



(12) 发明专利

(10) 授权公告号 CN 110499274 B

(45) 授权公告日 2021. 11. 23

(21) 申请号 201910880078.X

(22) 申请日 2019.09.18

(65) 同一申请的已公布的文献号
申请公布号 CN 110499274 A

(43) 申请公布日 2019.11.26

(73) 专利权人 清华大学
地址 100084 北京市海淀区清华大学

(72) 发明人 于慧敏 梁有向 焦松 唐玲珺

(74) 专利代理机构 北京众合诚成知识产权代理
有限公司 11246

代理人 陈波

(51) Int. Cl.

G12N 1/21 (2006.01)

G12P 13/02 (2006.01)

C12R 1/01 (2006.01)

(56) 对比文件

Miaomiao Wang et al..Improving stress tolerance and cell integrity of *Rhodococcus ruber* by overexpressing small-shock-protein Hsp16 of *Rhodococcus*.《*Journal of Industrial Microbiology & Biotechnology*》.2018,表1,第932-933页.

Xiaochao Xiong et al..Metabolic engineering of oleaginous bacterium *Rhodococcus jostii* RHA1 for producing biofuels and renewable chemicals from lignocellulosic biomass.《39th Symposium on Biotechnology for Fuels and Chemicals》.2017,全文.

审查员 赵鹏

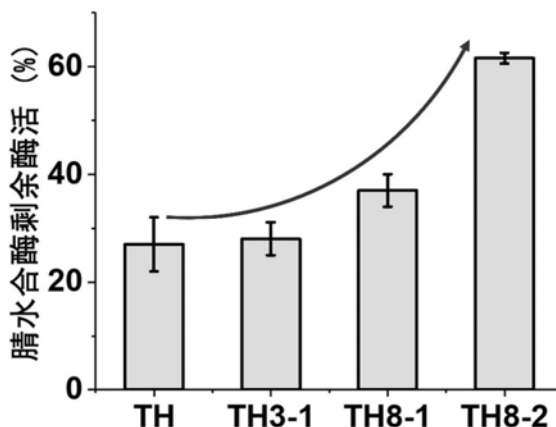
权利要求书1页 说明书10页
序列表8页 附图6页

(54) 发明名称

一种基因工程红球菌及其构建方法与应用

(57) 摘要

本发明公开了基因工程技术领域的一种基因工程红球菌及其构建方法与应用。其中的基因工程红球菌构建的方法,是基于CRISPR/Cas9的基因编辑方法,基因敲除效率高,能实现无痕敲除,以及基因叠加敲除,为红球菌的基因改造提供了高效的工具。以酰胺酶基因敲除为例,其敲除效率可达到75%。同时,利用上述编辑方法构建的低副产物合成、高催化稳定性红球菌,应用于催化丙烯腈水合生产丙烯酰胺或丙烯酸铵制备中,催化稳定性好,效率高,具有良好的应用前景。



1. 一种基因工程红球菌构建的方法,其特征在于,包括在宿主中过表达重组酶基因,并利用同源重组的方法敲除、插入或突变目的基因,构建基因工程菌;所述同源重组的方法包括将含有Cas9基因的载体、目标基因上下游同源臂和含有相应sgRNA表达盒的载体转入宿主;所述Cas9基因如SEQ ID NO.1所示;所述重组酶基因位于启动子Pa2控制下;所述重组酶基因为Che9c60&61;所述红球菌为红色红球菌*R. ruber* TH。

2. 根据权利要求1所述的方法,其特征在于,所述含相应sgRNA表达盒的载体为温敏型质粒,在特定温度下可快速丢失,进行多轮叠加基因编辑。

3. 一种基因工程红球菌构建的方法,其特征在于,包括在红球菌中过表达重组酶基因,并利用同源重组的方法敲除酰胺酶基因,并将突变脘水合酶基因或脘水解酶基因原位替换野生型脘水合酶基因,构建基因工程菌;所述同源重组的方法包括将含有Cas9基因的载体、目标基因上下游同源臂和含有相应sgRNA表达盒的载体转入宿主;所述Cas9基因如SEQ ID NO.1所示;所述重组酶基因位于启动子Pa2控制下;所述重组酶基因为Che9c60&61;所述红球菌为红色红球菌*R. ruber* TH。

一种基因工程红球菌及其构建方法与应用

技术领域

[0001] 本发明属于基因工程技术领域,具体涉及一种基因工程红球菌及其构建方法与应用。

背景技术

[0002] 红球菌是一种革兰氏阳性放线菌,具有丰富的代谢酶系、良好的有机溶剂耐受性,因此广泛应用于生物催化、生物修复和生物合成等方面。譬如,红色(赤)红球菌和紫红红球菌游离细胞可高表达脲水合酶,对极性溶剂丙烯腈和丙烯酰胺具有良好的耐受性,已成功用于丙烯酰胺、烟酰胺等化学品的生产(CN101892228A;CN104762338A);红平红球菌具有脱硫酶系,可应用于生物脱硫,在环保领域具有重要作用(CN107557106A);浑浊红球菌可降解、利用木质素生产油脂,其对芳香类化合物有丰富的代谢酶系,对木质素降解产物如有机酸、呋喃类、芳香类和卤代化合物有良好的耐受性,因此被视为木质纤维素资源利用的理想平台(DeLorenzo, et al., ACS Synth Biol 2018, 7 (2), 727-738.)。然而,红球菌的基因改造工具仍然比较匮乏,限制了红球菌的工程化改造。

[0003] 微生物法生产丙烯酰胺工艺最早由日本日东公司于1985年开发,其利用微生物如红球菌高表达脲水合酶,催化丙烯腈水合生成丙烯酰胺。该工艺具有选择性、收率高等特点,发展至今已有30多年的历史,并广泛实现产业化。尽管比化学法有明显优势,微生物法仍然存在一些问题,包括水合过程中副产物丙烯酸的生成(Ma, et al., Bioresource Technol 2010, 101 (1), 285-291.)、细胞催化稳定性仍需提高(Chen, et al., Journal of Biotechnology 164 (2012) 354-362)等。红球菌细胞的脲类代谢酶系包括脲水合酶、酰胺酶和脲水解酶,酰胺酶可将目标产物丙烯酰胺进一步反应生成丙烯酸,脲水解酶可将底物丙烯腈一步转化成丙烯酸,因此降低了丙烯酰胺的收率和纯度,从而增加生产成本。利用无脲类代谢酶系的宿主如大肠杆菌、谷氨酸棒状杆菌克隆表达脲水合酶,是避免副产物生成的有效方法,但存在酶活低、催化稳定性差等问题(CN1584024;Kang, et al., Appl Microbiol Biot 2014, 98 (10), 4379-4387.)。因此丙烯酰胺的微生物法生产中红球菌仍是主流生物催化剂。对红球菌的酰胺酶、脲水解酶进行基因敲除是产生副产物问题的重要解决方法。针对催化稳定性的问题,于慧敏等构建了高稳定性的脲水合酶突变体,在红球菌中使用质粒过表达该突变体,可以实现多批次约50%丙烯酰胺的水合生产(CN107177581A),但突变体的表达受质粒稳定性、基因组天然高表达野生型脲水合酶的影响,表达量有限,若将该突变体原位替换野生型脲水合酶,稳定性将进一步提高。上述问题及应对策略,对红球菌的基因编辑方法提出了要求。

[0004] 目前红球菌的基因组改造缺乏高效的编辑方法,主要依赖于借助自杀质粒的同源重组单交换和双交换。红球菌存在比较严重的非法重组现象(Desomer, et al., Molecular Microbiology 1991, 5 (9), 2115-2124.)。自杀质粒转化红球菌时可随机插入到基因组的各个位置而非特异性到目标基因处,因此导致敲除效率极低。CRISPR/Cas9 (the Clustered Regularly Interspaced Short Palindromic Repeats /CRISPR-associated nuclease

9) 是近几年发展起来的强大的基因编辑工具,已广泛应用于很多种动物、植物、微生物的基因组改造。Cas9是一种核酸酶,可与向导RNA (sgRNA) 结合并在sgRNA的引导下在目标基因位置引入DNA双链断裂,通过非同源性末端接合 (NHEJ) 或者同源重组修复 (HDR),可实现目标基因的失活 (Ran, et al., Nature Protocols 2013, 8 (11), 2281-2308.)。红球菌基因操作困难、同源重组效率低,增加了CRISPR/Cas9在红球菌中应用开发的难度,目前CRISPR/Cas9基因敲除在红球菌中的应用仍未见报道。

发明内容

[0005] 为了克服现有技术中的问题,本发明提供了一种基因工程红球菌构建的方法,包括在宿主中过表达重组酶基因,并利用同源重组的方法敲除、插入或突变目的基因,构建基因工程菌。

[0006] 上述方法中,所述重组酶基因过表达的方法为在宿主中通过载体转入重组酶基因和/或所述重组酶基因处于强启动子控制之下。

[0007] 上述方法中,所述同源重组的方法包括利用CRISPR/Cas9系统进行。

[0008] 上述方法中,所述同源重组的方法包括将含有Cas9基因的载体、目标基因上下游同源臂和含有相应sgRNA表达盒的载体转入宿主。

[0009] 上述方法中,所述含相应sgRNA表达盒的载体为温敏型质粒,在特定温度下可快速丢失,进行多轮叠加基因编辑。

[0010] 上述方法中,可选的,所述含有Cas9基因的载体为pNV18,优选pNV18-Pa2。

[0011] 上述方法中,所述目标基因上下游同源臂分别为500-1500bp,优选500-1000bp。

[0012] 上述方法中,所述Cas9基因如SEQ ID NO.1所示。

[0013] 上述方法中,所述重组酶基因为分枝杆菌噬菌体的重组酶基因,优选的,所述重组酶基因为Che9c60&61 (DeLorenzo, et al., ACS Synth Biol 2018, 7 (2), 727-738.)。

[0014] 上述方法中,可选的,所述重组酶基因的表达载体为pRCTc,优选pRCTc-Pa2。

[0015] 上述方法中,优选的,所述红球菌为*R. ruber* TH (Ma, et al., Bioresource Technol 2010, 101 (1), 285-291.)。

[0016] 进一步的,本发明还提供了一种基因工程红球菌构建的方法,包括在红球菌中过表达重组酶基因,并利用同源重组的方法敲除酰胺酶基因,并将突变脲水合酶基因或脲水解酶基因原位替换野生型脲水合酶基因,构建基因工程菌。

[0017] 上述方法中,可选的,所述突变脲水合酶基因为中国专利文献CN107177581A中所述的脲水合酶突变体。

[0018] 例如,上述方法中,具体可以包括,所述Cas9基因如SEQ ID NO:1所示,将Cas9基因连接到穿梭质粒pNV18.1 (Chiba, et al., Jpn J Infect Dis 2007, 60 (1), 45-47.), 启动子为Pa2。将过表达适用于红球菌宿主的Cas9基因的载体pNV18-Pa2-Cas9转化到含有重组酶的红球菌*R. ruber* TH (Che9c60&61), 得到同时表达重组酶和Cas9的基因工程红球菌*R. ruber* TH (Cas9+Che9c60&61); 然后将拟敲除的目标基因上下游同源臂和含相应sgRNA表达盒的载体同时转化红球菌*R. ruber* TH (Cas9+Che9c60&61), 得到实现基因编辑的基因工程红球菌。

