

1. 一种融合蛋白,其特征在于,所述融合蛋白具有式Ia或式Ib结构:

S-A-B-C (Ia)

S-C-B-A (Ib);

式中,

A为PabI元件;

B为无或连接肽;

C为eIF4G元件;

S为任意的信号肽;以及

各“-”为肽键。

2. 一种分离的多核苷酸,其特征在于,所述的多核苷酸编码权利要求1所述的融合蛋白。

3. 一种载体,其特征在于,所述载体含有权利要求2所述的多核苷酸。

4. 一种宿主细胞,其特征在于,所述宿主细胞含有权利要求3所述的载体或基因组中整合有权利要求2所述的多核苷酸。

5. 一种用于表达外源蛋白的体外蛋白质合成体系,其特征在于,所述反应体系包括:

(i) 酵母体外蛋白质合成体系,所述的合成体系含有(a) 酵母细胞提取物;(b) 任意的聚乙二醇;(c) 任意的的外源蔗糖;和(d) 任意的溶剂,所述溶剂为水或水性溶剂;以及

(ii) 权利要求1所述的融合蛋白。

6. 如权利要求5所述的体外蛋白质合成体系,其特征在于,所述反应体系还包括:(iii) 额外添加的eIF4G蛋白。

7. 如权利要求6所述的体外蛋白质合成体系,其特征在于,所述eIF4G蛋白由组成型或诱导性启动子诱导表达。

8. 一种生产权利要求1所述融合蛋白的方法,其特征在于,包括:

(i) 在适合表达的条件下,培养权利要求4所述的宿主细胞,从而表达出权利要求1所述的融合蛋白;和

(ii) 分离所述融合蛋白。

9. 一种权利要求1所述的融合蛋白的用途,其特征在于,用于制备提高体外蛋白质合成体系的体外蛋白合成能力的制剂。

10. 一种表达待表达的外源蛋白的方法,其特征在于,包括:

(i) 提供一酵母体外蛋白质合成体系,其中所述的合成体系中含有权利要求1所述的融合蛋白;和

(ii) 在适合表达蛋白的条件下,在所述外源蛋白的模板存在下,孵育所述酵母体外蛋白质合成体系,从而表达所述的外源蛋白。

新型融合蛋白的制备及其在提高蛋白质合成的应用

技术领域

[0001] 本发明涉及基因工程领域,具体地,涉及新型融合蛋白的制备及其在提高蛋白质合成的应用。

背景技术

[0002] 蛋白质是细胞中的重要分子,几乎参与了细胞所有功能的执行。蛋白的序列和结构不同,决定了其功能的不同。在细胞内,蛋白可以作为酶类催化各种生化反应,可以作为信号分子协调生物体的各种活动,可以支持生物形态,储存能量,运输分子,并使生物体运动。在生物医学领域,蛋白质抗体作为靶向药物,是治疗癌症等疾病的重要手段。

[0003] 在细胞中,蛋白质翻译的调节在应对营养缺失等外界压力,细胞发育与分化等很多过程中发挥重要作用。蛋白质翻译的四个过程包括翻译起始、翻译延伸、翻译终止和核糖体再循环,其中翻译起始是受调控最多的一个过程。在翻译起始阶段,核糖体小亚基(40S)结合(tRNA)^{iMet},并在翻译起始因子的作用下识别mRNA 5'末端。小亚基向下游移动,并在起始密码子(ATG)位置与核糖体大亚基(60S)结合,形成完整核糖体,并进入翻译延伸阶段。

[0004] 在快速分裂的酵母细胞中,蛋白的合成速率大约为13,000个/秒。在体内,蛋白的合成速率受到核糖体数目的限制,细胞的平均核糖体数目约为200,000个,mRNA分子的数目约为15,000-60,000个。

[0005] 目前,经常实验的商业化体外蛋白表达系统包括大肠杆菌系统(*E.coli* extract, ECE)、兔网织红细胞(Rabbit reticulocyte Lysate, RRL)、麦胚(Wheat germ extract, WGE)、昆虫(Insect cell extract, ICE)和人源系统。

[0006] 现有的商业化蛋白质体外合成体系中,原核系统的产量可以达到 ~0.5 mg/mL,费用约为 ~10 RMB/ μ g。真核系统中CHO系统的产量可以达到 ~0.7 mg/mL,费用约为 ~20 RMB/ μ g。因此无论自然界中细胞内还是细胞外的人造蛋白质合成体系都具有效率低,速度慢的特点,极大的限制了蛋白质合成的应用。

[0007] 因此,本领域迫切需要开发一种可以有效增强体外蛋白质合成效率的体外蛋白质合成体系。

发明内容

[0008] 本发明的目的在于提供一种可以有效增强体外蛋白质合成效率的体外蛋白质合成体系。

[0009] 本发明第一方面提供了一种融合蛋白,所述融合蛋白具有式Ia或式Ib结构:

S-A-B-C (Ia)

S-C-B-A (Ib);

式中,

A为PabI元件;

B为无或连接肽;

C为eIF4G元件；
S为任意的信号肽；以及
各“-”为肽键。

- [0010] 在另一优选例中,所述式Ia或Ib为从N端至C端的结构。
- [0011] 在另一优选例中,所述元件A包括野生型和突变型的PabI序列。
- [0012] 在另一优选例中,所述的PabI为来自酵母的PabI。
- [0013] 在另一优选例中,元件A具有SEQ ID NO.:1所示的序列或其活性片段,或者具有与SEQ ID NO.:1所示氨基酸序列 $\geq 85\%$ 同源性(优选地, $\geq 90\%$ 的同源性;等优选地 $\geq 95\%$ 的同源性;最优选地, $\geq 97\%$ 的同源性,如98%以上,99%以上)且具有与SEQ ID NO.:1序列相同活性的多肽。
- [0014] 在另一优选例中,所述元件C包括野生型和突变型的eIF4G序列。
- [0015] 在另一优选例中,所述的eIF4G为来自酵母的eIF4G。
- [0016] 在另一优选例中,元件C具有SEQ ID NO.:2所示的序列或其活性片段,或者具有与SEQ ID NO.:2所示氨基酸序列 $\geq 85\%$ 同源性(优选地, $\geq 90\%$ 的同源性;等优选地 $\geq 95\%$ 的同源性;最优选地, $\geq 97\%$ 的同源性,如98%以上,99%以上)且具有与SEQ ID NO.:2序列相同活性的多肽。
- [0017] 在另一优选例中,所述融合蛋白是重组蛋白,较佳地为酵母表达的重组蛋白。
- [0018] 在另一优选例中,所述酵母选自下组:克鲁维酵母、酿酒酵母、或其组合。
- [0019] 在另一优选例中,所述酵母选自下组:乳酸克鲁维酵母、马克斯克鲁维酵母、多布克鲁维酵母、或其组合。
- [0020] 在另一优选例中,所述元件A衍生自酵母的PabI蛋白。
- [0021] 在另一优选例中,所述元件C衍生自酵母的eIF4G蛋白。
- [0022] 在另一优选例中,所述的肽接头的长度为0-50氨基酸,较佳地为10-40个氨基酸,更佳地为15-25个氨基酸。
- [0023] 在另一优选例中,所述融合蛋白选自下组:
- (A) 具有SEQ ID NO:3所示氨基酸序列的多肽;(B) 具有与SEQ ID NO:3所示氨基酸序列 $\geq 80\%$ 同源性(优选地, $\geq 90\%$ 的同源性;等优选地 $\geq 95\%$ 的同源性;最优选地, $\geq 97\%$ 的同源性,如98%以上,99%以上)的多肽,且所述多肽具有提高外源蛋白表达效率的功能或活性;
- (C) 将SEQ ID NO:3中任一所示氨基酸序列经过1-15(较佳地,2-10,更佳地,3-8)个氨基酸残基的取代、缺失或添加而形成的,且具有提高外源蛋白表达效率的功能或活性的衍生多肽。
- [0024] 在另一优选例中,所述融合蛋白的氨基酸序列如SEQ ID NO.:3所示。
- [0025] 在另一优选例中,所述融合蛋白具有选自下组的一个或多个特性:
- (a) 提高外源蛋白表达效率;
- (b) 提高体外翻译效率。
- [0026] 在另一优选例中,所述外源蛋白选自下组:荧光素蛋白、或荧光素酶(如萤火虫荧光素酶)、绿色荧光蛋白、黄色荧光蛋白、氨酰 tRNA 合成酶、甘油醛-3-磷酸脱氢酶、过氧化氢酶、肌动蛋白、抗体的可变区域、萤光素酶突变、 α -淀粉酶、肠道菌素A、丙型肝炎病毒E2糖蛋白、胰岛素前体、干扰素 α A、白细胞介素-1 β 、溶菌酶素、血清白蛋白、单链抗体段(scFV)、甲

状腺素运载蛋白、酪氨酸酶、木聚糖酶、或其组合。

[0027] 本发明第二方面提供了一种分离的多核苷酸,所述的多核苷酸编码本发明第一方面所述的融合蛋白。

[0028] 在另一优选例中,所述的多核苷酸选自下组:DNA序列、RNA序列。

[0029] 在另一优选例中,所述的DNA序列选自下组:基因组序列、cDNA序列。

[0030] 在另一优选例中,所述的多核苷酸为mRNA或cDNA,并且所述多核苷酸具有式II所示结构:

A1-C1 (式II)

式中,

A1为编码上述A元件的核苷酸序列;

C1为编码上述C元件的核苷酸序列;

“-”为元件A1和元件C1之间的连接键。

[0031] 在另一优选例中,所述元件A1具有SEQ ID NO.:4所示的序列。

[0032] 在另一优选例中,所述元件C1具有SEQ ID NO.:5所示的序列。

[0033] 本发明第三方面提供了一种载体,所述载体含有本发明第二方面所述的多核苷酸。

[0034] 本发明第四方面提供了一种宿主细胞,所述宿主细胞含有本发明第三方面所述的载体或基因组中整合有本发明第二方面所述的多核苷酸。

[0035] 在另一优选例中,所述宿主细胞是通过转入本发明第三方面所述的表达载体或本发明第二方面所述的多核苷酸并经同源重组而形成的,从而在基因组或染色体中整合有本发明第一方面所述融合蛋白的编码序列。

[0036] 在另一优选例中,所述宿主细胞选自下组:克鲁维酵母、酿酒酵母、或其组合。

[0037] 在另一优选例中,所述宿主细胞选自下组:乳酸克鲁维酵母、马克斯克鲁维酵母、多布克鲁维酵母、或其组合。

[0038] 在另一优选例中,所述宿主细胞是乳酸克鲁维酵母。

[0039] 在另一优选例中,所述宿主细胞是乳酸克鲁维酵母的K.lactis Y1140菌株。

[0040] 本发明第五方面提供了一种用于表达外源蛋白的体外蛋白质合成体系,所述反应体系包括:

(i) 酵母体外蛋白质合成体系,所述的合成体系含有(a) 酵母细胞提取物;(b) 任选的聚乙二醇;(c) 任选的外源蔗糖;和(d) 任选的溶剂,所述溶剂为水或水性溶剂;以及

(ii) 本发明第一方面所述的融合蛋白。

[0041] 在另一优选例中,所述反应体系还包括:(iii) 额外添加的eIF4G蛋白。

[0042] 在另一优选例中,所述eIF4G蛋白由组成型或诱导性启动子诱导表达。

[0043] 在另一优选例中,所述组成型或诱导性启动子来源于酵母。

[0044] 在另一优选例中,所述酵母选自下组:克鲁维酵母、酿酒酵母、或其组合。

[0045] 在另一优选例中,所述组成型或诱导型启动子选自下组:pScTEF1、pScPGK1、pK1TEF1、pK1PGK1、pScADH1、pScTPI1、pScTDH3、pK1ADH1、pK1TPI1、pK1TDH3、或其组合。

[0046] 本发明第六方面提供了一种生产本发明第一方面所述融合蛋白的方法,包括:

(i) 在适合表达的条件下,培养本发明第四方面所述的宿主细胞,从而表达出本发明

第一方面所述的融合蛋白;和

(ii) 分离所述融合蛋白。

[0047] 本发明第七方面提供了一种本发明第一方面所述融合蛋白的用途,用于制备一表达外源蛋白的体外蛋白质合成体系,所述体外蛋白质合成体系用于提高外源蛋白的表达效率。

[0048] 在另一优选例中,所述反应体系还包括额外的eIF4G蛋白。

[0049] 本发明第八方面提供了一种本发明第一方面所述的融合蛋白的用途,用于制备提高体外蛋白质合成体系的体外蛋白合成能力的制剂。

[0050] 本发明第九方面提供了一种表达待表达的外源蛋白的方法,包括:

(i) 提供一酵母体外蛋白质合成体系,其中所述的合成体系中含有本发明第一方面所述的融合蛋白;和

(ii) 在适合表达蛋白的条件下,在所述外源蛋白的模板存在下,孵育所述酵母体外蛋白质合成体系,从而表达所述的外源蛋白。

[0051] 在另一优选例中,所述的融合蛋白为额外添加的。

[0052] 在另一优选例中,所述的融合蛋白与所述的酵母体外蛋白质合成体系中的其他蛋白为来自相同酵母的提取物。

[0053] 在另一优选例中,所述的方法是非诊断性和非治疗性的。

[0054] 在另一优选例中,所述步骤(ii)还包括步骤(iii):检测外源蛋白活性的表达活性Q1,并且在步骤(ii)相同条件下孵育野生型酵母菌株,检测所述外源蛋白的活性Q2,如果Q1显著高于Q2,则表明外源蛋白的表达效率显著提高。

[0055] 在另一优选例中,所述“显著高于”指 $Q1/Q2 \geq 2$,较佳地, ≥ 3 ,更佳地, ≥ 4 。

[0056] 在另一优选例中,所述酵母体外蛋白合成体系为基因改造的克鲁维酵母体外蛋白合成体系(优选乳酸克鲁维酵母体外蛋白合成体系)。

[0057] 在另一优选例中,所述外源蛋白的编码序列来自原核生物、真核生物。

[0058] 在另一优选例中,所述外源蛋白的编码序列来自动物、植物、病原体。

[0059] 在另一优选例中,所述外源蛋白的编码序列来自哺乳动物,较佳地灵长动物,啮齿动物,包括人、小鼠、大鼠。

[0060] 在另一优选例中,所述的外源蛋白的编码序列选自下组:编码荧光素蛋白、或荧光素酶(如萤火虫荧光素酶)、绿色荧光蛋白、黄色荧光蛋白、氨酰tRNA合成酶、甘油醛-3-磷酸脱氢酶、过氧化氢酶、肌动蛋白、抗体的可变区域的外源DNA、荧光素酶突变体的DNA、或其组合。

[0061] 在另一优选例中,所述外源蛋白选自下组:荧光素蛋白、或荧光素酶(如萤火虫荧光素酶)、绿色荧光蛋白、黄色荧光蛋白、氨酰tRNA合成酶、甘油醛-3-磷酸脱氢酶、过氧化氢酶、肌动蛋白、抗体的可变区域、荧光素酶突变、 α -淀粉酶、肠道菌素A、丙型肝炎病毒E2糖蛋白、胰岛素前体、干扰素 α A、白细胞介素-1 β 、溶菌酶素、血清白蛋白、单链抗体段(scFV)、甲状腺素运载蛋白、酪氨酸酶、木聚糖酶、或其组合。

