其中CS1能够被第一Cas切割以下调或失活第一Cas从其可表达的核苷酸序列的表达。

[0604] 94.项91的载体或组合,其与编码所述第一Cas(例如,Cas9)的可表达的核苷酸序列组合,其中第一Cas核苷酸序列由与第一载体组合的第二载体包含;且所述(b)的可表达的序列由第一载体或所述组合的另一载体(例如,所述第二载体)包含。

[0605] 95.项91至94中任一项的载体或组合,其中所述第一链包含第二前间隔区(PS2)和同源PAM (P2),其用于由第一或不同(第二)Cas核酸酶在第二切割位点(CS2)处切割PS2。

[0606] 96. 项95的载体或载体组合,其中当CS1和CS2被切割时,P1和P2通过可从载体切除的间插序列(IS)连接,其中IS包含调节元件、蛋白编码序列、外显子或标记序列。

[0607] 97.包含项96的载体或组合和第一和第二PCR引物的试剂盒,其中第一引物与IS的核苷酸序列互补,且第二引物与IS之外的核苷酸序列互补,其中所述引物能够检测第一载体中标记物的存在。

[0608] 98.项95的载体或组合,其中当CS1和CS2被切割时,P1和P2通过可从载体切除的间插序列(IS)连接,其中IS (i)包含P1和PS2序列,或(ii)包含P2和PS1序列,其中可通过序列(i)或(ii)的不存在检测间插序列的切除。

[0609] 99. 项96至98中任一项的载体或组合,其中第一载体包含分别侧接CS1和CS2(例如,紧邻CS1和CS2的5'和3'侧)的第一和第二同源臂(HA1和HA2),用于与预定靶核苷酸序列(分别为NS1和NS2)同源重组。

[0610] 100.项99的载体或组合,其中NS1和NS2是原核细胞基因组的序列,例如,是必需基因、毒力基因或抗生素抗性基因的序列(例如,在外显子或调节元件中或侧接外显子或调节元件)。

[0611] 101.项91至100中任一项的载体或组合,其用于治疗、预防或诊断人或非人动物中的疾病或病况,例如,当所述方法是在人或动物细胞中进行时,用于人或动物、人或动物细胞的基因疗法;或者,当所述方法在细菌细胞中进行时,用于治疗或预防人或动物中的细菌感染。

[0612] 本发明还提供以下段落:-

[0613] 1. 抑制原核宿主细胞的活力、生长或增殖的方法,所述方法包括

[0614] (a) 提供第一载体,其包含与PAM (P1)相邻的前间隔区序列(PS1),所述PAM (P1)与第一Cas核酸酶同源;

[0615] (b) 提供毒素或毒素前体;或包含编码毒素或毒素前体的序列的载体,其中所述毒素对原核细胞有毒;

[0616] (c) 将细胞与载体组合,其中将载体引入细胞,并在细胞中提供毒素;和

[0617] (d) 使用Cas核酸酶切割第一载体的PS1,其使用第一Cas核酸酶,其中切割促进细胞中的毒素活性,由此杀死细胞或减少细胞生长或增殖。

[0618] 2. 段落1的方法,其中所述切割切割抑制毒素表达或毒素活性的载体序列。

[0619] 3. 段落1或2的方法,其中所述切割切割毒素表达或毒素活性的阻遏物的表达或活性所需的载体序列。

[0620] 4.任一前述段落的方法,其中所述切割将毒素前体序列转化为编码毒素的序列,用于在细胞中表达毒素;任选地,其中PS1由编码毒素或前体的序列包含,或PS1由其调节元件(例如,启动子或阻遏物序列)包含。

[0621] 5. 任一前述段落的方法,其中所述切割引起毒素的活化物的表达或活性增加或调节编码毒素或前体的序列的启动子的活性增加。

[0622] 在一个实例中,所述活化物由第一载体编码。

[0623] 6.任一前述段落的方法,其中所述切割促进细胞中的毒素表达。

[0624] 7. 段落1、2或5的方法,其中第一载体包含编码抗毒素的核酸序列(AT),其中步骤(d)包括使用Cas核酸酶切割细胞中的第一载体以减少或消除细胞中的抗毒素,其中所述切割促进细胞中的毒素活性,由此杀死细胞。

[0625] 8.段落7的方法,其中所述切割使从编码抗毒素的核酸序列的表达失活或降低。

[0626] 9. 段落7或8的方法,其中所述第一载体的AT序列包含第一Cas切割位点或在一侧侧接第一Cas切割位点;且所述序列(AT)包含第二Cas切割位点或在另一侧侧接第二Cas切割位点,每个位点在相应的前间隔区中,所述前间隔区与相应的PAM的5'相邻,其中前间隔区和相应的PAM之一是PS1和P1,其中在步骤(d)中使用Cas切割来切割第一和第二位点,由此从第一载体切除抗毒素序列(或启动子序列),并减少或消除来自第一载体的抗毒素表达。

[0627] 10.段落1的方法,其中步骤(c)的产物是包含编码抗毒素的序列的细胞,其中第一载体包含抗-抗毒素剂的核酸序列(AAT),其中步骤(d)包括使用Cas核酸酶切割细胞中的第一载体以增加试剂或增加试剂在细胞中的抗-抗毒素活性,由此减少或消除细胞中的抗毒素表达或活性,其中所述切割促进细胞中的毒素活性,由此杀死细胞。

[0628] 11. 段落10的方法,其中所述抗毒素由细胞编码,例如由宿主细胞染色体或宿主附加体(例如质粒)核苷酸序列编码。

[0629] 12.段落10的方法,其中所述抗毒素由步骤(c)的载体编码。

[0630] 13.段落10至12中任一项的方法,其中所述试剂是核酸,例如DNA。

[0631] 在一个实例中,所述试剂(例如,DNA试剂片段)与抗毒素基因(例如,其外显子或调节元件)杂交以抑制或减少抗毒素表达。例如,所述DNA片段包含与抗毒素基因中(例如,在其调节元件或外显子中)的序列同源的一个或多个区域,由此将DNA片段的一个或多个序列插入抗毒素基因中,由此使抗毒素基因表达(插入可以有或没有基因序列的缺失)失活或减少。或者,所述试剂与编码抗毒素或其亚基的mRNA杂交。在一个实例中,所述片段编码指导RNA,其指导抗毒素基因(例如,其外显子或调节元件)的核酸酶切割(例如,Cas切割)以抑制或降低抗毒素表达。所述gRNA可以与宿主的内源性Cas同源,或者可以与载体编码的Cas(例如,Cas9)同源。在一个实例中,在DNA片段序列插入抗毒素基因中的情况下,插入的序列可以编码对细胞有毒的毒素或可以编码用于靶向至细胞中的抗毒素基因或任何其他核苷酸序列的Cas的指导RNA,例如宿主染色体或宿主附加体序列,例如宿主必需、毒力或抗生素抗性基因的序列。以这种方式,使抗毒素(例如抗生素)基因失活或通过插入DNA片段序列来降低活性,并且此外,插入的序列(例如,包括在细胞中发挥功能的启动子;或当序列与该内源启动子以功能性排列插入时,使用抗毒素基因的内源启动子)可以表达功能性指导RNA(例如,单一gRNA),其可以用于修饰细胞中的另外的位点(例如,其切割辅助细胞杀死的位点)。

[0632] 在本发明的这个和其它配置中,通常提供了一种方法,所述方法包括Cas切割已引入细胞(例如,原核或真核细胞)中的载体以从载体释放DNA片段。在一个实例中,通过同源重组将片段核苷酸序列插入宿主染色体或附加体核酸(例如DNA)(或引入载体中)。在一个实例中,所述插入下调(例如,失活)或上调(例如,活化)宿主核酸的基因。此外或可替代地,插入的序列可以表达所需的RNA或编码所需的蛋白(例如,标记物,其指示成功产生DNA片段并插入宿主核酸中,由此允许选择使用标记物检测的一种或多种宿主细胞的方法中的另外的步骤)。在另一个实施方案中,插入的序列编码细胞中表达的指导RNA(例如,单一指导RNA),用于靶向宿主染色体或附加体核酸中的相应位置(或者可替代地用于靶向引入的载体中的序列,例如,修饰所述载体携带的基因,例如上调或下调的选择标记物,用于提供通过Cas切割在细胞中成功产生DNA片段的指示)。因此,本发明的主题利用细胞中的Cas切割以释放DNA用于同源插入靶序列中,其中重要的是该插入上调或下调其靶序列,例如,以提供成功的Cas切割和插入靶标中或用于将具有敲入效用的序列并入靶标中用于产生所需的表达产物(蛋白或RNA,例如gRNA)的标记物。该表达的gRNA可用于细胞中的另外的宿主或载体修饰,例如用于指导细胞中Cas介导的(宿主或载体序列的)切割、活化或失活。在DNA片段包含在内源调节元件的影响下靶向插入序列的同源臂的情况下,这对于利用内源调节控制(例如,启动子控制或增强子控制)来控制插入序列的表达,例如,对于提供针对细胞(例如,生长和复制的细胞周期)调整的插入序列的时间表达,是有用的。

[0633] 14. 段落10至12中任一项的方法,其中所述试剂是蛋白。

[0634] 15.段落13的方法,其中第一载体是DNA载体且第一载体的AAT序列包含第一Cas切割位点或在一侧侧接第一Cas切割位点;其中所述序列(AAT)包含第二Cas切割位点或在另一侧侧接第二Cas切割位点,每个位点在相应的前间隔区中,所述前间隔区与相应的PAM的5'相邻,其中前间隔区和相应的PAM之一是PS1和P1,其中在步骤(d)中使用Cas切割来切割第一和第二位点,由此从第一载体释放抗-抗毒素剂,并通过将抗-抗毒素序列插入宿主细胞的染色体或附加体核酸(例如通过同源重组)来减少或消除细胞中的抗毒素表达。

[0635] 16. 段落9或15的方法,其中第一和第二位点被相同的Cas核酸酶种类(例如内源性Cas或由载体序列编码的Cas)切割。

[0636] 17. 段落9或15的方法,其中第一和第二位点被不同的Cas核酸酶种类切割(例如,其中两种Cas对细胞是内源的;一种是内源的,且另一种由载体序列编码;或者两者都由载体序列编码)。

[0637] 18.段落7至17中任一项的方法,其包括允许切割的第一载体序列(例如,抗毒素或AAT序列)与宿主细胞核酸重组以产生重组产物,其中从载体确实抗毒素核苷酸序列。

[0638] 19.段落18的方法,其中将宿主细胞序列插入载体中,例如,以替换抗毒素或AAT序列。

[0639] 20. 段落13或从属于段落13的段落15至19中任一项的方法,其包括允许DNA试剂与宿主细胞核酸重组以产生重组产物,其中将试剂序列插入抗毒素核苷酸序列或其调节元件中,由此减少或消除抗毒素表达或活性。

[0640] 21.段落20的方法,其中抗毒素核苷酸序列或调节元件序列从宿主细胞核酸缺失。

[0641] 22.任一前述段落的方法,其中第一Cas由细胞中的与PS1杂交的第一指导RNA(gRNA1,例如单一指导RNA)指导,由此Cas切割PS1。

[0642] 在一个实例中,编码gRNA1的序列由引入细胞的载体(例如,第一载体)包含。

[0643] 23.段落22的方法,其中所述方法包括在步骤(d)之前在细胞中提供gRNA1或在细胞中产生gRNA1的序列,任选地,其中所述序列由在步骤(c)中引入细胞的第一或另一载体包含。

[0644] 24.任一前述段落的方法,其中第一或各Cas核酸酶对细胞是内源的。

[0645] 25.段落1至23中任一项的方法,其中在步骤(d)之前将所述或各Cas核酸酶引入细胞中或将编码所述Cas的相应核苷酸序列(例如,由第一载体或另一载体包含)引入细胞中。

[0646] 26. 任一前述段落的方法,其中第一或各Cas是Cas9(例如,化酿脓链球菌或金黄色葡萄球菌Cas9)。

[0647] 27. 段落1至25中任一项的方法,其中第一或各Cas是大肠杆菌Cas。

[0648] 28.任一前述段落的方法,其中所述毒素是抗生素,例如抗生素药物。

[0649] 29.任一前述段落的方法,其中所述毒素和抗毒素是原核毒素/抗毒素模块(例如,I型或II型模块或质粒成瘾模块)的组分。

[0650] 例如,所述模块是I型模块,例如Hok-Sok模块。对于实施方案,所述模块是II型模块,例如HiCa-HicB模块。例如,所述模块是tad-ata型毒素-抗毒素模块。例如,所述模块是质粒成瘾模块。

[0651] 30. 转化宿主细胞的方法,所述方法包括

[0652] (a) 提供第一核酸载体,所述载体处于第一状态,其包含

[0653] i 与PAM (P1)相邻的前间隔区序列(PS1),所述PAM (P1)与第一Cas核酸酶同源;和

[0654] ii 第一核苷酸序列,其包含用于标记载体的第一状态的第一标记序列;

[0655] (b) 将细胞与载体组合,其中将载体引入细胞;和

[0656] (c) 使用Cas核酸酶切割第一载体的PS1,其使用第一Cas核酸酶,其中切割将载体转化为第二状态,其中使得标记序列无功能;和

[0657] (d) 任选地通过检测标记物在细胞中无功能来检测第二状态。

[0658] 31.段落30的方法,其中通过允许切割载体经历与细胞的染色体或附加型核酸的同源重组来使标记序列无功能。

[0659] 32.段落31的方法,其中将载体序列插入宿主核酸中。

[0660] 33.段落31或32的方法,其中将宿主核酸的序列插入载体中。

[0661] 34.段落33的方法,其中通过插入宿主序列在载体中破坏PS1和/或P1。

[0662] 35. 任一前述段落的方法,其中通过切割在载体中破坏PS1和/或P1。

[0663] 36. 段落30至35中任一项的方法,其中所述切割促进细胞中的毒素的细胞毒性活性,由此杀死细胞或减少细胞生长或增殖。

[0664] 37.段落36的方法,其中所述毒素是抗生素,例如抗生素药物。

[0665] 38.段落30至37中任一项的方法,其中所述切割从载体切除标记序列,由此使标记物无功能。

[0666] 39. 任一前述段落的方法,其中PS1包含由第一Cas切割的靶位点,且其中靶位点在标记序列的5’末端或侧接标记序列的5’侧。

[0667] 40.任一前述段落的方法,其中所述载体包含第二前间隔区(PS2)和同源PAM(P2),其在步骤(c)中由第一或不同(第二)Cas核酸酶切割,其中PS1和PS2的切割使得标记物在细胞中无功能。

[0668] 41. 段落40的方法,其中PS2包含被切割的靶位点,且其中靶位点在标记序列的3’末端或侧接标记序列的3’侧。

[0669] 从属于段落39的段落41的方法,其中所述标记序列从载体切除,且所述载体经历通过宿主核酸的序列同源重组插入载体的缺口修复。

[0670] 43.段落42的方法,其中PS1和/或P1被破坏。

[0671] 44.段落42或43的方法,其中PS2和/或P2被破坏。

[0672] 45.根据段落33中任一项、从属于段落33的段落33至44中任一项或根据段落42至44中任一项所述的方法,其中插入的宿主核酸序列是调节元件、蛋白编码序列、外显子或第二标记序列。

[0673] 46.段落45的方法,其中插入的序列对于从载体表达是有功能的。

[0674] 47. 段落46的方法,其中插入的序列是蛋白编码序列、外显子或标记序列,其与载体调节元件以功能性排列插入,由此所述调节元件控制宿主序列从载体的表达。

[0675] 48. 段落45至47中任一项的方法,其中所述宿主序列是第二标记序列,且所述方法包括检测载体中的第二标记序列;或检测第二标记物,其中第二标记物在步骤(c)之前在细胞中是检测不到的,且所述第二标记序列在步骤(c)之前在第一载体中是检测不到的。

[0676] 49.根据段落33中任一项、从属于段落34的段落33至48中任一项或根据段落42至48中任一项所述的方法，所述方法包括在所述宿主核酸的一个或两个末端或侧接所述宿主核酸的一个或两个末端使用Cas切割宿主。

[0677] 50. 根据段落33中任一项、从属于段落33的段落34至49中任一项或根据段落42至49中任一项所述的方法，其中从宿主染色体或附加体(例如质粒)缺失所述宿主核酸序列；任选地其中从染色体或附加体的缺失对细胞是致命的。

[0678] 51.任一前述段落的方法，其中Cas切割宿主核酸或将宿主核酸插入载体或第一链中下调细胞活力、生长或增殖(例如，杀死细胞)。

[0679] 52.任一前述段落的方法，其中Cas切割宿主核酸或将宿主核酸插入载体或第一链中上调细胞活力、生长或增殖。

[0680] 53.任一前述段落的方法，其中，其中所述方法在宿主细胞中进行，所述方法包括分离修饰的细胞、修饰的第一载体或第一链和/或修饰的宿主染色体或附加体核酸，其中所述宿主核酸序列已通过所述方法缺失。

[0681] 54.任一前述段落的方法，其中所述方法在细菌或古细菌细胞(例如大肠杆菌细胞)中进行。

[0682] 55.段落54的方法，其中所述细胞是人病原体物种的。

[0683] 56.段落54或55的方法，其中所述细胞是人微生物组物种(例如人肠道微生物组物种)的细胞。

[0684] 57.任一前述段落的方法，其用于人或动物医学治疗、预防或诊断。

[0685] 58.任一前述段落的方法，其中所述方法在体外进行。

[0686] 59. 段落58的方法，其中所述方法是重组工程改造方法。

[0687] 60.段落1至57中任一项的方法，其中所述方法在体内(例如，但不在人或人胚胎中)进行。

[0688] 61.段落1至57和60中任一项的方法，其中所述方法在人(例如，其中所述方法是化妆方法)、胚胎(例如，非人胚胎)、受精卵(例如，非人受精卵)、生殖细胞(例如，人或非人生殖细胞)或非人动物中进行。

[0689] 62.任一前述段落的方法，其中所述方法在不将外源Cas或编码外源Cas的序列引入细胞的情况下在一种或多种细胞中进行。

[0690] 63.任一前述段落的方法，其中所述方法在细胞中进行，且PS1和P1与用于进行步骤的切割的细胞内源的Cas同源。

[0691] 64.任一前述段落的方法，其中所述方法在细胞中进行以杀死原核细胞群体，其中所述细胞是相同物种或菌株的。

[0692] 65.段落64的方法，其中所述群体由以下包含：食品、食品成分或前体成分；饮料；水(例如，意欲用于人食用)；工业或环境物质(例如，油、石油产品、土壤或水路或水库；或用于回收或处理油、石油产品、土壤、水、食品、食品成分或前体、或饮料或前体的饮料成分的设备)。

[0693] 66.段落64或65的方法，其中所述群体由原核(例如，细菌)细胞的混合群体构成，且所述方法选择性杀死混合群体中的一些、但不是所有细胞，其中杀死的细胞是其中引入第一载体或第一链并通过Cas切割PS1的细胞。

[0694] 67.段落66的方法，其中所述第一载体或第一链包含编码与细胞中的PS1杂交的gRNA (gRNA1，例如单一gRNA)的序列或所述载体或链编码在细胞中形成gRNA1的crRNA，其任选地具有由宿主细胞序列编码的tracrRNA。

[0695] 68.段落66或67的方法，其中PS1和P1与被杀死的细胞的Cas核酸酶(例如，内源性Cas)同源，但不与未被杀死的群体中的细胞同源，其中PS1在杀死的细胞中被选择性切割。

[0696] 69.段落66至68中任一项的方法，其中混合群体是人或非人动物微生物组群体，例如肠道、阴道、腋窝或口腔微生物组群体。

[0697] 70.段落69的方法，其包括分离段落69的产物群体或所述产物的样品。

[0698] 71. 段落70的方法，其包括将所述样品或其后代样品移植至人体中用于治疗或预防疾病或病况(例如，肥胖症)。

[0699] 72.治疗或预防人或非人动物受试者中原核细胞(例如细菌)感染的方法，所述方法包括进行任一前述段落的方法以杀死受试者中的细菌细胞，由此治疗或预防感染。