[0019] 可选的,所述sgRNA序列如SEQ ID NO.2所示。

[0020] 可选的,进一步,将sgRNA表达盒序列连接到穿梭质粒pBNVCm上,启动子为不带有核糖体结合位点(RBS)的强组成型启动子PamiC(Jiao et al.,New Biotechnol 2018,44:41-49.)。所述穿梭质粒pBNVCm是以pNV18.1为骨架,替换复制子,替换抗性基因而得。所述的拟敲除目标基因上下游同源臂序列,通过PCR扩增得到,以红球菌*R.ruber* TH为模板,分别扩增目标基因上游、下游同源序列,通过overlap PCR将两段序列融合得到同源臂。

[0021] 进一步的,本发明还建立了利用CRISPR/Cas9进行多轮叠加基因编辑的方法构建基因工程菌。本发明中利用CRISPR/Cas9实现基因敲除的红球菌同时含有3个质粒,其中携带sgRNA的穿梭质粒pBNVCm为温敏型,37℃培养条件下可快速丢失,从而可引入靶向新的拟敲除基因的sgRNA质粒进行下一轮基因编辑。

[0022] 将同时携带三个质粒的红球菌接种到含四环素和卡那霉素的种子培养基中,37℃培养72小时,菌液稀释涂布于含四环素和卡那霉素的平板,37℃培养72小时,生长的菌落划线于含氯霉素的平板,若不生长说明已经丢失sgRNA质粒。丢失了sgRNA质粒的红球菌含有分别携带Cas9、重组酶的两个质粒,制备成感受态细胞,共转化新sgRNA质粒和拟敲除基因的同源臂,进行下一轮基因编辑。最后得到的菌株在不含抗生素的条件下培养,可同时丢失3个质粒,得到工业菌株进行性能评价。

[0023] 本发明还提供了上述任一方法构建的基因工程红球菌。

[0024] 同时,本发明还提供了上述基因工程红球菌在丙烯酰胺或丙烯酸铵制备中的应用。

[0025] 具体的,本发明利用CRISPR/Cas9基因编辑方法构建了一种低副产物合成、高催化稳定性红球菌,应用于丙烯酰胺的高效生产。其要点是无痕敲除红球菌中副产物相关的酰胺酶基因(CN101663389),得到重组红球菌*R.ruber* TH3-1,进一步将高稳定性的脲水合酶突变体(CN107177581A)替换基因组上的野生脲水合酶,得到的重组红球菌*R.ruber* TH8-2。所述重组红球菌TH8-2具有低副产物合成、高催化稳定性的特点,可用于生产高浓度丙烯酰胺,并实现多批次细胞回用

[0026] 另外,本发明进一步拓展CRISPR/Cas9的应用,提供了一种红球菌基因组高效表达外源酶(如脲水解酶)的方法。使用上述CRISPR/Cas9基因编辑方法,将脲水解酶基因(CN105420154A)替换基因组上的脲水合酶基因,利用原脲水合酶的启动子高效表达脲水解酶,且避免了质粒不稳定问题。

[0027] 本发明的有益效果:

[0028] 本发明建立了基于CRISPR/Cas9的红球菌基因编辑方法,基因敲除效率高,能实现无痕敲除,以及基因叠加敲除,为红球菌的基因改造提供了高效的操作工具。以酰胺酶基因敲除为例,其敲除效率可达到75%。进一步,本发明利用CRISPR/Cas9编辑方法构建的低副产物合成、高催化稳定性红球菌*R.ruber* TH8-2,应用于催化丙烯脲水合生产丙烯酰胺,催化副产物的酰胺酶活性降低60%,催化稳定性提高1.2倍,在高浓度丙烯酰胺生产过程中,产物浓度达到500g/L,可实现4批次细胞回用,副产物丙烯酸浓度降低80%,具有良好的应用前景。再进一步地,本发明提出的利用红球菌基因组表达外源酶的策略,借助于原脲水合酶的强启动子,基因组表达脲水解酶,其酶活相比于质粒表达提高了51%,且避免了质粒稳定性的问题。

附图说明

- [0029] 图1为载体pRCTc-Pa2-Che9c60&61的质粒图谱。
- [0030] 图2为载体pNV-Pa2-Cas9的质粒图谱。
- [0031] 图3为载体pBNVCm-BbsI-sgRNA的质粒图谱
- [0032] 图4为利用重组酶Che9c60&61辅助线性DNA片段敲除酰胺酶得到的平板菌落和PCR验证。泳道M为DNA分子量标准；泳道1~8为从平板挑取菌落的PCR产物，泳道WT为野生菌*R. ruber* TH的PCR产物。敲除型菌落的条带比野生菌的小200bp。
- [0033] 图5为基于CRISPR/Cas9的红球菌基因编辑方法的流程。
- [0034] 图6为利用CRISPR/Cas9无痕酰胺酶基因的菌落PCR验证。泳道M为DNA分子量标准；泳道1~16为从平板挑取菌落的PCR产物，泳道WT为野生菌*R. ruber* TH的PCR产物。敲除型菌落的条带比野生菌的小1000bp。
- [0035] 图7为利用CRISPR/Cas9将脲水合酶突变体原位替换野生型脲水合酶的重要突变点的测序结果。脲水合酶 β 亚基的215位氨基酸Asp被突变为TGT编码的Cys， α 亚基的133位氨基酸Pro被突变为TGT编码的Cys。
- [0036] 图8为重组红球菌的酰胺酶活性。
- [0037] 图9为重组红球菌的催化稳定性。
- [0038] 图10为重组红球菌多批次催化生产高浓度丙烯酰胺过程的产物浓度变化曲线。
- [0039] 图11为重组红球菌催化生产高浓度丙烯酰胺过程的副产物浓度变化曲线。
- [0040] 图12为红球菌基因组表达和质粒表达脲水解酶的活性比较。

具体实施方式

[0041] 下面结合附图和具体实施例对本发明作进一步说明。如未特别指明，实施例中所用的生化试剂均为市售试剂，实施例中所用的技术手段为本领域技术人员书中的常规手段。

[0042] 实施例1构建具有不同抗性基因、不同复制子的质粒载体pRCTc、pBNVCm

[0043] 本发明中所建立的CRISPR/Cas9基因编辑方法需要使用不同的质粒载体来表达不同的元件，因此需改造现有的质粒来满足实验需求。质粒pRCTc和pBNVCm为大肠杆菌-红球菌穿梭质粒，以质粒pNV18.1 (Chiba, et al., *Jpn J Infect Dis* 2007, 60 (1), 45-47.) 为基础，通过替换抗性基因和红球菌复制子改造而成。

[0044] 质粒pRCTc携带四环素抗性基因，复制子为pRC4，其构建过程如下：以红球菌*R. ruber* TH3 (W02009117843-A1) 的基因组为模板，使用引物fusion-Tc-F和fusion-Tc-R，PCR扩增出四环素抗性基因；以质粒pNV18.1为模板，使用引物pNV-F和pNV-R，扩增出不带红球菌复制子的pNV18.1骨架；使用Gibson连接试剂盒 (Vazayme公司) 使两个片段连接环化；连接产物转化大肠杆菌*E. coli* TOP10感受态细胞 (索莱宝公司)，涂布LB培养基 (蛋白胨10g/L，酵母粉5g/L，NaCl 10g/L，pH 7.0) 固体平板 (含8mg/L四环素)；挑选抗性克隆，小量提取质粒，得到重组质粒pNVTc。复制子pRC4由华大青兰生物科技 (无锡) 有限公司合成，两端带NheI酶切位点，用NheI酶切pRC4和pNVTc，纯化回收pRC4和去除红球菌复制子的pNVTc骨架，用T4 DNA连接酶进行连接环化，连接产物转化大肠杆菌*E. coli* TOP10感受态细胞 (索莱宝公司)，涂布LB培养基固体平板 (含8mg/L四环素)；挑选抗性克隆，小量提取质粒，含有

重组质粒pRCTc。

[0045] 质粒pBNVCm携带氯霉素抗性基因,复制子为pB264,其构建过程如下:以商用质粒pXMJ19为模板,使用引物fusion-Cm-F和fusion-Cm-R引物扩增氯霉素抗性基因的开放阅读框;以红球菌*R. ruber* TH3的基因组为模板,使用引物fusion-PamiC-F、PamiC-R扩增启动子PamiC;以质粒pNV18.1为模板,使用引物pNV-F和pNV-R扩增出不带红球菌复制子的pNV18.1骨架;使用Gibson连接试剂盒(Vazayme公司)使3个片段连接环化;连接产物转化大肠杆菌*E. coli* TOP10感受态细胞(索莱宝公司),涂布LB培养基固体平板(含5mg/L氯霉素);挑选抗性克隆,小量提取质粒,得到重组质粒pNVCm。复制子pB264由华大青兰生物科技(无锡)有限公司合成,两端带NheI酶切位点,用NheI酶切pB264和pNVCm,纯化回收pB264和去除红球菌复制子的pNVCm骨架,用T4DNA连接酶进行连接,连接产物转化大肠杆菌*E. coli* TOP10感受态细胞,涂布LB培养基固体平板(含5mg/L氯霉素);挑选抗性克隆,小量提取质粒,得到重组质粒pBNVCm。

[0046] 上述所用引物序列如下:

[0047] fusion-Tc-F: CAGAGTCCCCTCAGAAGAACTCAATCGTCACCCTTTCTCGG;

[0048] fusion-Tc-R: GTAAGGTTGGGAAGCCCTGCTCAGCGATCGGCTCGTTG;

[0049] pNV-F: GCAGGGCTTCCCAACCTTAC;

[0050] pNV-R: GTTCTTCTGAGCGGGACTCTG;

[0051] fusion-Cm-F:

[0052] CGAGTCTACTAAGGAGGGATCCATGGAGAAAAAATCACTGGATATACCACC;

[0053] fusion-Cm-R: CAGAGTCCCCTCAGAAGAACTTACGCCCCGCCCTGC;

[0054] PamiC-F:

[0055] GTAAGGTTGGGAAGCCCTGCAACAGGTGAGATTACGGAGAACGG;

[0056] PamiC-R:

[0057] TCTAGACTCCTTAGTGACTCGCCGGCGCTCATGACTGCAGTGAAGTCAACGTTGTGG。

[0058] 引物由苏州金唯智生物科技有限公司合成,用无菌水溶解并稀释至10 μ M使用。PCR扩增所用的预制缓冲液购自Vazayme公司。扩增反应体系为:

	质粒溶液或菌液	1 μ L
	正向引物	2.5 μ L
	反向引物	2.5 μ L
[0059]	Phanta Mix	25 μ L
	无菌水	19.9 μ L
	总体积	50 μ L

[0060] 热循环条件为94 $^{\circ}$ C 3min;94 $^{\circ}$ C 15s,60 $^{\circ}$ C 15s,72 $^{\circ}$ C 180s,35个循环;72 $^{\circ}$ C,10min。

[0061] 实施例2构建过表达重组酶Che9c60&61的红球菌并分析其对敲除效率的影响

[0062] 本实施例构建了携带重组酶的质粒pRCTc-Pa2-Che9c60&61(如附图1所示),转化到红球菌*R. ruber* TH得到过表达重组酶的*R. ruber* TH(Che9c60&61)。以酰胺酶的基因敲除

为例评价重组酶对敲除效率的影响,将带氯霉素抗性基因的酰胺酶基因同源臂转化 *R.ruber* TH (Che9c60&61) 和 *R.ruber* TH, 统计转化子个数和敲除效率。

