



(12) 发明专利申请

(10) 申请公布号 CN 114277447 A

(43) 申请公布日 2022. 04. 05

(21) 申请号 202111573668.1

(22) 申请日 2021.12.21

(71) 申请人 翌圣生物科技(上海)股份有限公司
地址 200120 上海市浦东新区天雄路166弄
1号楼402室

(72) 发明人 宋东亮 刘倩 黄成 侯策 王嫚
孙睿 江翱 陈晶晶 曹振

(74) 专利代理机构 苏州根号专利代理事务所
(普通合伙) 32276

代理人 朱华庆

(51) Int. Cl.

C40B 50/06 (2006.01)

C40B 40/06 (2006.01)

C12N 15/113 (2010.01)

C12N 9/22 (2006.01)

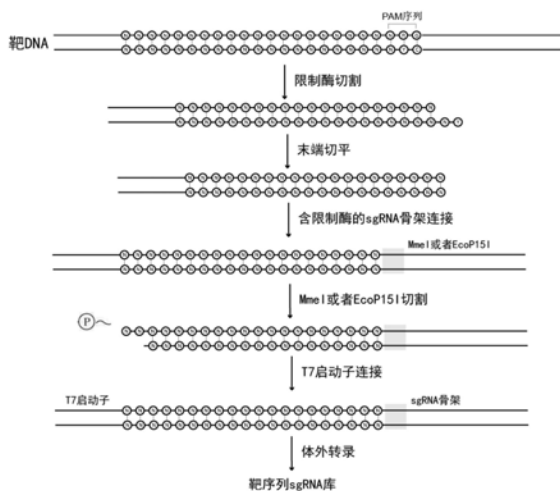
权利要求书2页 说明书9页 附图8页

(54) 发明名称

靶序列随机sgRNA全覆盖组的制备方法

(57) 摘要

本发明提供了一种靶序列随机sgRNA全覆盖组的制备方法,使用限制酶对目标序列的PAM区域进行切割;连接上sgRNA骨架;通过sgRNA骨架上的旁切活性限制酶位点进行靶向原间隔区域的序列获取;连接T7启动子;扩增获得带T7启动子的sgRNA文库;体外转录获得靶序列sgRNA文库。还公开了核糖体RNA的sgRNA文库制备方法以及去除RNA文库中核糖体RNA的方法,和在宿主基因组去除人全基因组的方法。本发明方法具有成本低、制作简单、覆盖均一、偏好性小、不受靶序列长度限制和无需大批量设计sgRNA等优点。



1. 一种靶序列随机sgRNA全覆盖组的制备方法,其特征在于其步骤包括:
 - (1) 采用识别PAM序列的限制酶切割样本DNA,并将末端切平;
 - (2) 步骤(1)获得的切平后的双链DNA3'端连接sgRNA骨架,其中sgRNA骨架带有旁切活性限制酶位点;
 - (3) 采用旁切活性限制酶切割步骤(2)的连接产物,获取原间隔区域DNA,并将其5'端磷酸化;
 - (4) 在原间隔区域DNA的5'端连接T7启动子序列;
 - (5) 扩增获得含T7启动子的sgRNA文库模板;
 - (6) 体外转录获得sgRNA文库。
2. 根据权利要求1所述的靶序列随机sgRNA全覆盖组的制备方法,其特征在于:步骤(1)中的限制酶为ScrFI、MspI、HpaII、BstNI、BfaI、DdeI中的一种或数种的混合物。
3. 根据权利要求1所述的靶序列随机sgRNA全覆盖组的制备方法,其特征为:步骤(1)中采用Mung Bean Nuclease将末端切平。
4. 根据权利要求1所述的靶序列随机sgRNA全覆盖组的制备方法,其特征为:步骤(2)的sgRNA骨架为两条单链DNA互补配对形成的双链DNA,其中正向序列为/rApp/-CGGTTGGAGCTAGAAATAGCAAGTCAACCTAACGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTT-/NH2C6/,反向序列为/NH2C6/-AAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCGTTAGGTTGACTTGCTATTTCTAGCTCCAACC-/ddG/,设有MmeI酶切位点;或者正向序列为/rApp/-CTGCTGGAGCTAGAAATAGCAAGTCAGCATAACGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTT-/NH2C6/,反向序列为/NH2C6/-AAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCGTTATGCTGACTTGCTATTTCTAGCTCCAGCA-/ddG/,设有EcoP15I酶切位点。
5. 根据权利要求6所述的靶序列随机sgRNA全覆盖组的制备方法,其特征为:步骤(2)中使用T4 DNA连接酶突变体K159L连接sgRNA骨架与双链DNA。
6. 根据权利要求1所述的靶序列随机sgRNA全覆盖组的制备方法,其特征为:步骤(3)的旁切活性限制酶为MmeI,步骤(4)中的T7启动子正向序列为/NH2C6/-TTCTAATACGACTCACTATAGGN,反向序列为/ddC/-CTATAGTGAGTCGTATTAGAA-/NH2C6/。
7. 根据权利要求1所述的靶序列随机sgRNA全覆盖组的制备方法,其特征为:步骤(3)的旁切活性限制酶为EcoP15I,步骤(4)中的T7启动子正向序列为/NH2C6/-TTCTAATACGACTCACTATAGG,反向序列为/ddN/-NCTATAGTGAGTCGTATTAGAA-/NH2C6/。
8. 根据权利要求1所述的靶序列随机sgRNA全覆盖组的制备方法,其特征为:步骤(4)采用T4 DNA连接酶进行连接。
9. 根据权利要求1所述的靶序列随机sgRNA全覆盖组的制备方法,其特征为:步骤(5)中使用正向序列为TTCTAATACGACTCACTATAGG,反向序列为AAAAGCACCGACTCGGTGCC的引物对进行文库扩增。
10. 根据权利要求1所述的靶序列随机sgRNA全覆盖组的制备方法,其特征为:步骤(6)中采用T7 RNA聚合酶进行转录,使用RNA回收磁珠回收sgRNA文库。
11. 一种核糖体RNA的sgRNA文库制备方法,其特征在于其步骤包括:
 - A、逆转录制备18S和28S全长cDNA;
 - B、PCR扩增获得18S和28S全长双链DNA;

C、采用权利要求1-10中任一项所述的方法,以18S和28S全长双链DNA为样本DNA,制得靶向覆盖18S和28S的sgRNA文库。

12. 一种去除RNA文库中核糖体RNA的方法,其特征在于其步骤包括:

A、将权利要求12制得的sgRNA文库与Cas9蛋白预组装;

B、采用预组装的Cas9-sgRNA切割RNA文库;

C、RNA文库扩增并测序。

13. 一种在宿主基因组去除人全基因组的方法,其特征在于其步骤包括:

A、采用权利要求1-10中任一项所述的方法,以人全基因组DNA为样本DNA,制得靶向覆盖人全基因组DNA的sgRNA文库;

B、制得的sgRNA文库与Cas9蛋白预组装;

C、采用预组装的Cas9-sgRNA切割宿主基因组DNA文库;

D、DNA文库扩增并测序。

靶序列随机sgRNA全覆盖组的制备方法

技术领域

[0001] 本发明专利涉及一种靶序列随机sgRNA全覆盖组的制备方法,属于生物技术领域。

背景技术

[0002] CRISPR基因编辑技术自从被开发出来就被广泛应用在基因治疗、体外诊断、基因捕获和靶基因去除等各个领域,并获得2020年诺贝尔生理医学奖,是一种高效且实用的技术。实用型CRISPR系统主要有两个部分组成,一个是具有两个核酸内切酶活性位点的Cas蛋白,负责切割靶位点DNA的两条链;另一个是具有与靶位点处DNA配对序列和Cas蛋白结合序列的引导RNA (sgRNA),负责募集Cas蛋白并引导Cas蛋白结合到互补配对的靶位点上。在CRISPR系统中,Cas蛋白先与sgRNA结合形成Cas-sgRNA复合物,并在DNA上进行检索。当检索到与sgRNA互补配对的区域(原间隔区域,ProtoSpacer),并且ProtoSpacer区域的3'端存在NGG序列(PAM序列)时,Cas蛋白将靶位点进行解旋,使得解开的双链DNA进入到Cas蛋白的DNA切割活性结构域,Cas蛋白对双链DNA进行切割,产生双链DNA断裂。断裂的DNA通过同源重组修复HR或者非同源末端连接NHEJ等DNA损伤修复方式完成靶基因的编辑。目前用于商业化应用的Cas蛋白主要有Cas9、Cas12、Cas13和Cas14及其变体,不同的Cas蛋白识别的PAM序列和要求,ProtoSpacer的长度也不同。因此,不同的CRISPR系统的具有不同的应用场景。

[0003] 除Cas蛋白的纯化和制备外,sgRNA的体外构建和合成也是CRISPR商业化应用的重要环节。常规的方法需要利用引物合成的方式合成含T7启动子的靶标sgRNA引物,再利用重叠PCR获得全长sgRNA骨架模板,并利用体外转录的方法获得需要的sgRNA。这种方法耗时短、成本低、可控性高,已经大规模商业化,成为sgRNA体外制备的主要形式,但仍存在通量低的问题。基因捕获或者去除通常需要对大区域的基因组进行捕获或者去除,需要覆盖长度达1M bp甚至整个完整的基因组DNA。这需要设计和合成千万级别条数的sgRNA。在sgRNA的设计合成上,无论成本还是技术都是很大的挑战。

发明内容

[0004] 本发明的目的是提供一种靶序列随机sgRNA全覆盖组的制备方法。

[0005] 本发明采取的技术方案为:

[0006] 一种靶序列随机sgRNA全覆盖组的制备方法,其特征在于其步骤包括:

[0007] (1) 采用识别PAM序列的限制酶切割样本DNA,并将末端切平,切平是指去除双链DNA上的3'和5'单链悬垂,以产生钝端;

[0008] (2) 步骤(1)获得的切平后的双链DNA3'端连接sgRNA骨架,其中sgRNA骨架带有旁切活性限制酶位点;

[0009] (3) 采用旁切活性限制酶切割步骤(2)的连接产物,获取原间隔区域DNA,并将其5'端磷酸化;

[0010] (4) 在原间隔区域DNA的5'端连接T7启动子序列;

[0011] (5) 扩增获得含T7启动子的sgRNA文库模板;

[0012] (6) 体外转录获得sgRNA文库。

[0013] 优选的,步骤(1)中的限制酶为ScrFI、MspI、HpaII、BstNI、BfaI、DdeI中的一种或数种的混合物。

[0014] 优选的,步骤(1)中采用Mung Bean Nuclease将末端切平。

[0015] 优选的,步骤(2)的sgRNA骨架为两条单链DNA互补配对形成的双链DNA,其中正向序列为/rApp/-CGGTTGGAGCTAGAAATAGCAAGTCAACCTAACGCTAGTCCGTTATCAACTTGAAAAAGTGGC ACCGAGTCGGTGCTTTT-/NH2C6/,反向序列为/NH2C6/-AAAAGCACCGACTCGGTGCCACTTTTTCAAGT TGATAACGGACTAGCGTTAGGTTGACTTGCTATTTCTAGCTCCAACC-/ddG/,设有MmeI酶切位点;或者正向序列为/rApp/-CTGCTGGAGCTAGAAATAGCAAGTCAGCATAACGCTAGTCCGTTATCAACTTGAAAAAGTGG GCACCGAGTCGGTGCTTTT-/NH2C6/,反向序列为/NH2C6/-AAAAGCACCGACTCGGTGCCACTTTTTCAA GTTGATAACGGACTAGCGTTATGCTGACTTGCTATTTCTAGCTCCAGCA-/ddG/,设有EcoP15I酶切位点。

[0016] 优选的,步骤(2)中使用T4 DNA连接酶突变体K159L连接sgRNA骨架与双链DNA。

[0017] 优选的,步骤(3)的旁切活性限制酶为MmeI,步骤(4)中的T7启动子正向序列为/NH2C6/-TTCTAATACGACTCACTATAGGN N,反向序列为/ddC/-CTATAGTGAGTCGTATTAGAA-/NH2C6/。

[0018] 优选的,步骤(3)的旁切活性限制酶为EcoP15I,步骤(4)中的T7启动子正向序列为/NH2C6/-TTCTAATACGACTCACTATAGG,反向序列为/ddN/-NCTATAGTGAGTCGTATTAGAA-/NH2C6/。

[0019] 优选的,步骤(4)采用T4 DNA连接酶进行连接。

[0020] 优选的,步骤(5)中使用正向序列为TTCTAATACGACTCACTATAGG,反向序列为AAAAGCACCGACTCGGTGCC的引物对进行文库扩增。