[0700] 73. 任一前述段落的方法，其中第一链或第一或各载体是dsDNA载体。

[0701] 74.任一前述段落的方法，其中第一链或第一或各载体是ssDNA载体。

[0702] 75. 任一前述段落的方法，其中第一链或第一载体在切割之前是环状的。

[0703] 76.任一前述段落的方法，其中第一链或第一载体是BAC、PAC、YAC或质粒。

[0704] 77. 段落1至75中任一项的方法，其中所述方法在细菌宿主细胞中进行，且所述第一载体是噬菌体(或第一链由噬菌体包含)，其能够感染宿主细胞用于引入第一载体或第一链，其中所述载体或链通过用噬菌体感染细胞来引入细胞。

[0705] 78.第一核酸链(例如，由载体、例如回收或插入载体包含) (任选地用于本文的任一方法)，其包含

[0706] (a) 插入序列(IS)，其紧邻或侧接第一同源臂(HA1)，所述第一同源臂(HA1)与第二链的预定靶核苷酸序列(NS1)互补，用于通过同源重组将IS插入第二链中；

[0707] (b) 与PAM (P1)相邻的第一前间隔区序列(PS1)，其中P1与第一Cas核酸酶同源，所述第一Cas核酸酶用于在第一切割位点(CS1)处切割PS1；

[0708] 其中CS1在HA1的末端或侧接HA1，

[0709] 其中当CS1被Cas切割时，HA1能够与NS1同源重组用于所述IS的插入。

[0710] 79. 段落78的核酸链，其中CS1在HA1的5'末端。

[0711] 80.段落79的核酸链，其中P1紧邻第一链中的n个核苷酸的种子序列(S1)的3'侧(其中n = 3-5个连续核苷酸，例如3个连续核苷酸)，其中S1将P1连接至CS1，且其中(i) NS1的n个核苷酸(例如，3个连续核苷酸)的最5'的序列不同于S1的序列和/或(ii)紧邻第二核酸链中的NS1的5'侧的13个连续核苷酸的序列不同于与紧邻第一核酸中CS1的5'侧的13个连续核苷酸的序列。

[0712] 在一个实例中，在(i)中，差异是1、2、3、4或5或所有所述连续核苷酸。例如，在NS1的第3个核苷酸处存在差异。例如，在NS1的第2和第3或第3和第4个核苷酸处存在差异，例如，当S1长度为3个核苷酸时。

[0713] 例如，在(ii)中，所述序列在序列上是相同的，除了在3、4、5、6、7、8、9、10、11或12个或更多个核苷酸位置处。

[0714] 81.段落80的核酸链，其中NS1的最5'的核苷酸不同于S1的最5'的核苷酸；任选地，

其中S1长度为3个连续核苷酸，且P1与Cas9(例如，化脓脓链球菌Cas9)同源。

[0715] 82. 段落78的核酸链，其中CS1侧接HA1的5'末端。

[0716] 83. 段落78的核酸链，其中CS1在HA1的3'末端。

[0717] 84. 段落83的核酸链，其中CS1紧邻第一链中的n个核苷酸的种子序列(S2)的5'侧(其中n = 3-5个连续核苷酸，例如3个连续核苷酸)，其中S2将P1连接至CS1，且其中(i) 紧邻NS1的3'侧的第二链的n个核苷酸(例如，3个连续核苷酸)的序列不同于S2的序列和/或(ii) HA1的13个连续核苷酸的最3'的序列不同于NS1的13个连续核苷酸的最3'的序列。

[0718] 85.段落84的核酸链，其中(i)的所述第二链序列的最5'的核苷酸不同于S2的最5'的核苷酸；任选地，其中S2长度为3个连续核苷酸，且P1与Cas9(例如，化脓脓链球菌Cas9)同源。

[0719] 86. 段落78的核酸链，其中CS1侧接HA1的3'末端。

[0720] 87. 段落78至86中任一项的核酸链，其中当通过HA1与NS1的重组将IS插入第二核酸链时，第一链不能重新产生PS1与P1的组合。

[0721] 88. 段落78至87中任一项的核酸链，其中当通过HA1与NS1的重组将IS插入第二核酸链时，第一链不能重新产生CS1。

[0722] 89.段落78至88中任一项的核酸链，其中IS包含供体序列，且HA1在IS的5'侧，其中当CS1被Cas切割时，HA1能够与NS1同源重组用于IS的插入。

[0723] 90. 段落78至88中任一项的核酸链，其中IS包含供体序列，且HA1在IS的3'侧，其中当CS1被Cas切割时，HA1能够与NS1同源重组用于IS的插入。

[0724] 91.段落89或90的核酸链，其中第一核酸包含另外的前间隔区和同源PAM，其中另外的前间隔区和PAM中的一个或两个由IS包含。

[0725] 92.段落91的核酸链，其中所述另外的PAM与Cas同源，所述Cas不同于与P1同源的Cas。

[0726] 93. 段落89至92中任一项的核酸，其中IS包含编码调节元件、蛋白编码序列、外显子或标记序列的序列。

[0727] 94.段落78至93中任一项的核酸链，其中IS包含

[0728] (e) 与第二链的第二预定序列(NS2)互补的第二同源臂(HA2)；

[0729] (f) 与第二PAM (P2)相邻的第二前间隔区序列(PS2)，其中P2与Cas核酸酶同源，所述Cas核酸酶用于在第二切割位点(CS2)处切割PS2，

[0730] 其中当CS2被Cas切割时，HA2能够与NS2同源重组用于IS的插入。

[0731] 95. 段落94的核酸链，其中P2与第一Cas同源。

[0732] 96. 段落94或95的核酸链，其中HA1和HA2能够在HA1和HA2之间没有任何序列的情况下且任选地在第一核酸链中不插入侧接HA1和HA2的任何序列的情况下一起插入。

[0733] 97.段落94、95或96的核酸链，其中NS1和NS2在第二核酸链中间隔开，其中HA1和HA2能够被插入第二链以缺失NS1和NS2之间的序列。

[0734] 98.段落94、95或96的核酸链，其中NS1和NS2在第二核酸链中不间隔开，其中HA1和HA2能够在不缺失NS1和NS2之间的序列的情况下被插入第二链。

[0735] 99.段落94至98中任一项的核酸链，其中IS包含HA1和HA2之间的供体序列。

[0736] 100. 段落99的核酸链，其中第一核酸包含另外的前间隔区和同源PAM，其中另外

的前间隔区和PAM中的一个或两个由IS包含。

[0737] 101.段落100的核酸链,其中所述另外的PAM与Cas同源,所述Cas不同于与P1同源的Cas。

[0738] 102. 段落100或101的核酸链,其中所述另外的PAM与Cas同源,所述Cas不同于与P2同源的Cas。

[0739] 103.段落94至102中任一项的核酸,其中IS包含编码调节元件、蛋白编码序列、外显子、预定限制性内切核酸酶位点、额外的Cas切割位点或标记序列的序列。

[0740] 104. 从属于段落79至82的段落94至103中任一项的核酸链,其中CS1紧邻第一核酸链中序列(TS1)的3'侧,其中P2紧邻第一链中种子序列(S2')的3'侧,其中S2'将P2连接至CS2,其中序列TS1-S2不是与P2同源的另外的前间隔区,其中TS1-S2能够在HA1与NS1重组和HA2与NS2重组以及从链中缺失CS1和CS2之间的序列后形成,由此TS1-S2不能被与P2同源的Cas切割。

[0741] 105. 从属于段落79至82的段落94至103中任一项的核酸链,其中CS1紧邻第一核酸链中序列(TS1)的3'侧,其中P2紧邻第一链中种子序列(S2')的3'侧,其中S2'将P2连接至CS2,其中序列TS1-S2是与P2同源的另外的前间隔区(PS3),其中PS3能够在HA1与NS1重组和HA2与NS2重组以及从链缺失CS1和CS2之间的序列后形成。

[0742] 106. 段落78至105中任一项的核酸链,其中P1紧邻种子序列(S1)的3'侧,其中S1将P1连接至CS1,其中IS能够插入第二核酸链中以形成与P1同源的前间隔区(PS4),其中PS4在其3'末端包含S1,且在其5'末端包含第二链核酸序列。

[0743] 107. 段落78至106中任一项的核酸链,其中所述链由能够转化细胞的多拷贝载体包含。

[0744] 108. 段落78至107中任一项的核酸链,其中所述链与指导RNA (gRNA1,例如单一指导RNA)或用于在体外或在细胞中产生gRNA1的可表达的核苷酸序列组合,其中gRNA1包含与PS1互补的核苷酸序列,用于指导Cas以切割PS1。

[0745] gRNA1与PS1、P1和第一Cas同源,其用于指导Cas至PS1用于切割PS1。

[0746] 109. 从属于段落94的段落78至108中任一项的核酸链,其中所述链与指导RNA(gRNA2,例如单一指导RNA)或用于在体外或在细胞中产生gRNA2的可表达的核苷酸序列组合,其中gRNA2包含与PS2互补的核苷酸序列,用于指导Cas以切割PS2。

[0747] gRNA1与PS2、P2和Cas同源,其用于指导Cas至PS2用于切割PS2。

[0748] 110. 从属于段落105的段落78至109中任一项的核酸链,其中所述链与指导RNA(gRNA3,例如单一指导RNA)或用于在体外或在细胞中产生gRNA3的可表达的核苷酸序列组合,其中gRNA3包含与PS3互补的核苷酸序列,用于指导Cas以切割PS3。

[0749] gRNA1与PS3、P3和Cas同源,其用于指导Cas至PS3用于切割PS3。

[0750] 111. 从属于段落130的段落78至110中任一项的核酸链,其中所述链与指导RNA(gRNA4,例如单一指导RNA)或用于在体外或在细胞中产生gRNA4的可表达的核苷酸序列组合,其中gRNA4包含与PS4互补的核苷酸序列,用于指导Cas以切割PS4。

[0751] gRNA1与PS4、P4和Cas同源,其用于指导Cas至PS4用于切割PS4。

[0752] 112.段落108至111中任一项的核酸链,其中用于产生gRNA的所述核酸链或各核酸链是编码crRNA的核苷酸序列,所述crRNA包含与各自PS互补的序列。

[0753] 113. 段落112的核酸链,其与tracrRNA或用于产生tracrRNA的可表达的核苷酸序列组合,其中所述tracrRNA包含与crRNA的序列互补的序列,且其中所述指导RNA包含所述crRNA和tracrRNA的序列。

[0754] 114. 段落78至113中任一项的核酸链,其中所述链与所述第二链组合。

[0755] 115.段落78至114中任一项的核酸链,其中所述第一链或所述组合在宿主细胞中。

[0756] 116.从属于段落112的段落115的核酸链,其中所述宿主细胞表达tracrRNA,其中所述tracrRNA包含与crRNA的序列互补的序列,且其中所述指导RNA包含所述crRNA和tracrRNA的序列。

[0757] 117.段落115或116的核酸链,其中所述细胞是细菌或古细菌细胞,例如大肠杆菌细胞。

[0758] 118.段落117的核酸链,其中所述细胞是人病原体物种的。

[0759] 119. 段落117或118的核酸链,其中所述细胞是人微生物组物种(例如人肠道微生物组物种)的细胞。

[0760] 120.段落78至119中任一项的核酸,其用于人或动物医学治疗、预防或诊断。

[0761] 121.段落78至120中任一项的核酸,其用于体外。

[0762] 122.段落78至121中任一项的核酸,其与重组工程改造试剂组合。

[0763] 123.段落78至120中任一项的核酸,其用于体内(例如,但不在人或人胚胎中)。

[0764] 124.段落78至123中任一项的核酸,其中第一核酸由第一载体包含。

[0765] 125.从属于段落108至113中任一项的段落78至124中任一项的核酸,其中第一核酸和编码所述gRNA的序列由第一载体或载体组合包含。

[0766] 126.段落113的核酸,其中编码tracrRNA的核酸由所述第一载体或载体组合包含。

[0767] 127. 段落125至126中任一项的核酸,其中第一核酸和编码gRNA1的所述序列(和任选地编码gRNA2的序列)由所述第一载体包含。

[0768] 128. 根据段落124所述的第一载体,其与第二载体组合,其中第二载体包含所述编码gRNA1的序列(和任选地编码gRNA2的序列)。

[0769] 129.段落128的载体组合,其中第二载体包含编码所述第一Cas(例如,Cas9)的可表达的核苷酸序列。

[0770] 130.段落128的载体组合,其与包含编码所述第一Cas(例如,Cas9)的可表达的核苷酸序列的第三载体组合。

[0771] 131.段落128的载体组合,其中第一载体包含编码所述第一Cas(例如,Cas9)的可表达的核苷酸序列。

[0772] 132.从属于段落108至113中任一项的段落78至131中任一项的核酸或组合,其中gRNA1和gRNA2由相同启动子转录,其中编码gRNA1和gRNA2的核苷酸序列由所述第一载体或所述组合的共同载体包含。

[0773] 133.段落78至132中任一项的核酸或组合,其中所述核酸或组合由包含内源CRISPR/Cas系统的细胞包含,且P1与所述系统的Cas核酸酶同源,用于切割PS1。

[0774] 134. 段落78至133中任一项的核酸或组合,其与所述第一Cas或编码所述Cas核酸酶的可表达的核苷酸序列组合。

[0775] 135.段落78至134中任一项的核酸或组合，其中HA1和HA2能够与单一菌株或细菌物种的独特序列(NS1和NS2)杂交。

[0776] 136. 段落135的核酸或组合，其中所述细菌物种或菌株在以下中发现：食品、食品成分或前体成分；饮料；水(例如，意欲用于人食用)；工业或环境物质(例如，油、石油产品、土壤或水路或水库；或用于回收或处理油、石油产品、土壤、水、食品、食品成分或前体、或饮料或前体的饮料成分的设备)；任选地，其中所述核酸或组合与这些中的任一种组合。

[0777] 137.段落135或136的核酸或组合，其中所述物种或菌株是人或非人动物微生物组物种或菌株，例如肠道、阴道、腋窝或口腔微生物组群体。

[0778] 138.抗生素(例如，抗细菌或抗古细菌)组合物，其包含段落78至137中任一项的核酸或组合。

[0779] 139.用于治疗或预防人中的疾病或病况(例如细菌感染或肥胖症)的药物，所述药物包含段落78至138中任一项的第一核酸、抗生素或组合。

[0780] 140.段落124的核酸，或从属于段落124的段落125至139中任一项的核酸或组合，其中第一载体或各载体是dsDNA载体。

[0781] 141.段落124的核酸，或从属于段落124的段落125至139中任一项的核酸或组合，其中第一载体或各载体是dsDNA载体。

[0782] 142.段落140或141的核酸或组合，其中第一载体或各载体是BAC、PAC、YAC或质粒。

[0783] 143.段落140、141或142的核酸或组合，其中所述第一载体在切割之前是环状的。

[0784] 144. 段落124的核酸，或从属于段落124的段落125至139中任一项的核酸或组合，其中第一载体是噬菌体，其能够感染宿主细胞，用于引入第一载体并切割其中的PS1。

[0785] 145. 任一前述段落的方法、核酸、组合、抗生素或药物，其中HA1远离CS1间隔不超过250nt。

[0786] 146. 任一前述段落的方法、核酸、组合、抗生素或药物，其中HA1包含至少15个连续核苷酸。

[0787] 147.任一前述段落的方法、核酸、组合、抗生素或药物，其中HA1的序列与NS1同基因。

[0788] 148. 任一前述段落的方法、核酸、组合、抗生素或药物，其中HA2远离CS2间隔不超过250nt。

[0789] 149. 任一前述段落的方法、核酸、组合、抗生素或药物，其中HA2包含至少15个连续核苷酸。

[0790] 150. 任一前述段落的方法、核酸、组合、抗生素或药物，其中HA2的序列与NS2同基因。

[0791] 151. 制备适于与第二核酸链同源重组的第一核酸链的方法，所述方法包括在核酸链中组合

[0792] (a) 插入序列(IS)，其包含第一同源臂(HA1)(或由第一同源臂(HA1)组成)，所述第一同源臂(HA1)与第二链的预定靶核苷酸序列(NS1)互补，用于通过同源重组将IS插入第二链中；

[0793] (b) 与PAM (P1)相邻的第一前间隔区序列(PS1)，其中P1与第一Cas核酸酶同源，所述第一Cas核酸酶用于在第一切割位点(CS1)处切割PS1；

[0794] 其中CS1在HA1的末端或侧接HA1，其中当CS1被Cas切割时，HA1能够与NS1同源重组用于所述IS的插入。

[0795] 152.段落151的方法，所述方法包括在核酸链中进一步组合

[0796] (c) 与第二链的第二预定序列(NS2)互补的第二同源臂(HA2)；

[0797] (d) 与第二PAM (P2)相邻的第二前间隔区序列(PS2)，其中P2与Cas核酸酶同源，所述Cas核酸酶用于在第二切割位点(CS2)处切割PS2，

[0798] 其中当CS2被Cas切割时，HA2能够与NS2同源重组用于IS的插入。

[0799] 153.段落151或152的方法，所述方法包括用第一Cas(例如，Cas9，任选地化酿脓链球菌或金黄色葡萄球菌Cas9)切割CS1。

[0800] 154.段落152或153的方法，所述方法包括用与P2同源的Cas核酸酶切割CS2。

[0801] 155.段落154的方法，所述方法包括用相同的Cas核酸酶切割CS1和CS2。

[0802] 156.段落153至155中任一项的方法，所述方法包括在HA1(和任选还有HA2)与第二核酸链之间进行同源重组，其中IS插入第二链中；和任选地在所述同源重组后分离或测序修饰的第二核酸链和/或第一核酸链。

[0803] 157.修饰核酸链的方法，所述方法包括

[0804] (a) 提供根据段落78至150中任一项所述的第一核酸链或通过段落151至156中任一项的方法产生的第一核酸链；

[0805] (b) 用第一Cas(例如，Cas9，任选地化酿脓链球菌或金黄色葡萄球菌Cas9)切割CS1；和

[0806] (c) 在HA1(和任选还有HA2)与第二核酸链之间进行同源重组，其中IS插入第二链中；和

[0807] (d) 任选地在所述同源重组后分离或测序修饰的第二核酸链和/或第一核酸链。

[0808] 158.段落157的方法，其中所述方法包括在所述同源重组之前切割(例如，使用Cas切割)NS1和NS2之间的第二核酸链。

[0809] 159.段落158的方法，其中用第一Cas核酸酶切割第二核酸链。

[0810] 160.段落157至159中任一项的方法，其中通过所述同源重组破坏由第一或第二核酸链包含的标记物，所述方法包括检测所述标记物的不存在以确定已发生同源重组。

[0811] 161.段落157至159中任一项的方法，其中通过所述同源重组形成由第一或第二核酸链包含的标记物，所述方法包括检测所述标记物的存在以确定已发生同源重组。

[0812] 162.段落157至161中任一项的方法，其中IS包含供体序列，且HA1在IS的5’侧，其中所述方法包括在HA1和NS1之间进行同源重组，其中IS插入第二核酸序列中。

[0813] 163.段落162的方法，其中另外的前间隔区和PAM由IS包含，且所述方法包括在将IS插入第二核酸链后使用另外的前间隔区的Cas切割，其中在切割位点CS3处切割另外的前间隔区。

[0814] 164.段落163的方法，其中使用与P1或P2不同源的Cas来切割另外的前间隔区。

[0815] 165.段落163或164的方法，其包括进行同源重组以在CS3或侧接CS3处插入或缺失另外的序列。

[0816] 166.段落157至165中任一项的方法，所述方法包括在HA1和HA2之间没有插入任何序列的情况下且任选地在第一核酸链中不插入侧接HA1和HA2的任何序列的情况下将HA1和