[0063] 质粒pRCTc-Pa2-Che9c60&61构建方法如下:重组酶Che9c60&61由华大青兰生物科技(无锡)有限公司合成,两段带有XbaI、KpnI酶切位点,用XbaI、KpnI酶切Che9c60&61基因片段;以红球菌*R.ruber* TH为模板,用引物HindIII-Pa2-F、XbaI-Pa2-R扩增出启动子Pa2,用HindIII、XbaI酶切;使用HindIII、KpnI酶切质粒pRCTc;然后使用T4 DNA连接酶连接以上3个片段,连接产物转化大肠杆菌*E.coli* TOP10感受态细胞,涂布LB培养基固体平板(含8mg/L四环素);挑选抗性克隆,小量提取质粒,含有重组质粒pRCTc-Pa2-Che9c60&61。

[0064] 将该质粒采用电穿孔的方式(Jiao, et al., 2018, New Biotechnology, 44:41-49)转入红球菌*R.ruber* TH, 转化子使用含6mg/L四环素的平板(葡萄糖10g/L, 酵母膏3g/L, NaCl 1g/L, $K_2HPO_4 \cdot 3H_2O$ 2g/L, $MgSO_4 \cdot 7H_2O$ 0.2g/L, 琼脂15g/L)进行筛选,得到基因工程红球菌*R.ruber* TH (Che9c60&61)。

[0065] 带氯霉素抗性基因的酰胺酶基因同源臂构建方法如下:以红球菌*R.ruber* TH为模板,用引物amiUp-F、Cm-amiUp-R扩增出酰胺酶基因上游的同源臂(602bp),序列如SEQ ID NO.3所示,用引物Cm-amiDown-F、amiDown-R扩增出酰胺酶基因下游同源臂(1018bp),序列如SEQ ID NO.4所示;以pBNVCm为模板,用引物Cm-F、Cm-R扩增出氯霉素抗性基因表达盒;使用overlap PCR融合3个片段,得到带抗性基因的同源臂amiarm-Cm^R,序列如SEQ ID NO.5所示。

[0066] 将1 μ g同源臂amiarm-Cm^R采用电穿孔的方式分别转化红球菌*R.ruber* TH和TH (Che9c60&61),转化子使用含6mg/L四环素和5mg/L氯霉素的红球菌平板培养基筛选,统计转化子个数,并挑取菌落,用引物P1、P2做PCR验证酰胺酶是否敲除成功。如附图4所示,野生菌*R.ruber* TH无法得到转化子,而重组酶Che9c60&61存在时转化子个数达196,且酰胺酶的敲除效率达100%,说明重组酶可以有效促进红球菌的同源重组,提高基因敲除效率。

[0067] 上述所用引物序列如下:

[0068] amiUp-F: CACTGTCCGTCGACGGTGAC;

[0069] Cm-amiUp-R:

[0070] CTCCGTAATCTCACCTGTTGTACCAGCACTTCAAAGACGTTTC;

[0071] Cm-amiDown-F:

[0072] AGTGGCAGGGCGGGGCGTAAACCCATCGCAGATCAGACTG;

[0073] amiDown-R: CATTCCGTTTCGACTGGTCC;

[0074] Cm-F: AACAGGTGAGATTACGGAGAAC;

[0075] Cm-R: TTACGCCCCGCCCTGCC;

[0076] P1: GGCTGGAGAAGGTTGGACG;

[0077] P2: AGTAGACCTACCACCGCG。

[0078] 实施例3建立基于CRISPR/Cas9的红球菌基因编辑方法,实现酰胺酶的无痕敲除

[0079] CRISPR/Cas9发挥作用需要两个重要元件,核酸酶Cas9和向导RNA (sgRNA),本实施例需构建它们的表达载体pNV-Pa2-Cas9(如附图2所示)和pBNVCm-BbsI-sgRNA(如附图3所示);同时为了实现基因敲除,还需要提供修复DNA双链断裂所需的同源臂。利用CRISPR/Cas9对红球菌进行基因编辑的流程如附图5所示,具体的细节如下文所述。

[0080] pNV-Pa2-Cas9的构建方法如下:以质粒pCAS9-mCherry (Addgene#80975) 为模板,用引物XbaI-Cas9-F、BamHI-Cas9-R扩增出Cas9基因,用XbaI、BamHI酶切Cas9基因和质粒pNV-Pa2 (CN109182241A),纯化回收后使用T4 DNA连接酶连接环化,连接产物转化大肠杆菌E.coli Top10,涂布LB培养基固体平板(含50mg/L卡那霉素);挑选抗性克隆,小量提取质粒,含有重组质粒pNV-Pa2-Cas9。

[0081] pBNVCm-BbsI-sgRNA的构建方法如下:以质粒pBNVCm为模板,用引物XbaI-PamiC-F-sgRNA、BbsI-PamiC-R扩增启动子PamiC (Jiao et al., New Biotechnol 2018, 44:41-49.);以质粒pJOE8999为模板,用引物BbsI-sgRNA-F、KpnI-sgRNA-R扩增sgRNA,序列如SEQ ID NO.2所示;通过overlap PCR融合启动子PamiC和sgRNA得到PamiC-BbsI-sgRNA;用XbaI、KpnI酶切质粒pBNVCm和PamiC-BbsI-sgRNA,纯化回收后用T4 DNA连接酶连接环化,连接产物转化大肠杆菌E.coli Top10,涂布LB培养基固体平板(含5mg/L氯霉素);挑选抗性克隆,小量提取质粒,含有重组质粒pBNVCm-BbsI-sgRNA,该质粒携带BbsI酶切位点,可通过Golden Gate Assembly添加靶向目标基因的引导序列。

[0082] 以酰胺酶为例,构建靶向酰胺酶的sgRNA质粒pBNVCm-sgRNA1。设计靶向的引导序列,合成引物sgRNA1-F、sgRNA1-R,两个引物有20个碱基相互配对,即引导序列N20-1;取100 μM sgRNA1-F 1 μL、100 μM sgRNA1-R 1 μL和去离子水8 μL混合均匀,94 °C 5min,然后以0.1 °C/min的速度降温至25 °C,得到两端带有粘性末端的引导序列N20-1;用BbsI酶切质粒pBNVCm-BbsI-sgRNA;用T4 DNA连接酶连接环化,连接产物转化大肠杆菌E.coli Top10,涂布LB培养基固体平板(含5mg/L氯霉素);挑选抗性克隆,小量提取质粒,得到重组质粒pBNVCm-sgRNA1。

[0083] 酰胺酶同源臂的构建:以红球菌R.ruber TH为模板,用引物amiUp-F、amiUp-R扩增酰胺酶上游同源臂,用引物amiDown-F、amiDown-R扩增酰胺酶下游同源臂,通过overlap PCR融合两片段得到酰胺酶同源臂amiarm,序列如SEQ ID NO.6所示。

[0084] 在实施例2的基础上,本实施例将质粒载体pNV-Pa2-Cas9转化红球菌R.ruber TH (Che9c60&61),得到R.ruber TH (Cas9+Che9c60&61)。进一步地,针对酰胺酶的基因敲除,取1 μg质粒pBNVCm-sgRNA1和1 μg同源臂片段amiarm,共转化R.ruber TH (Cas9+Che9c60&61),用含5mg/L氯霉素、6mg/L四环素和25mg/L卡那霉素的平板筛选转化子,并使用引物P1和P2进行菌落PCR验证,DNA电泳结果如附图6所示,挑取16个菌落进行PCR验证,酰胺酶的敲除效率达75%,说明了本发明建立的基于CRISPR/Cas9的红球菌编辑方法具有高效性。

[0085] 本实施例中携带sgRNA的质粒是温敏型的,在37 °C培养可快速丢失携带sgRNA的质粒,从而可引入新sgRNA质粒进行下一轮的基因编辑。将同时携带三个质粒的红球菌接种到含6mg/L四环素和25mg/L卡那霉素的红球菌种子培养基(葡萄糖20g/L,酵母膏1g/L,蛋白胨1g/L, $K_2HPO_4 \cdot 3H_2O$ 0.5g/L, KH_2PO_4 0.5g/L, $MgSO_4 \cdot 7H_2O$ 0.5g/L, pH 7.5),于37 °C、200rpm培养72h,取菌液适当稀释后涂布于含6mg/L四环素和25mg/L卡那霉素的平板上,37 °C培养72小时,生长的菌落划线于含5mg/L氯霉素的平板,若不生长说明已经丢失了sgRNA质粒,本实施例中该菌命名为R.ruber TH3-1 (Cas9+Che9c60&61),挑取并培养相应的菌落,用于下一轮基因编辑。

[0086] 完成最后一轮基因编辑后,可将三个质粒同时丢失,得到最终的基因工程菌,方法如下:将编辑成功的红球菌接种于不含抗生素的种子培养基中,于37 °C、200rpm培养72h,取

菌液适当稀释后涂布于不含抗生素的平板,37℃培养72小时,生长的菌落分别划线于含5mg/L氯霉素、6mg/L四环素或者25mg/L卡那霉素的平板,若在3种平板上均不生长说明3个质粒均已丢失,本实施例中该菌命名为*R. ruber* TH3-1,挑取相应菌落进行培养,用于性能评价。

[0087] 上述所用引物序列如下:

[0088] XbaI-Cas9-F:GCTCTAGAATGGACAAGAAGTACAGCATCGG

[0089] BamHI-Cas9-R:CGGGATCCCTAGTCGCCGCCAGCTG

[0090] XbaI-PamiC-F:GCTCTAGAAACAGGTGAGATTACGGAGAACG

[0091] BbsI-PamiC-R:AGGTCTTCTCGAAGACCCCGTTGTGGACACGTTACGAC

[0092] BbsI-sgRNA-F:GGGTCTTCGAGAAGACCTGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGC

[0093] KpnI-sgRNA-R:GGGGTACCGGCCGTCGACCCTATAGTG

[0094] sgRNA1-F:AACGGTGATCCTCATGGCGCTCGC

[0095] sgRNA1-R:AAACGCGAGCGCCATGAGGATCAC

[0096] amiUp-F:GCTCTAGAGTGACCACCAGCTTGGCG

[0097] amiUp-R:CTGCGATGGGTCTAGGACTCCTTAGTGACTCGCC

[0098] amiDown-F:TAAGGAGTCCTAGACCCATCGCAGATCAGACTGTTC

[0099] amiDown-R:CCCAAGCTTGTGGTGGGGCCGGACAG。

[0100] 实施例4脲水合酶突变体原位替换基因组的野生型脲水合酶

[0101] 实施例3得到敲除酰胺酶基因的红球菌*R. ruber* TH3-1 (Cas9+Che9c60&61),为了提高红球菌在丙烯酰胺生产过程中的催化稳定性,将脲水合酶突变体原位替换基因组上的野生型脲水合酶,得到红球菌*R. ruber* TH8-2,其构建方法如下:将针对脲水合酶的sgRNA质粒pBNVCm-sgRNA2和携带突变体基因的脲水合酶同源臂,共转化*R. ruber* TH3-1 (Cas9+Che9c60&61),按照实施例3的方法筛选转化子,并测序验证,重要突变位点的测序结果如图7所示,说明突变体替换成功。按实施例3的方法丢失质粒,得到不含质粒的最终菌株,即*R. ruber* TH8-2。