[0021] 优选的,步骤(6)中采用T7 RNA聚合酶进行转录,使用RNA回收磁珠回收sgRNA文库。

[0022] 本发明还公开了一种核糖体RNA的sgRNA文库制备方法,其特征在于其步骤包括:

[0023] A、逆转录制备18S和28S全长cDNA;

[0024] B、PCR扩增获得18S和28S全长双链DNA;

[0025] C、采用上述的方法,以18S和28S全长双链DNA为样本DNA,制得靶向覆盖18S和28S的sgRNA文库。

[0026] 本发明还公开了一种去除RNA文库中核糖体RNA的方法,其特征在于其步骤包括:

[0027] A、将上面制得的sgRNA文库与Cas9蛋白预组装;

[0028] B、采用预组装的Cas9-sgRNA切割RNA文库;

[0029] C、RNA文库扩增并测序。

[0030] 本发明还公开了一种在宿主基因组去除人全基因组的方法,其特征在于其步骤包括:

[0031] A、采用上述的方法,以人全基因组DNA为样本DNA,制得靶向覆盖人全基因组DNA的sgRNA文库;

[0032] B、制得的sgRNA文库与Cas9蛋白预组装;

[0033] C、采用预组装的Cas9-sgRNA切割宿主基因组DNA文库;

[0034] D、DNA文库扩增并测序。

[0035] 本发明的有益效果为：

[0036] 本发明公开了一种靶序列随机sgRNA制备方法RPTS (Random sgRNA Preparation of Target Sequence), 可以一次性制备随机覆盖目标区域的所有sgRNA组。RPTS的原理和流程如下：使用限制酶对目标序列的PAM区域进行切割；连接上sgRNA骨架；通过sgRNA骨架上的旁切活性限制酶位点进行靶向原间隔区域的序列获取；连接T7启动子；扩增获得带T7启动子的sgRNA文库；体外转录获得靶序列sgRNA文库。RPTS具有成本低、制作简单、覆盖均一、偏好性小、不受靶序列长度限制和无需大批量设计sgRNA等优点。本发明还公开了RPTS在核糖体RNA去除和宿主基因组去除上的应用流程，打破了CRISPR技术在这些领域上的应用限制。

附图说明

[0037] 图1正常sgRNA结合靶位点图。

[0038] 图2包含MmeI酶切位点的sgRNA改造骨架。

[0039] 图3包含EcoP15I酶切位点的sgRNA改造骨架。

[0040] 图4 RPTS原理和流程示意图。

[0041] 图5 RPTS技术中使用限制酶识别位点汇总。阴影表示该限制酶识别位点中包含PAM序列类型。

[0042] 图6 18S rRNA的DNA上包含RPTS技术中使用限制酶的识别位点分布。

[0043] 图7 18S和28S cDNA扩增结果，左侧条带为Marker，中间条带为28S，右边条带为18S。

[0044] 图8 18S/28S DNA和人全基因组DNA (右) 的RPTS文库扩增结果，左侧条带为Marker，中间条带为18S/28S DNA，右边条带为人全基因组DNA。

[0045] 图9 18S/28S DNA (左) 和人全基因组DNA (右) 的体外转录结果，左侧条带为Marker，中间条带为18S/28S DNA，右边条带为人全基因组DNA。

[0046] 图10 18S/28S RPTS库在CRISPR去除rRNA应用上的文库分布。浅色是未加RPTS库，深色是加入RPTS库。

[0047] 图11 RNA-seq验证18S/28S RPTS库在CRISPR去除rRNA应用上去除效果。

[0048] 图12人类全基因组RPTS库在CRISPR去除宿主基因组应用上的文库分布。浅色是未加RPTS库，深色是加入RPTS库。

[0049] 图13 DNA-seq验证人全基因组RPTS库在CRISPR去除宿主基因组应用上去除效果。

具体实施方式

[0050] 为了进一步描述使本发明的具体内容，以下结合实施例对本发明进行详细说明。实施例所涉及的操作方法及试剂为业内技术人员所熟知，应当理解，以下所描述的具体实施例仅仅用于对发明进行详细说明，但本发明的实施方式并不受下述实施例的限制。本实施例所使用的接头序列及修饰如表1所示，N为A、T、C、G中的任意碱基。

[0051] 表1接头序列及修饰

	序列名称	5'-3'	修饰	
	SEQ NO: 1	18S-R	TAATGATCCTTCCGCAGGTTACCTACG	
	SEQ NO: 2	18S-F	TACCTGGTTGATCCTGCCAGTAG	
	SEQ NO: 3	28S-R	TAATGATCCTTCCGCAGGTTACCTACG	
	SEQ NO: 4	28S-F	CGCGACCTCAGATCAGACGTG	
[0052]	SEQ NO: 5	sgM-F	/rApp/-CGGTTGGAGCTAGAAATAGCAA GTCAACCTAACGCTAGTCCGTTATCAAC TTGAAAAAGTGGCACCGAGTCGGTGCTT TT-/NH2C6/	5' 端被预腺苷 酰化修饰 rApp, 3' 端被氨基修 饰 NH2C6 封闭
	SEQ NO: 6	sgM-R	/NH2C6/-AAAAGCACCGACTCGGTGCCA CTTTTCAAGTTGATAACGGACTAGCG TTAGGTTGACTTGCTATTTCTAGCTCC AACC-/ddG/	5' 端被氨基修 饰 NH2C6 封 闭, 3' 端被双 脱氧核苷酸修 饰 ddG 封闭
	SEQ NO: 7	sgE-F	/rApp/-CTGCTGGAGCTAGAAATAGCAA GTCAGCATAACGCTAGTCCGTTATCAAC TTGAAAAAGTGGCACCGAGTCGGTGCTT TT-/NH2C6/	5' 端被预腺苷 酰化修饰 rApp, 3' 端被氨基修 饰 NH2C6 封闭
	SEQ NO: 8	sgE-R	/NH2C6/-AAAAGCACCGACTCGGTGCCA CTTTTCAAGTTGATAACGGACTAGCG TTATGCTGACTTGCTATTTCTAGCTCC AGCA-/ddG/	5' 端被氨基修 饰 NH2C6 封 闭, 3' 端被双 脱氧核苷酸修 饰 ddG 封闭
	SEQ NO: 9	T7M-F	/NH2C6/-TTCTAATACGACTCACTATAGGNN	5' 端被氨基修 饰 NH2C6 封闭
[0053]	SEQ NO: 10	T7M-R	/ddC/-CTATAGTGAGTCGTATTAG AA-/NH2C6/	5' 端被双脱氧 核苷酸修饰 ddC 封闭, 3' 端被 氨基修饰 NH2C6 封闭
	SEQ NO: 11	T7E-F	/NH2C6/-TTCTAATACGACTCACTATAGG	5' 端被氨基修 饰 NH2C6 封闭
	SEQ NO: 12	T7E-R	/ddN/-NCTATAGTGAGTCGTATTAG AA-/NH2C6/	5' 端被双脱氧 核苷酸修饰 ddN 封闭, 3' 端被氨基修饰 NH2C6 封闭
	SEQ NO: 13	sgPCR-F	TTCTAATACGACTCACTATAGG	
	SEQ NO: 14	sgPCR-R	AAAAGCACCGACTCGGTGCC	