HA2插入第二核酸中。

[0817] 167.段落157至166中任一项的方法,所述方法包括缺失NS1和NS2之间的序列。

[0818] 168.段落167的方法,其中所述方法在细胞(例如,细菌细胞,例如,人病原体细胞)中进行,且所述第二核酸链由细胞的染色体或附加体包含,其中所述序列缺失杀死细胞。

[0819] 169.段落157至166中任一项的方法,所述方法包括在缺失NS1和NS2之间的序列的情况下插入IS。

[0820] 170.段落157至169中任一项的方法,所述方法包括在切割CS1之前将第一核酸链与指导RNA (gRNA1,例如单一指导RNA)组合,其中gRNA1包含与PS1互补的核苷酸序列,用于指导Cas以切割PS1。

[0821] 171.段落157至170中任一项的方法,所述方法包括在切割CS1之前将第一核酸与指导RNA (gRNA2,例如单一指导RNA)组合,其中gRNA2包含与PS2互补的核苷酸序列,用于指导Cas以切割PS2。

[0822] 172.段落157至171中任一项的方法,其中第一链由第一载体包含,且在切割CS1之前,所述方法包括将载体引入宿主细胞,其中CS1在宿主细胞中切割,且在宿主细胞中进行与第二链的同源重组。

[0823] 173.段落172的方法,其中所述细胞是细菌或古细菌细胞,例如大肠杆菌细胞。

[0824] 174.段落172或173的方法,其中所述细胞是人病原体物种的。

[0825] 175.段落172、173或174的方法,其中所述细胞是人微生物组物种(例如人肠道微生物组物种)的细胞。

[0826] 176.段落173至175中任一项的方法,其中第一Cas(例如,Cas9)对细胞是内源的。

[0827] 177.段落157至176中任一项的方法,其包括在切割CS1之前将第一Cas引入细胞中或在切割CS1之前将编码第一Cas的可表达的核苷酸序列引入细胞并在细胞中进一步表达所述Cas。

[0828] 178.段落157至177中任一项的方法,其在体外进行。

[0829] 179.段落157至178中任一项的方法,其中所述方法包括重组工程改造以修饰核酸链。

[0830] 180.段落157至177中任一项的方法,其在体内(例如,但不在人或人胚胎中)进行。

[0831] 181.段落157至177和180中任一项的方法,其用于人或动物医学治疗、预防或诊断,其中治疗、预防或诊断人或动物中的疾病或病况。

[0832] 182.段落172的方法,其中所述细胞是真核细胞。

[0833] 183.段落182的方法,其包括在切割CS1之前将第一Cas引入细胞中或在切割CS1之前将编码第一Cas的可表达的核苷酸序列引入细胞并在细胞中进一步表达所述Cas。

[0834] 184.段落157至183中任一项的方法,其包括进一步将第一核酸包括入食品、食品成分或前体成分;饮料;水(例如,意欲用于人食用);工业或环境物质(例如,油、石油产品、土壤或水路或水库;或用于回收或处理油、石油产品、土壤、水、食品、食品成分或前体、或饮料或前体的饮料成分的设备)中。

[0835] 185.段落157至184中任一项的方法,其包括进一步将第一核酸包括入组合物中,用于将核酸引入细菌或古细菌物种或菌株,所述细菌或古细菌物种或菌株是人或非人动物微生物组物种或菌株,例如,肠道、阴道、腋窝或口腔微生物组群体。

[0836] 186.段落157至185中任一项的方法,其包括进一步将第一核酸包括入抗生素(例如,抗细菌或抗古细菌)组合物或包括入药物中用于治疗或预防人中的疾病或病况(例如细菌感染或肥胖症)。

[0837] 187.段落157至183中任一项的方法,其中所述方法在细胞中进行并且CS1的切割杀死细胞或减少细胞生长或增殖。

[0838] 188.段落187的方法,其中PS1由载体抗毒素核苷酸序列包含和/或通过所述同源重组从第一链缺失抗毒素序列。

[0839] 189.段落188的方法,其中PS1由第一链的抗-抗毒素试剂的核酸序列(AAT)包含或侧接,其中所述试剂能够消除或降低宿主细胞抗毒素的活性。

[0840] 190.段落157至183和187至189中任一项的方法,其中所述方法在细胞中进行,且CS1的切割促进细胞中毒素的细胞毒性活性。

[0841] 191.段落157至190中任一项的方法,其中第一核酸根据段落78至150中任一项。

[0842] 192.修饰细胞(例如,原核或真核细胞)中的核酸的方法,所述方法包括

[0843] (a) 提供细胞,其含有

[0844] i 闭合的环状核酸载体(例如,BAC),其中所述载体包含第一核酸链;

[0845] ii 第二核酸链,其中所述第二链由细胞的染色体包含;和

[0846] iii 编码Cas核酸酶的核苷酸序列和任选地编码核定位信号(NLS)的核苷酸序列;

[0847] (b) 使用Cas(例如,Cas9)核酸酶切割来切割细胞内载体的第一链以产生重组产生的载体核酸;和

[0848] (c) 在切割的第一链和第二链之间的细胞内使用同源重组,和在第一和第二链之间交换至少10、20、30、40、50、60、70、80、90、100、150、200、250或300kb的连续核苷酸的序列,由此形成至少一条修饰的链;和

[0849] (d) 任选地分离或测序修饰的链或细胞。

[0850] 193.段落192的方法,其中第一链中的同源臂用于所述同源重组,且所述Cas切割切割远离同源臂不超过1kb的链。

[0851] 194.段落192或193的方法,其中至少0.5kb的同源性用于所述同源重组。

[0852] 195.段落192至194中任一项的方法,其中所述交换包括第一或第二链的同源性介导修复(HDR)。

[0853] 196.段落192至194中任一项的方法,其中不切割第二链。

[0854] 197.段落192至196中任一项的方法,其中所述序列是至少100kb的连续核苷酸的序列。

[0855] 198.段落192至197中任一项的方法,其中在步骤(a)之前将载体作为超螺旋闭合环状载体引入细胞中。

[0856] 199.段落192至198中任一项的方法,其中所述细胞是人、小鼠、植物或酵母细胞。

[0857] 200.段落192至199中任一项的方法,其中所述序列是人序列。

[0858] 201.段落192至200中任一项的方法,其中所述序列从第二链回收并插入第一链中。

[0859] 202.段落192至201中任一项的方法,其中所述序列从第一链插入第二链中。

[0860] 203.用于在体外校正人细胞中的突变的基因疗法方法,所述方法包括在人细胞

(例如,iPS细胞或T细胞)中进行段落192至202中任一项的方法来校正所述突变,其中所述突变在步骤(a)中由第二链包含。

[0861] 204.产生人CAR-T细胞的方法,所述方法包括在人T细胞(或其前体细胞)中进行段落192至203中任一项的方法,其中所述序列是用于表达嵌合抗原受体的蛋白、亚基或结构域的序列。

[0862] 在一个实例中,段落192至204中任一项的方法中的第一链具有如本发明的任何配置、实施方案、实例或项中所述的第一链的特征,因此任何那些特征可以与段落192至204中任一项组合。在一个实例中,段落192至204中任一项的方法中的第二链具有如本发明的任何配置、实施方案、实例或项中所述的供体或另外的链的特征,因此任何那些特征可以与段落192至204中任一项组合。

[0863] 在本文的任何配置中,HA1和HA1来自与其中进行所述方法的细胞相同的物种或菌株(即,NS1和NS2的相同物种或菌株)。例如,所述物种是人、小鼠或大鼠,例如小鼠129或C57/BL6品系。

[0864] 在其中载体产物包含从供体链回收的含有RS的第一链的配置中,产物载体任选地包含侧接HA1和HA2的分别3'和5' (或在HA1和HA2的分别3'和5'末端或侧接HA1和HA2的分别3'和5'末端)的核酸酶(例如Cas)切割位点,用于实现产生具有重组产生的末端且包含RS的序列,其用于随后将RS的HR或SSR插入另外的链,例如靶细胞或核酸的另外的链中。在一个实例中,额外的切割位点用Cas核酸酶切割,所述Cas核酸酶不同于用于切割CS1和CS2的Cas,和/或在首次切割CS1/CS2后,用于切割额外位点的gRNA与第一链组合(例如,引入其中回收RS的细胞中)。

[0865] 序列文库、集合和使用

[0866] 本发明可用于产生回收序列(RS)的集合或文库。在一个实施方案中,本发明可用于产生通过本发明的回收方法可获得(或获得)且包含回收的序列(RS)的集合的多种载体,其中所述集合包含彼此不同的第一和第二RS序列,例如,编码作为彼此的变体的多种蛋白。在一个实施方案中,第一和第二RS的总大小为1至600kb,例如1至500、400、300、200、250、150或100kb,并且任选地每个序列大小为不超过250或300kb。例如,所述载体是BAC或YAC载体,其包含从细胞样品回收的基因组片段的集合。所述样品可以包含多种相同或不同类型的细胞,例如,获得自相同生物体(诸如啮齿动物、小鼠、大鼠、人、细菌物种或菌株、酵母物种或菌株)的细胞,或其中所述序列是相同的病毒种类或菌株的。在一个实例中,本发明可用于产生包含人或动物的基因组的多种回收的基因组片段的载体集合。以这种方式,本发明容易提供为个体人产生个性化集合的方式,其可以例如用于测序人类基因组的全部或部分和/或作为一种或多种基因组片段的来源用于在一个或多个人中的诊断、治疗和/或预防性用途,诸如用作序列来源用于插入细胞,用于产生转基因细胞用于所述诊断、治疗或预防性用途。例如,从本发明的载体对一个或多个回收的序列(RS)进行测序、分离和/或复制,并用于方法中,在所述方法中,将序列(或其修饰或修正的版本)体外插入受体细胞中,并且任选地将细胞施用于动物(例如,人或动物患者)以治疗或预防患者中的疾病或病况,或者将细胞发育成组织或血液样品用于所述施用。所述受体细胞可以是人或动物多能性细胞(例如,人iPS细胞)、干细胞、胚胎干细胞(例如,非人脊椎动物细胞,例如啮齿动物、大鼠或小鼠ES细胞)或免疫细胞(例如,人免疫细胞或造血干细胞)。以这种方式,例如,可以产生修饰的

人免疫细胞(例如,CAR-T或NK细胞),其被施用于人患者以治疗或预防患者中的疾病或病况。

[0867] 通过在回收载体中使用相对小的同源臂(例如,小至15-50bp)和/或通过使用其序列与供体核酸具有几个错配的臂(例如,其中每个臂的序列与其各自的NS1或2包含40至20%错配),这可用于从供体回收随机(非预定或未知)RS,例如,从生物体(例如,人、啮齿动物、小鼠或大鼠)的一个或多个细胞产生随机基因组片段的集合。

[0868] 有利地,本发明的载体包含侧接一个或多个已知序列的同源臂(例如HA1和HA2)的RS。这提供了几个优点。例如,所述同源臂可用于将RS插入受体细胞(例如,本文提及的细胞)中,有利地,所述插入在由同源臂指示的预定基因组位置。有用地,同源臂通过本发明的方法与RS组合提供,而不需要进一步的、费力的克隆或其他步骤来在RS即用于同源重组插入受体细胞之前添加同源臂。

[0869] 本发明的载体(各自包含侧接一个或多个已知序列的同源臂的RS)也可用于产生基因组片段(RS)的文库或集合,诸如来自植物、动物(例如人、啮齿动物、小鼠或大鼠)、酵母、细菌或病毒基因组的序列集合。在一个实施方案中,这可用于产生序列文库,例如哺乳动物(例如人或小鼠)文库,其可以容易地编目或排序。参考基因组序列(例如,人、细菌、小鼠或酵母参考基因组序列)是本领域可得的,这些可用于选择多个不同的HA1/HA2对,每对用于构建相应类型的回收载体(本发明的第一链)。然后将此类多种载体与获得自相应生物体(例如,植物,动物(例如,人、啮齿动物、小鼠或大鼠)、酵母或细菌细胞)的一种或多种细胞组合,并用于本发明的方法中以回收多个RS(由于在载体中使用不同的HA1/HA2对而回收不同的RS)。在一个实施方案中,将至少第一和第二回收载体引入同一细胞,其中所述载体包含不同的HA1/HA2对,其中从细胞的染色体或附加体回收第一和第二不同的RS。额外地或可替代地,多种不同的回收载体与多种细胞(例如,获得自相同生物体或相同物种或菌株,例如细菌或酵母或获得自相同器官或组织)组合,其中将载体引入细胞中并从细胞回收多个RS。在一个实施方案中,从载体分离、测序或复制一种或多种回收的RS。RS可以用于如本文公开的一种或多种另外的应用(例如,用于基因疗法)。

[0870] 有利地,因为可以使用参考序列(例如,个体(例如,人)或物种的基因组或染色体序列)的知识来设计HA1/HA2对,HA1/HA2对的相对基因组或染色体顺序或位置将是已知的,并且这使得可以参考参考基因组中的已知顺序或位置将回收载体(包括它们各自的RS)进行排序或编目或标记。这允许自动化或计算机辅助编译文库目录或数据库。因此,例如,可以确定哪两个载体包含在同一染色体上重叠、相邻、连续的RS和/或由获得细胞初始样品的生物体基因组中的相同基因座所包含的RS。这大大简化了生物体或细胞的基因组序列的排序或编目载体文库的组装。在一个实施方案中,可以独特地标记载体(例如,通过包含一个或多个常规核苷酸序列条形码),其标识参考基因组序列中的已知位置、染色体或顺序。例如,这对于使用下一代测序方法(诸如使用MiSeq™)的高通量测序是有用的。在一个实例中,每种载体的HA1和HA2中的一者或两者包含条形码标识符(并且限制性或核酸酶位点(例如,Cas切割位点)可以任选地包括在侧接条形码或同源臂处或包括在条形码或同源臂中以允许一旦标识已发生就释放RS)。通过使用不同的条形码来识别不同的位置(参考参考序列或基因组),可以标识回收的每种类型的RS并将其与其在参考中的相应载体和位置相关联。例如,本发明包括使用条形码或标识符来将回收的序列分类并对序列排序(例如,在计算机

芯片上以编译序列汇编，所述序列汇编包含根据归因于参考参考基因组序列的顺序排序的多个回收的序列，例如，在适当时，人、小鼠、酵母或细菌参考序列)。这比目前技术简单得多，所述目前技术涉及在体外从细胞剪切全基因组DNA(这进一步减少可以有效收集的片段的范围和大小)，纯化和分离个别序列，添加适体末端(例如，在体外使用PCR或Gibson组装)至序列，且然后将修饰的序列片段克隆至多个载体(诸如BAC)中以产生文库。在这种情况下，使用许多费力的步骤 - 包括序列的有害剪切，其引起损坏并降低方法的效率并限制获得的片段的大小。此外，对载体的排序、标识或编目则需要对许多载体插入物进行排序(这是耗时的)，随后为基于这种信息的这种排序等。在实施方案中，本发明是优越的，因为可以在细胞的较不苛刻的环境内的回收载体中捕获大片段，所述捕获直接进入载体(诸如环状载体)，其直接用于下游使用而无需进一步操作，RS在供体序列中的位置是已知的(并且可以如上所讨论容易地标识和编目)，并且RS有用地侧接可以用作标识符序列的已知序列(以帮助标识RS)和/或用于下游插入(例如，通过同源重组)受体序列或细胞基因组中。本发明有用地避免了依赖于PCR以回收序列和处理核酸用于插入载体的需要；因此，本发明避免了PCR衍生的核苷酸序列误差(例如，错误地改变供体RS序列的一个或多个核苷酸的插入、缺失和/或取代)的可能性。因此，本发明的产物载体和RS是直接有用的，不需要任何PCR步骤，并且也不需要扩增或培养步骤(尽管如果需要，可以将回收的RS载体扩增或培养包括在该方法中)。可以将本发明多重化，使得可以通过简单地将多个不同的回收载体与供体序列样品(例如，细胞群体)组合来进行多次RS回收。本发明也易于操作，因为其可以采用单一类型的回收载体前体(通用载体)。所述通用载体可以被设计为包括一个或多个克隆位点用于插入同源臂，所述同源臂被选择以靶向靶细胞基因组或靶核酸的期望和预定区域。提供具有一个或多个切割位点的通用载体，每个切割位点侧接或紧邻相应位点(“克隆位点”)用于插入同源臂。其实施方案在本文中参考Cas切割进行描述，但可以容易地修改以用不同方式(诸如指导的核酸酶(例如RNA指导的核酸酶)或限制性内切核酸酶)切割。

[0871] 在一个实施方案中，本发明提供了包含动物基因组、器官或组织的多种不同序列的载体(例如，环状载体)的集合，其中每种序列已经从使用本发明的方法回收的RS获得或复制。在一个实例中，所述动物是人或啮齿动物、大鼠或小鼠。在一个替代方案中，所述集合包含相同物种的多种动物(例如，一个、两个或更多个人)的此类序列。在一个实例中，由所述多种包含的序列与标识符(例如，核苷酸序列标签或条形码)相关联以标识相应序列，例如，以标识序列的一个、多个或全部来源(例如，人)、基因组位置和染色体位置。

[0872] 如上所解释，可以通过选择本发明中的同源臂容易地控制选择回收哪些RS序列。以这种方式，可以回收第一和第二RS，其彼此相邻，彼此连续，或者在供体序列中(例如在供体细胞或生物体基因组中)重叠。任选地，回收第三RS，其中第三RS与供体细胞或生物体基因组中的第一或第二RS相邻、连续或重叠。通过以这种方式回收(并且可选地标识、排序或编目)多个RS，可以使用相对大的序列片段以相对较少的步骤沿供体染色体或附加体行进(特别是其中RS被回收至环状载体、诸如BAC或YAC中)。例如，当多个回收的RS包含重叠的RS和/或在供体中连续的RS时，这尤其方便。在一个实施方案中，回收的RS包含重叠的或在供体中连续的RS，其可用于产生工程改造的序列(例如，DNA序列)或细胞基因组(例如，iPS或ES细胞基因组)的方法中。为此，本发明提供了在体外产生工程改造的序列(例如，DNA序列)或细胞基因组的方法，其中所述方法包括(a)提供受体核酸序列(例如，DNA序列，例如由所

述细胞包含)；(b)将第一插入序列插入受体序列(或其末端)；(c)将第二插入序列插入第一插入序列(或其末端)，使得插入序列部分重叠或紧紧连续，和(d)分离包含第一和第二插入序列的序列或基因组或细胞；和任选地(d)从所得细胞产生组织或非人脊椎动物；其中第一插入序列包含已使用根据本发明的方法回收至载体中的第一RS的序列；其中第二插入序列包含已使用根据本发明的方法回收至载体中的第二RS的序列，且其中RS在供体序列中紧紧连续或部分重叠，其用于本发明的方法中的所述回收。在一个实例中，每个所述插入序列包含由序列(RS)连接的相应间隔的第一(HA1)和第二(HA2)同源臂，其中第一插入序列的HA2在第一插入序列的RS的3'侧，第二插入序列的HA1在第二插入序列的RS的5'侧，且第二插入序列的HA1与第一插入序列的HA2相同或与之部分重叠或由第一插入序列的RS包含或第二插入序列的HA1包含与第一插入序列的HA2至少90％、95％或98％相同的序列。在一个实例中，插入第三插入序列，其中所述第三插入序列包含由序列(RS)连接的相应间隔的第一(HA1)和第二(HA2)同源臂，其中第三插入序列的HA1在其RS的5'侧，且第三插入序列的HA1与第二插入序列的HA2相同或与之部分重叠或由第二插入序列的RS包含。在一个实例中，使用其各自的HA1/HA2对通过同源重组插入每个插入序列。有利地，插入序列的HA序列无一与任何其他HA序列相同。