[0062] 应理解,在本发明范围内中,本发明的上述各技术特征和在下文(如实施例)中具体描述的各技术特征之间都可以互相组合,从而构成新的或优选的技术方案。限于篇幅,在此不再一一累述。

附图说明

[0063]

图1显示了pKM-CAS1.0-K1eIF4G的质粒图谱。

[0064] 图2显示了pKM-pScTEF1-K1eIF4G-DD的质粒图谱。

[0065] 图3显示了pKM-pScPGK1-K1eIF4G-DD的质粒图谱。

[0066] 图4显示了pKM-pK1TEF1-K1eIF4G-DD的质粒图谱。

[0067] 图5显示了pKM-pK1PGK1-K1eIF4G-DD的质粒图谱。

[0068] 图6显示了pKM-CAS1.0-K1TDH3-1的质粒图谱。

[0069] 图7显示了pKM-CAS1.0-K1TDH3-2的质粒图谱。

[0070] 图8显示了pKM-K1TDH3-1-F-K1eIF4G-DD的质粒图谱。

[0071] 图9显示了pKM-K1TDH3-2-F-K1eIF4G-DD的质粒图谱。

[0072] 图10显示了pKM-CAS1.0-K1Pab1的质粒图谱。

[0073] 图11显示了pKM-K1Pab1-K1eIF4G-DD的质粒图谱。

[0074] 图12显示了改造菌株体外翻译活性测定示意图。利用荧光蛋白的强度指示系统的蛋白表达能力。

具体实施方式

[0075]

经过广泛而深入的研究,通过大量筛选和摸索,首次意外地发现了一种具有式Ia或式Ib结构的融合蛋白,本发明的融合蛋白可大幅度的提高体外翻译效率。此外,本发明还发现,在eIF4G前插入组成型或诱导型启动子(如pScTEF1、pScPGK1、pK1TEF1、pK1PGK1、pScADH1、pScTPI1、pScTDH3、pK1ADH1、pK1TPI1、pK1TDH3等)可显著增强体外蛋白质的合成能力。

[0076] 并且,本发明人还发现,本发明的融合蛋白在提高体外翻译效率时,本发明的融合蛋白的元件eIF4G的表达量并未增加。

[0077] 具体地,在含有本发明的融合蛋白的酵母体外蛋白质合成体系中,所合成的荧光素酶活性的相对光单位值高达 1.50×10^9 ,在eIF4G前插入组成型或诱导型启动子所合成的荧光素酶活性的相对光单位值高达 1.57×10^9 ,远远高于野生型酵母菌株(如Y1140)合成的荧光素酶的相对光单位值(4.11×10^8)。在此基础上,本发明人完成了本发明。

[0078] eIF4F元件

真核生物中,多种翻译起始因子参与蛋白质翻译起始过程(表1)。其中eIF4F负责“帽子结构”的识别以及下游翻译起始因子和核糖体的招募。eIF4F由三个蛋白质亚基组成:eIF4E、eIF4G和eIF4A。eIF4E特异性结合“帽子结构”,将eIF4F锚定在mRNA 5'端非翻译区;eIF4A是一种RNA解旋酶;eIF4G则几乎是整个翻译起始过程的支架蛋白,能与多种翻译起始因子相互作用,在下游因子招募过程中具有重要作用。

[0079] 表1 酵母中翻译起始因子

翻译起始因子	亚基	基因	蛋白长度(AA)
eIF1		<i>SUI1</i>	108

eIF1A		<i>TIF11</i>	153
eIF2	α	<i>SUI2</i>	304
	β	<i>SUI3</i>	285
	γ	<i>GCD11</i>	527
eIF2B	α	<i>GCN3</i>	305
	β	<i>GCD7</i>	381
	γ	<i>GCD1</i>	578
	δ	<i>GCD2</i>	651
	ϵ	<i>GCD6</i>	712
eIF3	a	<i>RPG1/TIF32</i>	964
	b	<i>PRT1</i>	763
	c	<i>NIP1</i>	812
	g	<i>TIF35</i>	274
	i	<i>TIF34</i>	347
	j	<i>HCR1</i>	265
eIF4A		<i>TIF1</i>	395
		<i>TIF2</i>	395
eIF4B		<i>TIF3/STM1</i>	436
eIF4E		<i>CDC33</i>	213
eIF4G		<i>TIF4631</i>	952
		<i>TIF4632</i>	914
eIF5		<i>TIF5</i>	405
eIF5B		<i>FUN12</i>	1002

在本发明中,通过在eIF4G前插入酵母来源(如酿酒酵母、克鲁维酵母等)的组成型或诱导型启动子(如pScTEF1、pScPGK1、pK1TEF1、pK1PGK1、pScADH1、pScTPI1、pScTDH3、pK1ADH1、pK1TPI1、pK1TDH3等),从而增强体外蛋白质的合成能力。

[0080] 在一优选实施方式中,所述eIF4G的核苷酸序列如SEQ ID NO.:5所示;所述eIF4G的蛋白序列如SEQ ID NO.:2所示。

[0081]

ATGGGCGAACCTACATCCGATCAGCAACCAGCTGTTGAAGCTCCAGTTGTGCAGGAGGAGACAACCAGTTCTCCGCA
 AAAAAACAGTGGATATGTCAAGAATACTGCTGGAAGCGGTGCTCCTAGAAATGGGAAATATGATGGTAACAGGAAGA
 ACTCTAGGCCTTATAACCAAAGAGGTAACAACAACAATAATAATGGTCTTCCTCGAATAAGCACTATCAAAAGTAT
 AACCAACCAGCGTACGGTGTCTGCGGGATACATCCGAACTACGGCGTATCGGCAGAGTACAACCCTCTGTACTA
 TAACCAGTACCAACAGCAGCAACAGCTGTACGCTGCTGCTTACCAGACTCCAATGAGCGGACAAGGTTATGTCCCC
 CAGTAGTGTCTCCAGCTGCTGTTTCAGCTAAACCAGCGAAGGTTGAGATTACTAACAAGTCTGGTGAACACATAGAT
 ATTGCTTCCATTGCTCATCCACATACTCATTCTCATTCTCAATCTCATTTCGCGTGCAGTTCAGTAGTGTGCGCTCC
 AGCTAACGTTACCGTCGCTGCTGCTGTATCATCCTCTGTGTCTCCATCAGCTTCTCCAGCTGTCAAAGTACAGAGCC
 CTGCTGCTAATGGTAAGGAACAATCTCCAGCTAAGCCTGAAGAACCAAAGAAGGACACTTTAATTGTGAACGATTTC
 TTGGAACAAGTTAAAAGACGCAAGGCTGCTTTAGCTGCTAAGAAGGCTGTGCAAGAGAAGGGTCTGAGGAACCGAA

GGAATCTGTCGTTGGAAGTACTGACTGATGCAAGCGTTGATACTAAGACAGGGCCTACAGCCACTGAATCTGCCAAGT
 CTGAAGAAGCTCAATCAGAATCACAAGAAAAGACTAAGGAAGAGGCTCCAGCTGAGCCAAAACCATTGACTTTGGCC
 GAAAAATTGAGACTTAAGAGGATGGAAGCTGCAAAGCAAGCTTCTGCTAAGACCGAGGAACTAAAGACTGAAGAATC
 TAAGCCTGAAGAAAACAAAGACCGAGGAGCTAAAGACTGAAGAATCTAAGCCTGAAGAAAACAAAGACCGAGGAGCTAA
 AGACTGAAGAAAACAAAGTCCGAGGAACTAAAGACTGAAGAACCCTAAGGCGGAAGAATCAAAGGCGGAAGAACCAAG
 CCTGAAGAACCAAAAGACCGAGGAAACCGACGACTGAACAACCAAGTCAGATGAACCAAGTCGGAAGAATCAAAAAC
 TGAAGAGCCAAAACCGAGGTATTAAGACTGAAGAACCAAAATCGGAAGAATCAAAGCCTGCAGAACCAAGACTG
 AAGAAAACAGCAACTGAAGAAAACAGCAACTGAAGCAAACGCCGAAGAAGGTGAACCGGCTCCTGCTGGTCCCGTTGAA
 ACTCCTGCTGATGTTGAAAACAAACCTCGAGAAGAGGCTGAAGTTGAAGACGATGGAAAGATTACCATGACCGATTT
 CCTACAGAAGTTGAAAGAGGTTTCTCCAGTTGATGATATTTATTCCTTCCAATACCCAAGTGACATTACGCCTCCAA
 ATGATAGATATAAAAAGACAAGCATTAAATATGCATACGGACCTGATTTCTTGTATCAGTTCAAAGAAAAGGTGCGAT
 GTTAAATACGATCCAGCGTGGATGGCTGAAATGACGAGTAAAATGTGCATCCCTCCTAAGAAGCCTGGTTCAAGCGG
 AAGAGGCGAAGATAGATTTAGTAAGGGTAAGGTTGGATCTCTAAGAAGTGAAGGCAGATCGGGTTCCAGGTCCAAC
 CGAAGAAGAAGTCAAAGAGGGATGATAGAAAATCTAATAGATCATACTTCCAGAAAGGACCGTGAAGATTGAGA
 GAGGAAGAAGTCAAGAGCCAAAGGTTGAGGTTGCCCATTTGGTCCAAGTGCTAATAGATGGGTTCTAAATCTAA
 GATGAAGAAAACAGAAGTCAAGTTAGCTCCAGACGGAACAGAACTTTACGACGCGGAAGAAGCATCAAGAAAGATGA
 AGTCATTGCTGAATAAATTGACATTAGAAAATGTTGAACTTATTTCTGATGATATCATGAAGATCGCTAACCAATCT
 AGATGGGAAGAAAAGGGTGAGACTTTGAAGATTGTGCATCCAACAAATTTTCAATAAGGCCTGCGATGAACCTCATTG
 GTCATCAATGTACGCGCAATTATGTGGTAAGGTCGTTAAAGACTTAGATGATAGCATTAAAGACTCAGAAAACCCAG
 ATAAGACTGGTTCTCACTTGGTTTTGCATTACTTAGTCCAAAGATGTCAAACCTGAATTCCAAACAGGATGGACTGAT
 CAACTACCTACAAACGAAGACGGTACTCCTCTACAACCTGAAATGATGTCCGATGAATACTATAAGATGGCTGCCG
 TAAGAGAAGAGGTTTGGGTTTGGTTTCGTTTCATTGGTTTCTTGTACCGTTCGAACTTATTGACTTCCAGAATGGTCT
 TCTTCTGTTTCAAGAGACTAATGAAGGATATTCAAAACCTCCTACTGAAGATACTCTAGAGTCTGTATGTGAACTT
 TTGAAAACAATTGGTGAACAGTTCGAAGGTGCTCGTATTCAAGTACTGCAGAAGCTGTGATTGAGGGTTCAAGCTT
 GCTAGACACACTATTCGACCAAAATAAAGAACGTGATCGAAAATGGTGACATCTCCAGCAGAATCAAGTTTAAAGTTGA
 TCGACATTGTGCAACTAAGAGAAAAGAGGAACTGGAATAGTAAAAATAAGAACGATGGTCCAAAGACCATTGCTCAA
 ATTCACGAAGAAGAAGCCTTGAAGAGGGCTTTGGAGGAAAGAGAAAAGAGATCGCCATGGGTCCAGAGGTGG
 TTCCAGACGATGAATAGCGAGAGAAAACCTTCTAGAAAGAGATTTCTCCTCTCATTCTCACAGTACAATCAAAATA
 GAGACGGTTTTCACTACTACCAGATCGTCATCAGTGAGATATTCTGAGCCAAAGAAGGAAGAACAAGCTCCAACCTCA
 ACTAAATCTTCTGGTGGCGCTGCCAACATGTTTGATGCATTGATGGATGCCGAAGATGATTAA (SEQ ID NO. :
 5)

MGEPTSDQQPAVEAPVVQEETTSSPQKNSGYVKNTAGSGAPRNGKYDGNRKNRPNYQRGNNNNNNGSSSNKH
 YQKYNQPAYGVSAGYIPNYGVSAEYNPLYNQYQQQQQLYAAAYQTPMSGQGYVPPVSPAASAKPAKVEITNKS
 EHIDIASIAHPHTSHSQSHSRAVPVVSPPANVTVAAVSSSVSPSASPAAVKVQSPAANGKEQSPAKPEEPKDLI
 VNDFLEQVKKRKAALAAKKAEEKGPPEPKESVVGTDASVDTKTGPTATESAKSEEAQSESQEKTKEEPAEPKP
 LTLAEKLRKMEAAKQASAKTEELKTEESKPEETKTEELKTEESKPEETKTEELKTEETKSEELKTEEPKAEESKA
 EEPKPEEPKTEEPTTEQPKSDEPKSEESKTEEPKTEVLKTEEPKSEESKPAEPKTEETATEETATEANAEEGEPAPA
 GPVETPADVETKPREEAIVEDDGKI TMTDFLQKLKEVSPVDDIYSFQYPSDI TPPNDRYKKTISKYAYGPDFLYQFK
 EKVDVKYDPAWMAEMTSKIVIPPKPGSSGRGEDRFSGKGVGSLRSEGRSGSRSNSKKKSKRDDRKSNRSYTSRKDR

ERFREEVEEPKVEVAPLVPSANRWVPKSKMKKTEVKLAPDGTELYDAEEASRKMKSLLNKLTFEMFEPISDDIMKI
 ANQSRWEEKGETLKIVIQQIFNKACDEPHWSSMYAQLCGKVVKDLDDSIKDSETPDKTGSHLVHLVQRCQTEFQT
 GWTDQLPTNEDGTPLQPEMMSDEYYKMAAAKRRGLGLVRFIGFLYRSNLLTSRMVFFCFKRLMKDIQNSPTEDTLES
 VCELLETIGEQFEGARIQVTAEAIEGSSLLDTLFDQIKNVIENGDISSRIKFKLIDIVELREKRNWNSKNKNDGPK
 TIAQIHEEEALKRALEERERERDRHGSRGGSRMNSERNSSRRDFSSHSHSNQNRDGFTTTRSSSVRYSEPKKEEQ
 APTPTKSSGGAANMFDALMDAEDD (SEQ ID NO.:2)

Pab1元件 (Pab1蛋白)

Pab1是一个71kDa的RNA结合蛋白,由4个RRM (RNA recognition motif 1-4) 结构域和1个MLLE结构域组成。每个RRM结构域中都包含2个保守的RNP结构 (RNP1/2),负责与RNA的结合。

[0082] 在一优选实施方式中,所述Pab1的核苷酸序列如SEQ ID NO.:4所示;所述Pab1的蛋白序列如SEQ ID NO.:1所示。

[0083]