[0054] 实施例1:sgRNA骨架的设计和RPTS的流程

[0055] 本实施例改造了传统sgRNA的骨架,融入了酶切位点MmeI或EcoP15I,但不改变sgRNA的结构和Cas9的结合。改造后的sgRNA序列和结构见图1-图3所示。

[0056] 测试RPTS的流程,流程如图4所示:

[0057] (1) 识别PAM序列的限制酶切割:

[0058] 表2限制酶酶切体系

[0059]	组分	用量
	PUC19 plasmid DNA	1 μ g
	10 \times rCutSmart buffer	8 μ L

ScrFI (NEB)	5U
MspI (NEB)	5U
HpaII (NEB)	5U
BstNI (NEB)	5U
BfaI (NEB)	5U
DdeI (NEB)	5U
补水至	80 μ L

[0060] 37 $^{\circ}$ C反应过夜。反应结束后,加入160 μ L Ampure DNA beads (Beckman)回收片段化的DNA。45 μ L水洗脱。

[0061] 表3末端切平体系

[0062]

组分	用量
上述回收的DNA	43 μ L
10 \times Mung Bean Nuclease Reaction Buffer	5 μ L
Mung Bean Nuclease (NEB)	20U
补水至	50 μ L

[0063] 30 $^{\circ}$ C反应2h。反应结束后,加入100 μ L Ampure DNA beads (Beckman)回收片段化的DNA。42 μ L水洗脱。

[0064] (2) 含旁切活性限制酶酶切位点的sgRNA骨架连接。

[0065] 骨架退火:sgM-F、sgM-R、sgE-F和sgE-R用100mM NaCl溶液溶解至100 μ M,取10 μ LsgM-F和10 μ L sgM-R于PCR管中,取10 μ L sgE-F和10 μ L sgE-R于另一PCR管中,在同一反应体系中,MmeI-sgRNA骨架和EcoP15I-sgRNA骨架二选一。95 $^{\circ}$ C反应5min,每分钟降低1 $^{\circ}$ C。反应结束后,用水将退火好的骨架稀释至10 μ M。

[0066] 骨架连接:

[0067] 表4骨架连接体系

[0068]

组分	用量
上述回收的DNA	40 μ L
10 μ M sgRNA骨架接头	5 μ L
10 \times T4 Ligase Reaction Buffer	5 μ L
T4 DNA Ligase (K159L)	2000U
补水至	50 μ L

[0069] 20 $^{\circ}$ C反应1h。反应结束后,加入22.5 μ L Ampure DNA beads回收DNA,去除掉未连接的接头。22 μ L水洗脱。

[0070] (3) MmeI或EcoP15I酶切

[0071] 表5 MmeI酶切体系

[0072]

组分	用量
上述回收的含MmeI-sgRNA骨架的DNA	20 μ L
10 \times rCutSmart buffer	3 μ L
MmeI	5U

补水至	30 μ L
-----	------------

[0073] 表6 EcoP15I酶切体系

组分	用量
上述回收的含 EcoP15I-sgRNA 骨架的 DNA	20 μ L
10 \times rCutSmart buffer	3 μ L
10 \times ATP	3 μ L
EcoP15I	20 U

补水至	30 μ L
-----	------------

[0076] 37 $^{\circ}$ C 反应2h。反应结束后,加入60 μ L Ampure DNA beads回收DNA。42 μ L水洗脱。

[0077] (4) T7启动子连接。

[0078] 接头退火:T7M-F、T7M-R、T7E-F和T7E-R用100mM NaCl溶液溶解至100 μ M,取10 μ L T7M-F和10 μ L T7M-R于PCR管中,取10 μ L T7E-F和10 μ L T7E-R于另一PCR管中,根据上面的sgRNA骨架的类型选用合适的T7接头。95 $^{\circ}$ C 反应5min,每分钟降低1 $^{\circ}$ C。反应结束后,用水将退火好的接头稀释至10 μ M。

[0079] 接头连接:

[0080] 表7接头连接体系

组分	用量
上述回收的DNA	40 μ L
10 μ M对应的T7接头	5 μ L
10 \times T4 Ligase Reaction Buffer	5 μ L
T4 DNA Ligase	2000U
补水至	50 μ L

[0082] 20 $^{\circ}$ C 反应1h。反应结束后,加入22.5 μ L Ampure DNA beads回收DNA,去除掉未连接的接头。22 μ L水洗脱。

[0083] (5) sgRNA库扩增

[0084] 表8 sgRNA文库扩增体系

组分	用量
上述回收的DNA	20 μ L
10 μ M sgPCR-F/R	5 μ L
Phusion High-Fidelity PCR Master Mix	25 μ L
补水至	50 μ L

[0086] 在98 $^{\circ}$ C 变性3min后,通过98 $^{\circ}$ C 10s变性、60 $^{\circ}$ C 20s退火和72 $^{\circ}$ C 10s延伸进行文库循环扩增。扩增产物使用70 μ L Ampure DNA beads进行回收,22 μ L DEPC水洗脱。

[0087] (6) sgRNA库体外转录

[0088] 表9 sgRNA文库体外转录体系

组分	用量
上述回收的DNA	1 μ g

10×Transcription Buffer	2μL
CTP/GTP/ATP/UTP (100mM each)	2μL
T7 RNA Polymerase Mix (Yeasen)	2μL
补水至	20μL

[0090] 37℃反应4h后,加入10U DNase I (TAKARA),37℃反应1h。加入50μL Ampure RNA beads (Beckman) 回收RNA。琼脂糖凝胶电泳检测sgRNA大小。

[0091] 如图5所述,我们设计的识别PAM序列 (NRG) 的限制酶组合包含6种限制性内切酶,这6种限制酶位点包括了Cas9蛋白的两个PAM序列。在DNA上,约每64bp就存在这样的限制酶位点,因此制备的sgRNA可以随机覆盖到整个基因组上。图6展示了18S rRNA上这6中限制酶位点的分布(阴影部分)。