[0873] 本发明在涉及序列回收的各个方面可大量应用于医学、法医学、诊断学(例如，疾病的产前诊断、疾病易感性、疾病相关突变、遗传、谱系和亲子关系)、考古学、系统发生、人类学、食品生产、油和石油生产、环境应用(例如，修复)、农业和其他工业环境中。在一个实例中，本发明提供了鉴定核酸的细胞或来源的方法，其中所述细胞包含供体核酸序列或其中所述来源包含所述供体，其中所述方法包括本文中用于将RS回收至载体中的方法，并且还包括对回收的RS进行测序并使用该序列来鉴定细胞或来源；或还包括使用核酸探针来探测序列并鉴定细胞或来源。例如，将所述序列或探针与参考序列进行比较以进行所述鉴定。在法医学实例中，所述参考可以是人(例如，人犯罪嫌疑人)或微生物的参考序列。在诊断或质量控制实例中，所述参考可以是微生物(例如，细菌、古细菌、酵母或病毒)的参考序列。质量控制实例可以用于工业过程，例如食品或饮料生产或储存过程；油或石油回收或加工过程；或农业过程的QC。在一个实例中，回收的序列和参考序列相差一个或多个SNP(例如，对于人序列，人SNP)，并且该方法包括确定所述差异。在一个实例中，回收的序列和参考序列共享一个或多个SNP(例如，对于人序列，人SNP)，并且所述方法包括确定回收的和参考序列中所述SNP的存在。有用的参考序列数据库的实例是1000 Genomes数据库和Ensembl。在一个实例中，本发明提供了动物(例如人)的产前诊断或确定动物(例如人)的亲子关系的方法，该方法包括提供来自动物的一种或多种细胞并进行根据本发明的回收RS序列的方法，其中将包含所述第一链的回收载体引入一个或多个所述细胞，其中每种载体包含对应于在动物的亲本(或疑似亲本)之一中发现的序列的HA1/HA2。在一个实例中，所有HA1/HA2对仅对应于亲本之一上的序列。在一种确定亲子关系的方法中，当一个或多个回收的序列由所述亲本(第一亲本)、而不是另一对(怀疑的)亲本(第二亲本)包含时，所述一对亲本被确定为动物(例如，人)的亲本，例如，如RS和所述第一亲本而不是所述第二亲本中的SNP匹配所指示。在一个实例中，RS不是Y染色体序列。在一个实例中，RS是X染色体序列(在这种情况下，本发明的方法可以鉴定遗传自一个亲本但不遗传自另一个亲本的RS或SNP的存在(例如，甚至在从每个亲本遗传X染色体的雌性儿童中)。在本发明的诊断或任何其他方法中，任

选地，回收或其他载体包含连接HA1和HA2的序列，其中在引入载体或从中分离载体的细胞中未发现该序列。因此，所述载体是非天然存在的，因为它将基因组序列(例如HA1和HA2)与不来自相同基因组的序列(例如，细菌或酵母载体序列)组合。在一个实例中，本发明的回收方法用于确定宿主细胞的基因组中的基因序列拷贝数。

[0874] 在一个实例中，供体链或细胞是动物(例如脊椎动物、哺乳动物、人、啮齿动物、小鼠或大鼠)的器官或血液的核酸或细胞。例如，所述供体链或细胞来自动物(例如脊椎动物、哺乳动物、人、啮齿动物、小鼠或大鼠)的肝、肾、脑、皮肤、肺、乳房、前列腺、G1组织、脾、骨髓、胰腺、胆囊、嘴、舌、咽喉、CNS组织、肠、直肠、阴道、阴囊、睾丸、眼睛、耳朵、鼻子、毛发、毛囊、尿道、卵巢或子宫。例如，供体链或细胞来自动物(例如脊椎动物、哺乳动物、人、啮齿动物、小鼠或大鼠)的造血细胞、造血干细胞、脐带血、婴儿血液、成体血液、淋巴细胞、T细胞、B细胞、NK细胞、树突细胞、巨噬细胞、血小板产生细胞或单核细胞。

[0875] 在本文中，当使用RS序列并将其插入另外的核酸链中时，所述插入可以本身是回收的RS(即，本发明的回收方法的直接产物)，或者可以是回收的RS的拷贝的插入(根据需要，任选地在其突变或修饰之后)。

[0876] 在载体(第一链)上使用在供体链或细胞基因组中未发现或罕见发现的切割位点可以是有利的。例如，可以设计Cas切割以仅靶向载体而不靶向基因组。当使用TALEN或锌指切割时也可以设计特异性。限制性内切核酸酶可用于切割载体，其中内切核酸酶识别载体中在细胞(供体链)基因组中未发现、在待回收的RS序列中未发现或仅在细胞(供体链基因组)中罕见发现的基序。一个实例是NotI，其在8个碱基对序列的第一个GC之后切割；具有7个和8个碱基对识别序列的限制酶也可以是罕见的切割内切核酸酶。罕见的切割内切核酸酶的一个实例是大范围核酸酶I-Sce I，其为识别18碱基对序列的归巢内切核酸酶。用于实现切割的一种或多种组分可以作为蛋白、RNA(例如，mRNA)引入细胞中，或者可以由引入细胞的第一链或载体所包含的核苷酸序列编码，其中序列在细胞中表达并产生组分。组分核酸序列可以与一种或多种核定位信号(NLS)组合以定位至细胞核。例如，Cas酶可以作为蛋白引入，作为在细胞中表达的编码Cas的mRNA引入，或由DNA编码(例如，通过第一链或引入的载体编码)用于在细胞中转录和表达相应的Cas。或者，可以利用细胞(当细胞是细菌或古细菌细胞时)表达的内源性Cas以实现切割；在这种情况下，将一种或多种cRNA(或单一指导RNA)编码序列引入细胞中(例如，作为mRNA或由载体或第一链编码)，并且表达的RNA可操作以指导内源性Cas以实现本发明的切割。

[0877] 应理解的是，本文描述的具体实施方案以举例说明的方式显示，而不是对本发明的限制。在不脱离本发明的范围的情况下，本发明的主要特征可以用于各个实施方案中。本领域技术人员将认识到或仅使用常规研究就能够确定本文描述的具体过程的多个等同方案。此类等同方案被认为在本发明的范围内并被权利要求所涵盖。说明书中提及的所有出版物和专利申请表明本发明所属的领域中的技术人员的技术水平。所有出版物和专利申请(以及其美国对应申请)通过引用并入本文，其程度如同具体地且单独地表明每个单独的出版物或专利申请通过引用并入。当在权利要求和/或说明书中与术语“包括(comprising)”联合使用时，词语“一(a)”或“一(an)”的使用可以指“一个(one)”，但其也与“一个或多个”、“至少一个”和“一个或多于一个”的含义一致。在权利要求中使用术语“或”用来指“和/或”，除非明确指出其仅指代供选择的方案，或供选择的方案是相互排斥的，但本发明支持仅指

代供选择的方案和“和/或”的定义。贯穿该申请，术语“约”用来表示这样的值，其包括装置、用于确定该值的方法或研究对象中存在的变化的误差的固有变化。

[0878] 如在本说明书和权利要求(多项权利要求)中所用，词语“包含 (comprising)”(和包含(comprising)的任何形式，如“包含(comprise)”和“包含(comprises))”，“具有(having)”(和具有(having)的任何形式，如“具有(have)”和“具有(has))”，“包括(including)”(和包括(including) 的任何形式，如“包括(includes)”和“包括(include))”或“含有(containing)”(和含有(containing)的任何形式，如“含有(contains)”和“含有(contain))”是包括性或开放式的，并且不排除另外的、未列举的要素或方法步骤。

[0879] 如本文使所用的术语“或其组合”或类似术语是指在该术语之前的所列举项目的所有排列和组合。例如，“A、B、C或其组合”旨在包括以下中的至少一种：A、B、C、AB、AC、BC或ABC，并且如果在具体上下文中顺序是重要的，则还包括BA、CA、CB、CBA、BCA、ACB、BAC或CAB。继续该实例，明确包括的是包含一个或多个项目或条目的重复的组合，如BB、AAA、MB、 BBC、AAABCCCC、CBBAAA、CABABB等等。技术人员将理解通常在任意组合中的项目或条目的数量没有限制，除非从上下文中显而易见。

[0880] 本公开的任何部分可以与本公开的任何其他部分组合阅读，除非另外从上下文显而易见。

[0881] 根据本发明，本文公开和请求保护的所有组合物和/或方法可以在没有过度实验的情况下形成和实施。虽然本发明的组合物和方法就优选的实施方案进行描述，但对本领域技术人员将显而易见的是，在不背离本发明的构思、精神和范围的情况下，可以对本文所述的组合物和/或方法和方法的步骤或步骤的顺序施加变化。对本领域技术人员显而易见的所有这样相似的替代和修改被视为在所附权利要求限定的本发明的精神、范围和构思内。

[0882] 在以下非限制性实施例中更详细地描述了本发明。

实施例

[0883] 实施例1：在真核细胞中组合闭合的环状载体与原位Cas切割用于交换大核苷酸序列 - 人细胞基因疗法应用

[0884] 这可以借助以下可能的基因疗法程序来描述。举例说明一个方面(图1)并且可以通过构建闭合的环状载体(诸如DNA BAC)的群体来实施，其中所述群体包含超螺旋载体。为了能够进行阳性-阴性选择，所述载体各自可以包含载体骨架中的潮霉素抗性基因和如下所述的HPRT基因。可以将环状载体工程改造为包含同源臂和化酿脓链球菌Cas9切割位点(即，如本文所述的HA2、PS2、CS2、P2、PS1、CS1、P1、HA1)。PS1和PS2将是相同的，如同P1和P2一样(各自都是TGG序列)，由此使得CS1和CS2两者均被Cas9切割。编码单一(嵌合)指导RNA(gRNA1)的序列将包括在载体中。CS1将远离HA1 10个核苷酸，且CS2将远离HA2 10个核苷酸。HA1和HA2可以一起包含3kb的核苷酸序列。因此，Cas切割可以产生切割载体，其中将切掉连接HA1和HA2的序列，由此切割载体当切除连接序列时各自失去HPRT基因序列。所得切割载体在HA2的5’侧具有HA1。

[0885] 闭合的环状载体可以在体外电穿孔至已经从患有疾病的人受试者获得的肝细胞

中。所述细胞包含具有引起疾病的SNP的基因序列。HA1与在SNP的3’侧的序列NS1 100％互补，且HA2与在SNP的5’侧的序列NS2 100％相同。电穿孔将包括将Cas9载体引入细胞，所述Cas9序列与U6、T7或人启动子和NLS可操作连接，使得所述序列能够进入细胞核。使用表达的gRNA1的细胞内的载体的切割切割将切除HPRT，并且通过HA1/HA2与NS1/NS2的HR的间隙修复将回收包含SNP的101kb基因序列(RS)，由此产生修饰的载体。可以在阳性-阴性选择中使用潮霉素和G418攻击检测正确的回收。

[0886] 正确修饰的载体可以与体外重组工作一起使用以校正RS中的SNP，由此产生闭合的环状插入载体的群体，每个载体包含两个同源臂之间的校正的RS序列。iPS细胞衍生自人受试者，并且将使用潮霉素和G418检测将这些电穿孔以引入Cas9载体和插入载体。细胞内的载体的Cas9切割切割这样的位点，所述位点将已被工程改造至插入载体中，其侧接同侧臂的任一侧的10个核苷酸，所述同源臂侧接校正的RS(切割在载体骨架中而不在RS中)。切割通过第二gRNA实现，所述第二gRNA将由插入载体携带的核苷酸序列表达。切割释放DNA片段，所述DNA片段各自包含与NS1和NS2 100％互补的同源臂侧接的校正的RS，由此iPS细胞中片段和缺陷基因之间的HR通过将校正的RS插入细胞基因组中来纠正SNP缺陷。

[0887] 从校正的iPS细胞发育出适当的分化细胞(例如肝组织移植物)，并且这些细胞用于人受试者的基因疗法以治疗或预防疾病。

[0888] 实施例2:利用内源性Cas核酸酶的抗生素方法

[0889] 这可以借助以下可能的程序来描述。该方面可以通过构建能够感染不希望的化脓性细菌群体的工程改造的噬菌粒群体来进行。将噬菌体基因组工程改造为包含同源臂和化酿脓链球菌Cas9切割位点(即，如本文所述的HA2、PS2、CS2、P2、PS1、CS1、P1、HA1)。PS1和PS2将是相同的，如同P1和P2一样(各自都是TGG序列)，由此使得CS1和CS2两者均被Cas9切割。编码单一(嵌合)指导RNA(gRNA1)的序列将包括在基因组中。CS1将远离HA1 10个核苷酸，且CS2将远离HA2 10个核苷酸。HA1和HA2可以一起包含3kb的核苷酸序列。因此，Cas9切割可以产生切割的噬菌体基因组，其中将切掉连接HA1和HA2的序列，由此从噬菌体载体切除切割的连接序列。所得切割载体在HA2的5’侧具有HA1。

[0890] 将工程改造的噬菌体施用于患有化酿脓链球菌的咽喉感染的人受试者，使得所述噬菌体感染细菌并将噬菌体核酸引入细胞。HA1与在化酿脓链球菌的必需基因的序列的3’侧的序列NS1 100％互补;HA2与基因序列的5’侧的序列NS2 100％相同。将靶向Hip基因，因为这是必要的(Appl Environ Microbiol. 2011 Jul; 77(13): 4422-4428; doi: 10.1128/AEM.00554-11;”The Histone-Like Protein Hlp Is Essential for Growth of Streptococcus pyogenes: Comparison of Genetic Approaches To Study Essential Genes”; Julia V. Bugrysheva等人)。电穿孔将排除将Cas9载体引入细胞中，因为将使用从噬菌体基因组携带的序列表达的单一指导RNA(gRNA1)利用功能性内源性化酿脓链球菌Cas9。使用表达的gRNA1的细胞内的噬菌体载体的Cas9切割将切除连接HA1和HA2的序列，且HA1/HA2与NS1/NS2的HR的间隙修复将回收*hlp*基因序列(RS)，由此从靶细菌细胞的基因组切除该序列。因此，正确的回收将导致细胞死亡，因为使得*hlp*失活，由此降低人中感染的严重程度或消除感染。

[0891] 在替代应用中，类似的方法可用于靶向人受试者的肠道微生物组中的混合群体中的一种或多种细菌物种。这重新平衡了肠道微生物组。在另一种应用中，微生物组的样品在

体外与感染拟杆菌属物种的噬菌体组合，所述噬菌体是根据本发明的回收载体，并且因此可以杀死靶向的拟杆菌。这改变了群体中的厚壁菌:拟杆菌比率。将重新平衡群体的样品鼻内移植至供体或不同的人受试者中以治疗或预防肥胖。

[0892] 在一种替代应用中，使用类似的方法靶向硫还原细菌(SRB)，其中所述细菌由油田包含或由油回收或加工设备包含。以这种方式，可以杀死SRB，由此减少油田或油回收或加工设备的微生物影响的腐蚀。

[0893] 实施例3:载体形式

[0894] 图2A和B显示用于本发明中的单臂载体。同源臂HA1长度为至少100kb，并且与靶核酸中的靶序列NS1至少90％相同。如所示，在指导RNA (gRNA1)存在的情况下的Cas切割产生与HA1接近(小于250nt)的重组末端，其对于与靶标的有效HR是有用的。例如，HR可用于将HA1中含有的SNP或PAM位点插入靶核酸中。HR可用于插入HA1外部的载体序列，例如，其中所述序列在HA1的100nt内，用于使HR插入靶标的机会最大化。

[0895] 图3A显示一种变型，其中环状载体包含两个Cas切割位点(例如，用作根据本发明的回收载体或插入载体)。切割可以用不同的gRNA(如所示)或相同的gRNA。切割也可以用两种不同的Cas核酸酶(同时或依次使用)以切割所述切割位点。

[0896] 图3B显示一种变型，其中已经在同源臂中鉴定PAM位点，将其工程改造成所述臂，或已经从靶核酸或细胞的序列选择所述臂以含有适当放置的PAM。该变型中的Cas切割在同源臂中提供重组产生的末端以促进有效的HR。如所示，可以包括标记物(例如，HPRT)以指示连接第二载体中的切割位点的序列的成功切除。

[0897] 图4显示标记片段的释放，以在回收载体中产生两个重组产生的末端。在这种情况下，起始载体不依赖于必须找到具有PAM的合适臂或将PAM工程改造成臂。相反，PAM反而在前间隔区序列中提供，其可以更容易地设计，而不是受限于与靶序列的合适同源性的需要。如图4B中所示，与供体链的HR回收含有所需RS的供体片段。回收的片段可以是大或非常大的(例如，大于100kb)，因为其被原位回收至闭合的环状载体中(例如，在其中发生HR的细胞内)。一旦在闭合的环状载体中，RS可以容易地用于体外后续基因工程改造(例如，重组工程改造)方法中或进行纯化。如图4C中所示，闭合的环状载体可用于基因疗法中或使用HR在体外或在任何原核或真核细胞中修饰靶核酸，其中含有RS的大或非常大的序列可凭借闭合的环状载体携带来管控。在一个替代方案中，可以在回收RS之前或之后将SSR位点(例如，如所示的不相容的lox位点)工程改造至载体中(例如，通过重组工程改造来工程改造)。SSR位点可用于与靶链中的位点重组以将RS插入后者中。如图4D中所示，这允许RS的RMCE插入以及还有将改变特异性引入靶标中。如果产物靶链中的lox位点侧接piggyBac位点，则PB转座酶的瞬时表达将干净地切除SSR位点，仅留下RS。当在ES细胞或iPS细胞(例如人iPS细胞或小鼠ES细胞)中进行时，这是有用的。

[0898] 图5A显示重组产生的核酸片段的释放，所述重组产生的核酸片段可用于随后的与靶标的HR过程(图5B)，用于插入插入序列(例如，调节元件或外显子序列)。如图5A中所示，任选的标记物可以包括在释放片段中。还如所示，从载体中切除该片段可以产生由PS2 (a)的5'序列与PS1(S1)的3'序列形成的新标记序列。在该实例中，序列aS1存在于修饰的载体中，但不存在于未修饰的载体中。aS1序列可以通过PCR检测或可以表达以产生可检测的标记蛋白(例如，抗生素抗性)。在图5B中，显示靶标(第二链)中的序列替换 - 这里缺失序列

(矩形框)并将IS插入其位置。在图5C中,载体同源臂与在靶链中间隔(由包括缺失的序列(矩形框)的序列间隔)的序列互补。没有插入IS。例如,这可以用于使靶中的基因失活(例如,当靶标由病原性细菌包含并且缺失的序列是必需、毒力或抗生素抗性基因的序列时)。

[0899] 图6A和6B显示插入载体的应用,其中Cas切割在靠近第一同源臂的一个位点处。通过包括具有在臂之间的插入序列(IS)的第二臂(HA2)来提供HR中靶向的效率和特异性。在本实施例中,发生靶序列的替代(图6B)。在图7B中,进行插入(即,不是替代,其涉及靶序列缺失)。

[0900] 图8A说明本发明的载体与Cas载体的共转化。当P1和P2相同且PS1和PS2也相同时,gRNA1作用于CS1和CS2(可以省略gRNA2 DNA)。当靶细胞包含与载体前间隔区和PAM同源的内源性Cas时,可以省略Cas载体。图8B显示其中在本发明的载体中包括Cas9序列的变型。

[0901] 图9A显示本发明的前体载体。将克隆位点工程改造至载体中,随后用于落入一个或多个同源臂(HA1和/或HA2)中。这在图9B和9C中使用具有不相容的SSR(lox位点)的RMCE来说明。这可以在用产物载体进行靶核酸链或细胞的修饰之前在体外进行。图9D和9E显示使用限制性内切核酸酶的变型。