ATGTCTGATATTACTGAAAAAAGCTGCTGAGCAATTGGAAAACTGCAGATCAACGATGATCAGCAACCAGCTCAATC
 TGCCAGTGCTCCATCCACTTCTGCTTCTGAAAGCGAAGCTTCTTCTGTTTCTAAGGTTGAAAACAACAACGCTTCAT
 TGTACGTTGGTGAATTGGATCCAAACATTACTGAAGCATTGTTGTACGATGTGTTTTACCATTGGGTCCAATTTCC
 TCGATCCGTGTTTGTCTGATGCCGTACCAAGGCTCGTTAGGTTACGCTTACGTTAACTATACTGATTACGAAGC
 TGGTAAGAAAGCTATTCAAGAATTGAACTATGCTGAAATCAACGGTAGACCATGTAGAATTATGTGGTCCGAACGTG
 ACCCAGCTATCAGAAAAGAAGGTTCTGGTAACATTTTCATCAAGAACTGCACCCAGCCATTGACAACAAGGCTTGT
 CATGAAACTTTCTCCACTTTCCGTGAAGTCTTGCTTGTAAAAGTTGCTTTAGATGAGAATGGAACTCTAGAGGCTT
 CGGTTTTCGTTCAATTTCAAGGAAGAATCCGATGCTAAGGATGCTATTGAAGCCGTCAACGGTATGTTGATGAACGGTT
 TGGAAGTTTACGTTGCCATGCACGTTCCAAGAAGGACCGTATCTCCAAGTTGGAAGAAGCCAAGGCTAACTTCACC
 AACATTTACGTCAAGAACATTGACGTTGAAACCACTGACGAAGAGTTCGAACAGTTGTTCTCCCAATACGGTGAAAT
 TGTCTCTGCTGCTTTGGAAAAGGATGCTGAGGGTAAGCCAAAGGTTTCGGTTTCGTTAACTTTGTTGACCACAACG
 CCGCTGCCAAGGCCGTTGAAGAGTTGAACGGTAAGGAATTCAAGTCTCAAGCTTTGTACGTTGGCAGAGCTCAAAG
 AAGTACGAACGTGCTGAAGAATTGAAGAAAACAATACGAACAATACCGTTTGGAAAAATTGGCTAAGTTCCAAGGTGT
 TAACTTGTTTCATCAAGAATTGGACGATTCCATCGATGACGAAAAATTGAAGGAAGAATTCGCCCCATACGGTACCA
 TCACCTCTGCTAGAGTCATGAGAGACCAAGAGGGTAACTCTAAGGTTTCGGTTTCGTTTCTTCTCTCCAGAA
 GAAGCTACCAAGGCTATGACCGAAAAGAACCAACAAAATTGTTGCCGTAAGCCATTGTACGTTGCCATTGCTCAAAG
 AAAGGATGTCAGAAGATCCCAATTGGCTCAACAAAATTCAAGCCAGAAACCAATCAGATTCACAACAACAGCAACAAC
 AACAAGCTGCTGCCGCTGCTGCTGGTATGCCAGGCCAATACATGCCACAAATGTTCTATGGTGTATGGCCCCAAGA
 GGTTTCCCAGGTCCAAACCCAGGTATGAACGGCCCAATGGGTGCCGTTATCCAAAGAACGGTATGGTCCCACCACC
 ACAACAATTTGCTGGTAGACCAAACGGTCCAATGTACCAAGGTATGCCACCTCAAACCAATTCCTCAAGACACCAAC
 AACAACACTACATCCAACAACAAAAGCAAAGACAAGCCTTGGGTGAACAATTGTACAAGAAGGTCAGTGCCAAGATT
 GACGACGAAAACGCCGCTGGTAAGATCACCGGTATGATCTTGATCTACCACCACAGCAAGTCATCCAATTGTTGGA
 CAACGACGAACAATTTGAACAGCAATTCCAAGAAGCCTTAGCTGCTTACGAAAACCTCAAGAAGGAACAAGAAGCTC
 AAGCTTAA (SEQ ID NO.:4)

MSDITEKTAEQLENLQINDDQQPAQSASAPSTSASESEASSVSKVENNNASLYVGELDPNI TEALLYDVFSP
 GPISSIRVCRDAVTKASLGYAYVNYTDYEAGKKAIQELNYAEINGRPCRIMWSEDPAIRKKGSGNIFIKNLHPAID

NKALHETFSTFGEVL SCKVALDENGNSRFGFVHFKEESDAKDA IEAVNGMLMNGLEVVAMHVPK KDRI SKLEEAK
ANFTNIYVKNI DVETTDEEFELFSQYGEIVSAALEKDAEGKPKGFGFVNFVDHNAAA KAVEELNGKEFKSQALYVG
RAQKKYERAEELKKQYEQYRLEKLAKFQGVNLF IKNLDDSIDDEKLKEEFAPYGTITSARVMRDQEGNSKGFVFCF
SSPEEATKAMTEKNQQIVAGKPLYVAIAQRKDVRRSQAQQIQARNQIRFQQQQQQAAAAAAGMPGQYMPQMFYGV
MAPRGFPGPNMNGPMGAGIPKNGMVP PPQQFAGRPNGPMYQGMPPQNQFPRHQQQHYIQQQKQRQALGEQLYKKV
SAKIDDENAAGKITGMILDLPPQQVIQLLDNDEQFEQQFQEALAAAYENFKKEQEAQA (SEQ ID NO.:1)

融合蛋白

如本文所用,术语“本发明融合蛋白”、“本发明PabI-eIF4G融合蛋白”以及“PabI-eIF4G融合蛋白”可互换使用,指PabI元件与eIF4G元件融合形成的融合蛋白。在本发明的融合蛋白中,PabI元件与eIF4G元件之间可以含有或不含有连接肽或柔性接头。此外,所述融合蛋白可以含有或不含有起始的Met;可以含有或不含有信号肽;以及含有或不含有标签序列(如6His等)。

[0084] 在一优选实施方式中,本发明所述的融合蛋白具有上述的式Ia或式Ib结构。优选地,本发明的融合蛋白的氨基酸序列如SEQ ID NO.:3所示。

[0085]

MSDI TEKTAEQLENLQINDDQQPAQSASAPSTSASESEASSVSKVENNANSLYVGELDPNI TEALLYDVF SPLGPIS
SIRVCRDAVTKASLGAYVNYTDYEAGKKAI QELNYAE INGRPCRIMW SERDPAIRKKGSGNIFIKNLHPAIDNKAL
HETFSTFGEVL SCKVALDENGNSRFGFVHFKEESDAKDA IEAVNGMLMNGLEVVAMHVPK KDRI SKLEEAKANFT
NIYVKNI DVETTDEEFELFSQYGEIVSAALEKDAEGKPKGFGFVNFVDHNAAA KAVEELNGKEFKSQALYVGRAQK
KYERAEELKKQYEQYRLEKLAKFQGVNLF IKNLDDSIDDEKLKEEFAPYGTITSARVMRDQEGNSKGFVFCFSSPE
EATKAMTEKNQQIVAGKPLYVAIAQRKDVRRSQAQQIQARNQIRFQQQQQQAAAAAAGMPGQYMPQMFYGVMAPR
GFPGPNMNGPMGAGIPKNGMVP PPQQFAGRPNGPMYQGMPPQNQFPRHQQQHYIQQQKQRQALGEQLYKKVSAKI
DDENAAGKITGMILDLPPQQVIQLLDNDEQFEQQFQEALAAAYENFKKEQEAQAGGGGSGGGGSTQDEVQGP HAGKST
VGGGGSGEPTSDQQPAVEAPVVQEETTSSPQKNSGYVKN TAGSGAPRNGKYDGNRKN SRPYNQRGNNNNNNGSSSNK
HYQKYNQPAYGVSAGYIPNYGVS AEYNPLYNQQYQQQQQLYAAAYQTPMSGQGYVPPVVS PAAVSAKPAKVEITNKS
GEHIDIASIAHPHTSHSQSHSRAPVVSPPANVTVA AVSSSVSPSASP AVKVQSPAANGKEQSPAKPEEPK DTL
IVNDFLEQVKRKAALAAKAVEEKGPEEPKESVVGTD TDASVDTKTGPTATESAKSEE AQSESQEKTKEEAPAEPK
PLTLAEKRLRMEAAKQASAKTEELKTEESKPEETKTEELKTEESKPEETKTEELKTEETKSEELKTEEPKAEESK
AEEPKPEEPKTEEPTTEQPKSDEPKSEESKTEEPKTEVLKTEEPKSEESKPAEPKTEETATEETATEANAEEGEPAP
AGPVETPADVETKPREAEVEDDGKITMTDFLQKLKEVSPVDDIYSFQYPSDITPPNDRYK KTSIKYAYGPDFLYQF
KEKVDVKYDPAWMAEMTSKIVIPPKKPGSSGRGEDRF SKGKVGSLRSEGRSGRSNSKKKSKRDDRKSNRSYTSRKD
RERFREEEVEEPKVEVAPLVPSANRWVPKSKMKKTEVKLAPDGT ELYDAEEASRKMKSLLNKLTLEMFEPISDDIMK
IANQSRWEEKGETLKI VIQQIFNKACDEPHWSSMYAQLCGKVVKDLDDSIKDSETPDKTGSHLVLHYLVQRCQTEFQ
TGWTDQLPTNEDGTPLQPEMMSDEYYKMAAAKRRGLGLVRF IGFLYRSNLLTSRMVFFCFKRLMKDIQNSPTEDTLE
SVCELLETIGE QFEGARIQVTAEA VIEGSSLLDTLFDQIKNVIENGDISSRIKFKLIDIVELREKRNWNSKNKNDGP
KTIAQIHEEEALKRALEERERERDRHGSRRGSRMNSERNSSRRDFSSHSHSHNQNRDGF TTRSSSVRYSEPKKEE
QAPTPTKSSGGAANMFDALMDAEDD (SEQ ID NO.:3)

在本发明中,本发明的融合蛋白可显著提高无细胞的、体外蛋白质合成体系(尤其是酵母体外蛋白质合成体系)的体外蛋白质合成能力。

[0086] 体外蛋白质合成体系

本发明融合蛋白可用于提高体外蛋白质合成体系的蛋白质合成能力。

[0087] 一种典型的体外蛋白质合成体系是酵母体外蛋白质合成体系。

[0088] 酵母 (yeast) 兼具培养简单、高效蛋白质折叠、和翻译后修饰的优势。其中酿酒酵母 (*Saccharomyces cerevisiae*) 和毕氏酵母 (*Pichia pastoris*) 是表达复杂真核蛋白质和膜蛋白的模式生物, 酵母也可作为制备体外翻译系统的原料。

[0089] 克鲁维酵母 (*Kluyveromyces*) 是一种子囊孢子酵母, 其中的马克斯克鲁维酵母 (*Kluyveromyces marxianus*) 和乳酸克鲁维酵母 (*Kluyveromyces lactis*) 是工业上广泛使用的酵母。例如乳酸克鲁维酵母是一种能够以乳酸作为其唯一的碳源和能源的酵母。与其他酵母相比, 乳酸克鲁维酵母具有许多优点, 如超强的分泌能力, 良好的大规模发酵特性、食品安全的级别及同时具有蛋白翻译后修饰的能力等, 其作为宿主系统表达药用蛋白也已显示出巨大的潜力。

[0090] 在本发明中, 酵母体外蛋白质合成体系不受特别限制, 一种优选的酵母体外蛋白质合成体系为克鲁维酵母表达系统 (更佳地, 乳酸克鲁维酵母表达系统)。

[0091] 在一优选实施方式中, 本发明的酵母体外蛋白质合成体系为基因改造后的酸克鲁维酵母表达系统。

[0092] 在一优选实施方式中, 本发明提供了一种体外的无细胞的蛋白合成体系, 所述合成体系包括:

- (a) 酵母细胞提取物;
- (b) 聚乙二醇;
- (c) 任选的外源蔗糖; 和
- (d) 任选的溶剂, 所述溶剂为水或水性溶剂。

[0093] 在另一优选例中, 所述聚乙二醇选自下组: PEG3000、PEG8000、PEG6000、PEG3350、或其组合。

[0094] 在另一优选例中, 所述聚乙二醇包括分子量 (Da) 为200-10000的聚乙二醇, 更佳地, 分子量为3000-10000的聚乙二醇。

[0095] 在另一优选例中, 所述蛋白合成体系中, 组分 (a) 的浓度 (v/v) 为20%-70%, 更佳地, 30-60%, 更佳地, 40%-50%, 以所述蛋白合成体系的总体积计。

[0096] 在另一优选例中, 所述蛋白合成体系中, 组分 (b) 的浓度 (w/v, 例如g/ml) 为0.1-8%, 更佳地, 0.5-4%, 更佳地, 1-2%。

[0097] 在另一优选例中, 所述蛋白合成体系中, 组分 (c) 的浓度为0.03-40wt%, 更佳地, 0.08-10wt%, 更佳地, 0.1-5wt%, 以所述蛋白合成体系的总重量计。

[0098] 在一特别优选的实施方式中, 本发明提供的体外蛋白合成体系包括: 酵母细胞提取物, 4-羟乙基哌嗪乙磺酸, 醋酸钾, 醋酸镁, 腺嘌呤核苷三磷酸 (ATP), 鸟嘌呤核苷三磷酸 (GTP), 胞嘧啶核苷三磷酸 (CTP), 胸腺嘧啶核苷三磷酸 (TTP), 氨基酸混合物, 磷酸肌酸, 二硫苏糖醇 (DTT), 磷酸肌酸激酶, RNA酶抑制剂, 荧光素, 荧光素酶DNA, RNA聚合酶。

[0099] 在本发明中, RNA聚合酶没有特别限制, 可以选自一种或多种RNA聚合酶, 典型的RNA聚合酶为T7 RNA聚合酶。

[0100] 在本发明中, 所述酵母细胞提取物在体外蛋白合成体系中的比例不受特别限制,

通常所述酵母细胞提取物在体外蛋白质合成蛋白合成体系中所占体系为20-70%，更佳地，30-60%，更佳地，40-50%。

[0101] 在本发明中，所述的酵母细胞提取物不含完整的细胞，典型的酵母细胞提取物包括用于蛋白翻译的核糖体、转运RNA、氨酰tRNA合成酶、蛋白质合成需要的起始因子和延伸因子以及终止释放因子。此外，酵母提取物中还含有一些源自酵母细胞的细胞质中的其他蛋白，尤其是可溶性蛋白。

[0102] 在本发明中，所述的酵母细胞提取物所含蛋白含量为20-100mg/ml， 较佳为50-100mg/ml。所述的测定蛋白含量方法为考马斯亮蓝测定方法。

[0103] 在本发明中，所述的酵母细胞提取物的制备方法不受限制，一种优选的制备方法包括以下步骤：

- (i) 提供酵母细胞；
- (ii) 对酵母细胞进行洗涤处理，获得经洗涤的酵母细胞；
- (iii) 对经洗涤的酵母细胞进行破细胞处理，从而获得酵母粗提物；
- (iv) 对所述酵母粗提物进行固液分离，获得液体部分，即为酵母细胞提取物。