[0092] 实施例2:18S rRNA或28S rRNA的RPTS制备。

[0093] 在本实施例中,我们利用RPTS制备了覆盖18S rRNA或28S rRNA的随机sgRNA库。具体实施方式如下:

[0094] (1) DNA片段的获取

[0095] 表10逆转录体系

[0096]

组分	用量
293细胞RNA	1μg
10μM逆转录引物18S-R或者28S-R	1μL
10mM dNTPs	1
75℃反应5min	
5×FS Buffer	4μL
0.1M DTT	1μL
SuperScript IV (Thermo)	2μL
总体积	20μL

[0097] 42℃15min,50℃15min,55℃15min,50℃15min,55℃15min,70℃15min。

[0098] 表11 PCR扩增体系

[0099]

组分	用量
上述逆转录产物	1μL
10μM 18S-F/R或者28S-F/R	5μL
Phusion High-Fidelity PCR Master Mix	25μL
加水至	50μL

[0100] 在98℃变性3min后,通过98℃10s变性、60℃20s退火和72℃3min延伸进行文库循环扩增。扩增产物使用35μL Ampure DNA beads进行回收,22μL DEPC水洗脱。

[0101] RPTS制备sgRNA库:按照实施例1的方法制备18S sgRNA库或28S sgRNA库。

[0102] (3) RNA文库制备。使用翌圣生物的双模式RNA建库试剂盒 (12252) 进行RNA文库构建。连接DNA接头后,使用0.6×Ampure DNA beads回收文库。

[0103] (4) CRISPR去除18S和28S DNA。

[0104] 表12 sgRNA库与Cas9蛋白预组装体系:

[0105]	组分	用量
	18SsgRNA库和28S sgRNA库	2-10 μ g
	Cas9 (NEB)	0.2-0.5 μ g
	500mM氯化钠	1 μ L
	总体积	5 μ L

[0106] 37 $^{\circ}$ C 30min。

[0107] 表13 CRISPR切割体系

[0108]	组分	用量
	Cas9-sgRNA库	5 μ L
	RNA文库	1-100ng
	10 \times NEB buffer 3.1	1 μ L
	总体积	10 μ L

[0109] 37 $^{\circ}$ C 30-90min, 90 $^{\circ}$ C 10min。

[0110] (5) 文库扩增

[0111] 表14文库扩增体系

[0112]	组分	用量
	上述反应体系	10 μ L
	Index Primer F/R (Yeasen, 12610)	5 μ L
	2 \times Canase PCR mix	25 μ L
	补水至	50 μ L

[0113] 在98 $^{\circ}$ C变性3min后,通过98 $^{\circ}$ C 10s变性、60 $^{\circ}$ C 30s退火和72 $^{\circ}$ C 30s延伸进行文库循环扩增。扩增产物使用45 μ L Ampure DNA beads进行回收,22 μ L DEPC水洗脱。

[0114] 制备好的文库经过Qsep100质检后,在Illumina的NovaSeq 6000平台上进行测序并分析。

[0115] 18S和28S RPTS结果如图7、图8和图9所示,RPTS技术能够成功构建18S或28S的sgRNA库。RNA建库结果和测序结果如图10和图11所示,RPTS方法制备的sgRNA文库能够有效去除18S和28S rRNA。

[0116] 实施例3:人全基因组的RPTS制备。

[0117] 在本实施例中,我们利用RPTS制备了覆盖人全基因组的随机sgRNA库。具体实施方式如下:

[0118] (1) RPTS制备sgRNA库:按照实施例1的方法制备sgRNA库,DNA使用人基因组DNA标准品NA12878 (Coriell)。

[0119] (2) DNA文库制备。DNA使用人基因组DNA标准品NA12878 (Coriell) 和大肠杆菌基因组按照100:1比例混合的DNA标准品混合物。使用翌圣生物的一步法建库试剂盒(12204)进行DNA文库构建。连接DNA接头后,使用0.6 \times Ampure DNA beads回收文库。

[0120] (3) CRISPR去除人基因组DNA

[0121] 表15 sgRNA库与Cas9蛋白预组装体系:

[0122]	组分	用量
--------	----	----

人基因组RPTS库	2-10 μ g
Cas9 (NEB)	0.2-0.5 μ g
500mM氯化钠	1 μ L
总体积	5 μ L

[0123] 37 $^{\circ}$ C 30min。

[0124] 表16 CRISPR切割体系

组分	用量
Cas9-sgRNA库	5 μ L
RNA文库	1-100ng
10 \times NEB buffer 3.1	1 μ L
总体积	10 μ L

[0126] 37 $^{\circ}$ C 30-90min, 90 $^{\circ}$ C 10min。

[0127] (5) 文库扩增

[0128] 表17文库扩增体系

组分	用量
上述反应体系	10 μ L
Index Primer F/R (Yeasen, 12610)	5 μ L
2 \times Canace pro PCR mix	25 μ L
补水至	50 μ L

[0130] 在98 $^{\circ}$ C变性3min后,通过98 $^{\circ}$ C 10s变性,60 $^{\circ}$ C 30s退火和72 $^{\circ}$ C 30s延伸进行文库循环扩增。扩增产物使用45 μ L Ampure DNA beads进行回收,22 μ L DEPC水洗脱。

[0131] 制备好的文库经过Qsep100质检后,在Illumina的NovaSeq 6000平台上进行测序并分析。

[0132] DNA建库结果和测序结果如图11和图12所示,RPTS方法制备的sgRNA文库能够有效去除DNA建库过程中的人类宿主基因组DNA。

[0133] 综上,本发明公开了一种靶序列随机sgRNA制备方法RPTS (Random sgRNA Preparation of Target Sequence),可以一次性制备随机覆盖目标区域的所有sgRNA组。RPTS的原理和流程如下:使用限制酶对目标序列的PAM区域进行切割;连接上sgRNA骨架;通过sgRNA骨架上的旁切活性限制酶位点进行靶向原间隔区域的序列获取;连接T7启动子;扩增获得带T7启动子的sgRNA文库;体外转录获得靶序列sgRNA文库。RPTS具有成本低、制作简单、覆盖均一、偏好性小、不受靶序列长度限制和无需大批量设计sgRNA等优点。本发明还公开了RPTS在核糖体RNA去除和宿主基因组去除上的应用流程,打破了CRISPR技术在这些领域上的应用限制。