[0102] 针对脲水合酶的sgRNA质粒pBNVCm-sgRNA2的构建方法与pBNVCm-sgRNA1类似,其不同点在于引物更换为sgRNA2-F和sgRNA2-R。携带突变体基因的脲水合酶同源臂的构建方法如下:以红球菌*R. ruber* TH为模板,用引物nhhUp-F、nhhUp-R扩增脲水合酶上游同源臂,序列如SEQ ID NO.7所示,用引物nhhDown-F、nhhDown-R扩增脲水合酶下游同源臂,序列如SEQ ID NO.8所示;脲水合酶突变体(CN107177581A)基因由华大青兰生物科技(无锡)有限公司合成,且两端各添加了20bp与上、下游同源臂重叠的序列;使用overlap PCR融合3个片段得到脲水合酶同源臂nhhUp-NHase^M-nhhDown。

[0103] 引物序列如下:

[0104] sgRNA2-F:AACGCGCATGACCGGATACGGAC;

[0105] sgRNA2-R:AAACGTCCGTATCCGGTCATGCCG;

[0106] nhhUp-F:CACACACCGGGTCACCG;

[0107] nhhUp-R:TTCCTCATTCCTTTCATCGGAGCTG;

[0108] nhhDown-F:GTGAAGACACACTCACTGATCGGC;

[0109] nhhDown-R:CCGGCAACAGTCGCGGAG。

[0110] 改造后红球菌的脲水合酶表达及酰胺酶活性、催化稳定性评价：

[0111] 为了对比基因组表达和质粒表达脲水合酶突变体的差别，在实施例3的基础上，将质粒pNV-SBMDB (CN107177581A) 转化红球菌*R. ruber* TH3-1，得到质粒表达脲水合酶突变体的红球菌TH8-1。本发明有以下野生或改造红球菌*R. ruber* TH、TH3-1、TH8-1和TH8-2，本实施例对4株红球菌进行评价性能，包括副产物生成（即酰胺酶活性）、催化稳定性，评价方法如下文所述。

[0112] 将红球菌*R. ruber* TH、TH3-1、TH8和TH8-2接种于种子培养基中，28℃、200rpm培养60h，按10%接种量转接至发酵培养基（葡萄糖20g/L，酵母膏5g/L，尿素6g/L， $K_2HPO_4 \cdot 3H_2O$ 2.28g/L， KH_2PO_4 0.866g/L， $MgSO_4 \cdot 7H_2O$ 1g/L，谷氨酸钠0.5g/L， Co^{2+} 50ppm），于28℃、200rpm培养48h，将菌液离心后去上清液收获菌体，用原体积的去离子水重悬，菌液 OD_{460} 约为50。

[0113] 酰胺酶活性测定：取500μL菌液与500μL 20%丙烯酰胺溶液混匀，40℃静置反应1小时，离心后取500μL上清液与500μL 4%的乙酰胺（内标）溶液混匀，采用气相色谱仪Trace 1300 (Thermo, 美国) 测定丙烯酰胺的浓度。气相色谱操作条件为：聚乙二醇高分子毛细管柱PEG-20M (30m×0.25mm×2μm)，进样口为SPL，温度260℃；FID检测器，温度260℃；柱温190℃；载气为氮气，分压为108kpa；分流进样，进样量0.4μL，分流比为50:1。酰胺酶活性：1U表示测量条件下催化1μmol丙烯酰胺反应所需要的酶量。

[0114] 脲水合酶活性测定：取100μL菌液与4.5mL去离子水混匀，加入200μL丙烯脲，在28℃反应5min，加入200μL的3mol/L盐酸终止反应。离心后取500μL上清液与500μL 4%的乙酰胺（内标）溶液混匀，进行气相色谱分析。脲水合酶活性：1U表示测量条件下催化生产1μmol丙烯酰胺所需要的酶量。

[0115] 催化稳定性评价：取20mL菌液于100mL三角瓶中，以0.4mL/min的速率流加16mL丙烯脲，由于反应剧烈放热，故三角瓶置于冰水浴中。40分钟后反应结束，将反应液离心后去掉上清液收获菌体，用去离子水重悬至20mL。取100μL菌液测定脲水合酶的剩余活性，以初始酶活为100%，计算相对酶活。

[0116] 如附图8所示，改造菌*R. ruber* TH8-2的副产物酶—酰胺酶活性为3.6U/mL，比野生菌*R. ruber* TH降低了60%。在催化稳定性评价中，丙烯脲水合反应结束后，TH8-2脲水合酶剩余酶活为62%，而野生菌仅剩27%，TH8-1剩余37%，如附图9所示，说明突变体原位替换基因组的野生型脲水合酶，可大幅度提高红球菌的催化稳定性，且效果显著优于质粒表达突变体的方式。

[0117] 实施例6重组红球菌多批次催化生产高浓度丙烯酰胺

[0118] 取400mL红球菌*R. ruber* TH、TH8-2的细胞悬液（菌浓度为1.5gdcw/L）放置于1000mL三口瓶中，冰浴条件下进行水合反应。边搅拌边滴加丙烯脲，滴加速度以控制反应温度为18-25℃来调节，当丙烯酰胺浓度达到50%时，停止滴加丙烯脲。反应得到的丙烯酰胺通过中空纤维膜与菌体分离，所得菌体回收并继续下一批水合反应，如此重复。反应过程中隔一段时间取样测定反应液中的丙烯酰胺浓度和丙烯酸浓度。如图10和图11所示，重组红球菌*R. ruber* TH8-2可以完成4批次500g/L高浓度丙烯酰胺生产，第一批次中副产物丙烯酸的浓度仅0.5g/L，而改造前的细胞只能完成1批400g/L丙烯酰胺生产，丙烯酸浓度达到2.54g/L，说明重组红球菌具有更加优良的水合催化效果，包括低副产物合成和高催化稳定

性。

[0119] 实施例7红球菌基因组表达外源酶如脲水解酶

[0120] 进一步地,借助CRISPR/Cas9基因编辑方法,本发明提出一种基因组表达外源酶的策略,将外源酶如脲水解酶基因整合到红球菌的基因组上,替换脲水合酶,利用其强启动子表达外源酶。其方法如下:靶向脲水合酶的sgRNA质粒pBNVCm-sgRNA2和携带脲水解酶基因的脲水合酶同源臂,同源臂序列如SEQ ID NO.9所示,共转化红球菌R.ruber TH(Cas9+Che9c60&61),然后按照实施例3的方法筛选转化子,并测序验证。正确的转化子按实施例3丢失质粒得到最终改造菌R.ruber THdN::Nit。

[0121] 携带外源酶的脲水合酶同源臂构建如下:脲水合酶上、下游同源臂的获取同实施例4;脲水解酶基因序列(CN105420154A)由华大青兰生物科技(无锡)有限公司合成,两端各有20bp与上、下游同源臂重叠的序列;将上、下游同源臂与脲水解酶基因使用overlap PCR进行融合,得到携带脲水解酶的同源臂nhhUp-Nit-nhhDown。

[0122] 为了对比基因组表达和质粒表达效果,将质粒pNV-Nit(CN105420154A)转化红球菌R.ruber TH,得到质粒表达脲水解酶的红球菌R.ruber TH(Nit)。

[0123] 按照专利CN105420154A培养改造菌R.ruber THdN::Nit和R.ruber TH(Nit)并测定脲水解酶酶活。如附图12所示,R.ruber THdN::Nit的酶活为56U/mL,比R.ruber TH(Nit)高51%,说明了利用基因组的脲水合酶强启动子可以实现脲水解酶的高表达。相比于质粒表达方式,利用红球菌基因组上的脲水合酶强启动子表达外源酶,酶活高、无质粒稳定性问题,具有很好的应用潜力。

[0001]	序列表	
[0002]	<110> 清华大学	
[0003]	<120> 一种基因工程红球菌及其构建方法与应用	
[0004]	<160> 9	
[0005]	<170> SIPOSequenceListing 1.0	
[0006]	<210> 1	
[0007]	<211> 4107	
[0008]	<212> DNA	
[0009]	<213> 人工序列(Artificial Sequence)	
[0010]	<400> 1	
[0011]	atggacaaga agtacagcat cggcctggac atcggcacca acagcgtggg ctgggccgtg	60
[0012]	atcaccgacg agtacaaggt gcccagcaag aagttcaagg tgctgggcaa caccgaccgc	120
[0013]	cacagcatca agaagaacct gatcggcgcc ctgctgttcg acagcggcga gaccgccgag	180
[0014]	gccaccgcc tgaagcgac cgcccggcg cgctacacc gccgcaagaa ccgcatctgc	240
[0015]	tacctgcagg agatctcag caacgagatg gccaaagtgg acgacagctt cttccaccgc	300
[0016]	ctggaggaga gcttctgtt ggaggaggac aagaagcacg agcgccacc catcttcggc	360
[0017]	aacatcgtgg acgaggtggc ctaccagag aagtaccca ccatctacca cctgcgcaag	420
[0018]	aagctggtgg acagcaccga caaggccgac ctgcgcctga tctacctggc cctggccca	480
[0019]	atgatcaagt tccgcggcca cttcctgatc gagggcgacc tgaaccccga caacagcgac	540
[0020]	gtggacaagc tgttcatcca gctggtgcag acctacaacc agctgttcga ggagaacccc	600
[0021]	atcaacgcca gcggcgtgga cgccaaggcc atcctgagcg cccgcctgag caagagccgc	660
[0022]	cgctggaga acctgatgc ccagctgcc ggcgagaaga agaacggcct gttcggcaac	720
[0023]	ctgatgccc tgagcctggg cctgaccccc aacttcaaga gcaacttca cctggccgag	780
[0024]	gacgccaagc tgcagctgag caaggacacc tacgacgacg acctggacaa cctgctggcc	840
[0025]	cagatcggeg accagtacgc cgacctgtc ctggccgcca agaacctgag cgacgccatc	900
[0026]	ctgctgagcg acatctcgc cgtgaacacc gagatcacca aggccccct gagegccagc	960
[0027]	atgatcaagc gctacgacga gcaccaccag gacctgacc tgctgaaggc cctggtgcgc	1020
[0028]	cagcagctgc ccgagaagta caaggagatc ttcttcgacc agagcaagaa cggctacgcc	1080
[0029]	ggctacatcg acggcggcgc cagccaggag gagttctaca agttcatcaa gcccatcctg	1140
[0030]	gagaagatgg acggcaccga ggagctgctg gtgaagctga accgcgagga cctgctgcgc	1200
[0031]	aagcagcga ccttcgacaa cggcagcatc ccccaccaga tccacctggg cgagctgcac	1260
[0032]	gccatcctgc gccgccagga ggacttctac cccttctga aggacaaccg cgagaagatc	1320
[0033]	gagaagatcc tgaccttcg catcccctac tacgtcggcc ccctggcccg cggcaacagc	1380
[0034]	cgcttcgect ggatgaccg caagagcgag gagaccatca ccccctggaa cttcgaggag	1440
[0035]	gtggtggaca agggcggcag cgcccagagc ttcacgagc gcatgaccaa cttcgacaag	1500
[0036]	aacctgcca acgagaaggt gctgccaag cacagcctgc tgtacgagta cttcaccgtg	1560
[0037]	tacaacgagc tgaccaaggt gaagtacgtg accgagggca tgcgcaagcc cgccttctg	1620
[0038]	agcggcgagc agaagaagge catcgtggac ctgctgttca agaccaaccg caaggtgacc	1680
[0039]	gtgaagcagc tgaaggagga ctacttcaag aagatcgagt gcttcgacag cgtggagatc	1740
[0040]	agcggcgtgg aggaccgctt caacgccagc ctgggcacct accacgacct gctgaagatc	1800
[0041]	atcaaggaca aggacttctt ggacaacgag gagaacgagg acatcctgga ggacatcgtg	1860