[0104] 在本发明中，所述的固液分离方式不受特别限制，一种优选的方式为离心。

[0105] 在一优选实施方式中，所述离心在液态下进行。

[0106] 在本发明中，所述离心条件不受特别限制，一种优选的离心条件为5000-100000 × g，较佳地，8000-30000 × g。

[0107] 在本发明中，所述离心时间不受特别限制，一种优选的离心时间为0.5min-2h，较佳地，20min-50min。

[0108] 在本发明中，所述离心的温度不受特别限制，优选的，所述离心在1-10℃下进行，较佳地，在2-6℃下进行。

[0109] 在本发明中，所述的洗涤处理方式不受特别限制，一种优选的洗涤处理方式为采用洗涤液在pH为7-8（较佳地，7.4）下进行处理，所述洗涤液没有特别限制，典型的所述洗涤液选自下组：4-羟乙基哌嗪乙磺酸钾、醋酸钾、醋酸镁、或其组合。

[0110] 在本发明中，所述破细胞处理的方式不受特别限制，一种优选的所述的破细胞处理包括高压破碎、冻融（如液氮低温）破碎。

[0111] 所述体外蛋白质合成体系中的核苷三磷酸混合物为腺嘌呤核苷三磷酸、鸟嘌呤核苷三磷酸、胞嘧啶核苷三磷酸和尿嘧啶核苷三磷酸。在本发明中，各种单核苷酸的浓度没有特别限制，通常每种单核苷酸的浓度为0.5-5mM，较佳地为1.0-2.0mM。

[0112] 所述体外蛋白质合成体系中的氨基酸混合物可包括天然或非天然氨基酸，可包括D型或L型氨基酸。代表性的氨基酸包括（但并不限于）20种天然氨基酸：甘氨酸、丙氨酸、缬氨酸、亮氨酸、异亮氨酸、苯丙氨酸、脯氨酸、色氨酸、丝氨酸、酪氨酸、半胱氨酸、蛋氨酸、天冬酰胺、谷氨酰胺、苏氨酸、天冬氨酸、谷氨酸、赖氨酸、精氨酸和组氨酸。每种氨基酸的浓度通常为0.01-0.5mM，较佳地0.02-0.2mM，如0.05、0.06、0.07、0.08 mM。

[0113] 在优选例中，所述体外蛋白质合成体系还含有聚乙二醇或其类似物。聚乙二醇或其类似物的浓度没有特别限制，通常，聚乙二醇或其类似物的浓度（w/v）为0.1-8%，较佳地，0.5-4%，更佳地，1-2%，以所述蛋白合成体系的总重量计。代表性的PEG例子包括（但并不限于）：PEG3000，PEG8000，PEG6000和PEG3350。应理解，本发明的体系还可包括其他各种分

子量的聚乙二醇(如PEG200、400、1500、2000、4000、6000、8000、10000等)。

[0114] 在优选例中,所述体外蛋白质合成体系还含有蔗糖。蔗糖的浓度没有特别限制,通常,蔗糖的浓度为0.03-40wt%,较佳地,0.08-10wt%,更佳地,0.1-5wt%,以所述蛋白合成体系的总重量计。

[0115] 一种特别优选的体外蛋白质合成体系,除了酵母提取物之外,还含有以下组分:22 mM,pH为7.4的4-羟乙基哌嗪乙磺酸,30-150 mM醋酸钾,1.0-5.0 mM醋酸镁,1.5-4 mM核苷三磷酸混合物,0.08-0.24 mM的氨基酸混合物,25 mM磷酸肌酸,1.7 mM二硫苏糖醇,0.27 mg/mL磷酸肌酸激酶,1%-4%聚乙二醇,0.5%-2%蔗糖,8-20ng/ μ l萤火虫荧光素酶的DNA,0.027-0.054 mg/mL T7 RNA聚合酶。

[0116] 在一优选实施方式中,本发明的酵母体外蛋白质合成体系还含有:(a)本发明所述的融合蛋白,即PabI-eIF4G融合蛋白。

[0117] 在一优选实施方式中,本发明的酵母体外蛋白质合成体系还包括eIF4G蛋白;其中,本发明的eIF4G蛋白通过来源于酵母(如酿酒酵母、克鲁维酵母等)组成型或诱导型的启动子(如pScTEF1、pScPGK1、pK1TEF1、pK1PGK1、pScADH1、pScTPI1、pScTDH3、pK1ADH1、pK1TPI1、pK1TDH3等)进行诱导表达。

[0118] 在本发明中,含有本发明融合蛋白的酵母体外蛋白质合成体系可显著增强体外蛋白质的合成能力。此外,将本发明的融合蛋白和eIF4G蛋白联合使用的酵母体外蛋白质合成体系具有更高的体外蛋白质的合成能力。

[0119] 一类优选的酵母体外蛋白质合成体系描述于本发明人的在先申请CN201710125619.9中。在本文中,该专利文献通过引用方式全部纳入本文。该文献中的酵母体外蛋白质合成体系未采用本发明所述的融合蛋白。

[0120] 典型地,本发明的酵母体外蛋白质合成体系含有(a) 酵母细胞提取物;(b) 任选的聚乙二醇;(c) 任选的外源蔗糖;和(d) 任选的溶剂,所述溶剂为水或水性溶剂;以及(ii)本发明所述的融合蛋白。

[0121] 在另一优选例中,所述无细胞的蛋白合成体系还包括选自下组的一种或多种组分:

- (e1) 用于合成RNA的底物;
- (e2) 用于合成蛋白的底物;
- (e3) 镁离子;
- (e4) 钾离子;
- (e5) 缓冲剂;
- (e6) RNA聚合酶;
- (e7) 能量再生系统。

[0122] 在另一优选例中,所述蛋白合成体系中,组分(e1)的浓度为0.1-5mM,较佳地,0.5-3mM,更佳地,1-1.5mM。

[0123] 在另一优选例中,所述的酵母细胞提取物为对酵母细胞的水性提取物。

[0124] 在另一优选例中,所述酵母细胞提取物不含酵母内源性的长链核酸分子。

[0125] 在另一优选例中,所述的合成RNA的底物包括:核苷单磷酸、核苷三磷酸、或其组合。

[0126] 在另一优选例中,所述的合成蛋白的底物包括:1-20种天然氨基酸、以及非天然氨基酸。

[0127] 在另一优选例中,所述镁离子来源于镁离子源,所述镁离子源选自下组:醋酸镁、谷氨酸镁、或其组合。

[0128] 在另一优选例中,所述钾离子来源于钾离子源,所述钾离子源选自下组:醋酸钾、谷氨酸钾、或其组合。

[0129] 在另一优选例中,所述能量再生系统选自下组:磷酸肌酸/磷酸肌酸酶系统、糖酵解途径及其中间产物能量系统、或其组合。

[0130] 在另一优选例中,所述无细胞的蛋白合成体系还包括(f1) 人工合成的tRNA。

[0131] 在另一优选例中,所述缓冲剂选自下组:4-羟乙基哌嗪乙磺酸、三羟甲基氨基甲烷、或其组合。

[0132] 在另一优选例中,所述无细胞的蛋白合成体系还包括(g1) 外源的用于指导蛋白质合成的DNA分子。

[0133] 在另一优选例中,所述的DNA分子为线性的。

[0134] 在另一优选例中,所述的DNA分子为环状的。

[0135] 在另一优选例中,所述的DNA分子含有编码外源蛋白的序列。

[0136] 在另一优选例中,所述的编码外源蛋白的序列包括基因组序列、cDNA序列。

[0137] 在另一优选例中,所述的编码外源蛋白的序列还含有启动子序列、5'非翻译序列、3'非翻译序列。

[0138] 本发明的主要优点包括:

(a) 本发明首次通过基因改造技术,借助高效的细胞转化平台,对细胞内基因进行改造,从而提高了翻译系统的蛋白合成效率。

[0139] (b) 本发明首次发现一种融合蛋白,本发明的融合蛋白可显著增强体外蛋白质的合成能力。

[0140] (c) 本发明首次发现,在eIF4G前插入组成型或诱导型的启动子(如pScTEF1、pScPGK1、pK1TEF1、pK1PGK1、pScADH1、pScTPI1、pScTDH3、pK1ADH1、pK1TPI1、pK1TDH3等)可显著增强体外蛋白质的合成能力。

[0141] (d) 本发明首次通过CRISPR-Cas9基因编辑技术改造eIF4G,从而显著增强体外蛋白质的合成能力。

[0142] (e) 本发明首次发现,本发明的融合蛋白在提高体外翻译系统的效率时,本发明的融合蛋白的元件eIF4G的表达量并未增加。

[0143] 实施例1 通过基因改造提高蛋白质合成的理论模型

本发明通过CRISPR-Cas9基因编辑技术,对*K. lactis*中翻译起始因子eIF4G和Pab1进行优化,以提高无细胞体外翻译系统的效率。

[0144] 实施例2 通过CRISPR-Cas9对翻译起始因子进行改造,提高体外翻译系统效率

2.1 通过CRISPR-Cas9技术在翻译起始因子*K1eIF4G*前加入强启动子

2.1.1 *K1eIF4G*序列检索及CRISPR gRNA序列确定

eIF4G是翻译起始过程中的重要因子。目前的报导中,未见有通过基因编辑技术优化内源eIF4G表达,从而提高体外翻译活性的案例。本发明依据实施例1中的理论模型,通过

CRISPR-Cas9基因编辑技术,对翻译起始因子K1eIF4G的表达进行改造,以提高无细胞体外翻译系统的效率。

[0145] 基于*S. cerevisiae*酵母中*eIF4G*基因序列。在NCBI数据库中以*eIF4G*基因进行BLAST比对分析,确定乳酸克鲁维酵母中*eIF4G*同源基因序列,命名为*K1eIF4G*(位于染色体A的421863...424928)。在这里以此基因尾部插入一段标记DNA为例,其他目标基因或插入位置、序列均可采用类似方法操作。

[0146] ii. 在*K1eIF4G*基因起始密码子,搜索临近的PAM序列(NGG),并确定gRNA序列。gRNA选择的原则为:GC含量适中,本发明的标准为GC含量为40%-60%;避免poly T结构的存在。最终,本发明确定的优化的*K1eIF4G* gRNA序列为CGGTTTTTCAAAGCAGATAT(SEQ ID NO.: 6),位于染色体A的424927...424936位点。

[0147] 介导的在*K1eIF4G*前插入强启动子质粒构建

为了实现*K1eIF4G*的过量表达,本发明通过CRISPR-Cas技术在*K1eIF4G*基因前分别插入pScTEF1、pScPGK1、pK1TEF1和pK1PGK1启动子。质粒构建及转化方法如下:

i. CRISPR质粒构建:

使用引物PF1:CGGTTTTTCAAAGCAGATATGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCG(SEQ ID NO.: 7),PR1:GCTCTAAAACATATCTGCTTTGAAAAACCGAAAGTCCCATTCCGCCACCCG(SEQ ID NO.: 8),以pCAS质粒为模板,进行PCR扩增。将扩增产物17 μ L混合,加入1 μ L DpnI,2 μ L 10 \times digestion buffer,37 $^{\circ}$ C温浴3 h。将DpnI处理后产物10 μ L加入100 μ L DH5 α 感受态细胞中,冰上放置30 min,42 $^{\circ}$ C热激45 s后,加入1 mL LB液体培养基37 $^{\circ}$ C振荡培养1 h,涂布于Kan抗性LB固体培养,37 $^{\circ}$ C倒置培养至单克隆长出。挑取5个单克隆在LB液体培养基中振荡培养,PCR检测阳性并测序确认后,提取质粒保存,命名为pKM-CAS1.0-K1eIF4G(图1)。

[0148] 供体 DNA质粒构建及扩增

为了便于线性供体DNA的保存及扩增,本发明首先将供体DNA插入到pMD18质粒中,然后通过PCR扩增得到线性供体DNA序列。

[0149] 以乳酸克鲁维酵母基因组DNA为模板,以引物PF2:GAGCTCGGTACCCGGGGATCCTCTAGA GATAATAAAATTTCAACCTTTAAGCCATTGA ATTTTACCATTACG(SEQ ID NO.: 9)和PR2:GCCAAGCTTGCATGCCTGCAGGTCGAC GATCTTGTTAGTAATCTCAACCTTCGCTGG(SEQ ID NO.: 10)进行PCR扩增;以pMD18质粒为模板,以引物pMD18-F:ATCGTCGACCTGCAGGCATG(SEQ ID NO.: 11)和pMD18-R:ATCTCTAGAGGATCCCCGGG(SEQ ID NO.: 12)进行PCR扩增。将两次扩增产物各8.5 μ L混合,加入1 μ L DpnI,2 μ L 10 \times 消化缓冲液,37 $^{\circ}$ C温浴3 h。将DpnI处理后产物10 μ L加入100 μ L DH5 α 感受态细胞中,冰上放置30 min,42 $^{\circ}$ C热激45 s后,加入1 mL LB液体培养基37 $^{\circ}$ C振荡培养1 h,涂布于Amp抗性LB固体培养,37 $^{\circ}$ C倒置培养至单克隆长出。挑取5个单克隆在LB液体培养基中振荡培养,PCR检测阳性并测序确认后,提取质粒保存,命名为pKM-K1eIF4G-DD。

[0150] 以pKM-K1eIF4G-DD质粒为模板,以引物PF3:ATGGGCGAACCTACATCCGATC(SEQ ID NO.: 13)和PR3:ATCTGCTTTGAAAAACCGCTCTTCTCTC(SEQ ID NO.: 14)进行PCR扩增;以*S. cerevisiae*酵母基因组DNA为模板,以引物PF4:AGAGAGAAAGAGCGGTTTTTCAAAGCAGATCCACACACCATAGCTTCAAATGTTTCTAC(SEQ ID NO.: 15)和PR4:TGGTTGCTGATCGGATGTAGGTTCCGCCATCT

TAGATTAGATTGCTATGCTTTCTTTCTAATGAGC (SEQ ID NO.: 16) 进行PCR扩增 (pScTEF1 启动子扩增); 以 *S. cerevisiae* 酵母基因组DNA为模板, 以引物PF5: AGAGAGAAAGAGCGGTTTTTCAAAGCAGATAGACGCGAATTTTTCGAAGAAGTACC (SEQ ID NO.: 17) 和PR5: AGCTTCAACAGCTGGTTGCTGATCGGATGTTAGGTTTCGCCATTGTTTTATATTTGTTGTAAAAAGTAGATAATTACTTCCTTGATGATC (SEQ ID NO.: 18) 进行PCR扩增 (pScPGK1 启动子扩增); 以乳酸克鲁维酵母基因组DNA为模板, 以引物PF6: AGAGAGAAAGAGCGGTTTTTCAAAGCAGATGAGCCTGTCCAAGCAAATGCC (SEQ ID NO.: 19) 和PR6: TGGTTGCTGATCGGATGTAGGTTTCGCCATTTTTAATGTTACTTCTCTGCAGTTAGGGAAC (SEQ ID NO.: 20) 进行PCR扩增 (pK1TEF1 启动子扩增); 以乳酸克鲁维酵母基因组DNA为模板, 以引物PF7: AGAGAGAAAGAGCGGTTTTTCAAAGCAGATGTTCTCATCACTAGAAGCCGAAGT (SEQ ID NO.: 21) 和PR7: AGCTTCAACAGCTGGTTGCTGATCGGATGTAGGTTTCGCCATTTTTATTAATTCTTGATCGATTTTTTTGTTATTTCTGAAGTAACTCT (SEQ ID NO.: 22) 进行PCR扩增 (pK1PGK1 启动子扩增)。将PF3/PR3扩增产物分别与PF4/PR4、PF5/PR5、PF6/PR6和PF7/PR7扩增产物混合, 分别构建pKM-pScTEF1-K1eIF4G-DD、pKM-pScPGK1-K1eIF4G-DD、pKM-pK1TEF1-K1eIF4G-DD和pKM-pK1PGK1-K1eIF4G-DD (如图2, 3, 4, 5)。具体步骤为: 两种PCR产物各8.5 μ L混合, 加入1 μ L DpnI, 2 μ L 10 \times 消化缓冲液, 37 $^{\circ}$ C温浴3 h。将DpnI处理后产物10 μ L加入100 μ L DH5 α 感受态细胞中, 冰上放置30 min, 42 $^{\circ}$ C热激45 s后, 加入1 mL LB液体培养基37 $^{\circ}$ C振荡培养1 h, 涂布于Amp抗性LB固体培养基, 37 $^{\circ}$ C倒置培养至单克隆长出。挑取5个单克隆在LB液体培养基中振荡培养, PCR检测阳性并测序确认后, 提取质粒保存。