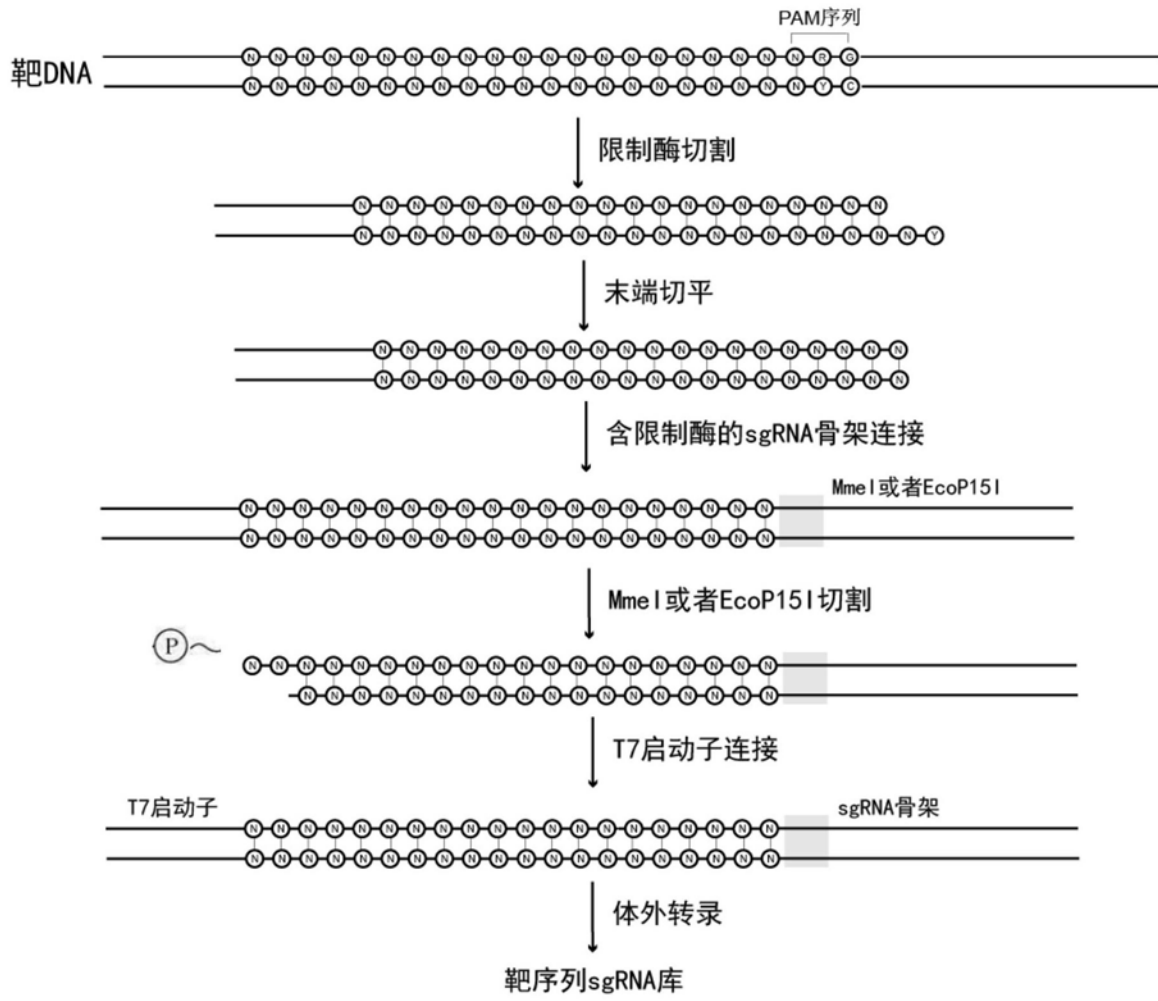


图4

ScrF1



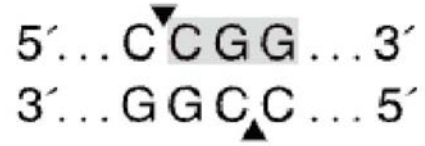
Msp1



BstN1



HpaII



DdeI



BfaI



图5

TA CCTGG TTGATCCTGCCAGTAGCATATGCTTGTCTCAAAGATTAAGCCATGCATGTCTAAGTACGCACGGC
 CGGTACAGTGAAACTGCGAATGGCTCATTAAATCAGTTATGGTTCCTTTGGTCGCTCGCTCCTCTCCTACTT
 GGATAACTGTGGAATTCTAGAGCTAATACATGCCGACGGGCGCTGACCCCTTCGCGGGGGGATGCGT
 GCATTTATCAGATCAAACCAA CCGGTCAGCCCTCT CCGGCC CCGGCCGGGGGGCGGGCG CCGGCGG
 CTTTGGTGACTCTAGATAACCTCGGGCCGATCGCACGCCCCCGTGCGGGCGACGACCCATTGGAACGTC
 TGCCCTATCAACTTTCGATGGTAGTCGCCGTGCCTACCATGGTGACCACGGGTGACGGGGAATCAGGGTTC
 GATTCCGGGAGAGGGAGCCCTGAGAAACGGCTACCACATCCAAGGAAGGCAGCAGGCGCGCAAATTACCCA
 CTCCCGAC CCGGGGAGGTAGTGACGAAAAATAACAATACAGGACTCTTTCGAGGCCCTGTAATTGGAATG
 AGTCCACTTTAAATCCTTTAACGAGGATCCATTGGAGGGCAAGTCTGGTGCCAGCAGCCGCGTAATTCCA
 GCTCCAATAGCGTATATTAAGTTGCTGCAGTTAAAAGCTCGTAGTTGGATCTTGGGAGCGGGCGGGCG
 GTCGCGCGGAGGCGAGCCACCGCCCGTCCCGCCCTTGCCTCTCGGCGCCCCCTCGATGCTCTTAGCT
 GAGTGTCCCGCGGGCCCGAAGCGTTTACTTTGAAAAAATTAGAGTGTTCAAAGCAGGCCCGCCGCT
 GGATACCGCAGCTAGGAATAATGGAATAGGACCGCGGTTCTATTTTGTGGTTTTCGGAACTGAGGCCATG
 ATTAAGAGGGACGG CCGGGGGCATTGATTGCGCCGCTAGAGGTGAAATCTTGGA CCGGCGCAAGACG
 GACCAGAGCGAAAGCATTGCCAAGAATGTTTTCAATCAAGAACGAAAGTCGGAGGTTTCAAGACGA
 TCAGATACCGTCGTAGTTCCGACCATAAACGATGCCGAC CCGGCGATGCGGCGGCGTTATTCCCATGACCC
 G CCGGGCAGCTT CCGGGAAACCAAAGTCTTTGGTT CCGGGGGGAGTATGGTTGCAAAGCTGAAACTTAA
 AGGAATTGACGGAAGGGCACCACCAGGAGTGGAGCCTGCGGCTTAATTTGACTCAACACGGGAAACCTCA
 CCCGGCCCCGACACGGACAGGATTGACAGATTGATAGCTCTTCTCGATTCCGTGGGTGGTGGTGCATGG
 CCGTTCTTAGTTGGTGGAGCGATTTGTCTGGTTAATTCGATAACGAACGAGACTCTGGCATGCTAACTAGT
 TACGCGACCCCGAGCGGTGCGCGTCCCCAACTTCTTAGAGGGACAAGTGCGTTTACGCCACCCGAGAT
 TGAGCAATAACAGGTCTGTGATGCCCTTAGATGT CCGGGGCTGCACGCGCGCTACACTGACTGGCTCAGC
 GTGTGCCCTACCCTACG CCGGCAGGCGCGGTAACCCGTTGAACCCATTCTGTATGGGGATCGGGGATTG
 CAATTATCCCATGAACGAGGAATCCAGTAAGTGCGGGTCATAAGCTTGCCTGATTAGTCCCTGCC
 CTTTGTACACACCGCCCGTCTACTACCGATTGGATGGTTTGTAGGAGCCCTCGGATCGGCCCCG CCGGG
 GTCGGCCACGGCCCTGGCGGAGCGCTGAGAAAGACGGTCGAACTTGACTATCTAGAGGAAGTAAAAGTC
 GTAACAAGTTTTCCGTAGGTGAACCTGCGGAAGGATCATTAA