[0042]	ctgaccctga ccctgttcga ggaccgcgag atgatcgagg agcgcctgaa gacctacgcc	1920
[0043]	cacctgttcg acgacaaggt gatgaagcag ctgaagcgcc gccgctacac cggctggggc	1980
[0044]	cgctgagcc gcaagctgat caacggcatc cgcgacaagc agagcggcaa gaccatcctg	2040
[0045]	gacttcctga agagcgacgg cttcgccaac cgcaacttca tgcagctgat ccacgacgac	2100
[0046]	agcctgacct tcaaggagga catccagaag gccaggtga gcggccaggc cgacagcctg	2160
[0047]	cacgagcaca tcgccaacct ggccggcagc cccgccatca agaagggcat cctgcagacc	2220
[0048]	gtgaaggtgg tggacgagct ggtgaagtg atgggccgcc acaagcccga gaacatcgtg	2280
[0049]	atcgagatgg cccgcgagaa ccagaccacc cagaagggcc agaagaacag ccgcgagcgc	2340
[0050]	atgaagcgca tcgaggaggg catcaaggag ctgggcagcc agatcctgaa ggagcacccc	2400
[0051]	gtggagaaca cccagctgca gaacgagaag ctgtacctgt actacctgca gaacggccgc	2460
[0052]	gacatgtacg tggaccagga gctggacatc aaccgcctga gcgactacga cgtggaccac	2520
[0053]	atcgtgcccc agagcttct gaaggacgac agcatcgaca acaaggtgct gaccgcgagc	2580
[0054]	gacaagaacc gcgcaagag cgacaactg cccagcgagg aggtggtgaa gaagatgaag	2640
[0055]	aactactggc gccagctgct gaacgccaag ctgatcacc agcgcgaagt cgacaacctg	2700
[0056]	accaagggcg agcgcggcgg cctgagcgag ctggacaagg ccggcttcat caagcgccag	2760
[0057]	ctggtggaga cccgccagat caccaagcac gtggcccaga tcttgacag ccgcatgaac	2820
[0058]	accaagtacg acgagaacga caagctgatc cgcgaggtga aggtgatcac cctgaagagc	2880
[0059]	aagctggtga gcgacttccg caaggacttc cagttctaca aggtgcgcga gatcaacaac	2940
[0060]	taccaccacg cccacgacgc ctacctgaac gccgtggtgg gcaccgccct gatcaagaag	3000
[0061]	taccccaagc tggagagcga gttcgtgtac ggcgactaca aggtgtacga cgtgcgcaag	3060
[0062]	atgatcgcca agagcgagca ggagatcggc aaggccaccg ccaagtactt cttctacagc	3120
[0063]	aacatcatga acttcttcaa gaccgagatc accctggcca acggcgagat ccgcaagcgc	3180
[0064]	ccctgatcag agaccaacgg cgagaccggc gagatcgtgt gggacaaggc ccgcgacttc	3240
[0065]	gccaccgtgc gcaaggtgct gagcatgccc caggtgaaca tcgtgaagaa gaccgaggtg	3300
[0066]	cagaccggcg gcttcagcaa ggagagcatc ctgcccagc gcaacagcga caagctgatc	3360
[0067]	gcccgaaga aggactggga cccaagaag tacggcggct tcgacagccc caccgtggcc	3420
[0068]	tacagcgtgc tgggtggtgg caaggtggag aagggaaga gcaagaagct gaagagcgtg	3480
[0069]	aaggagctgc tgggcatcac catcatggag cgcagcagct tcgagaagaa ccccatcgac	3540
[0070]	ttcttgagg ccaagggcta caaggagtg aagaaggacc tgatcatcaa gctgcccag	3600
[0071]	tacagcctgt tcgagctgga gaacggccgc aagcgcagc tggccagcgc cggcgagctg	3660
[0072]	cagaagggca acgagctggc cctgcccagc aagtacgtga acttctctgta cctggccagc	3720
[0073]	cactacgaga agctgaaggc cagccccgag gacaacgagc agaagcagct gttcgtggag	3780
[0074]	cagcacaagc actacctgga cgagatcatc gagcagatca gcgagttcag caagcgcgtg	3840
[0075]	atcctggccg acgccaacct ggacaagtg ctgagcgctt acaacaagca ccgcgacaag	3900
[0076]	cccatccgcg agcaggccga gaacatcatc cacctgttca ccctgaccaa cctgggcgcc	3960
[0077]	ccgcccct tcaagtactt cgacaccacc atcgaccgca agcgcctacac cagcaccaag	4020
[0078]	gagtgctgg acgccacct gatccaccag agcateaccg gcctgtacga gaccgcgac	4080
[0079]	gacctgagcc agctgggcgg cgactga	4107
[0080]	<210> 2	
[0081]	<211> 197	
[0082]	<212> DNA	
[0083]	<213> 人工序列(Artificial Sequence)	

[0084]	<400> 2	
[0085]	gttttagagc tagaaatagc aagttaaaat aaggctagtc cgttatcaac ttgaaaaagt	60
[0086]	ggcaccgagt cgggtctttt tactccatct ggatttggtc agaacgctcg gttgccgceg	120
[0087]	ggcgtttttt atctaaagct taggccagct cgaagactg ggccttttta atacgactca	180
[0088]	ctatagggtc gacggcc	197
[0089]	<210> 3	
[0090]	<211> 602	
[0091]	<212> DNA	
[0092]	<213> 人工序列(Artificial Sequence)	
[0093]	<400> 3	
[0094]	cactgtccgt cgacggtgac caccagcttg gcgcgggctg cgtcgatgctg tgagcgcagc	60
[0095]	gcgtgcgagg agaaccgcc gaagaccacc gagtgggtca ggcccaggcg cgcgcaggcg	120
[0096]	agcatcgtca cgatgcctc ggggatcctc ggcatgtaga tcgccaccgc gtcaccggcg	180
[0097]	accaaaccga gggcggtgaa ggtgttcgcc gccttgaca cctcggcgag cagatcgttg	240
[0098]	taggtcagcg cccgcgaatc gccgggctcg ccctcgaagt ggatcgcgac ccggtcgcg	300
[0099]	ttgccggcga gcacgtccg gtcgacgag ttgtaggcga cgttgagctg cccgccgacg	360
[0100]	aaccacttcg ccaccggcgc ctccgaccag tccagcacct cgggtccacgg cgtggcccag	420
[0101]	tccagccggc gggcctcgt cgcccagaac ccgagccggt cggcgtccgc ggcggcctgc	480
[0102]	aactcgggac cggcgggtgc ggtggcgacg aactctcgg acggcggata cgtcgtggct	540
[0103]	ggagaaggtt ggacggctgt catgataggc acctttctc aacgtctttg aagtgtgtgt	600
[0104]	ac	602
[0105]	<210> 4	
[0106]	<211> 1018	
[0107]	<212> DNA	
[0108]	<213> 人工序列(Artificial Sequence)	
[0109]	<400> 4	
[0110]	accatcgcga gatcagactg ttcgaccctg gcgacaaa aacctcggcg tggaatcgga	60
[0111]	ttgagacggt accggccgca aaggaccact cgggcaatcg cggatcgaac ctctttgct	120
[0112]	tccccaccgc aaagccgagt cgactgcaat atcgggtgctt cgaggaatct cggggacaga	180
[0113]	gcagtgcacg gtcgatccg taccagaaa gcgaccgca cgtagcccc gcaactctgc	240
[0114]	ccgccaggga taccgtatc gaagtacac tgccagatcc gagcagtccc acccccgatg	300
[0115]	cgacgcatca cggggtcctc gaaccagcag gacccttccc ctccccattc gtctttcagg	360
[0116]	agatcacct atgaccactg ccggcacttc gaagcagaca cgttggcctc tcaaaggcca	420
[0117]	gtggttcgac atttgcagtt gcatcctgcc ctgcggctgc accatggcac agcctcccac	480
[0118]	cgatggggtg tgctacggca ccctggtgta tcagatcgac gaggggtact tcggcgagct	540
[0119]	ggatctgtcc ggtctgacgg tggtcacgat cggcgagatc aagagcgaga acttgtggga	600
[0120]	ttccagcaag cccgtcgaag ggatctacga tctcatcctc gacgagcagc ccgataccga	660
[0121]	tcagcgcgat gcgctcgagc gtctgtggac aggccaggag ggtgggtgga tcgccaacct	720
[0122]	cgtcggcctg ttggggacgg tgcgccagct ggagtacgcc ccgatcagct gtcggatcga	780
[0123]	aggggacctg gcgcgggtga gcatcgacgt gccgggtcgg gtcagcggcg cggtcgacgc	840
[0124]	cctgtccggc cccaccacgc ctccaggaca acgctacag acgttcaatc cgcccggatc	900
[0125]	ggaaaccggc gggctgccgg cgacctgggg agttctctac aagatggaac tggtggactt	960

[0126]	cgggtttctc ggagagtgga cggccaagtc cagcaaacac attccgttcg actggtcc	1018
[0127]	<210>	5
[0128]	<211>	2397
[0129]	<212>	DNA
[0130]	<213>	人工序列(Artificial Sequence)
[0131]	<400>	5
[0132]	cactgtccgt cgacggtgac caccagcttg gcgcgggctg cgtcgatgag tgagcgcage	60
[0133]	gcgtgcgagg agaaccgcc gaagaccacc gagtgggtca ggcccaggcg cgcgcaggcg	120
[0134]	agcatcgtca cgatcgctc ggggatcatc ggcatgtaga tcgccaccgg gtcaccggcg	180
[0135]	accaaacca gggcggtgaa ggtgttcgcc gccttggaca cctcggcgag cagatcgttg	240
[0136]	taggtcagcg cccgcgaatc gccgggctcg ccctcgaagt ggatcgcgac ccggtcgcgg	300
[0137]	ttgccggcga gcacgtgccc gtcgacgag ttgtaggcga cgttgagctg cccgccgacg	360
[0138]	aaccacttcg ccaccggcgc ctcggaccag tccagcacct cggtcacagg cgtggcccag	420
[0139]	tccagccggc gggcctgctg cgcccagaac ccgagccggt cggcgtccgc ggcggcctgc	480
[0140]	aactcgggac cggcggtggc ggtggcgacg aactctcgg acggcggata cgtcgtggct	540
[0141]	ggagaagggt ggacggtcgt catgatagc acctttctc aacgtctttg aagtgtgtgt	600
[0142]	acaacagggt agattacgga gaacgggct tgtggccgtc cctgtcgtgt cgtaacgtgt	660
[0143]	ccacaacggt gcagttcact gcagtcatga gcgcccggcg agtcactaag gagggatcca	720
[0144]	tggagaaaaa aatcactgga tataccaccg ttgatatac ccaatggcat cgtaaagaac	780
[0145]	atthttaggc atthcagtc gttgctcaat gtacctataa ccagaccgtt cagctggata	840
[0146]	ttacggcctt tttaaagacc gtaaagaaaa ataagcaca gttttatccg gcctttatcc	900
[0147]	acattcttgc ccgctgatg aatgctcacc cggagtccg tatggcaatg aaagacgggtg	960
[0148]	agctggtgat atgggatagt gttcacctt gttacaccgt tttccatgag caaactgaaa	1020
[0149]	cgttttcacc gctctggagt gaataccacg acgatttccg gcagtttcta cacatatatt	1080
[0150]	cgcaagatgt ggcgtgttac ggtgaaaacc tggcctatth ccctaaaggg thtattgaga	1140
[0151]	atatgttttt cgtctcagcc aatccctggg tgagtttacc cagttttgat ttaaactgtg	1200
[0152]	ccaatatgga caactcttc gccccgtht tcaccatggg caaatattat acgcaaggcg	1260
[0153]	acaagggtgt gatgccgtg gcgattcagg ttcacatgac cgtctgtgat ggcttccatg	1320
[0154]	tcggcagaat gcttaatgaa ttacaacagt actgcgatga gtggcagggc ggggcgtaaa	1380
[0155]	cccatcgcag atcagactgt tcgaccgtg gcgacaaaa acctcggcgt ggaatcggat	1440
[0156]	tgagacggta ccggcccaa aggaccactc gggcaatcgc ggatcgaacc tccttgcctt	1500
[0157]	ccccaccgca aagccgagtc gactgcaata tcggtgcttc gaggaatctc cgggacagag	1560
[0158]	cagtgcacgg tcgatccgt acccagaaag cgaccgcgac gtagcccccg caactctgcc	1620
[0159]	cgccagggat accgctatcg aagtcacact gccagatccg agcagtccca cccccgatgc	1680
[0160]	gacgcacac ggggtcatcg aaccagcagg acccttccc tccccattcg tctttcagga	1740
[0161]	gatcacccta tgaccactgc cggcacttcg aagcagacac gttggcatct caaaggccag	1800
[0162]	tggttcgaca tttgagttg catcctgccc tgcggctgca ccatggcaca gcctcccacc	1860
[0163]	gatgggggtg gctacggcac cctggtgtat cagatcgacg aggggtactt cggcgagctg	1920
[0164]	gatctgtccg gtctgacggt ggtcacgac ggcgagatca agagcgagaa cttgtgggat	1980
[0165]	tccagcaagc ccgtcgaagg gatctacgat ctcatcatcg acgagcgagc cgataccgat	2040
[0166]	cagcgcgatg cgctcgagcg tctgtggaca ggccaggagg gtgggtggat cgccaacctc	2100
[0167]	gtcggcctgt tggggacggt gcgccagctg gactacgccc cgatcgagtg tcggatcga	2160