[0151] 乳酸克鲁维酵母转化及阳性鉴定

i. 将乳酸克鲁维酵母菌液在YPD固体培养基上划线并挑取单克隆, 于25 mL 2 \times YPD液体培养基中振荡培养过夜, 取2 mL菌液于50 mL液体2 \times YPD培养基中继续振荡培养2-8 h。20 $^{\circ}$ C条件下3000 g离心5 min收集酵母细胞, 加入500 μ L无菌水重悬, 同样条件下离心收集细胞。配制感受态细胞溶液 (5% v/v甘油, 10% v/v DMSO) 并将酵母细胞溶解于500 μ L该溶液中。分装50 μ L至1.5 mL离心管中, -80 $^{\circ}$ C保存。

[0152] 将感受态细胞置于37 $^{\circ}$ C融化15-30 s, 13000 g离心2 min并去除上清。配制转化缓冲液: PEG 3350 (50% (w/v)) 260 μ L, LiAc (1.0 M) 36 μ L, carrier DNA (5.0 mg/mL) 20 μ L, Cas9/gRNA质粒15 μ L, 供体DNA 10 μ L, 加入无菌水至最终体积360 μ L。热激后, 13000 g离心30 s去除上清。加入1 mL YPD液体培养基, 培养2-3 h, 吸取200 μ L涂布于固体YPD (200 μ g/mL G418) 培养基, 培养2-3天至单菌落出现。

[0153] 在乳酸克鲁维酵母转化后的平板上挑取10-20个单克隆, 置于1 mL YPD (200 μ g/mL G418) 液体培养基中振荡培养过夜, 以菌液为模板, 以CRISPR Insertion Check引物对, 对相应样品进行PCR检测。PCR结果阳性并经测序鉴定的菌株, 确定为阳性菌株。

[0154] 2.2 通过CRISPR-Cas9技术将*K1eIF4G*与高表达基因融合

2.2.1 *K1TDH3*序列检索及CRISPR gRNA序列确定

在酿酒酵母 (*S. cerevisiae*) 中, TDH3以四聚体形式存在, 参与糖酵解途径中的催化反应。其启动子pTDH3是基因工程中广泛使用的一种持续型强启动子。为了实现乳酸克鲁维酵母中*K1eIF4G*的足量表达, 并在执行翻译起始功能时形成局部高浓度, 本发明将*K1eIF4G*基因连接到乳酸克鲁维酵母*TDH3*基因ORF 3'端。

[0155] i. 基于*S. cerevisiae*酵母中*TDH3*基因序列。在NCBI数据库中以*TDH3*基因进行

BLAST比对分析,确定乳酸克鲁维酵母中TDH3同源基因序列。经比对发现,在乳酸克鲁维酵母基因组中存在两个TDH3同源基因,在本发明中分别命名为*K1TDH3-1*(位于染色体A的1024297...1025292)和*K1TDH3-2*(位于染色体F的1960417...1961406)。在这里以此基因尾部插入一段标记DNA为例,其他目标基因或插入位置、序列均可采用类似方法操作。

[0156] ii.在*K1TDH3*基因终止密码子附近搜索PAM序列(NGG),并确定*K1TDH3*gRNA序列(*K1TDH3-1*:CTTGTTGCTAAGAAGTAAAG(SEQ ID NO.: 23),位于染色体A的1024272...1024291位点,*K1TDH3-2*:CTCTGAAAGAGTTGTCGATT(SEQ ID NO.: 24)位于染色体F的1960378...1960397位点)。

[0157] 介导的将*K1eIF4G*整合到*K1TDH3*位点质粒构建
CRISPR质粒构建

i.对*K1TDH3-1*,使用引物PF8:CTTGTTGCTAAGAAGTAAAGGTTTTAGAGCTAGAAATAGCAAGTTAAAAT(SEQ ID NO.: 25),PR8:GCTCTAAAACCTTTAGTTCTTAGCAACAAGAAAGTCCCATTCGCCACCCG(SEQ ID NO.: 26),以pCAS质粒为模板,进行PCR扩增。将扩增产物17 μL混合,加入1 μL DpnI,2 μL 10 × 消化缓冲液,37 °C温浴3 h。将DpnI处理后产物10 μL加入100 μL DH5α感受态细胞中,冰上放置30 min,42 °C热激45 s后,加入1 mL LB液体培养基37 °C振荡培养1 h,涂布于Kan抗性LB固体培养,37 °C倒置培养至单克隆长出。挑取5个单克隆在LB液体培养基中振荡培养,PCR检测阳性并测序确认后,提取质粒保存,命名为pKM-CAS1.0-K1TDH3-1(如图6)。

[0158] ii.对*K1TDH3-2*,使用引物PF9:CTCTGAAAGAGTTGTCGATTGTTTTAGAGCTAGAAATAGCAAGTTAAAAT(SEQ ID NO.: 27),PR9:GCTCTAAAACAATCGACAACCTTTTCAGAGAAAGTCCCATTCGCCACCCG(SEQ ID NO.: 28),以pCAS质粒为模板,进行PCR扩增。将扩增产物17 μL混合,加入1 μL DpnI,2 μL 10 × digestion buffer,37 °C温浴3 h。将DpnI处理后产物10 μL加入100 μL DH5α感受态细胞中,冰上放置30 min,42 °C热激45 s后,加入1 mL LB液体培养基37 °C振荡培养1 h,涂布于Kan抗性LB固体培养,37 °C倒置培养至单克隆长出。挑取5个单克隆在LB液体培养基中振荡培养,PCR检测阳性并测序确认后,提取质粒保存,命名为pKM-CAS1.0-K1TDH3-2(如图7)。

[0159] 供体 DNA质粒构建及扩增

为了便于线性供体DNA的保存及扩增,本发明首先将供体DNA插入到pMD18质粒中,然后通过PCR扩增得到线性供体DNA序列。

[0160] i.对*K1TDH3-1*,以乳酸克鲁维酵母基因组DNA为模板,以引物PF10:GAGCTCGGTACCCGGGATCCTCTAGAGATCATCCACTCCATCACCGCTACCCAA(SEQ ID NO.: 29)和PR10:GCCAAGCTTGCATGCCTGCAGGTCGACGATCAACGTCCCCATCTACAAGAGC(SEQ ID NO.: 30)进行PCR扩增;以pMD18质粒为模板,以引物pMD18-F:ATCGTCGACCTGCAGGCATG(SEQ ID NO.: 31)和pMD18-R:ATCTCTAGAGGATCCCGGG(SEQ ID NO.: 32)进行PCR扩增。将两次扩增产物各8.5 μL混合,加入1 μL DpnI,2 μL 10 × 消化缓冲液,37 °C温浴3 h。将DpnI处理后产物10 μL加入100 μL DH5α感受态细胞中,冰上放置30 min,42 °C热激45 s后,加入1 mL LB液体培养基37 °C振荡培养1 h,涂布于Amp抗性LB固体培养,37 °C倒置培养至单克隆长出。挑取5个单克隆在LB液体培养基中振荡培养,PCR检测阳性并测序确认后,提取质粒保存,命名为pKM-K1TDH3-1-DD。

[0161] 以pKM-K1TDH3-1-DD为模板,以引物PF11:GATGCATTGATGGATGCCGAAGATGATTAAGA

GGTTGATGTAATTGATATTTTCTGATAAAATTACTATTG (SEQ ID NO.: 33) 和PR11: AGCTGGTTGCTGATCGGATGTAGGTTCCGCCAGATCCACCTCCTTCCACGTTTGTGGTCTTGATCCACCTCCACCGTTCTTAGCAACAAGTTCGACCAAATCG (SEQ ID NO.: 34) 进行扩增;以*K. lactis*基因组DNA为模板,以引物PF12: GGCGAACCTACATCCGATCAGC (SEQ ID NO.: 35) 和PR12: TTAATCATCTTCGGCATCCATCAATGC (SEQ ID NO.: 36) 进行扩增。将两次扩增产物各8.5 μ L, 1 μ L DpnI, 2 μ L 10 \times digestion buffer混合, 37 $^{\circ}$ C温浴3 h。将DpnI处理后产物10 μ L加入100 μ L DH5 α 感受态细胞中,冰上放置30 min, 42 $^{\circ}$ C热激45 s后,加入1 mL LB液体培养基37 $^{\circ}$ C振荡培养1 h,涂布于Amp抗性LB固体培养, 37 $^{\circ}$ C倒置培养至单克隆长出。挑取5个单克隆在LB液体培养基中振荡培养, PCR检测阳性并测序确认后,提取质粒保存,命名为pKM-K1TDH3-1-F-K1eIF4G-DD (图8)。

[0162] 以pKM-K1TDH3-1-F-K1eIF4G-DD质粒为模板,以引物M13-F: GTAAAACGACGGCCAGT (SEQ ID NO.: 37) 和M13-R: CAGGAAACAGCTATGAC (SEQ ID NO.: 38) 进行扩增,得到线性供体DNA。

[0163] ii. 对*K1TDH3-2*,以乳酸克鲁维酵母基因组DNA为模板,以引物PF13: GAGCTCGGTACCCGGGATCCTCTAGAGATGAAGCTTTGATGACTACCGTTC (SEQ ID NO.: 39) 和PR13: GCCAAGCTTGCATGCCTGCAGGTCGACGATGTCTATTGTATCGGAAGAAGTGTCA (SEQ ID NO.: 40) 进行PCR扩增;以pMD18质粒为模板,以引物pMD18-F: ATCGTCGACCTGCAGGCATG (SEQ ID NO.: 41) 和pMD18-R: ATCTCTAGAGGATCCCCGGG (SEQ ID NO.: 42) 进行PCR扩增。将两次扩增产物各8.5 μ L混合,加入1 μ L DpnI, 2 μ L 10 \times 消化缓冲液, 37 $^{\circ}$ C温浴3 h。将DpnI处理后产物10 μ L加入100 μ L DH5 α 感受态细胞中,冰上放置30 min, 42 $^{\circ}$ C热激45 s后,加入1 mL LB液体培养基37 $^{\circ}$ C振荡培养1 h,涂布于Amp抗性LB固体培养, 37 $^{\circ}$ C倒置培养至单克隆长出。挑取5个单克隆在LB液体培养基中振荡培养, PCR检测阳性并测序确认后,提取质粒保存,命名为pKM-K1TDH3-2-DD。

[0164] 以pKM-K1TDH3-2-DD为模板,以引物PF14: GATGCATTGATGGATGCCGAAGATGATTAAATTACTCTTTTAAGTTAACGAACGCTTTTGTGATGAG (SEQ ID NO.: 43) 和PR14: AGCTGGTTGCTGATCGGATGTAGGTTCCGCCAGATCCACCTCCTTCCACGTTTGTGGTCTTGATCCACCTCCACCAGCAACGTGCTCAACtAAgTCaACgACcCTTTCAGAGTAACCGTATTCGTTATCG (SEQ ID NO.: 44) 进行扩增;以乳酸克鲁维酵母DNA为模板,以引物PF15: GGCGAACCTACATCCGATCAGC (SEQ ID NO.: 45) 和PR15: TTAATCATCTTCGGCATCCATCAATGC (SEQ ID NO.: 46) 进行扩增。将两次扩增产物各8.5 μ L, 1 μ L DpnI, 2 μ L 10 \times 消化缓冲液混合, 37 $^{\circ}$ C温浴3 h。将DpnI处理后产物10 μ L加入100 μ L DH5 α 感受态细胞中,冰上放置30 min, 42 $^{\circ}$ C热激45 s后,加入1 mL LB液体培养基37 $^{\circ}$ C振荡培养1 h,涂布于Amp抗性LB固体培养, 37 $^{\circ}$ C倒置培养至单克隆长出。挑取5个单克隆在LB液体培养基中振荡培养, PCR检测阳性并测序确认后,提取质粒保存,命名为pKM-K1TDH3-2-F-K1eIF4G-DD (图9)。

[0165] 以pKM-K1TDH3-2-F-K1eIF4G-DD质粒为模板,以引物M13-F: GTAAAACGACGGCCAGT (SEQ ID NO.: 47) 和M13-R: CAGGAAACAGCTATGAC (SEQ ID NO.: 48) 进行扩增,得到线性供体DNA。

[0166] 乳酸克鲁维酵母转化及阳性鉴定

i. 将乳酸克鲁维酵母菌液在YPD固体培养基上划线并挑取单克隆,于25 mL 2 \times YPD液体培养基中振荡培养过夜,取2 mL菌液于50 mL液体2 \times YPD培养基中继续振荡培养2-8

h. 20 °C条件下3000 g离心5 min收集酵母细胞,加入500 μL无菌水重悬,同样条件下离心收集细胞。配制感受态细胞溶液(5% v/v甘油,10% v/v DMSO)并将酵母细胞溶解于500 μL该溶液中。分装50 μL至1.5 mL离心管中,-80 °C保存。

[0167] 将感受态细胞置于37 °C融化15-30 s,13000 g离心2 min并去除上清。配制转化缓冲液:PEG 3350 (50% (w/v)) 260 μL,LiAc (1.0 M) 36 μL,carrier DNA (5.0 mg/mL) 20 μL,Cas9/gRNA质粒15 μL,供体DNA 10 μL,加入无菌水至最终体积360 μL。热激后,13000 g离心30 s去除上清。加入1 mL YPD液体培养基,培养2-3 h,吸取200 μL涂布于固体YPD(200 μg/mL G418)培养基,培养2-3天至单菌落出现。

[0168] ii. 在乳酸克鲁维酵母转化后的平板上挑取10-20个单克隆,置于1 mL YPD(200 μg/mL G418)液体培养基中振荡培养过夜,以菌液为模板,以CRISPR Insertion Check引物K1TDH3-1-CICF1(K1TDH3-1序列内引物):CTTCTACTGCTCCAATGTTTCGTCGTT(SEQ ID NO.: 49)和引物K1TDH3-2-CICF1(K1TDH3-2序列内引物):TTAACGAAGACAAGTACAACGGTGA(SEQ ID NO.: 50)分别与进行PCR扩增,分别与K1eIF4G-CICR2(K1eIF4G序列内引物):TTCTCTTCGACAGCCTTCTTAGCAG(SEQ ID NO.:51)配对进行PCR,对K1TDH3-1和K1TDH3-2位点的K1eIF4G插入进行检测,PCR结果阳性并经测序鉴定的菌株确定为阳性菌株。