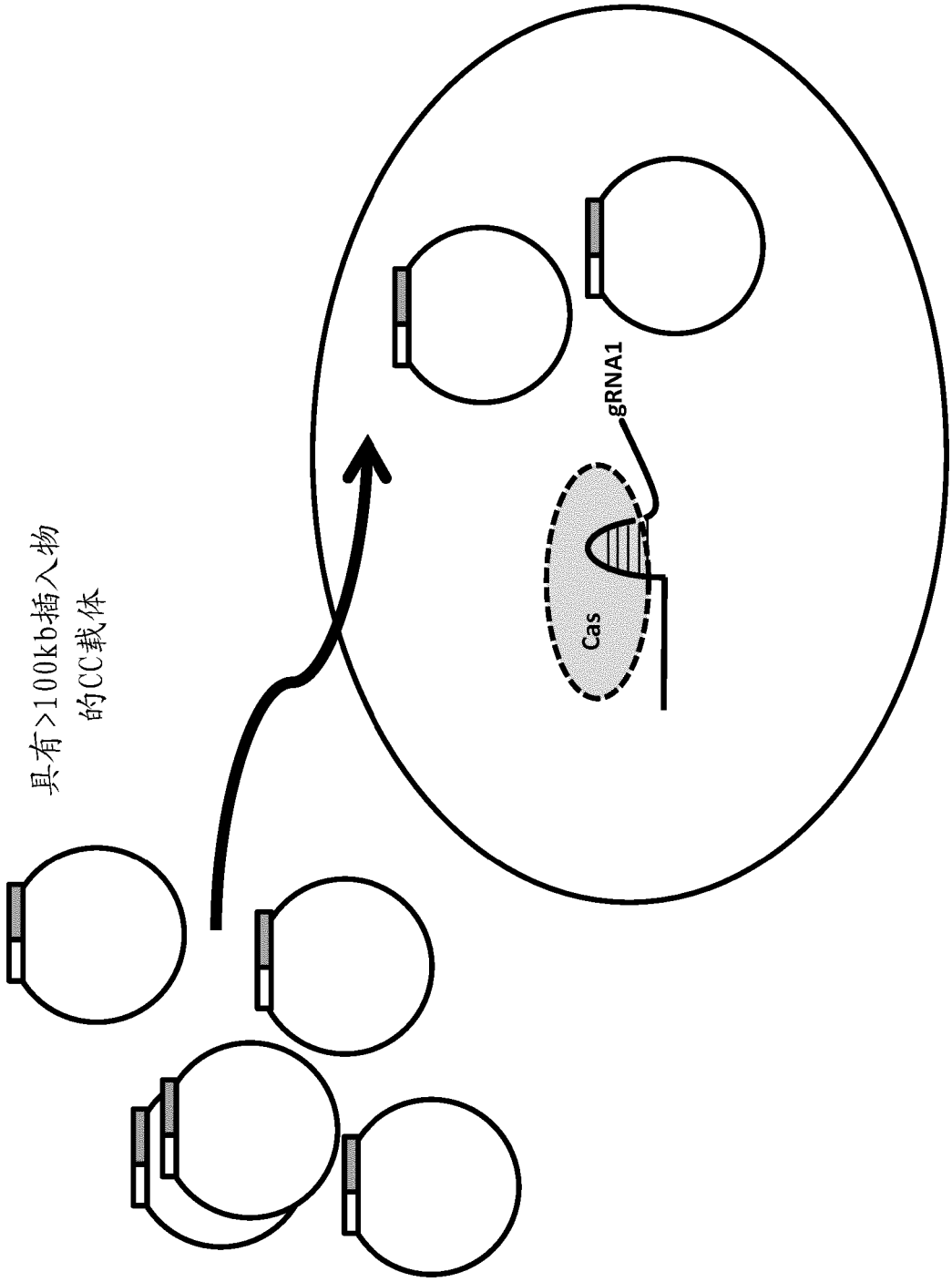


图 1

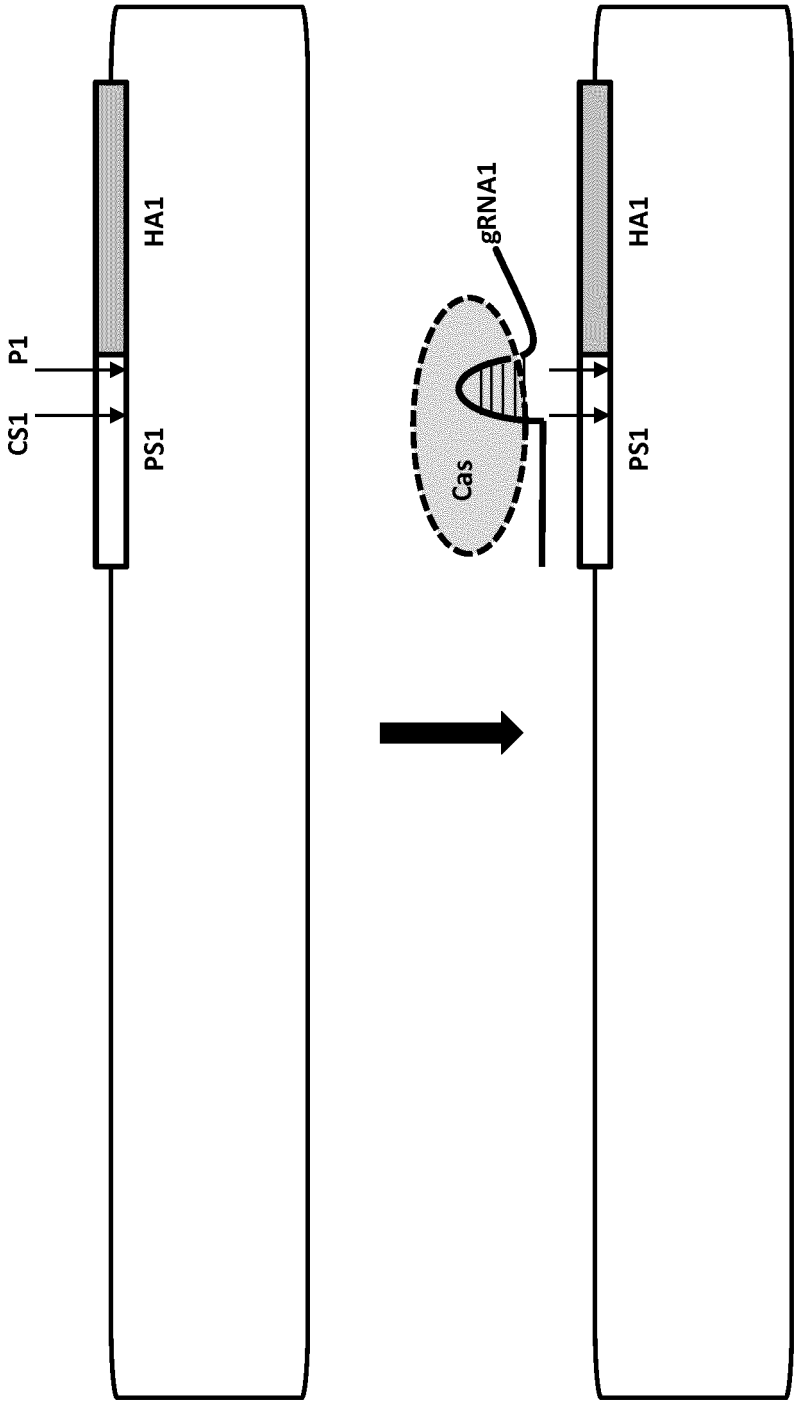


图 2A

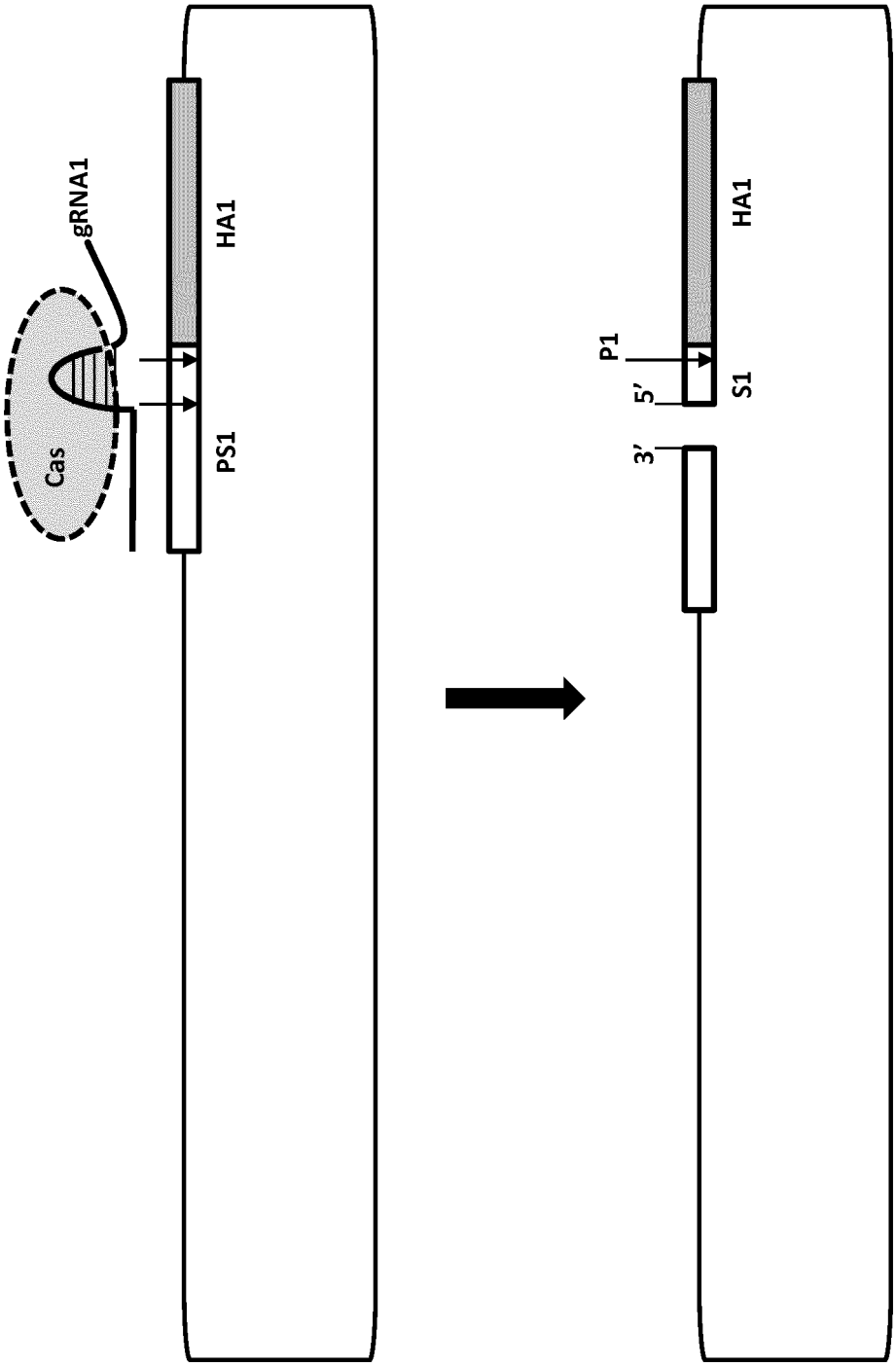


图 2B

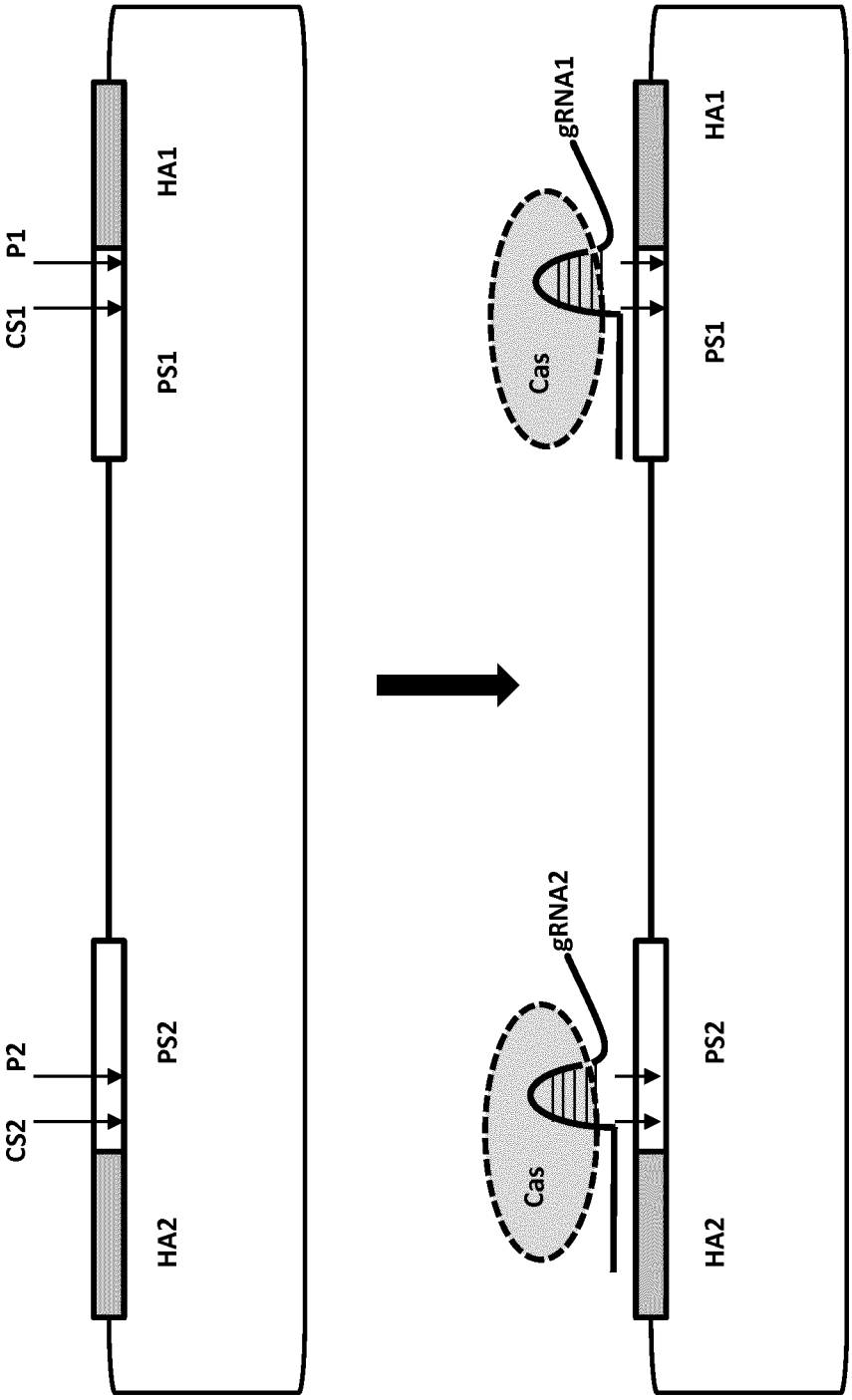


图 3A

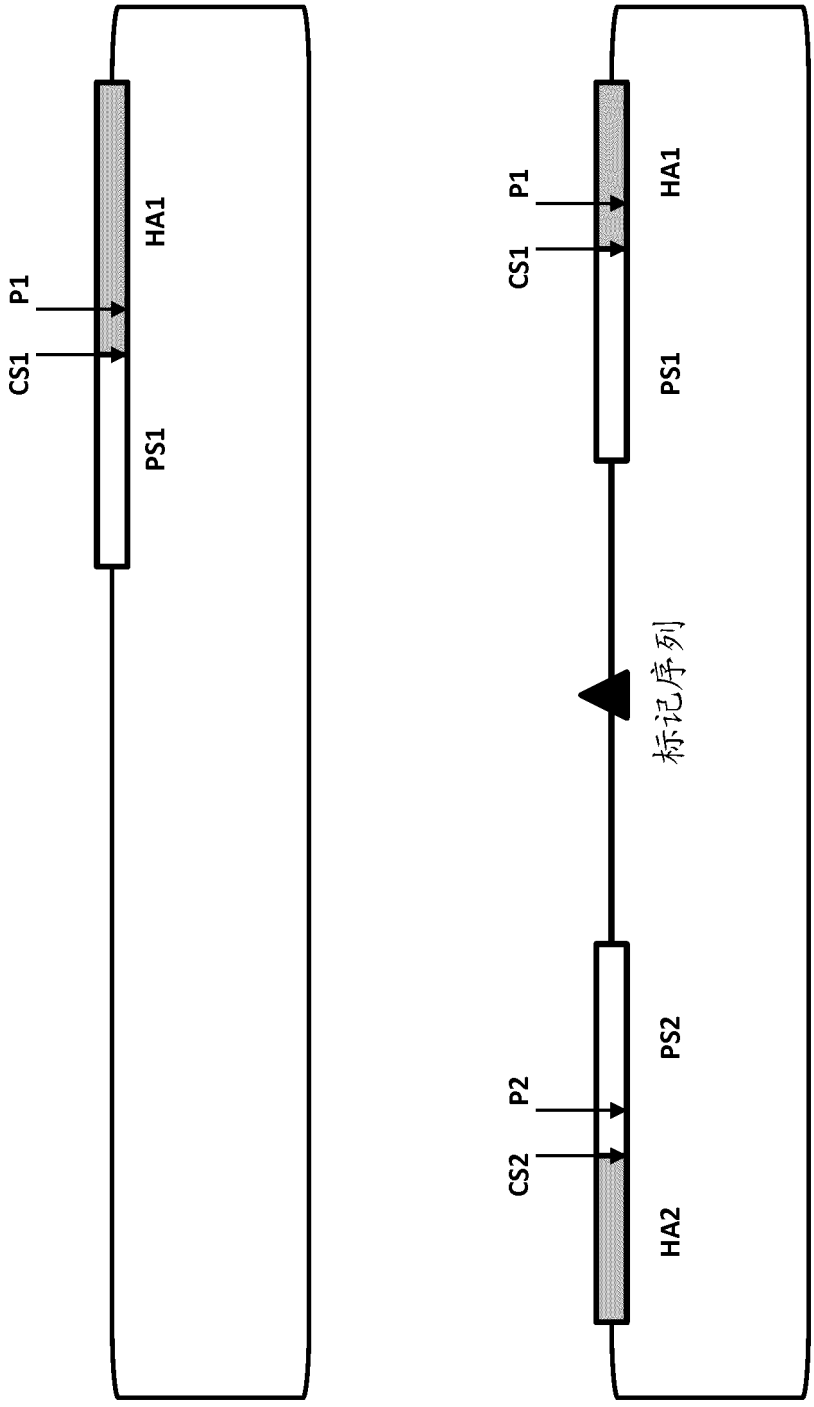


图 3B

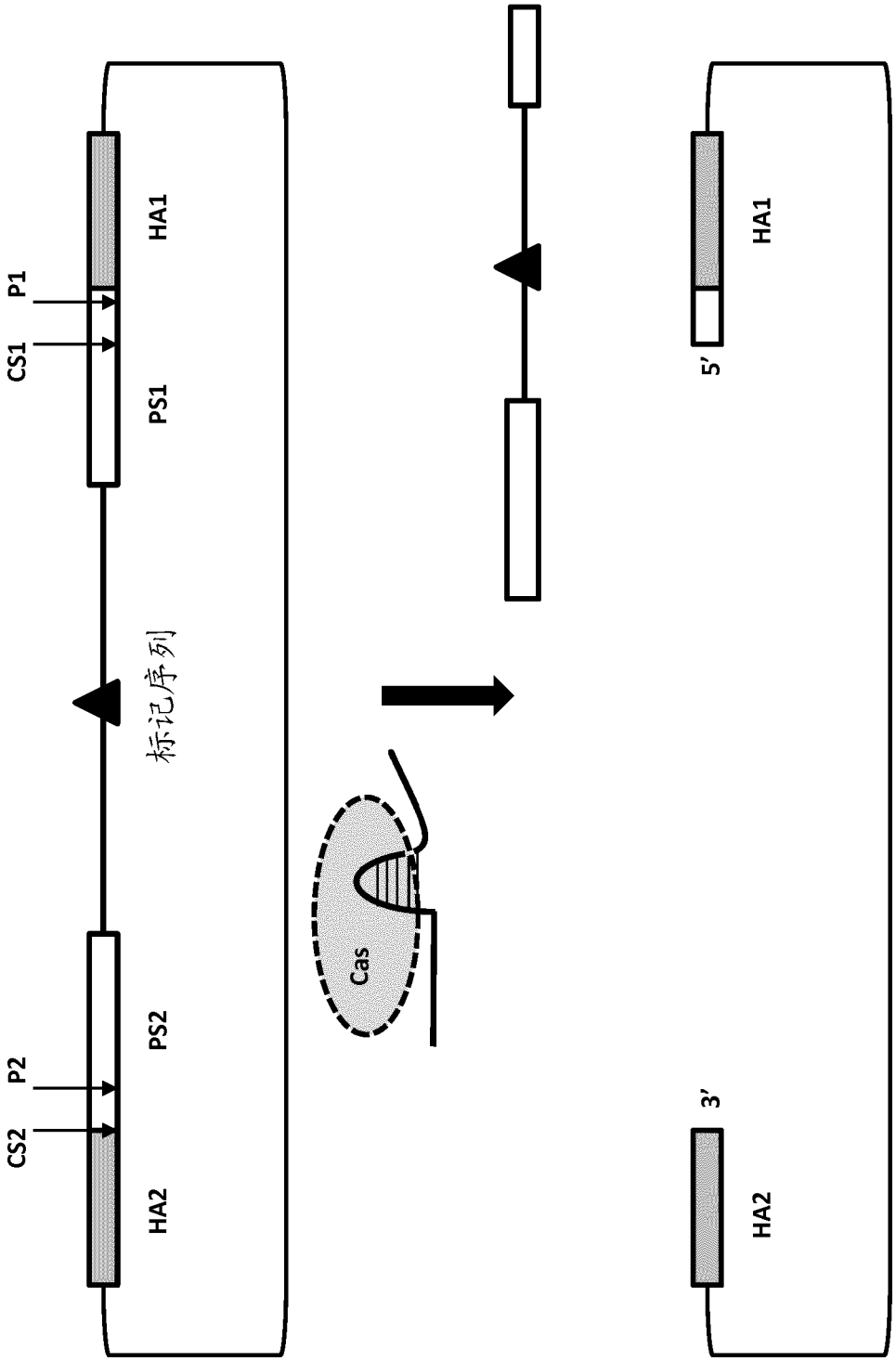


图 4A

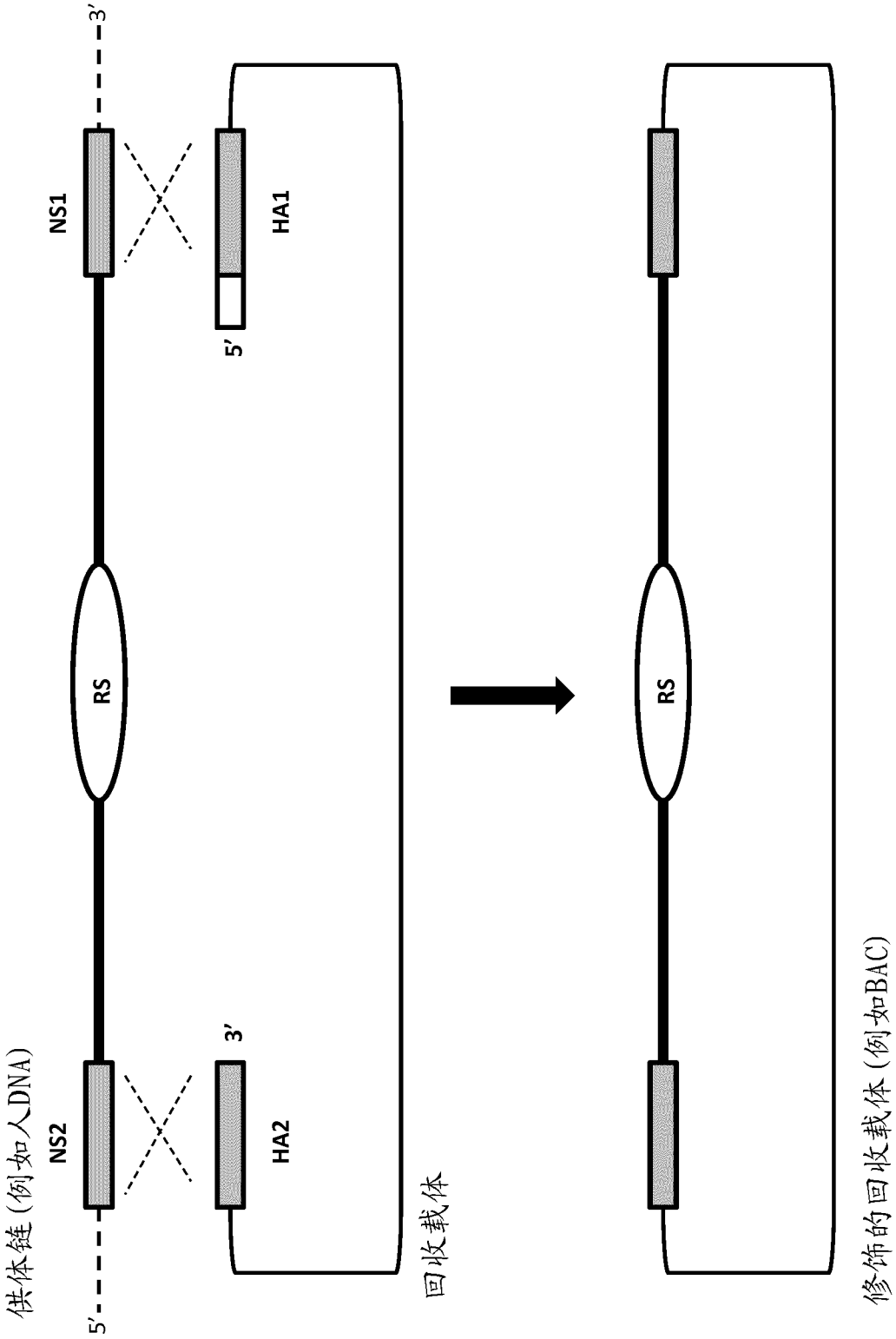


图 4B

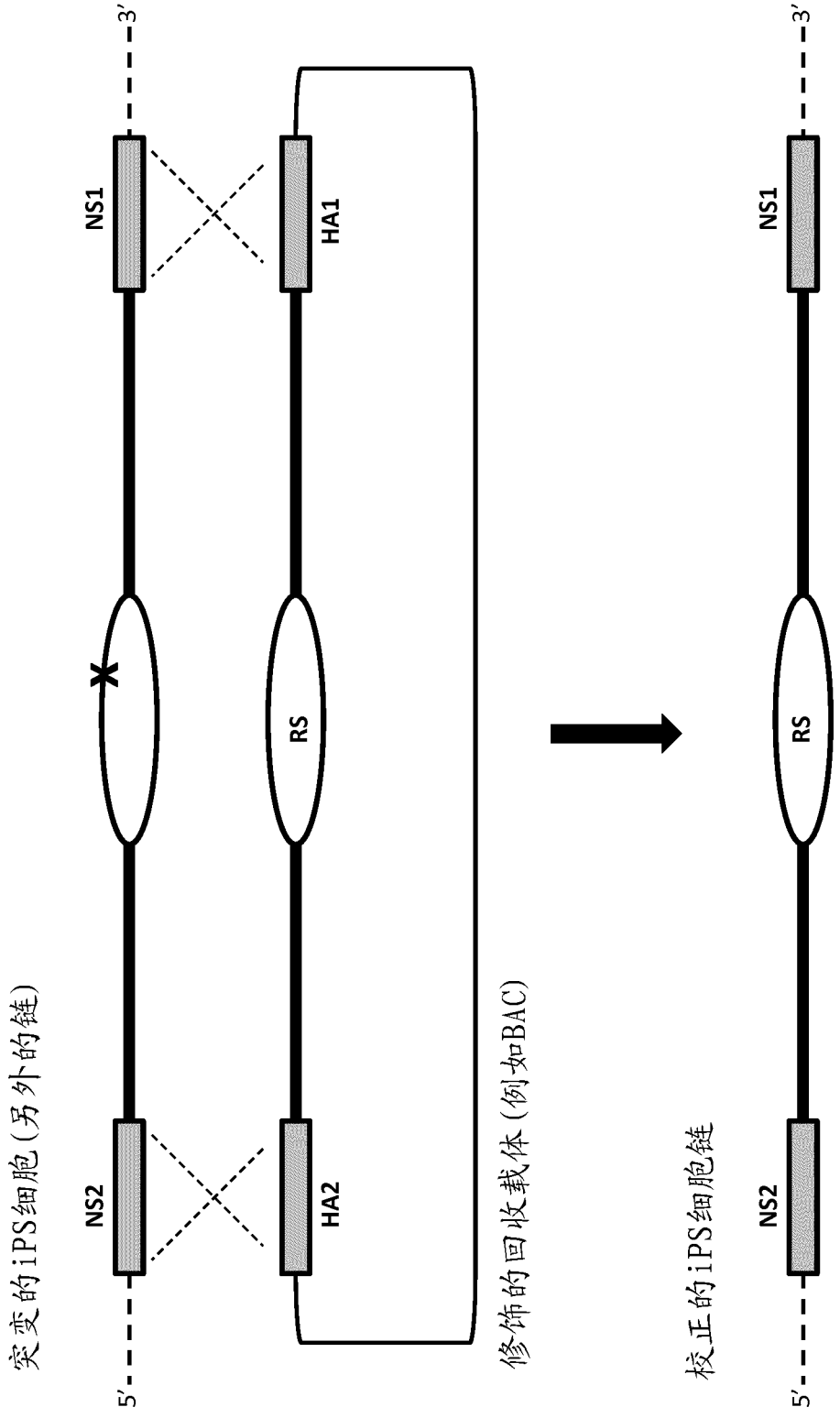