[0168]	ggggacctgg cgcggtggag catcgactg cgggtcggg tcagcggcgc ggtcgcgcc	2220
[0169]	ctgtccggcc ccaccacgcc tccaggacaa cgcgtacaga cgttcaatcc gcccgatcg	2280
[0170]	gaaaccggcg ggtcgcggc gacctgggga gttcctacga agatggaact ggtggacttc	2340
[0171]	gggtttctcg gagagtggac ggccaagtcc agcaaacaca ttccgttcga ctggtcc	2397
[0172]	<210> 6	
[0173]	<211> 1694	
[0174]	<212> DNA	
[0175]	<213> 人工序列(Artificial Sequence)	
[0176]	<400> 6	
[0177]	gtgaccacca gcttggcgcg ggcgtcgtcg atgcgcgagc gcagcgcgtg cgaggagaac	60
[0178]	ccgccgaaga ccaccgagtg ggtcaggccc aggcgcgcgc aggcgagcat cgtcacgac	120
[0179]	gcctcgggga tcatcgcat gtagatgcc acccggtcac cggcgaccaa acccagggcg	180
[0180]	gtgaaggtgt tcgccctt ggacacctg gcgagcagat cgtttaggt cagcggcgc	240
[0181]	gaatcggcg gctcgcctc gaagtggatc gcgaccggc cggcgttgc ggcgagcacg	300
[0182]	tgccggtcga cgcagttgta ggagcgtg agctgccgc cgacgaacca cttgccacc	360
[0183]	ggcgctcgg accagtccag cacctcgtc cacggcgtgg ccagtccag ccggcgggc	420
[0184]	tgcgtcggc agaaccgag ccggtcggc tccgcggcg cctgcaactc gggaccggcg	480
[0185]	gtggcggtag cgacgaactc tgcggacggc ggatacgtc tggctggaga aggttgacg	540
[0186]	gtcgtcatga taggcacctt ttctcaactc ctttgaagt ctggtacatc cgcagcggc	600
[0187]	gatgctcaga gaatacatg tgcctaacg aagtaaagat ccacggaggt ggacgtgca	660
[0188]	aggaacggc cctgcctatc gctgtgaaca ggtgagatta cggagaacgg ggcttgtggc	720
[0189]	cgtccctgtc gtgtcgtaac gtgtccaca cgttgcagt catgcaatg ggaacactc	780
[0190]	aagtccgaag caaacgtcgg gtcatgagc cccggcagc cactaaggag tcttagacc	840
[0191]	atcgcagatc agactgttcg acccgtggc accaaaaacc tcggcgtgga atcggattg	900
[0192]	gacggtaccg gccgcaagg accactcgg caatcgcgga tcgaacctc ttgccttccc	960
[0193]	caccgcaag ccgagtcgac tgcaatatc gtgcttcgag gaatctccg gacagagcag	1020
[0194]	tgcacggtc atgccgtacc cagaaagcga ccgcgacgta gccccgca ctctgccgc	1080
[0195]	caggatacc gctatcgaag tcaactgcc agatccgagc agtcccacc ccgatgcgac	1140
[0196]	gcatacggg gtatcgaac cagcaggacc ctccctcc ccattcgtct ttcaggagat	1200
[0197]	caccctatga cactgccg cacttcgaag cagacagtt ggcattcaa aggccagtgg	1260
[0198]	ttgacattt gcagttgat cctgccctc ggctgcacca tggcacagc tcccaccgat	1320
[0199]	ggggtgtgct acggcaccct ggtgtatcag atcgcagagg ggtacttcg cgagctggat	1380
[0200]	ctgtccggtc tgacggtggt cagatcggc gagatcaaga gcgagaactt gtgggattc	1440
[0201]	agcaagccc tcgaaggat ctacgatctc atcatcgac agcagccga taccgatcag	1500
[0202]	cgcatgcgc tcgagctct gtggacaggc caggagggtg ggtgatcgc caacctcgtc	1560
[0203]	ggcctgttg ggacggtgc ccagctggag tacgcccga tcgagtgtc gatcgaagg	1620
[0204]	gacctggcgc ggtggagcat cgacgtccg ggtcgggtca gcggcggcgc cgacgcctg	1680
[0205]	tccggcccca ccac	1694
[0206]	<210> 7	
[0207]	<211> 919	
[0208]	<212> DNA	
[0209]	<213> 人工序列(Artificial Sequence)	

[0210]	<400> 7	
[0211]	cacacaccgg gtcaccggag ccctgtacgc cgccgcgcg cagggtctgc tcaccctcgc	60
[0212]	cgaccagggc tatcagggca cggcatcgg tatccacatg cccacgaaag ccctgccga	120
[0213]	cggcaacacc ctcgacaccg acaccgtctg ccgcaacatg ctgctgacca gtctgcggtg	180
[0214]	cctcgcgaa cgtgccgtg cgctgtcac cacgcgatgg aaggcactcg acaggatcac	240
[0215]	cctgtcccc aaaaggatcg gttccatcac caaagcgcg ctcgtactca cgcaattcga	300
[0216]	gcacgcaggc cgttactgag aaaacctcaa tgctgatcac cctgccgccc ttggacgacc	360
[0217]	acggttgcta cgagtgtcg gagccaacca taggcatcat gcgatcgccg gagtcttcat	420
[0218]	cctgttttg gatgvcagg attaacacat ctacacattg acatccgttc cgatgtgaag	480
[0219]	taaaaattgt cacgtagggc ggcaggcga gtctgcagct cgaacatcga aggggtgggag	540
[0220]	ccgagagatc ggagacgcag acaccggag ggaacctagc ctccggacc gatgvcgtgc	600
[0221]	ctggcaacgc ctcaagattc agcgaacgc attcaatctt gttacttcca gaaccgaatc	660
[0222]	acgtcccct agtvtgccc gagvcgccc gaacgcagg atggtatcca tgcgcccctt	720
[0223]	ctcttttga acgagaaccg gccgctacag ccgaccgga gacactgtga cgccgttcaa	780
[0224]	cgattgtgt gctgtgaagg attcactcaa gccactgat atcgccattc cgttgccgga	840
[0225]	acatttgaca cttctcct acgagtagaa gccagctgga cccctcttg agcccagctc	900
[0226]	cgatgaaagg aatgaggaa	919
[0227]	<210> 8	
[0228]	<211> 815	
[0229]	<212> DNA	
[0230]	<213> 人工序列(Artificial Sequence)	
[0231]	<400> 8	
[0232]	gtgaagacac actcactgat cggctcccgg cgactgggac cgccgcaccg ccccgcgaca	60
[0233]	atggcgagct tgtattcacc gagccttggg aagcaacggc attcggggtc gccatcgcgc	120
[0234]	tttcggatca gaagtcgtac gaatgggagt tcttccgaca gcgtctcatt cactccatcg	180
[0235]	ctgaggccaa cggttgcgag gcatactacg agagctggac aaaggcgcgc gaggccagcg	240
[0236]	tggtcgactc ggggctgatc agcgaagatg agatccgcga gcgcatggaa tcgatggcca	300
[0237]	tcacgcactg acatcccctg cgtttccatc cagcagcagt gcgggcgtac cccgacgggg	360
[0238]	ctgagccgac ggggtacgcc cgcacttcat caatgacgtt ctgaggcaca cagcgcagag	420
[0239]	tcgagctgag tgctcagaa cgtcatgacg gtggttcta attcggctcg gtgggtactg	480
[0240]	agctcgcgga aggtaacgcg gtgacgctgt aggcgttcat ggcaagtggg actccggtgc	540
[0241]	gccgagcctg aggtgttcca tatggtatac cgcttcgtcc aacaacgacg cgacatgcga	600
[0242]	gtcgtagagg ctgtacacca cgctacgacc actgcgttcg ccgatcacca accgcagcgc	660
[0243]	tcgcaacagg cggagctggt gaaaaaccgc aggttgttcc ataccgacag cctcggcgag	720
[0244]	ctcggtgact ccgcacggtc cttgccgcaa cctggccaaa atcaacagtc gattcgggtga	780
[0245]	cgccaatgcc tgcaaagtct ccgcgactgt tgccc	815
[0246]	<210> 9	
[0247]	<211> 2835	
[0248]	<212> DNA	
[0249]	<213> 人工序列(Artificial Sequence)	
[0250]	<400> 9	
[0251]	cacacaccgg gtcaccggag ccctgtacgc cgccgcgcg cagggtctgc tcaccctcgc	60