[0169] 通过CRISPR-Cas9技术将K1eIF4G与其互作蛋白融合

2.3.1 K1Pab1序列检索及CRISPR gRNA序列确定

如前文所述,Pab1蛋白与eIF4G蛋白在翻译起始过程中存在相互作用。本发明通过CRISPR-Cas9基因编辑技术,将K1Pab1和K1eIF4G融合,促进两者的相互作用,以提高体外翻译效率。

[0170] 基于Pab1序列,得到乳酸克鲁维酵母中K1Pab1基因序列(位于染色体C的1553322... 1555100)。在K1Pab1基因终止密码子附近搜索PAM序列(NGG),并确定gRNA序列。gRNA选择的的原则为:GC含量适中,本发明的标准为GC含量为40%-60%;避免poly T结构的存在。最终,本发明确定的K1Pab1 gRNA序列为TGCTTACGAAAACCTCAAGA(SEQ ID NO.: 52),位于染色体C的1555058...1555077位点。

[0171] 介导的将K1eIF4G整合到K1Pab1位点质粒构建

i. CRISPR质粒构建

使用引物PF16: TGCTTACGAAAACCTCAAGAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTC CG(SEQ ID NO.: 53),PR16:GCTCTAAAACCTTGAAGTTTTTCGTAAGCAAAAAGTCCCATTCCGCCACCCG(SEQ ID NO.: 54),以pCAS质粒为模板,进行PCR扩增。将扩增产物17 μL混合,加入1 μL DpnI,2 μL 10× digestion buffer,37 °C温浴3 h。将DpnI处理后产物10 μL加入100 μL DH5α感受态细胞中,冰上放置30 min,42 °C热激45 s后,加入1 mL LB液体培养基37 °C振荡培养1 h,涂布于Kan抗性LB固体培养,37 °C倒置培养至单克隆长出。挑取5个单克隆在LB液体培养基中振荡培养,PCR检测阳性并测序确认后,提取质粒保存,命名为pKM-CAS1.0-K1Pab1(图10)。

[0172] 供体 DNA质粒构建及扩增

为了便于线性供体DNA的保存及扩增,首先将供体DNA插入到pMD18质粒中,然后通过PCR扩增得到线性供体DNA序列。

[0173] 以乳酸克鲁维酵母基因组DNA为模板,以引物PF17:GAGCTCGGTACCCGGGGATCCTCTAG

AGATCCGGTAAGCCATTGTACGTTGCCAT (SEQ ID NO.: 55) 和PR17:GCCAAGCTTGCATGCCTGCAGGTCG ACGATCAGTATACCGTCCATGTTGATGACT (SEQ ID NO.: 56) 进行PCR扩增;以pMD18质粒为模板,以引物pMD18-F:ATCGTCGACCTGCAGGCATG (SEQ ID NO.: 57) 和pMD18-R:ATCTCTAGAGGATCCCCGGG (SEQ ID NO.: 58) 进行PCR扩增。将两次扩增产物各8.5 μ L混合,加入1 μ L DpnI,2 μ L 10 \times 消化缓冲液,37 $^{\circ}$ C温浴3 h。将DpnI处理后产物10 μ L加入100 μ L DH5 α 感受态细胞中,冰上放置30 min,42 $^{\circ}$ C热激45 s后,加入1 mL LB液体培养基37 $^{\circ}$ C振荡培养1 h,涂布于Amp抗性LB固体培养,37 $^{\circ}$ C倒置培养至单克隆长出。挑取5个单克隆在LB液体培养基中振荡培养,PCR检测阳性并测序确认后,提取质粒保存,命名为pKM-K1Pab1-DD。

[0174] 以pKM-K1Pab1-DD为模板,以引物PF18: GATGCATTGATGGATGCCGAAGATGATTAACCTG ATTTTTGGACCTTGATCTTCATCTTGTC (SEQ ID NO.: 59) 和PR18: CTTGAACTTCATCTTGAGTTGAACC TCCACCTCCAGATCCACCTCCACCAGCTTGAGCTTCTGTTCtTTtTTaAAaTTcTCGTAAGCAGCTAAGGCTTC (SEQ ID NO.: 60) 进行扩增;以乳酸克鲁维酵母DNA为模板,以引物PF19: GTGGAGGTTCAACT CAAGATGAAGTTCAAGGTCCACATGCTGGTAAGTCTACTGTTGGTGGAGGTGGATCTGGCGAACCTACATCCGATCA GC (SEQ ID NO.: 61) 和PR19: TTAATCATCTTCGGCATCCATCAATGC (SEQ ID NO.: 62) 进行扩增。将两次扩增产物各8.5 μ L,1 μ L DpnI,2 μ L 10 \times digestion buffer混合,37 $^{\circ}$ C温浴3 h。将DpnI处理后产物10 μ L加入100 μ L DH5 α 感受态细胞中,冰上放置30 min,42 $^{\circ}$ C热激45 s后,加入1 mL LB液体培养基37 $^{\circ}$ C振荡培养1 h,涂布于Amp抗性LB固体培养,37 $^{\circ}$ C倒置培养至单克隆长出。挑取5个单克隆在LB液体培养基中振荡培养,PCR检测阳性并测序确认后,提取质粒保存,命名为pKM-K1Pab1-K1eIF4G-DD (图11)。

[0175] 以pKM-K1Pab1-K1eIF4G-DD质粒为模板,以引物M13-F: GTAAAACGACGGCCAGT (SEQ ID NO.: 63) 和M13-R: CAGGAAACAGCTATGAC (SEQ ID NO.: 64) 进行扩增,得到线性供体DNA。

[0176] 乳酸克鲁维酵母转化及阳性鉴定

i. 将乳酸克鲁维酵母菌液在YPD固体培养基上划线并挑取单克隆,于25 mL 2 \times YPD液体培养基中振荡培养过夜,取2 mL菌液于50 mL液体2 \times YPD培养基中继续振荡培养2-8 h。20 $^{\circ}$ C条件下3000 g离心5 min收集酵母细胞,加入500 μ L无菌水重悬,同样条件下离心收集细胞。配制感受态细胞溶液(5% v/v甘油,10% v/v DMSO)并将酵母细胞溶解于500 μ L该溶液中。分装50 μ L至1.5 mL离心管中,-80 $^{\circ}$ C保存。

[0177] 将感受态细胞置于37 $^{\circ}$ C融化15-30 s,13000 g离心2 min并去除上清。配制转化缓冲液:PEG3350 (50% (w/v)) 260 μ L,LiAc (1.0 M) 36 μ L,carrier DNA (5.0 mg / mL) 20 μ L,Cas9/gRNA质粒15 μ L,供体DNA 10 μ L,加入无菌水至最终体积360 μ L。热激后,13000 g离心30 s去除上清。加入1 mL YPD液体培养基,培养2-3 h,吸取200 μ L涂布于固体YPD (200 μ g/mL G418) 培养基,培养2-3天至单菌落出现。

[0178] 在乳酸克鲁维酵母转化后的平板上挑取10-20个单克隆,置于1 mL YPD (200 μ g/mL G418) 液体培养基中振荡培养过夜,以菌液为模板,以引物K1PAB1-CICF1 (K1PAB1序列内引物):TCTCCAGAAGAAGCTACCAAGGCTA (SEQ ID NO.: 65) 和引物K1eIF4G-CICR2 (K1eIF4G序列内引物):TTCTCTCGACAGCCTTCTTAGCAG (SEQ ID NO.: 66) 进行PCR扩增,对K1PAB1位点K1eIF4G插入进行PCR检测,PCR结果阳性并经测序鉴定的菌株,确定为阳性菌株。

[0179] 实施例3 改造菌株体外翻译活性测定

将基因改造后的乳酸克鲁维酵母菌株制备成体外蛋白质合成体系,并加入萤火虫荧光素酶(Firefly Luciferase, Fluc)基因DNA模板以测定改造菌株的蛋白翻译能力。将上述反应体系置于25-30 °C的环境中,静置孵育约2-6 h。反应结束后,在96孔白板或者384孔白板中加入等体积的Fluc底物荧光素(luciferin),立即放置于Envision 2120多功能酶标仪(Perkin Elmer),读数,检测Fluc活性,相对光单位值(Relative Light Unit, RLU)作为活性单位。

[0180] 在改造的结构当中,*KleIF4G*前插入启动子pK1PGK1的结构

pK1PGK1::KleIF4G,及*KleIF4G*连接到*K1Pab1* C端的结构

K1Pab1-KleIF4G,都表现出比野生型酵母菌株Y1140更强的体外蛋白质合成能力。其编码合成的Fluc蛋白放出的相对光单位值分别达到 1.57×10^9 和 1.50×10^9 ,而野生型酵母菌株Y1140合成的Fluc蛋白的相对光单位值仅有 4.11×10^8 。这表明对*KleIF4G*的改造能够有效增强酵母体外蛋白质合成体系合成蛋白质的效率(图12)。

[0181] 具体的效果如表1所示。

[0182] 表1

NO.	Data1	Data2	Data3	活性	稀释倍数	最终活性(RLU)
pK1PGK1_KleIF4G	33216610	28584890	32598650	31466717	50	1.57×10^9
pK1TEF1_KleIF4G	6685609	10189150	8594529	8489763	50	4.24×10^8
pScPGK1_KleIF4G	4130719	8605461	4555399	5763860	50	2.88×10^8
pScTEF1_KleIF4G	8230202	6415045	7578242	7407830	50	3.70×10^8
K1TDH3_1_KleIF4G	788821	751243	941381	827148.3	50	4.13×10^7
K1TDH3_2_KleIF4G	8676941	6496592	7904461	7692665	50	3.85×10^8
K1PAB1_KleIF4G	22155330	33507550	34075530	29912803	50	1.50×10^9
Y1140	10925600	6729764	6997436	8217600	50	4.11×10^8
NC	707	965				

NC:表示negative control,阴性对照。

[0183] 上述实验结果表明:通过对乳酸克鲁维酵母*KleIF4G*基因的相关改造,本发明的融合蛋白能够显著增强酵母体外蛋白质合成体系产生蛋白质的效率。

[0184] 在本发明提及的所有文献都在本申请中引用作为参考,就如同每一篇文献被单独引用作为参考那样。此外应理解,在阅读了本发明的上述讲授内容之后,本领域技术人员可以对本发明作各种改动或修改,这些等价形式同样落于本申请所附权利要求书所限定的范围。