3′
NS1
HA1
RS
NS2
HA2
5′
突变的iPS细胞(另外的链)
修饰的回收载体(例如BAC)
校正的iPS细胞链
3′
NS1
RS
NS2
5′

图 4C

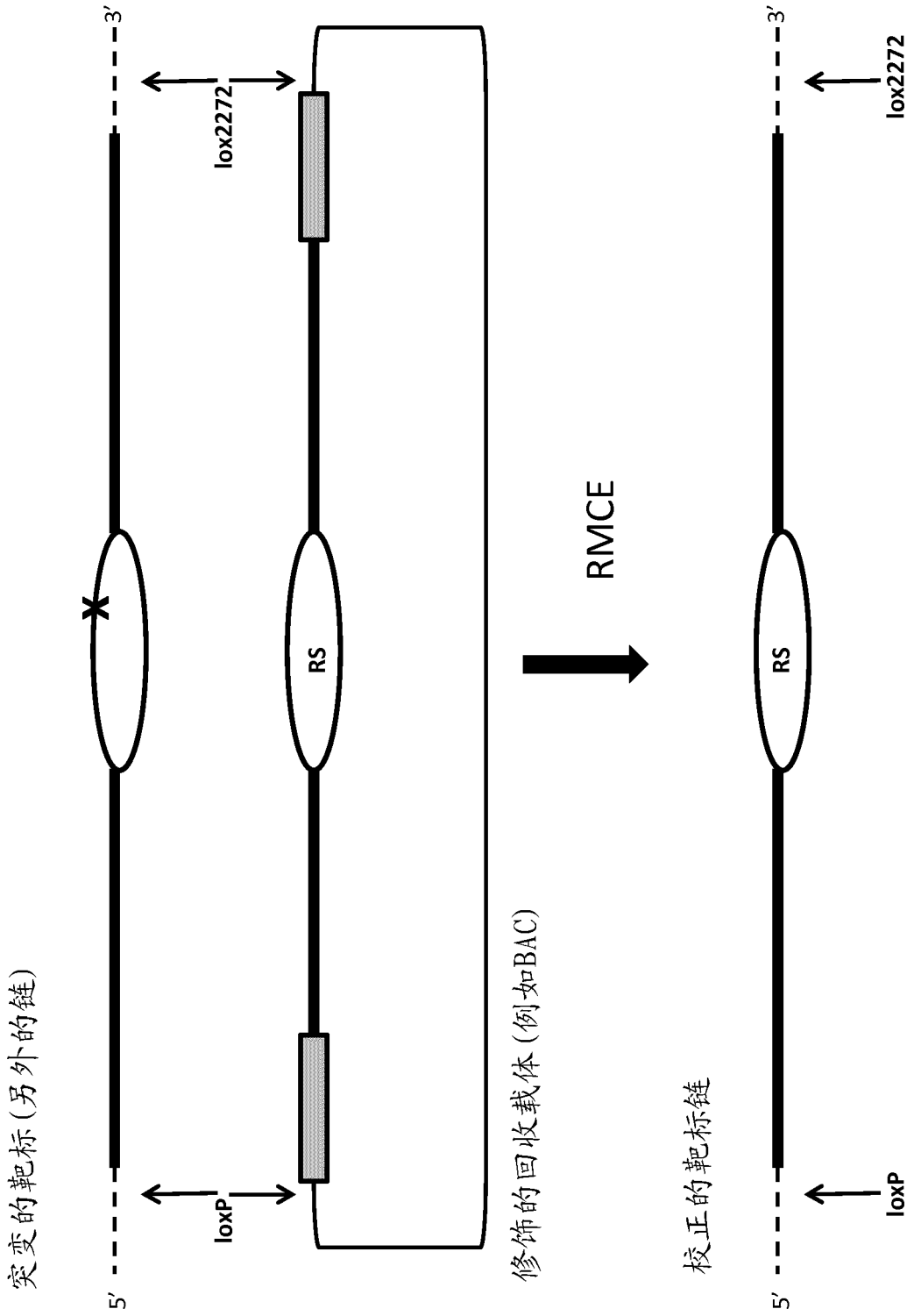

突变的靶标(另外的链)
3′
5′
lox2272
loxP
RS
修饰的回收载体(例如BAC)
RMCE
校正的靶标链
3′
5′
lox2272
loxP
RS

图 4D

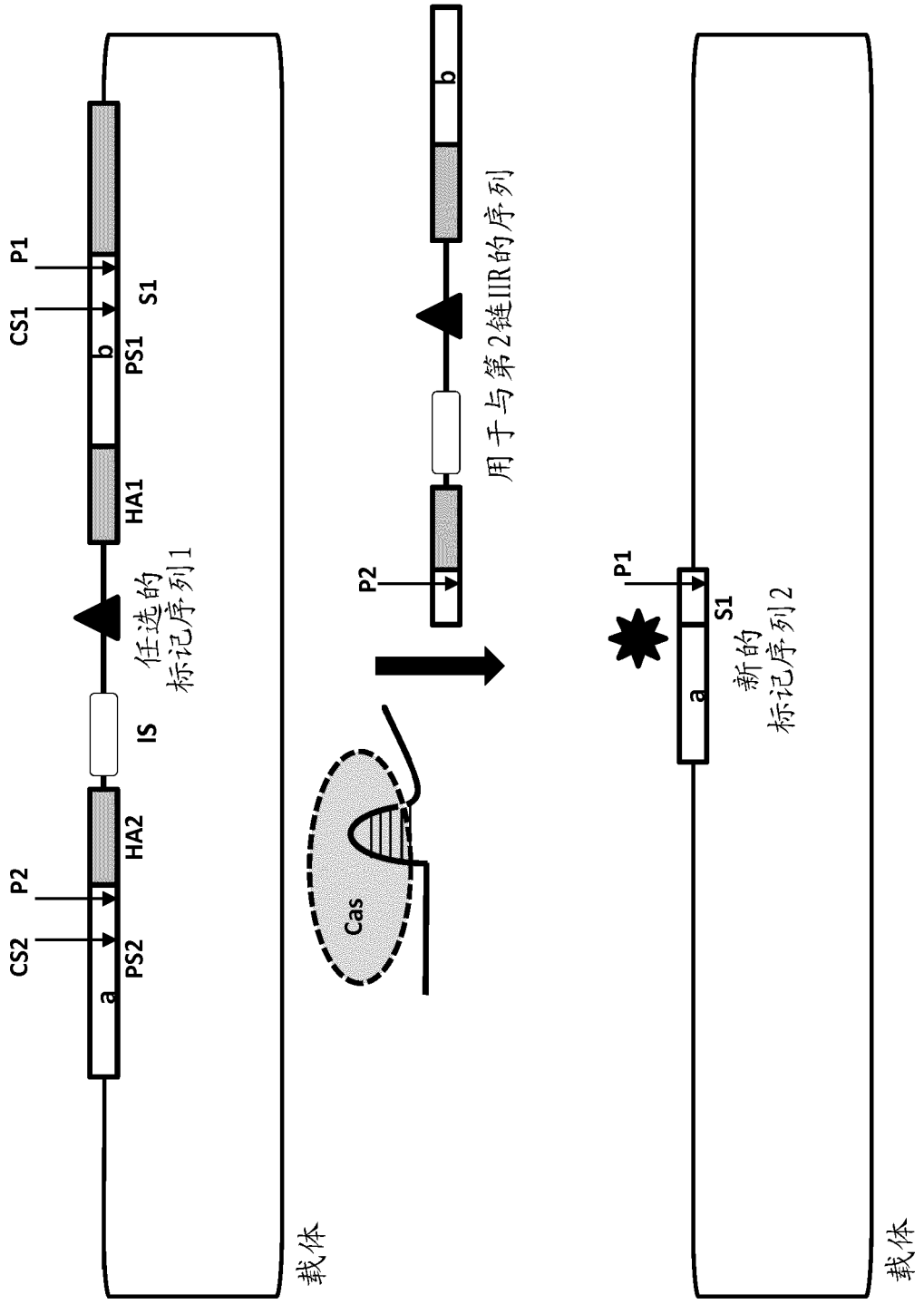

P1
CS1
b
S1
PS1
HA1
任选的
标记序列1
IS
HA2
P2
CS2
a
PS2
载体
Cas
P2
b
用于与第2链IIR的序列
P1
S1
a
新的
标记序列2
载体