图6

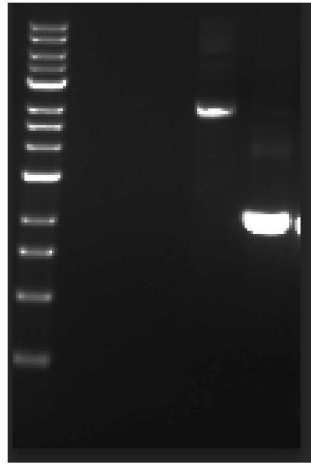


图7

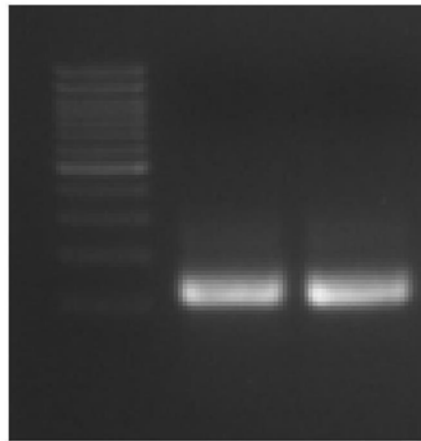


图8

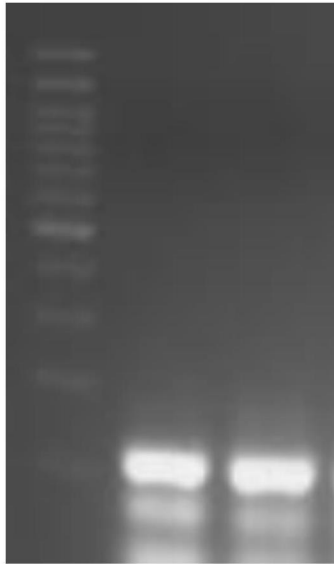


图9

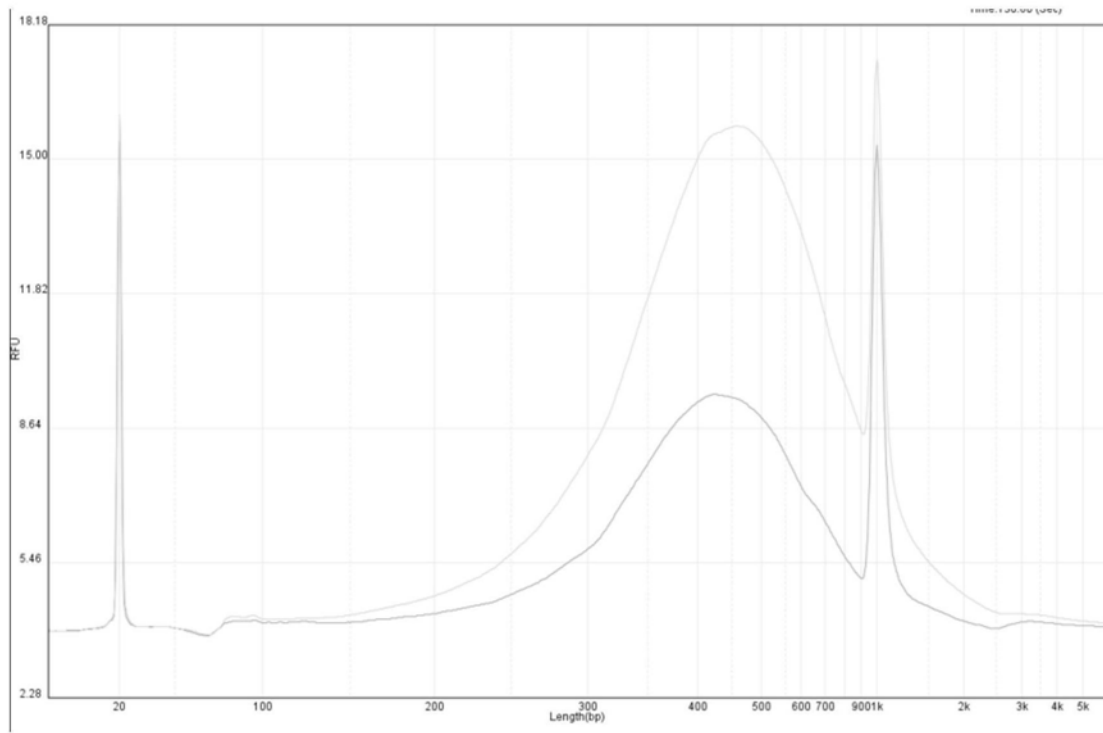