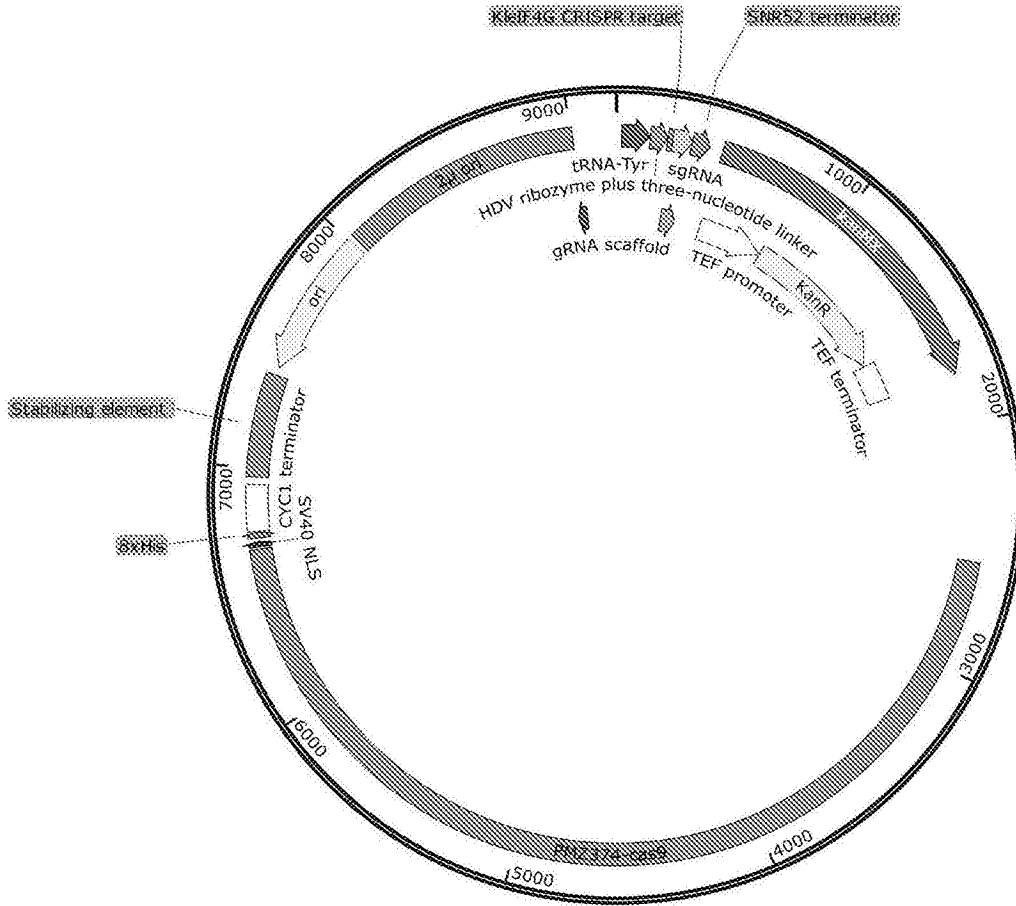


图1

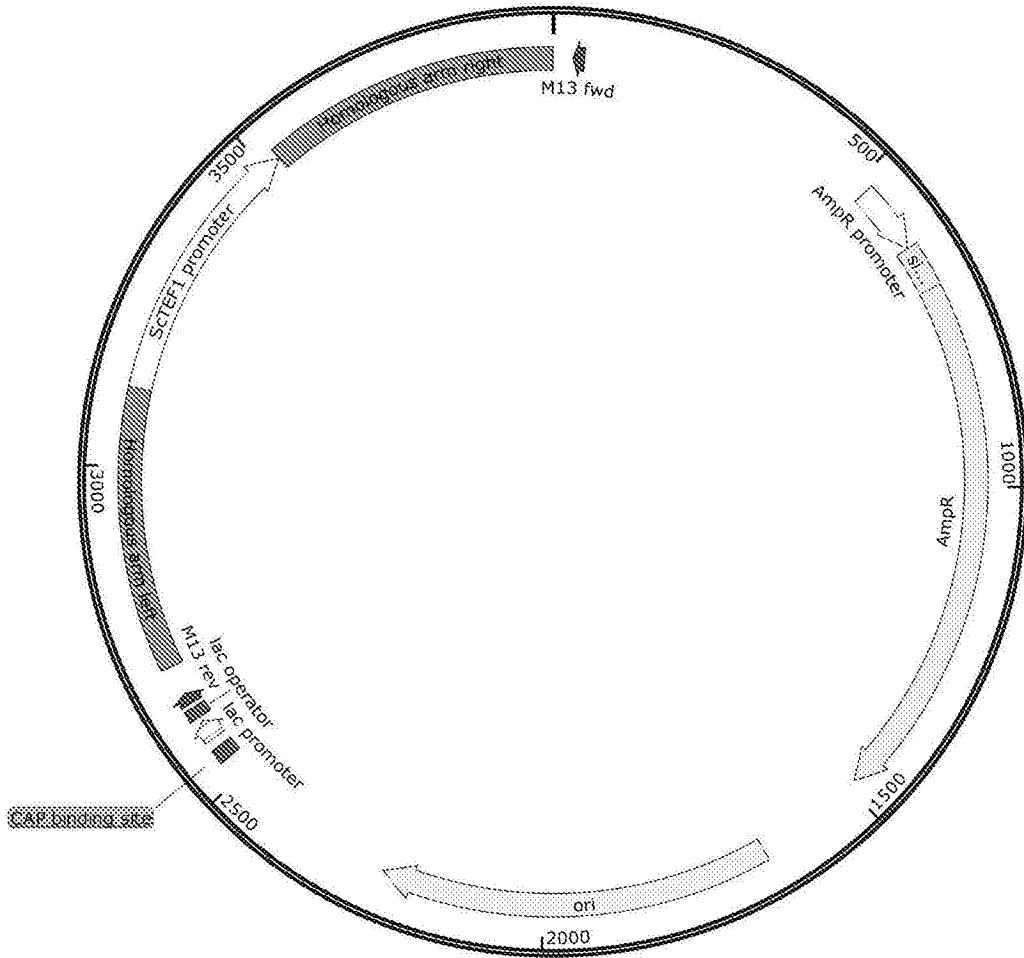


图2

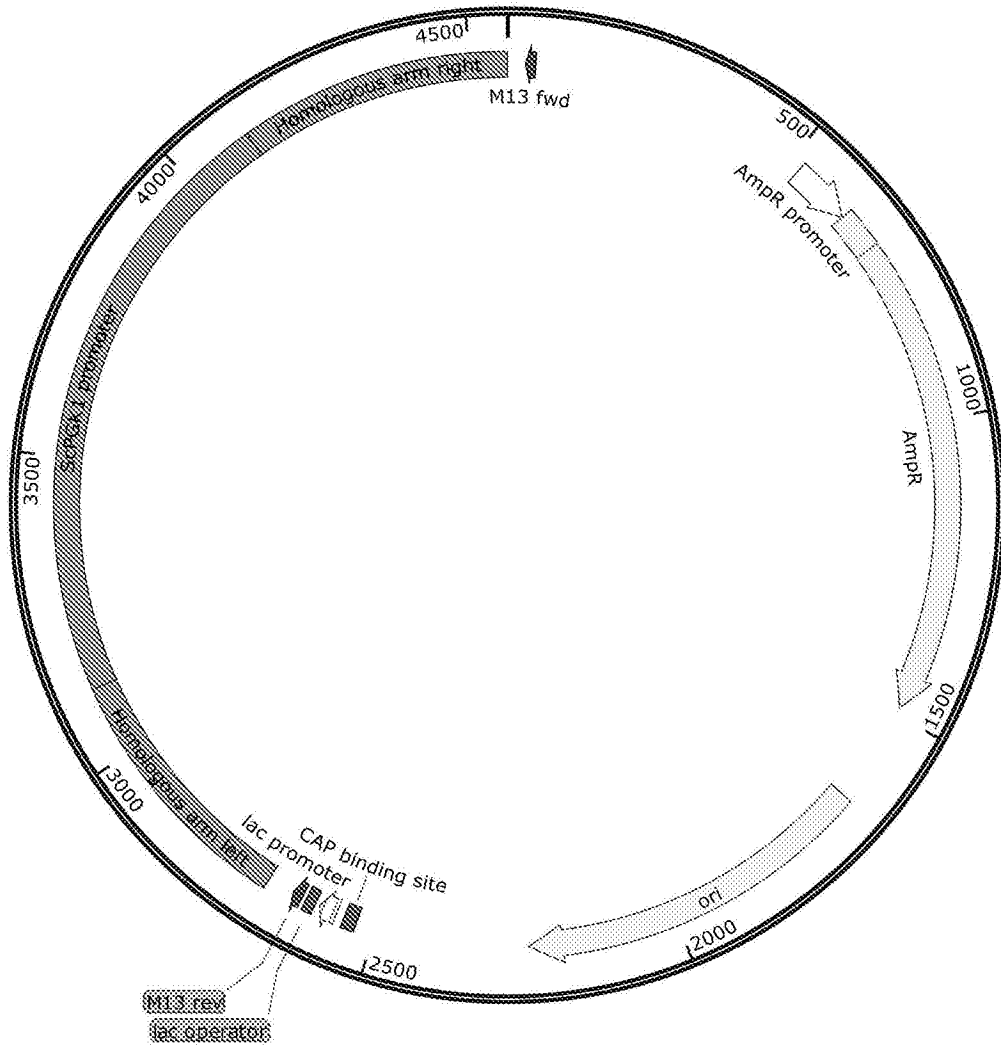


图3

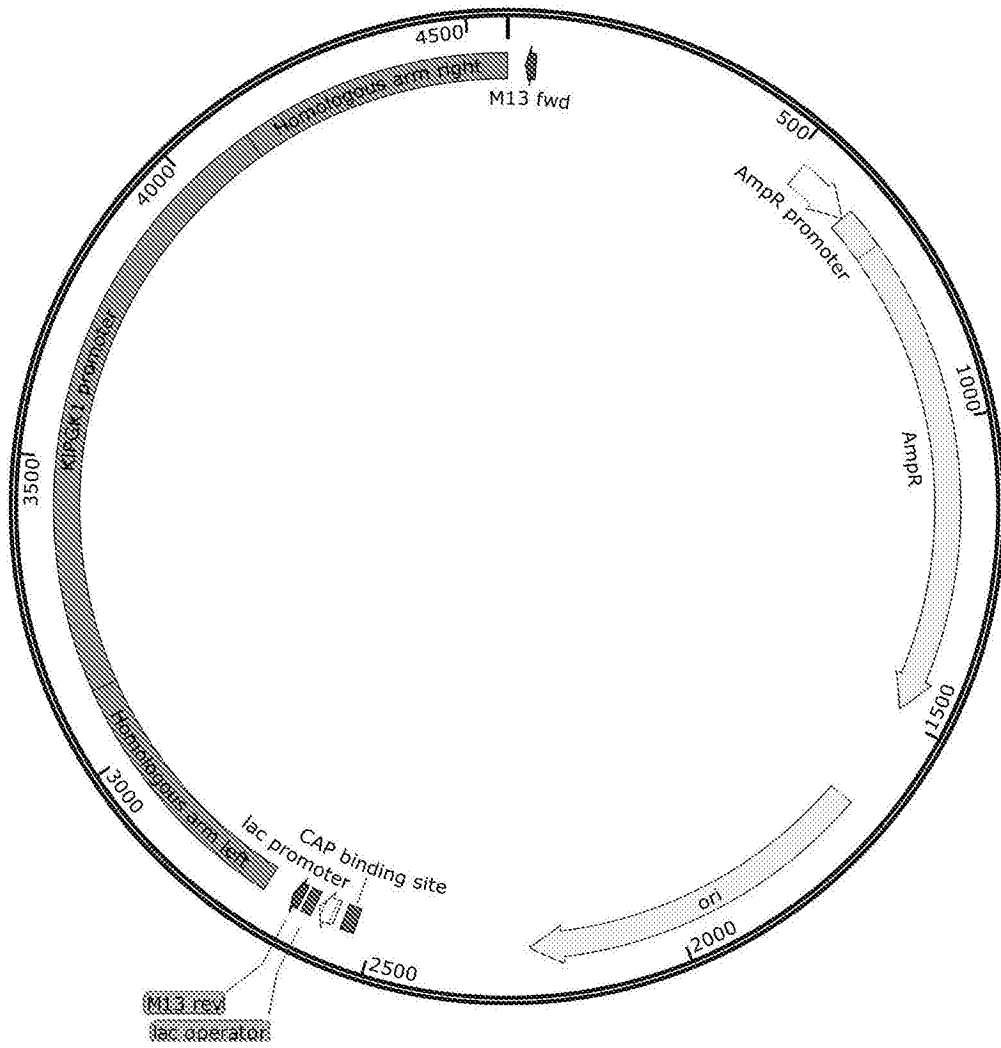


图4

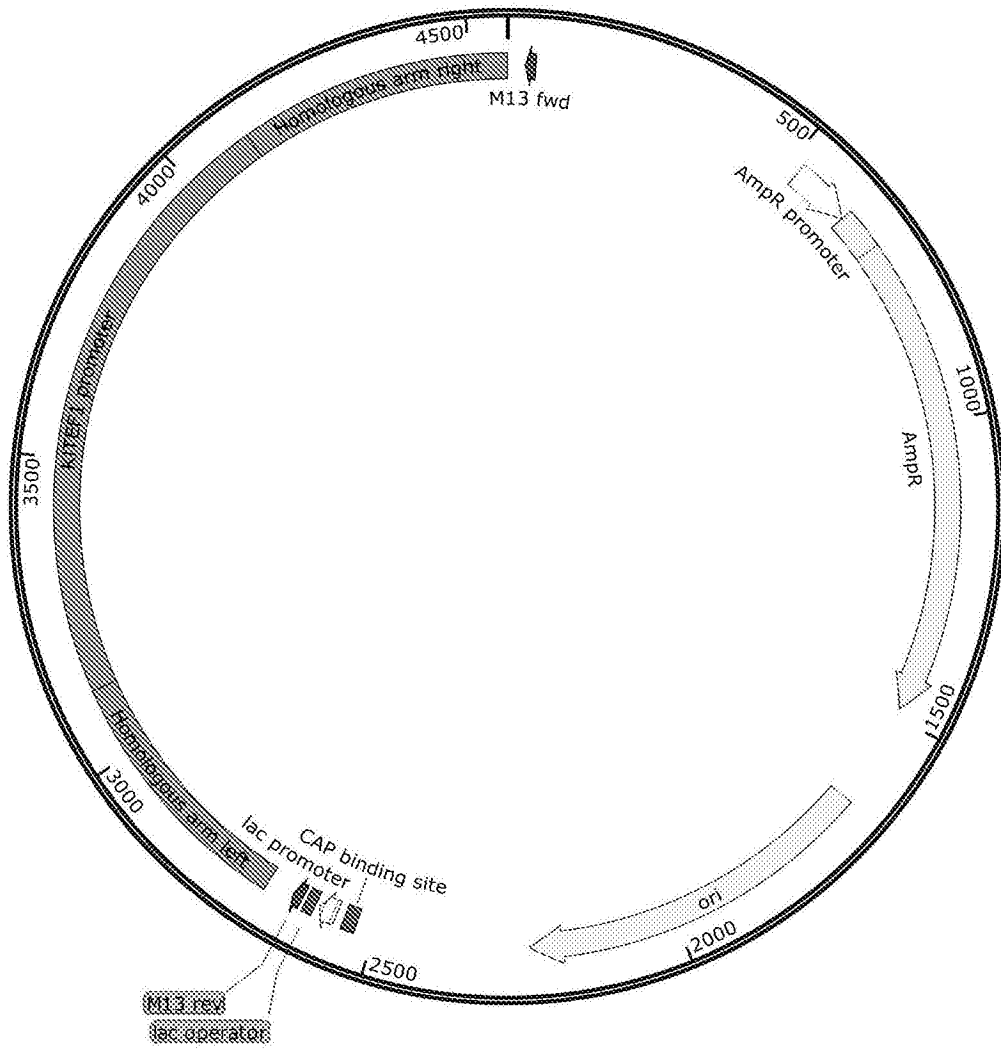


图5

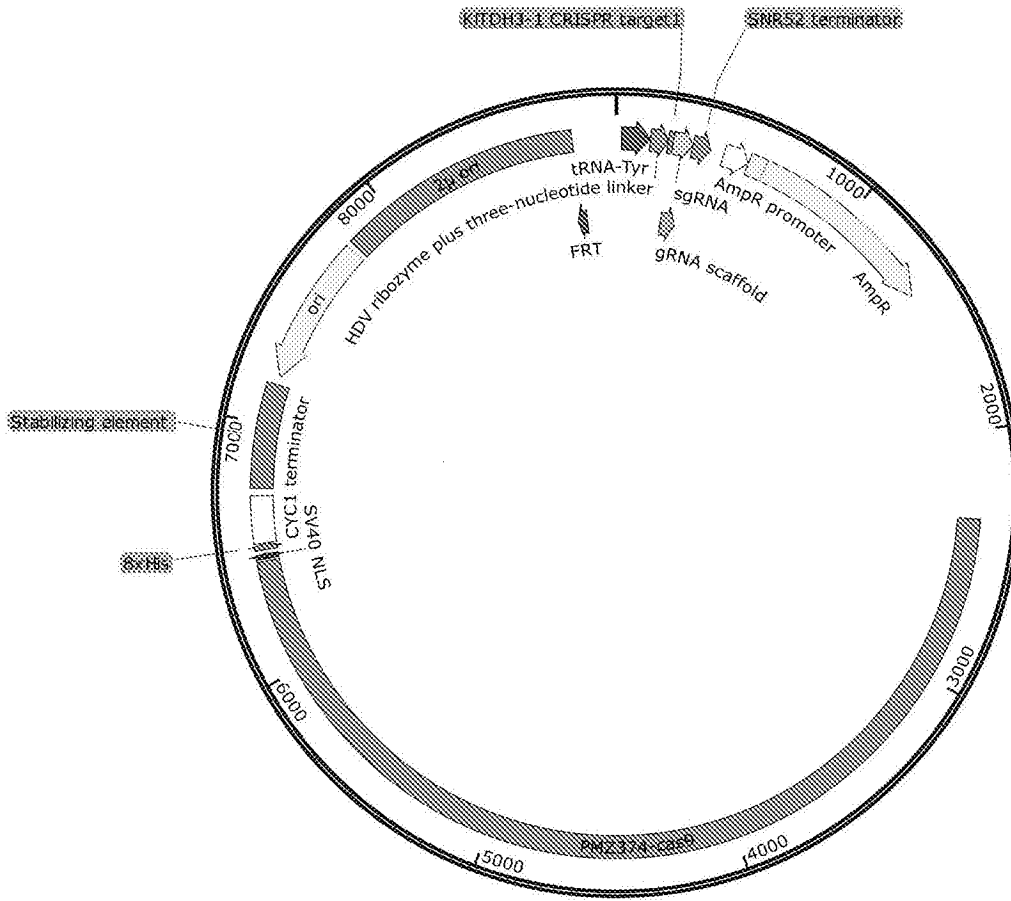