图 5A

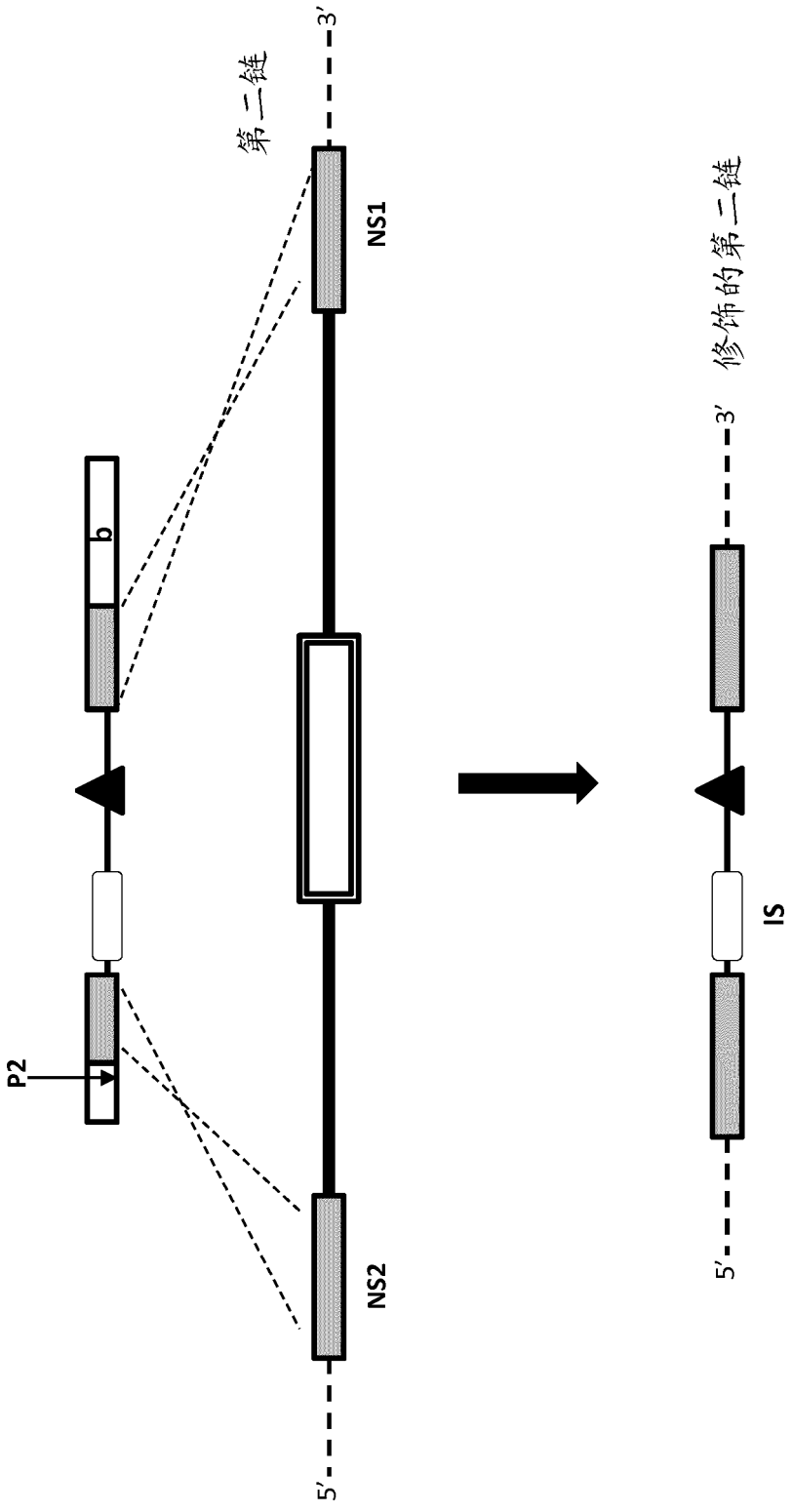


图 5B

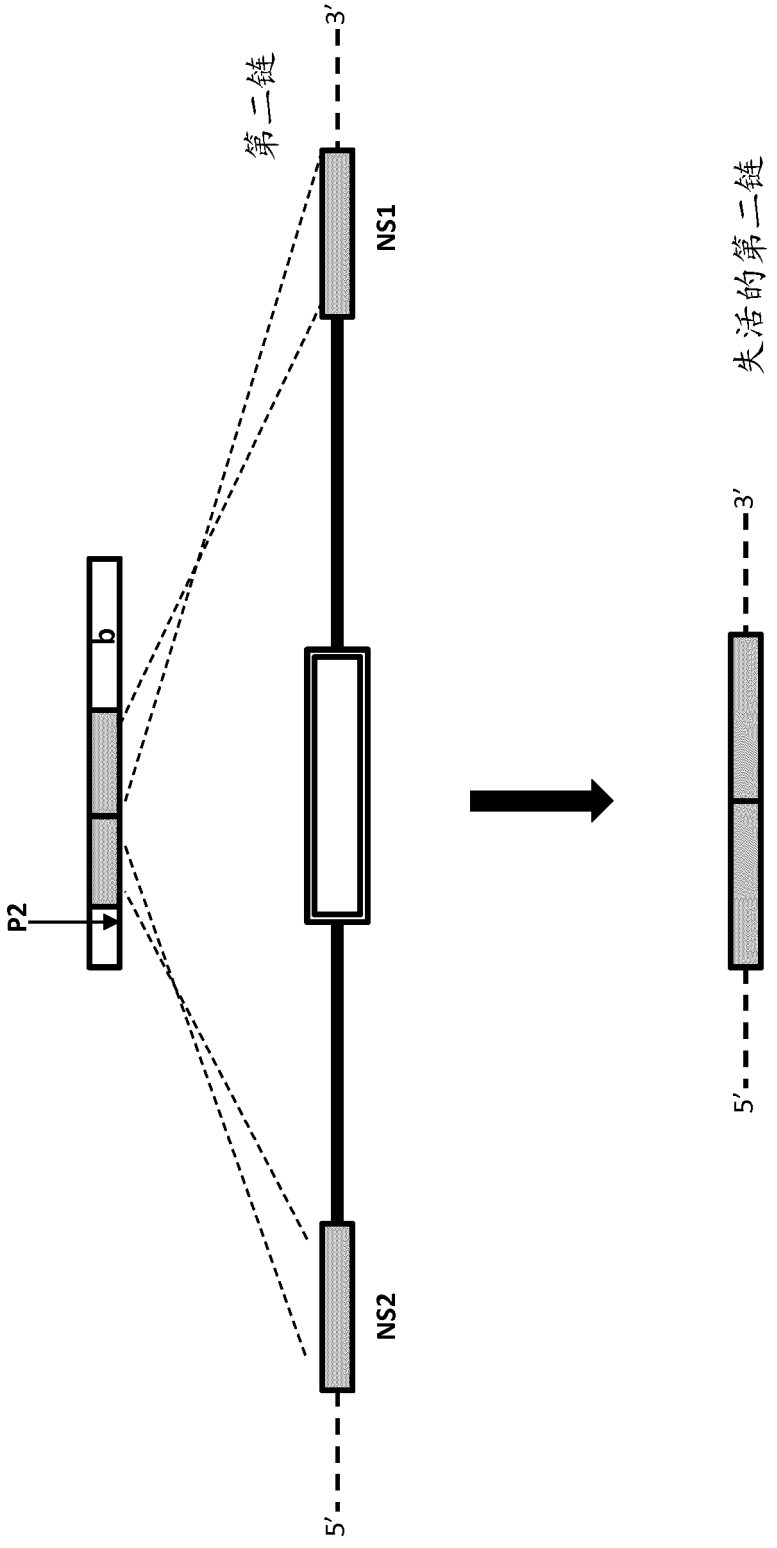


图 5C

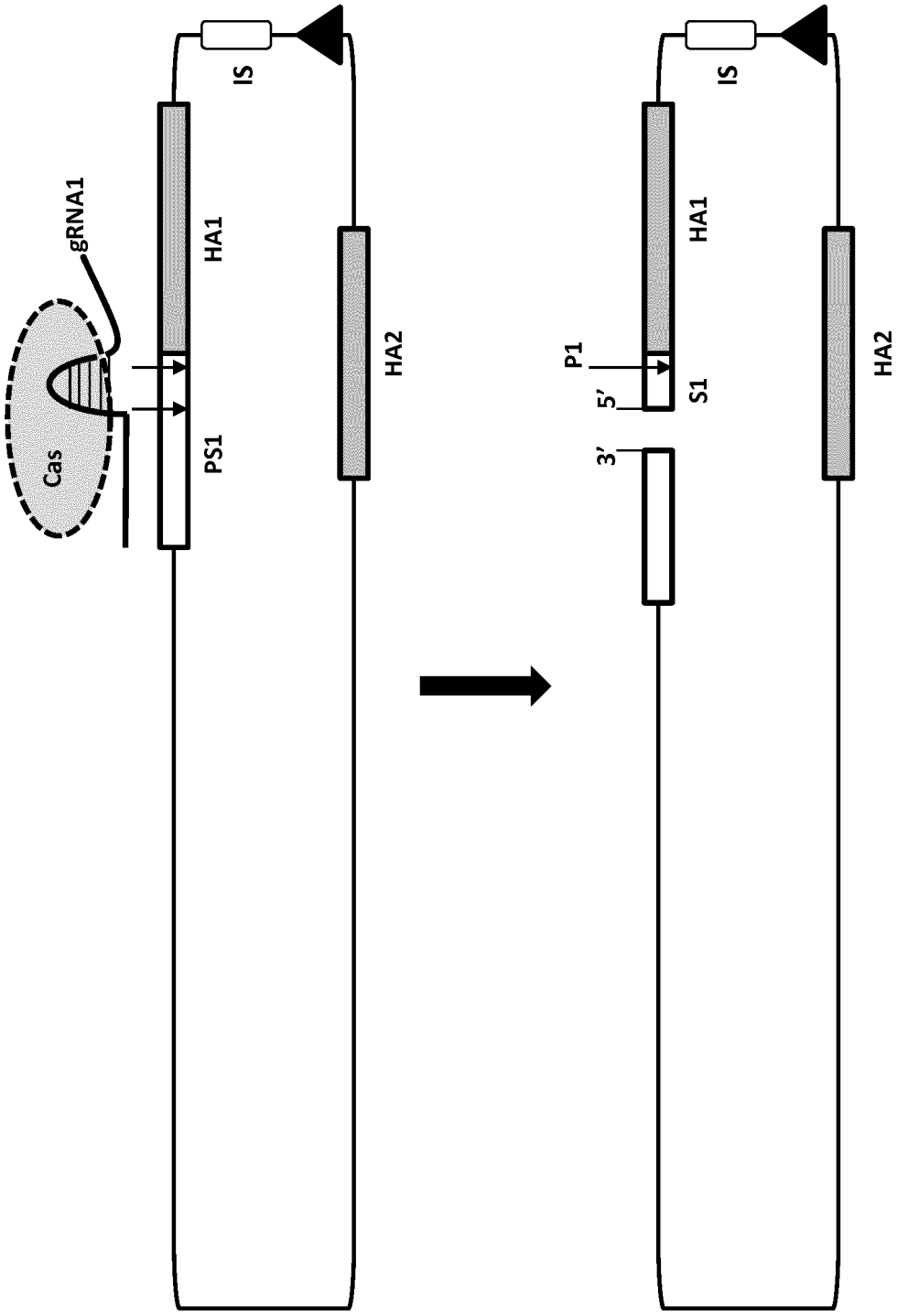


图 6A

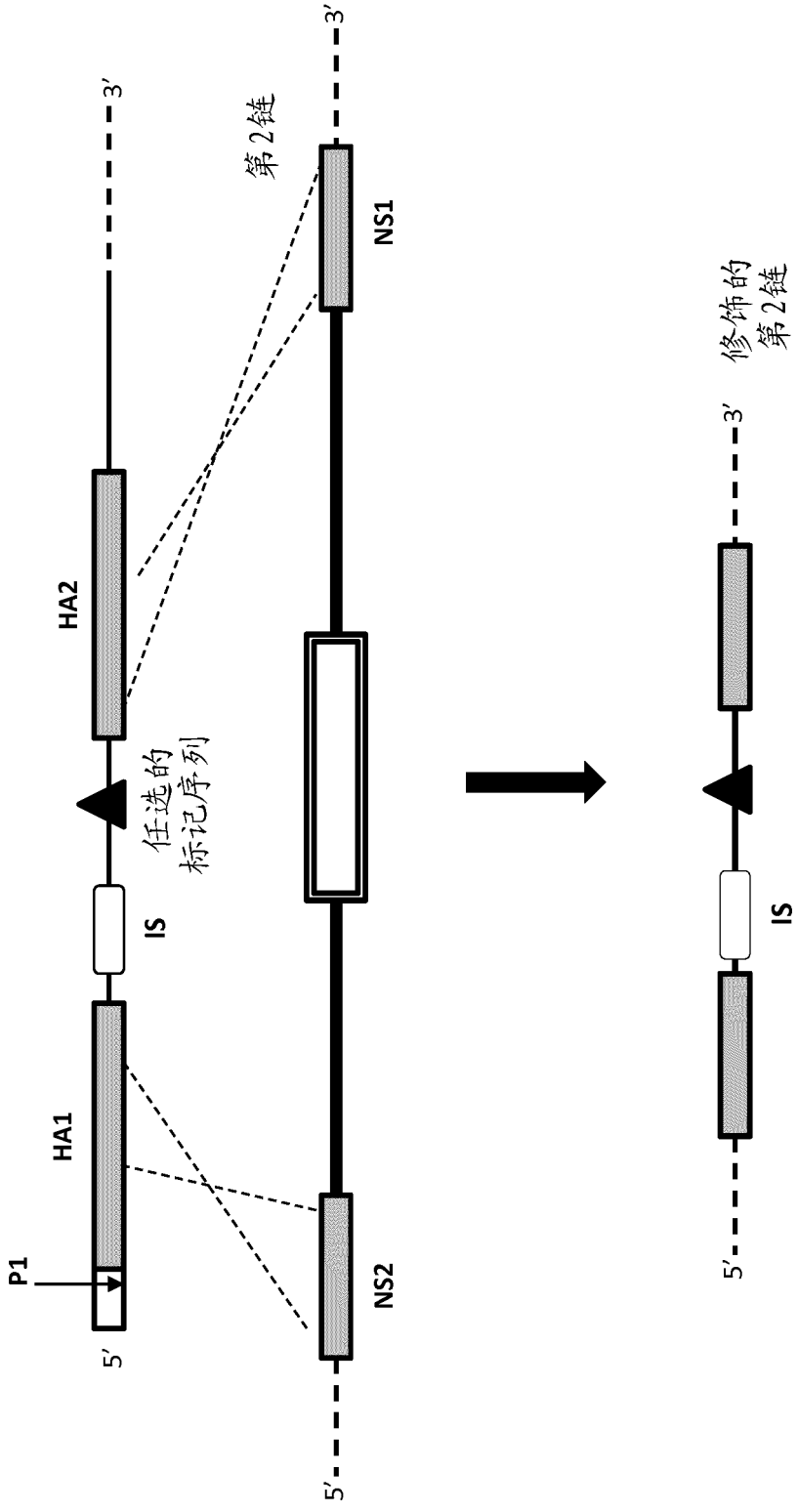


图 6B

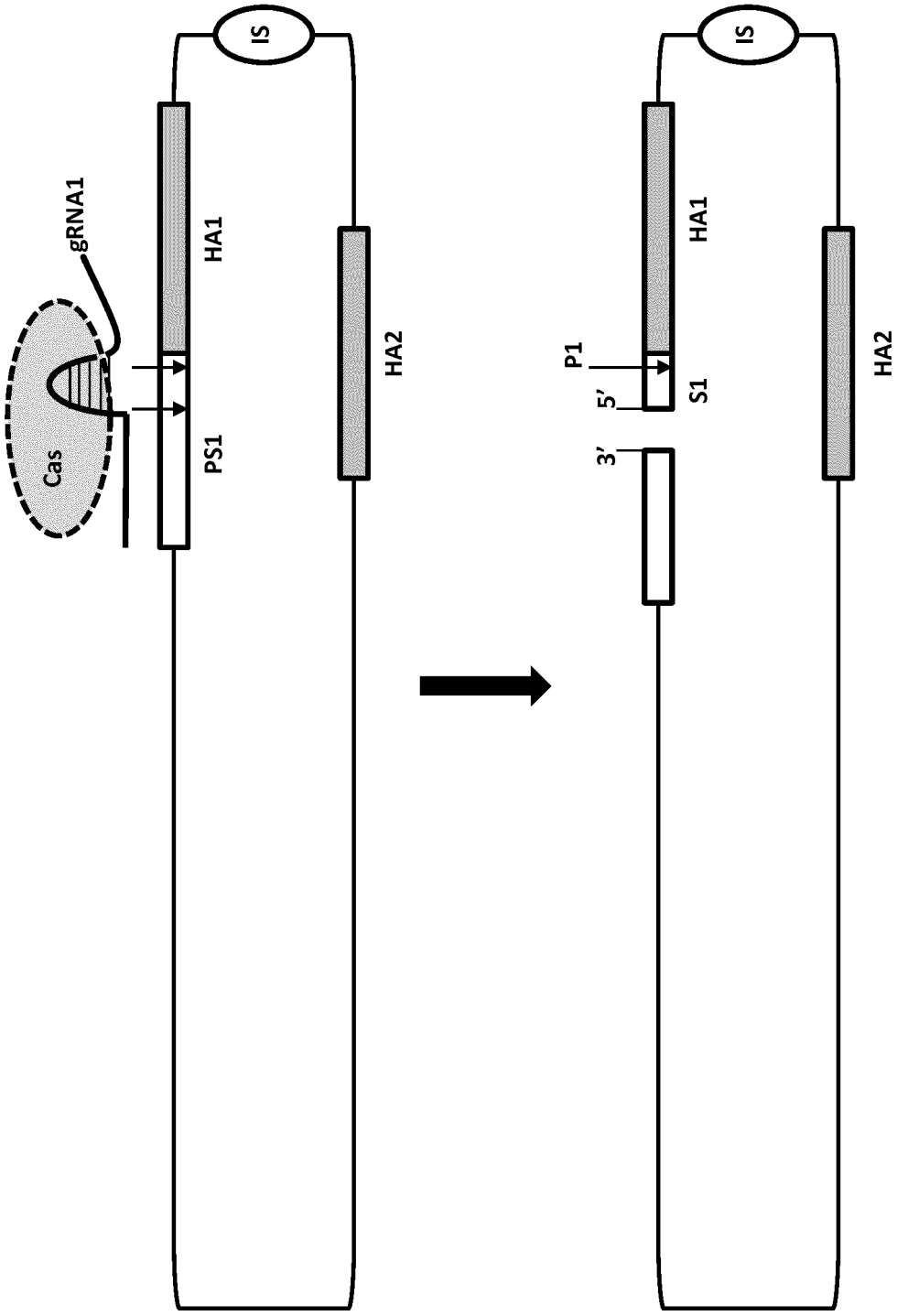


图 7A

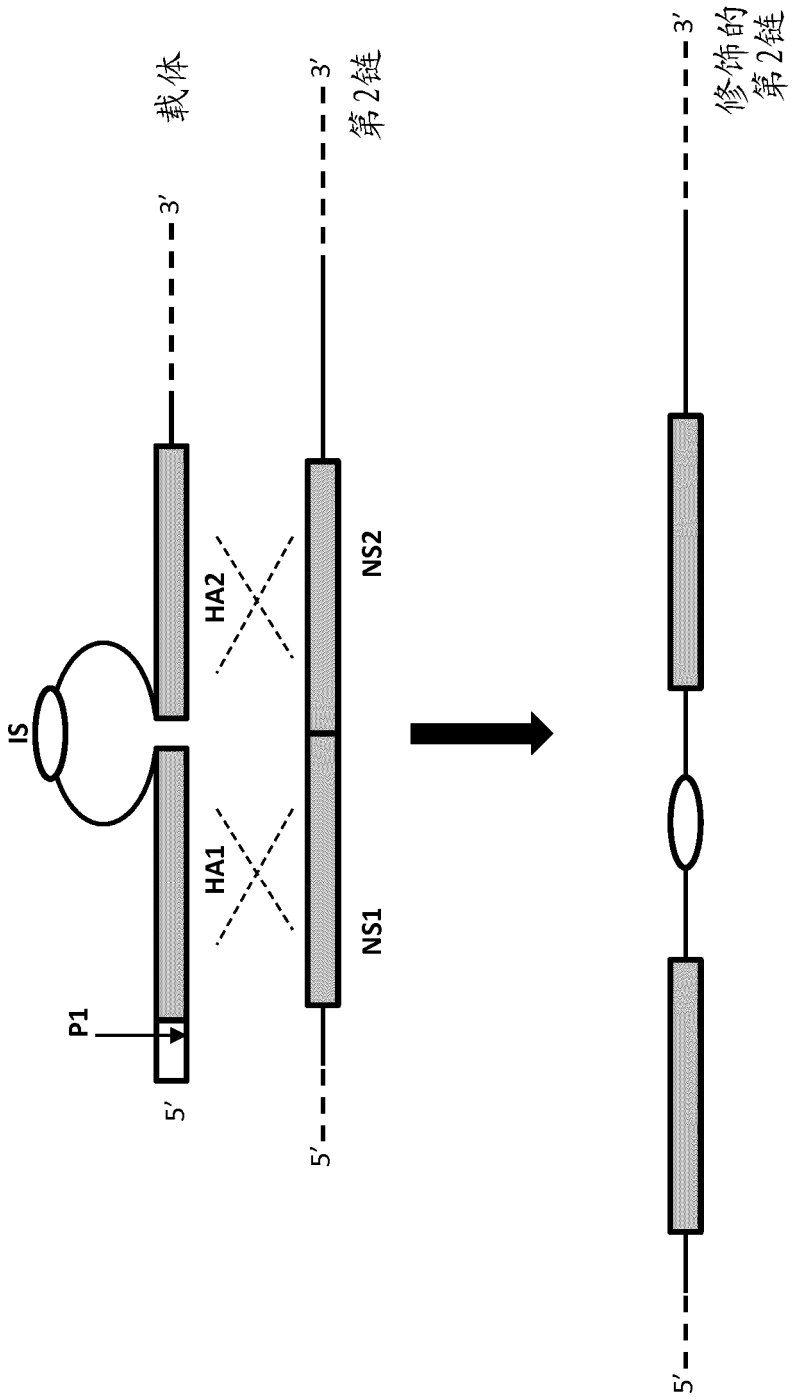


图 7B

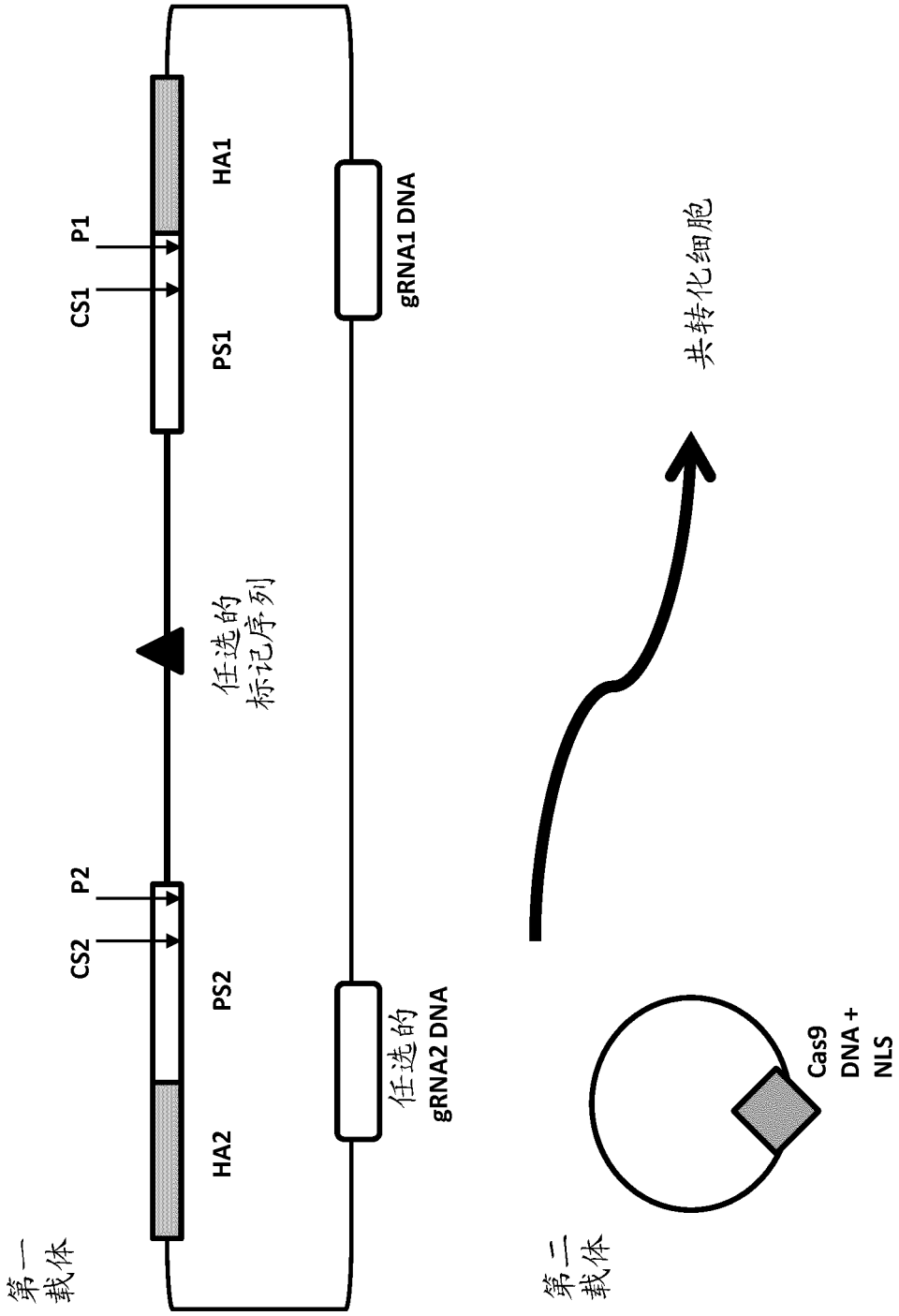

第一载体
HA2
PS2
CS2
P2
任选的gRNA2 DNA
任选的标记序列
HA1
PS1
CS1
P1
gRNA1 DNA
第二载体
Cas9 DNA + NLS
共转化细胞

图 8A

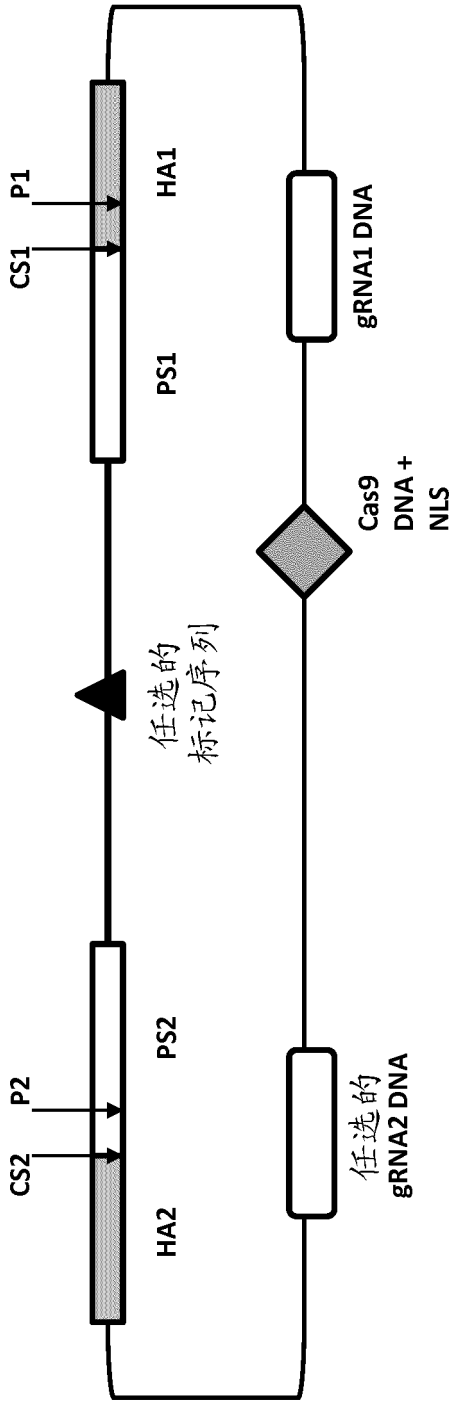

P1
CS1
HA1
PS1
gRNA1 DNA
Cas9
DNA +
NLS
任选的
标记序列
P2