[0252]	cgaccagggc taccagggca ccggcatcgg tatccacatg cccacgaaag cccctgccga	120
[0253]	eggcaacacc ctgacaccg acaccgtctg ccgcaacatg ctgctgacca gtctgcgggtg	180
[0254]	cctcgccgaa cgtgccgtg cgtctctcac caccgatgg aaggcactcg acaggatcac	240
[0255]	cctgtgcccc aaaaggatcg gttccatcac caaagcggcg ctctactca cgcaattcga	300
[0256]	gcacgcaggc cgttactgag aaaacctcaa tgctgatcac cctgccgccg ttggacgacc	360
[0257]	acggttgcta cgagtgtcg gagccaacca taggcatcat gcgatgccg gagtcttcat	420
[0258]	cctgttttgg gatgcgcagg attaacacat ctacacattg acatccgttc cgatgtgaag	480
[0259]	taaaaattgt cacgtagggc ggcaggcgaa gtctgcagct cgaacatcga aggggtggag	540
[0260]	ccgagagatc ggagacgcag acaccggag ggaacctagc ctcccggacc gatgcgtgtc	600
[0261]	ctggcaacgc ctcaagattc agcgcgaagcg attcaatctt gttacttcca gaaccgaatc	660
[0262]	acgtccccgt agtgtgcggg gagagcggc gaacgcaggg atggtatcca tgcgccctt	720
[0263]	ctcttttga acgagaaccg gccgctacag ccgaccgga gacactgtga cgccgttcaa	780
[0264]	cgattgttgt gctgtgaagg attcactcaa gccactgat atgccattc cgttgccgga	840
[0265]	acatttgaca ctttccct acgagtagaa gccagctgga cccctcttg agcccagtc	900
[0266]	cgatgaaagg aatgaggaaa tggtcgaata cacaacaca ttcaaagttg ctgcgggtgca	960
[0267]	ggcacagcct gtgtggttcg acgcggccaa aacggtcgac aagaccgtgt ccatcatcgc	1020
[0268]	ggaagcagcc cggaacgggt gcgagctcgt tgcgtttccc gaggtattca tcccgggta	1080
[0269]	cccgtaccac atctgggtcg acagcccgt cgccggaatg gcgaagttcg ccgtgcgcta	1140
[0270]	ccacgagaat tcctgacga tggacagccc gcacgtacag cggttgctcg atgcccccg	1200
[0271]	cgaccacaac atcgccgtag tgggtgggaat cagcagcgg gatggcggca gcttgatcat	1260
[0272]	gaccagctc atcatcgac ccgatgggca actggtcgcc cgacccgca agctcaagcc	1320
[0273]	caccacgtc gagcgttcg tatacggaga aggaaacggc tcggatatct ccgtgtacga	1380
[0274]	catgcctttc gcacggcttg gcgcgctcaa ctgctgggag catttccaga cgctaccaa	1440
[0275]	gtacgcaatg tactcgatgc acgagcaggt gcacgtcgcg agctggcctg gcatgtcgt	1500
[0276]	gtaccagccg gagtccccg cattcggtgt cgatgccag ctacggcca cgcgtatgta	1560
[0277]	cgactcgag ggacaaacct tcgtggtctg caccaccag gtggtcacac cggaggccca	1620
[0278]	cgagtcttc tgcgagaacg aggaacagcg aaagttgatc ggccgaggcg gaggtttcgc	1680
[0279]	gcgcatcatc gggcccagc gccgcgatct cgcaactcct ctgccgaag atgaggaggg	1740
[0280]	gatctctac gccgacatc atctgtctgc gatcacttg gcgaagcagg ccgctgacc	1800
[0281]	cgtgggccac tactcacggc cggatgtgct gtcgtgaaac ttcaaccagc gccgcaccac	1860
[0282]	gcccgtcaac acccacttt ccaccatcca tgccacgcac acgttcgtgc cgcagttcgg	1920
[0283]	ggcactcgac ggcgtccgtg agctcaacgg agcggacgaa cagcgcgat tgcctccac	1980
[0284]	acattccgac gagacggacc gggcgacagc ctccatctga gtgaagacac actcactgat	2040
[0285]	cggtccccg cgactgggac cccgcaccg ccccgacaca atggcgagct tgtattcacc	2100
[0286]	gagccttggg aagcaacggc attcggggtc gccatcgcgc tttcggatca gaagtcgtac	2160
[0287]	gaatgggagt tcttccgaca gcgtctcatt cactccatcg ctgaggccaa cggttgcgag	2220
[0288]	gcatactac agagctggac aaaggcgtc gagccagcg tggtcgactc ggggctgatc	2280
[0289]	agcgaagatg agatccgca gcgcatggaa tcgatggcca tcatcgactg acatcccctg	2340
[0290]	cgtttccatc cagcagcagt gcggcgctac cccgacggg ctgagccgac ggggtacgcc	2400
[0291]	cgacttcat caatgacgtt ctgaggcaca cagcgcagag tcgagctgag tgcctcagaa	2460
[0292]	cgatcatgac gtggttcccta attcggctcg gtgggtactg agctcgcgga aggtaacgcg	2520
[0293]	gtgacgctgt aggcgttcat ggcaagtggg actccggtgc gccgagcctg aggtgttcca	2580

[0294]	tatggatac cgcttcgtcc aacaacgacg cgacatgcga gtcgtagagg ctgtacacca	2640
[0295]	cgctacgacc actgcgttcg ccgatcacca accgcagcgc tcgcaacagg cggagctggt	2700
[0296]	gagaaaccgc aggttggtcc ataccgacag cctcggecag ctcggtgact ccgcacggtc	2760
[0297]	cttgccgcaa cctggccaaa atcaacagtc gattcgggtga cgccaatgcc tgcaaagtct	2820
[0298]	ccgcgactgt tgccg	2835