图10

RNA-seq数据分析

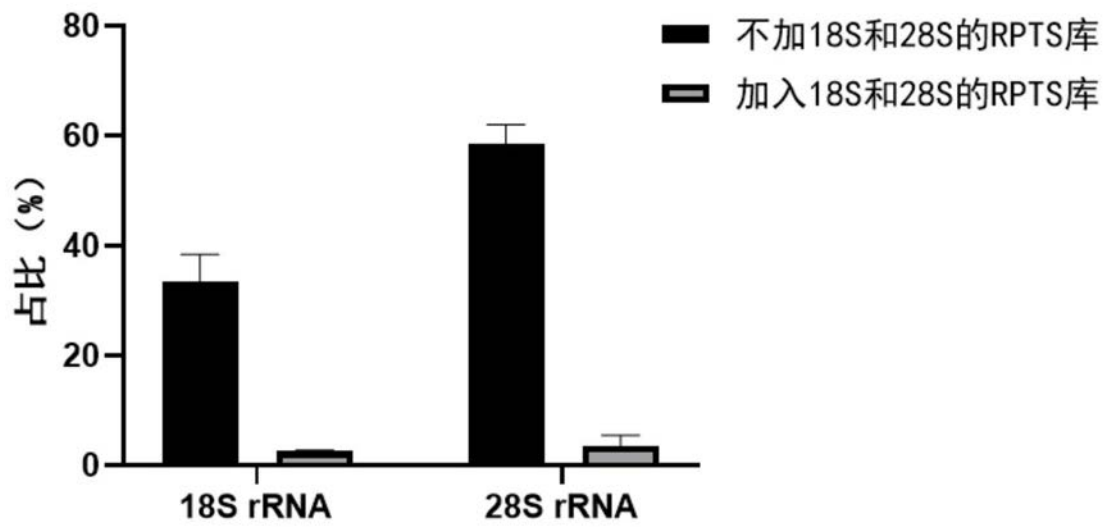


图11

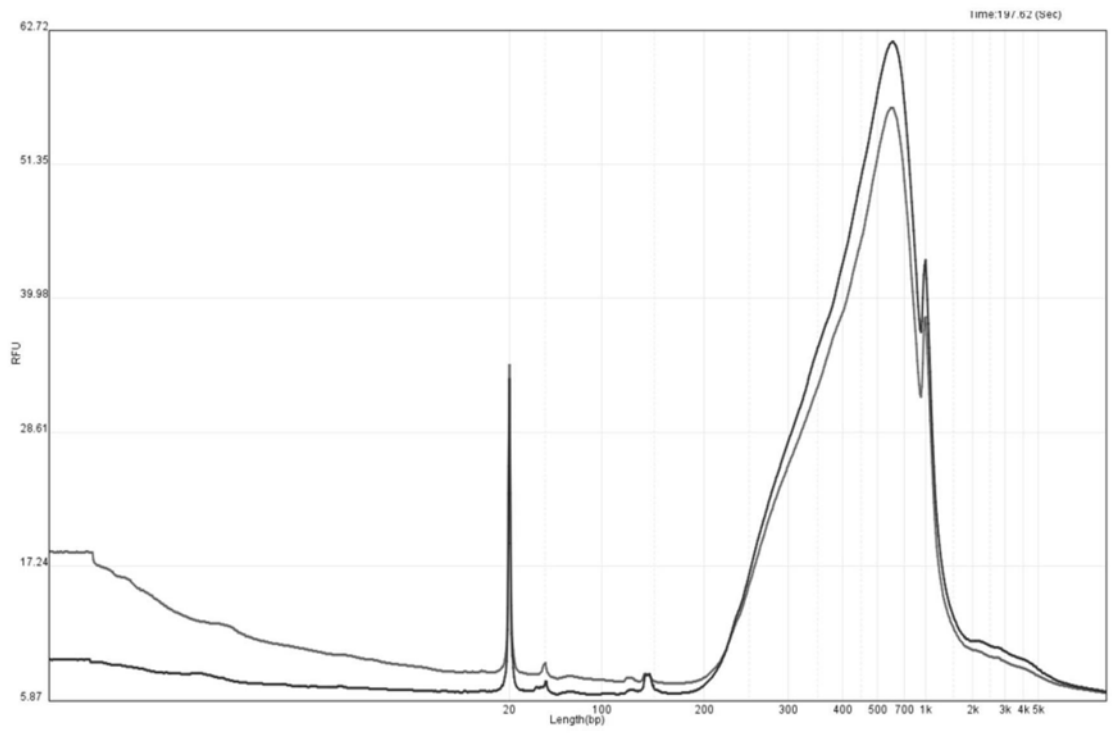


图12

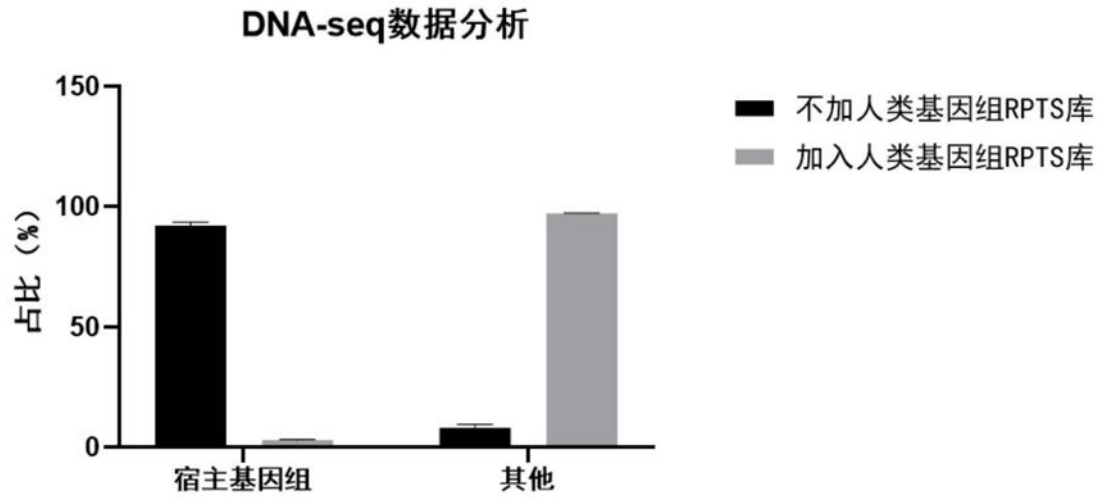


图13