CS2
PS2
HA2
任选的
gRNA2 DNA

图 8B

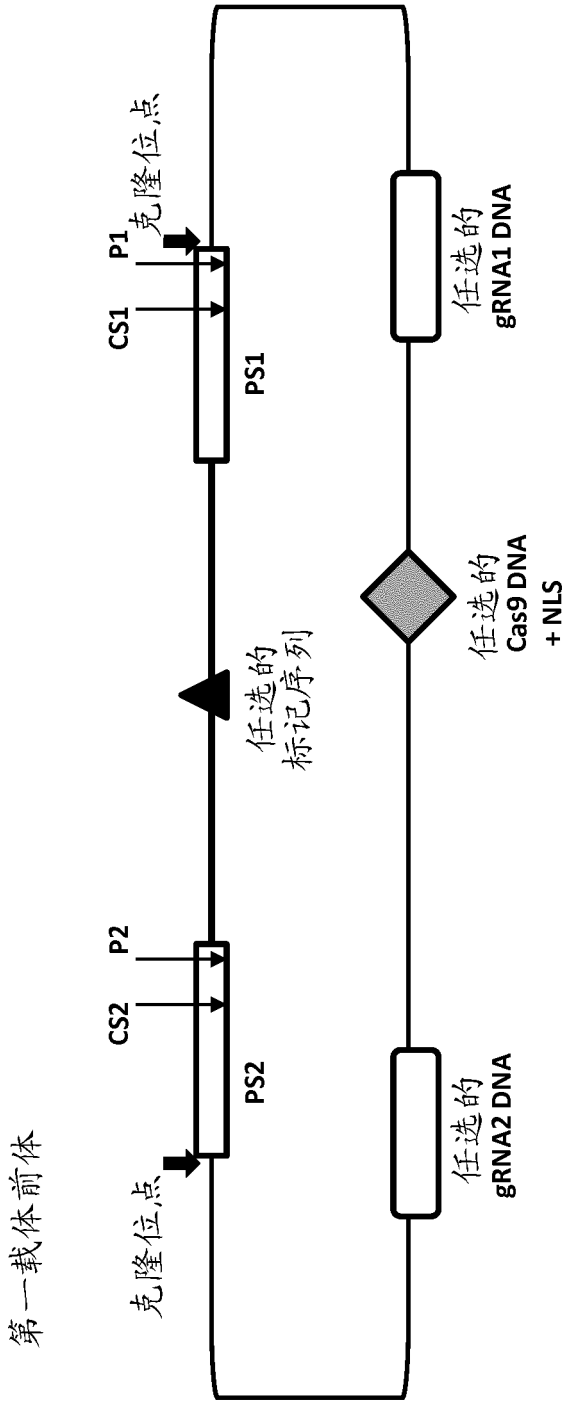

第一载体前体
克隆位点
P1
CS1
PS1
任选的
gRNA1 DNA
任选的
Cas9 DNA
+ NLS
任选的
标记序列
P2
CS2
PS2
克隆位点
任选的
gRNA2 DNA

图 9A

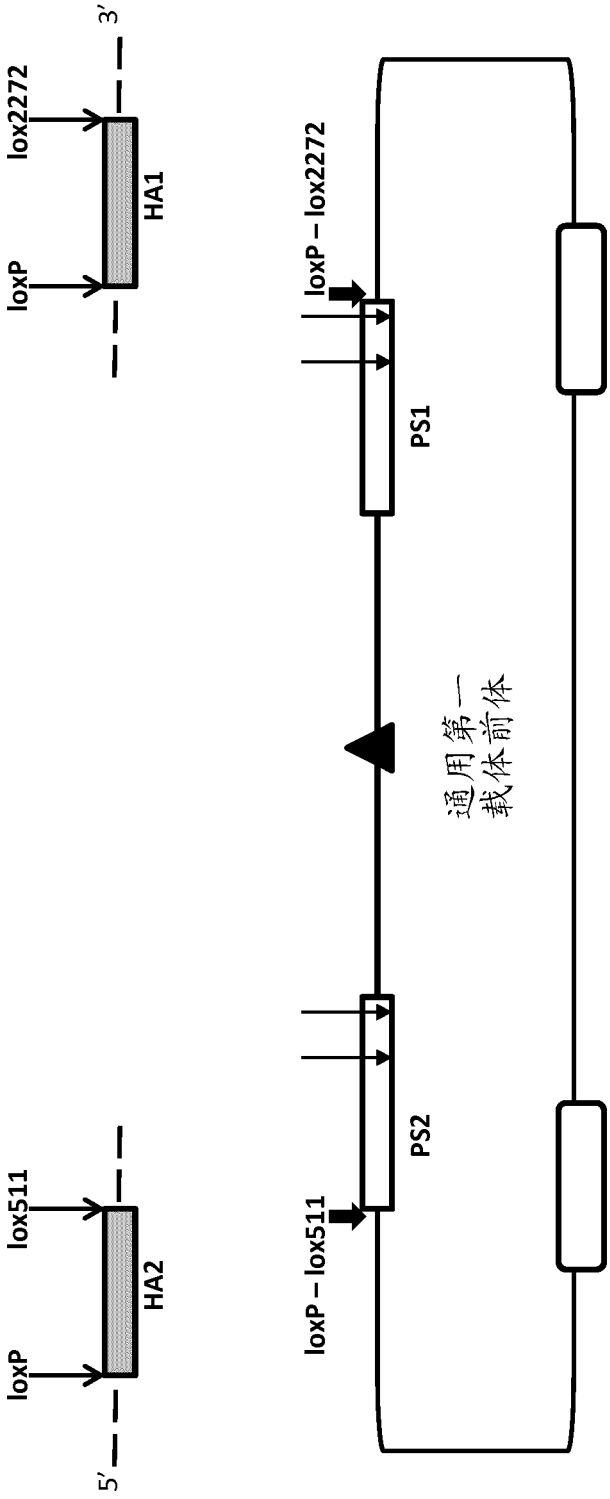

3′
lox2272
HA1
loxP
loxP – lox2272
PS1
通用第一载体前体
PS2
loxP – lox511
lox511
HA2
loxP
5′

图 9B

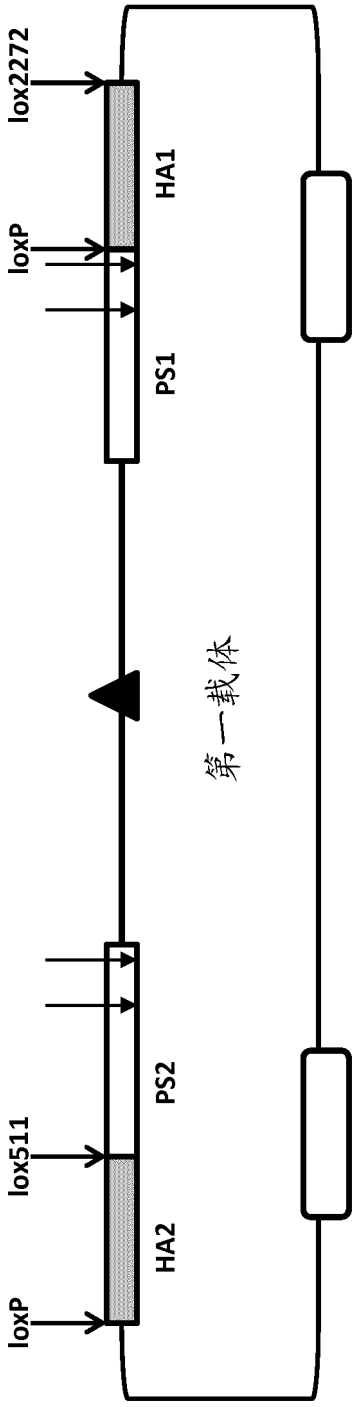


图 9C

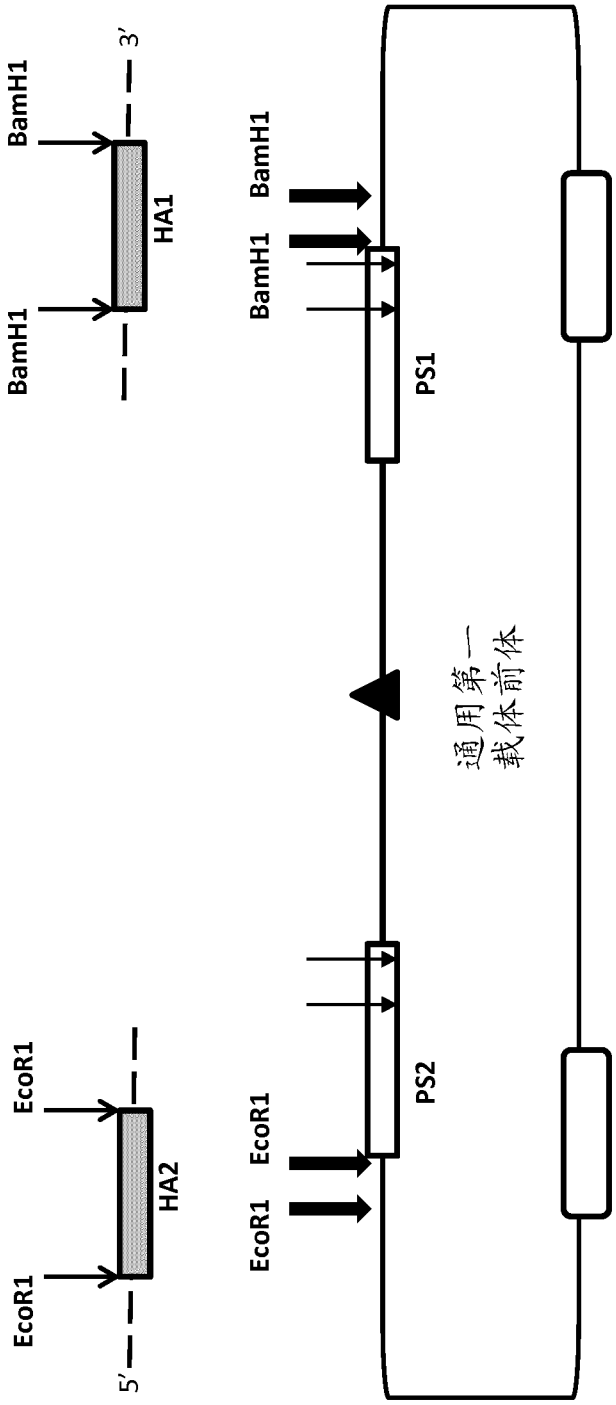

3′
BamH1
HA1
BamH1
BamH1
BamH1
PS1
通用第一
载体前体
EcoR1
HA2
EcoR1
EcoR1
EcoR1
PS2
5′

图 9D

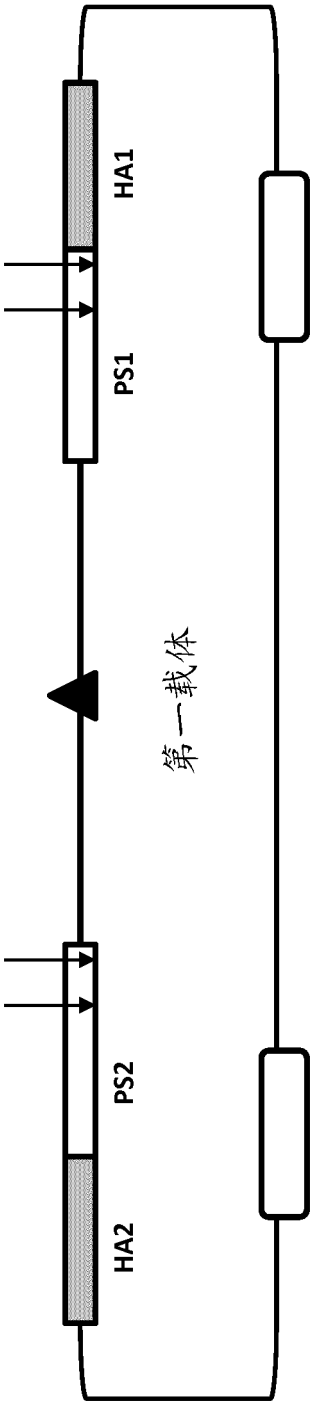


图 9E