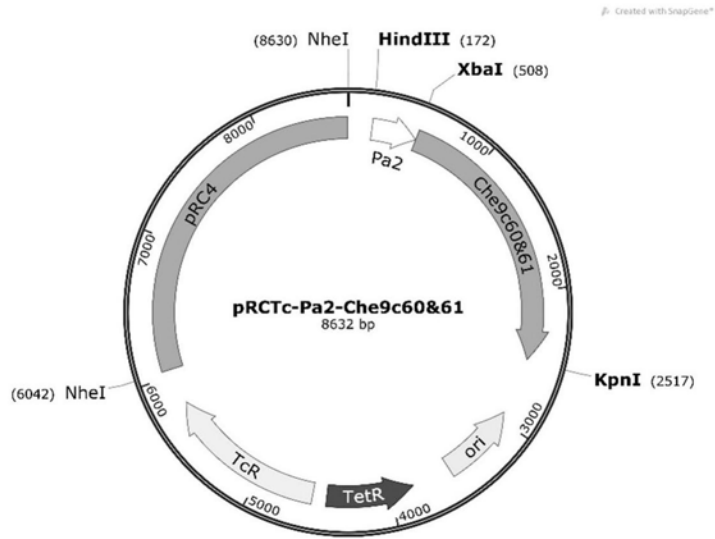


图1

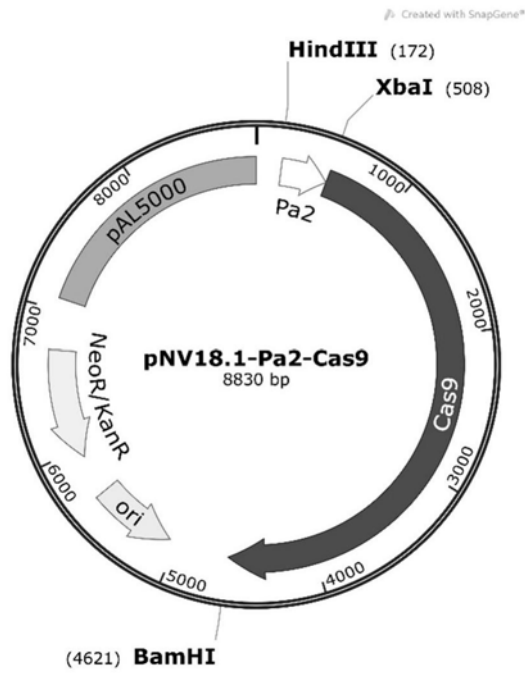


图2

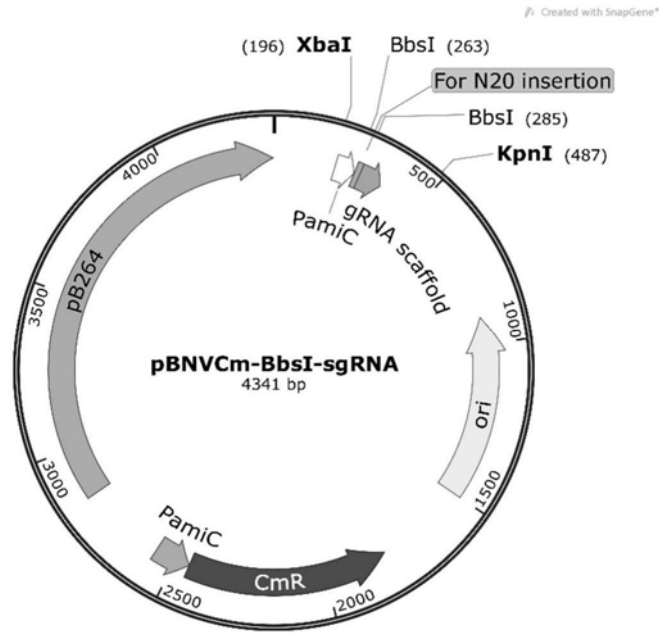


图3

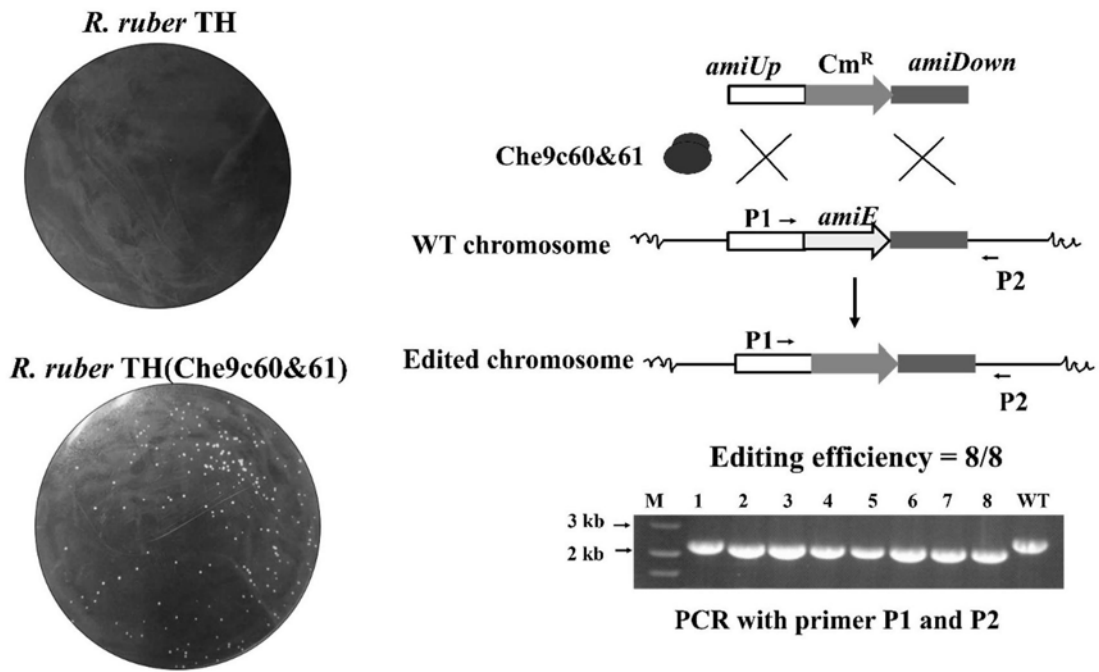


图4

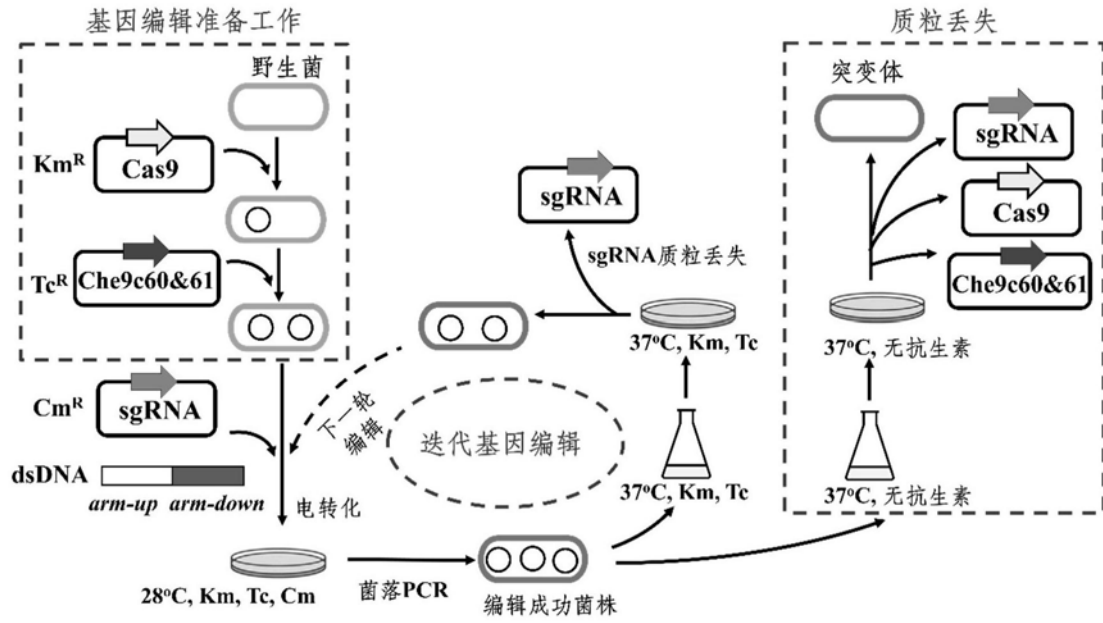


图5

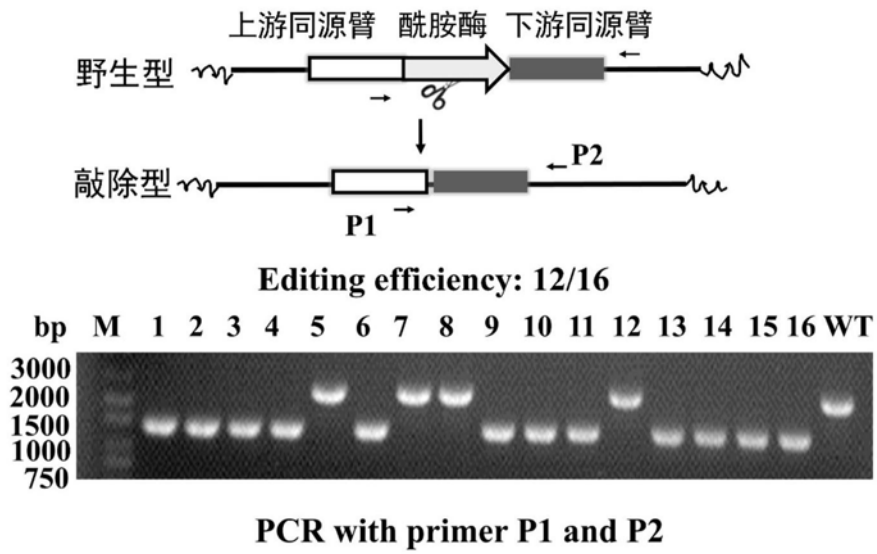


图6

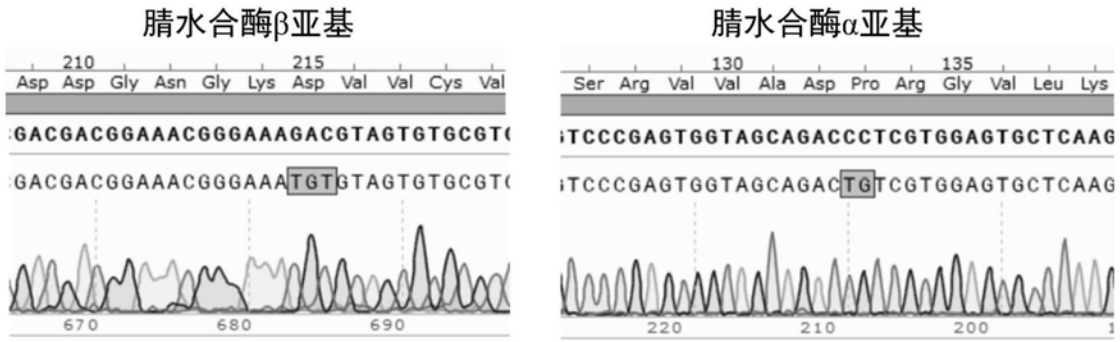


图7

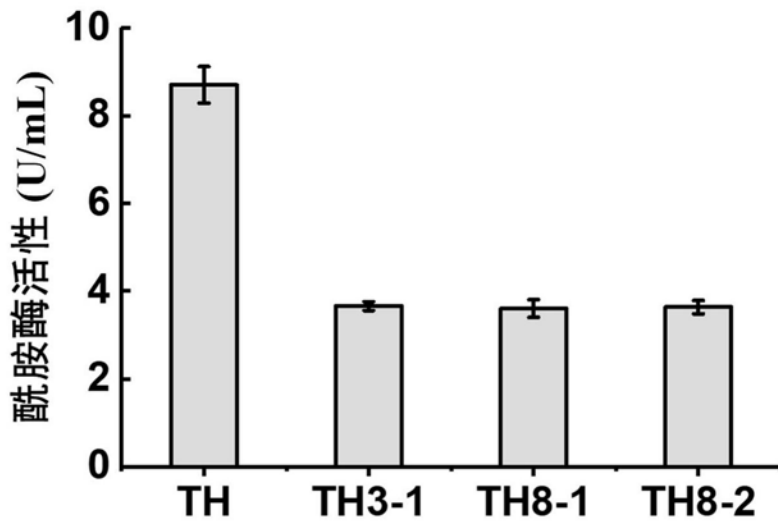


图8

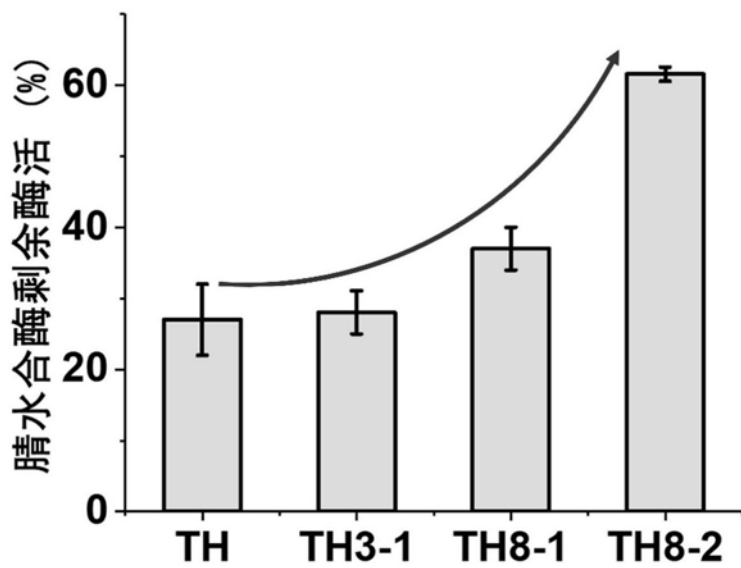


图9

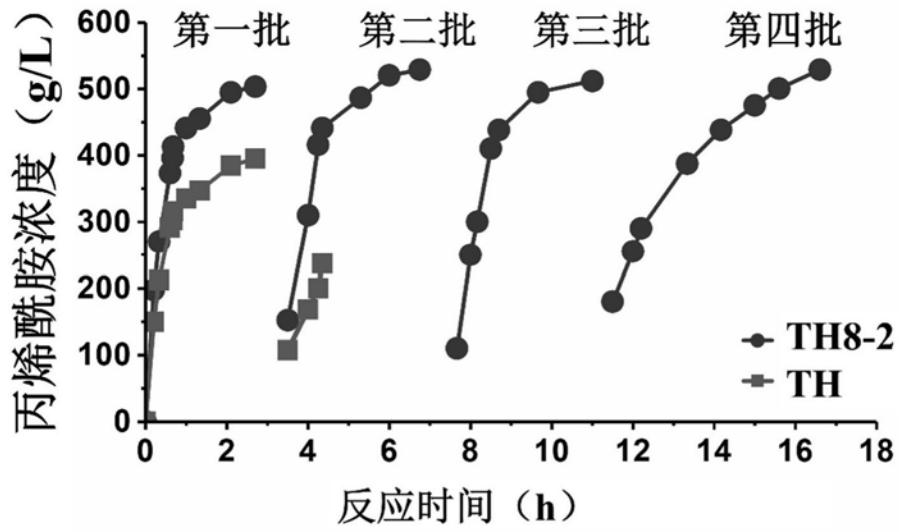


图10

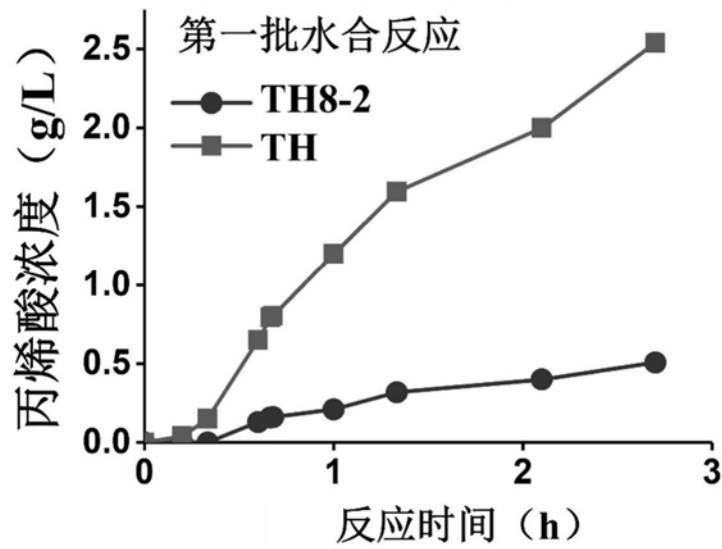


图11

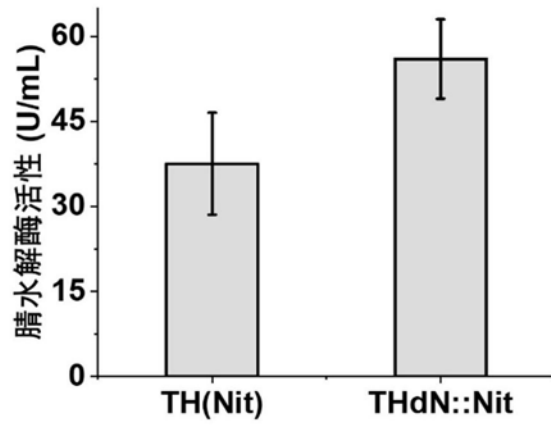


图12