图6

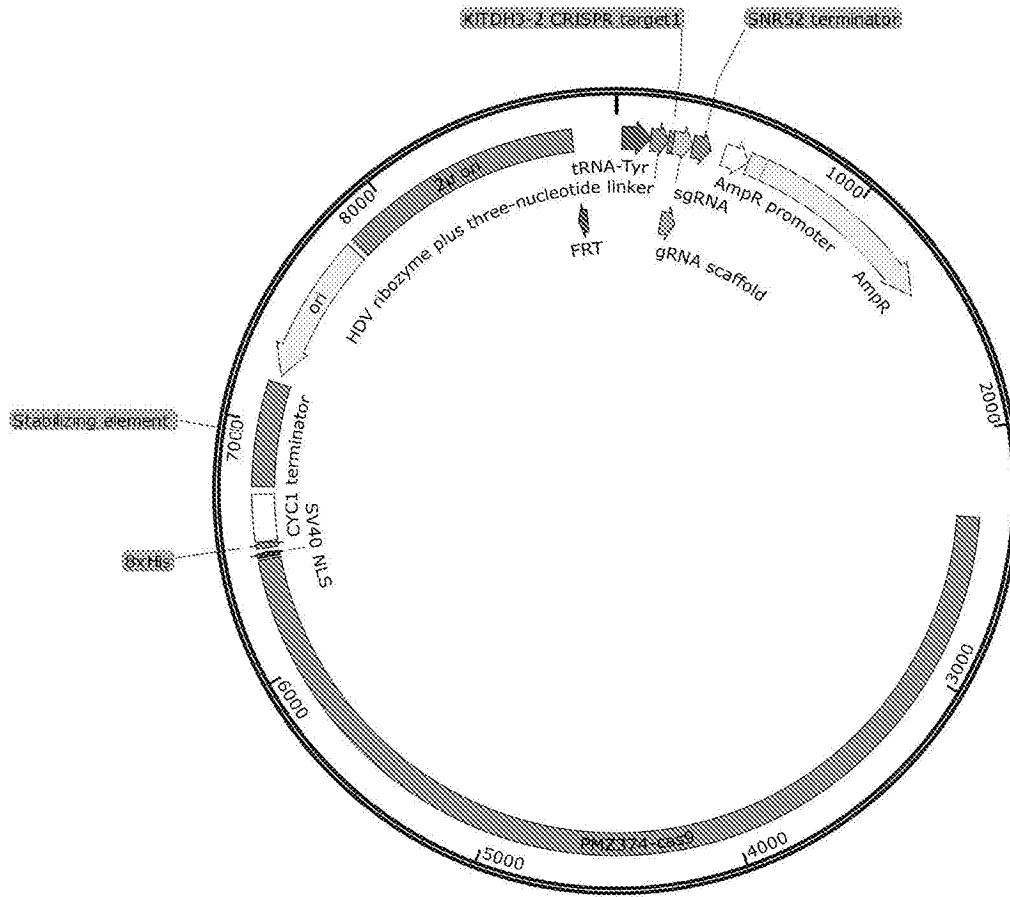


图7

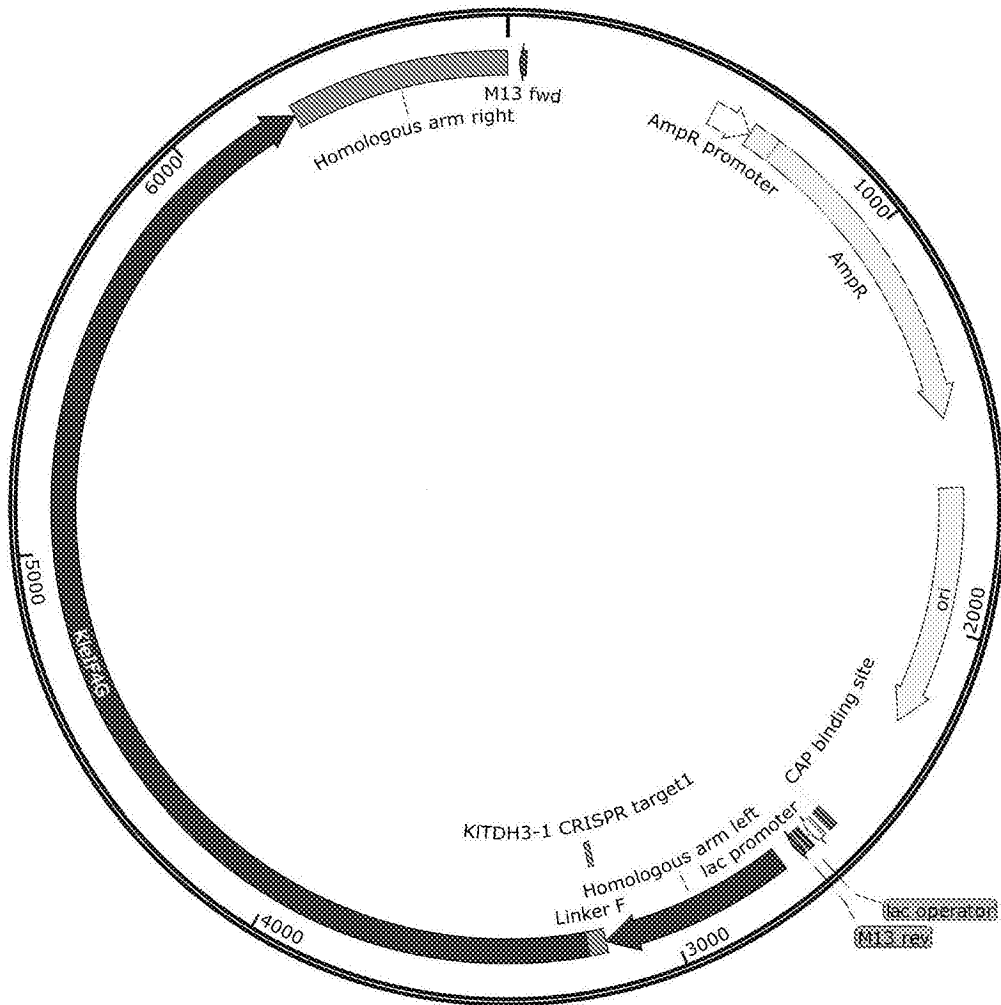


图8

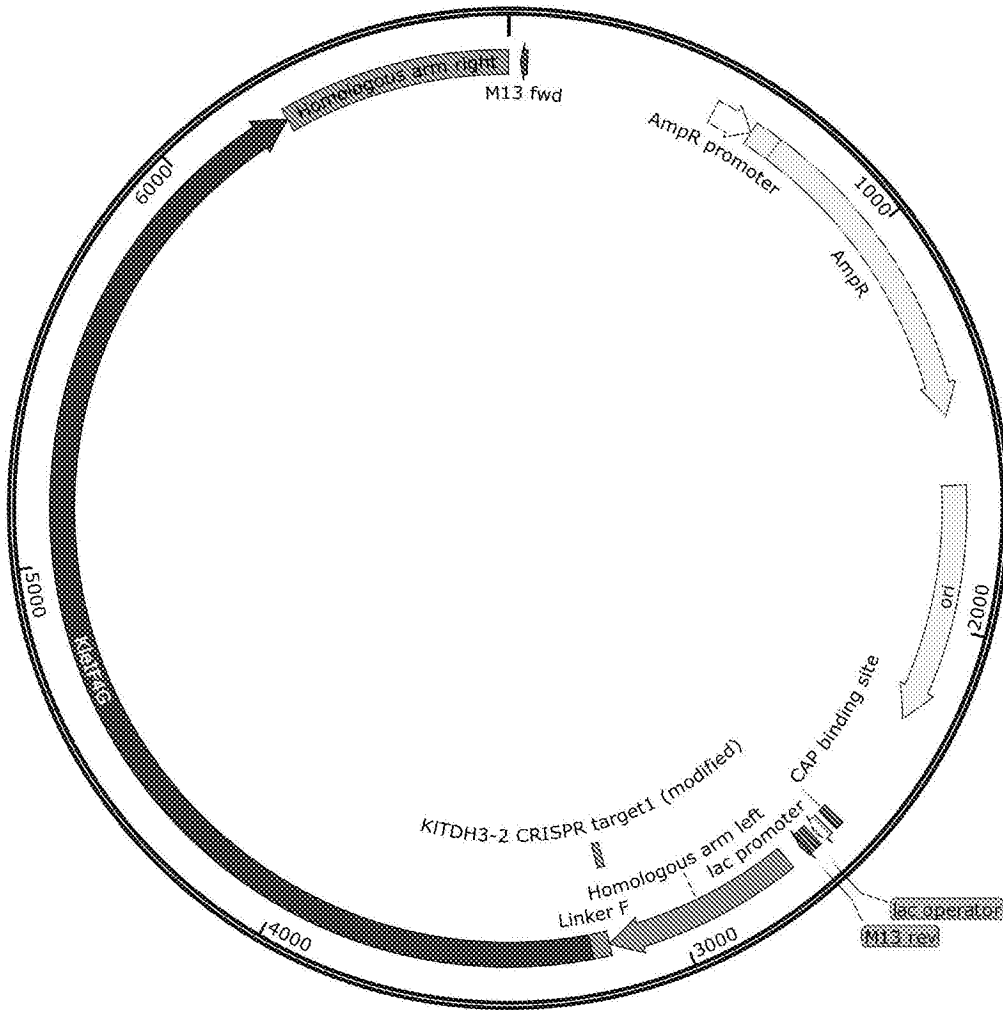


图9

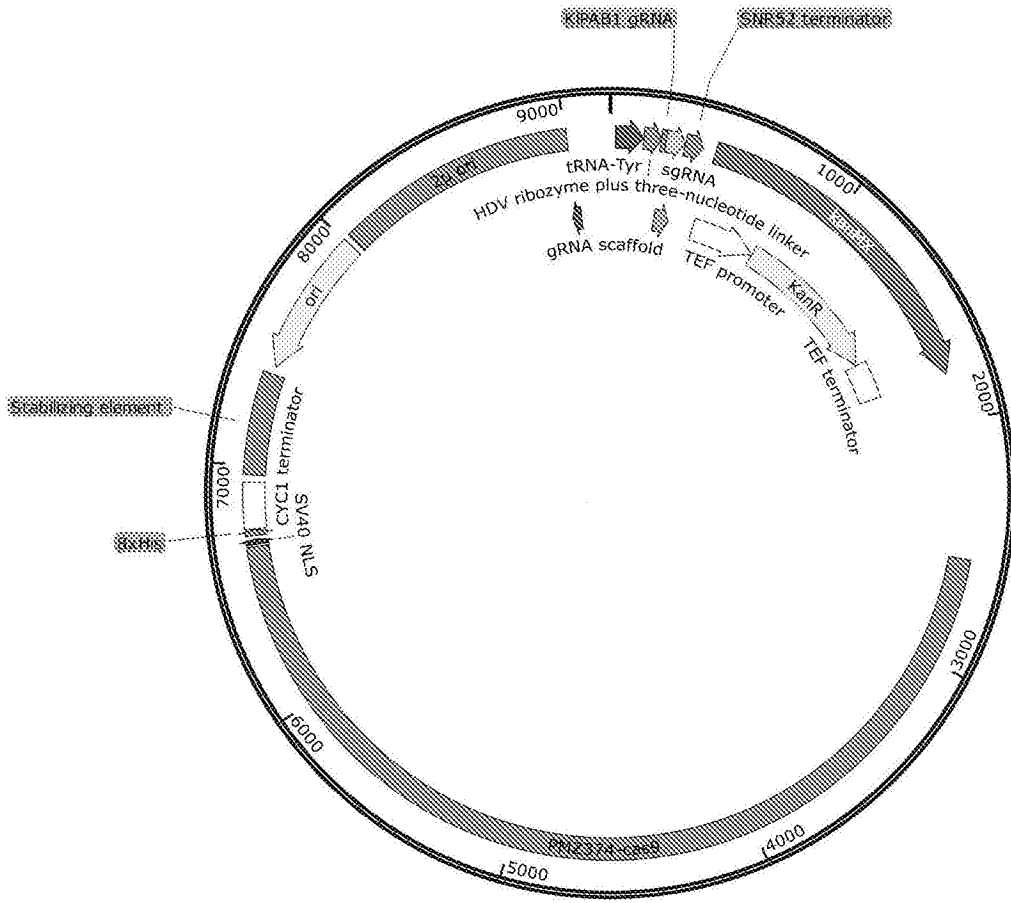


图10

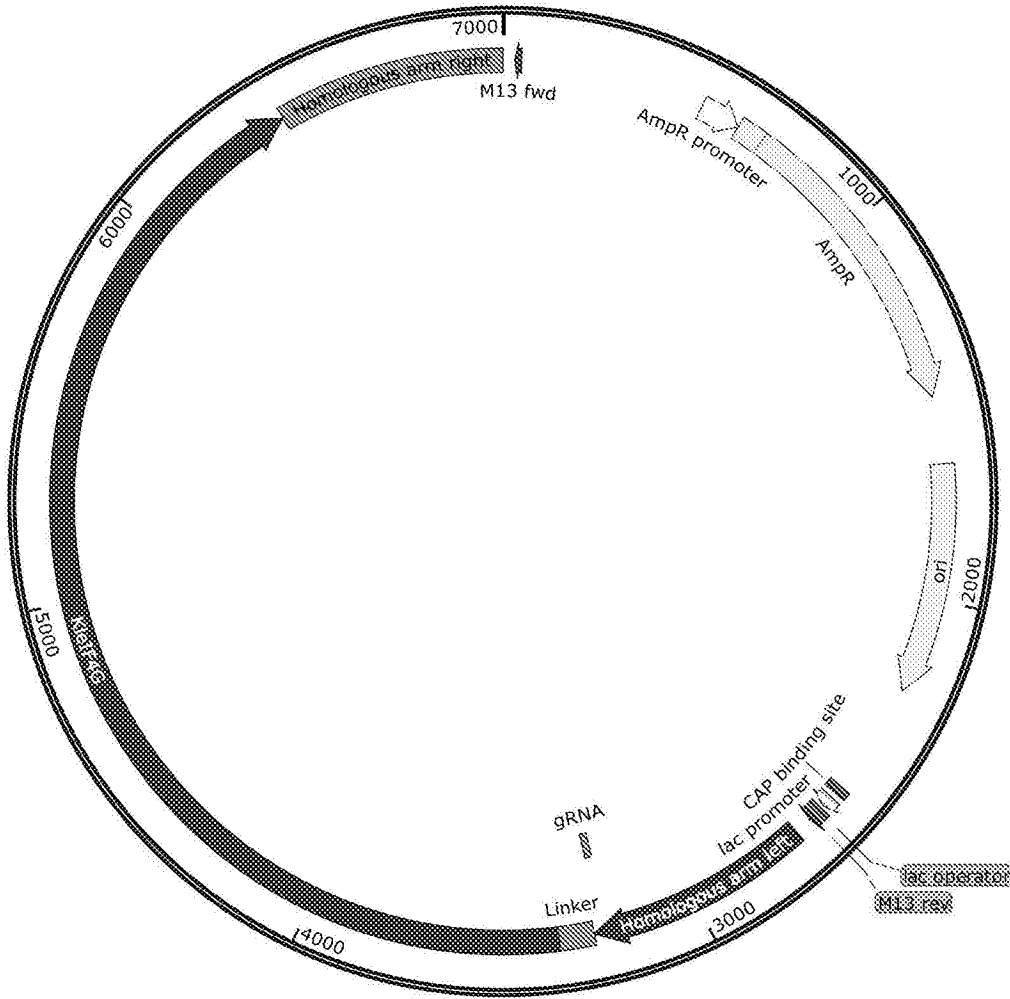


图11

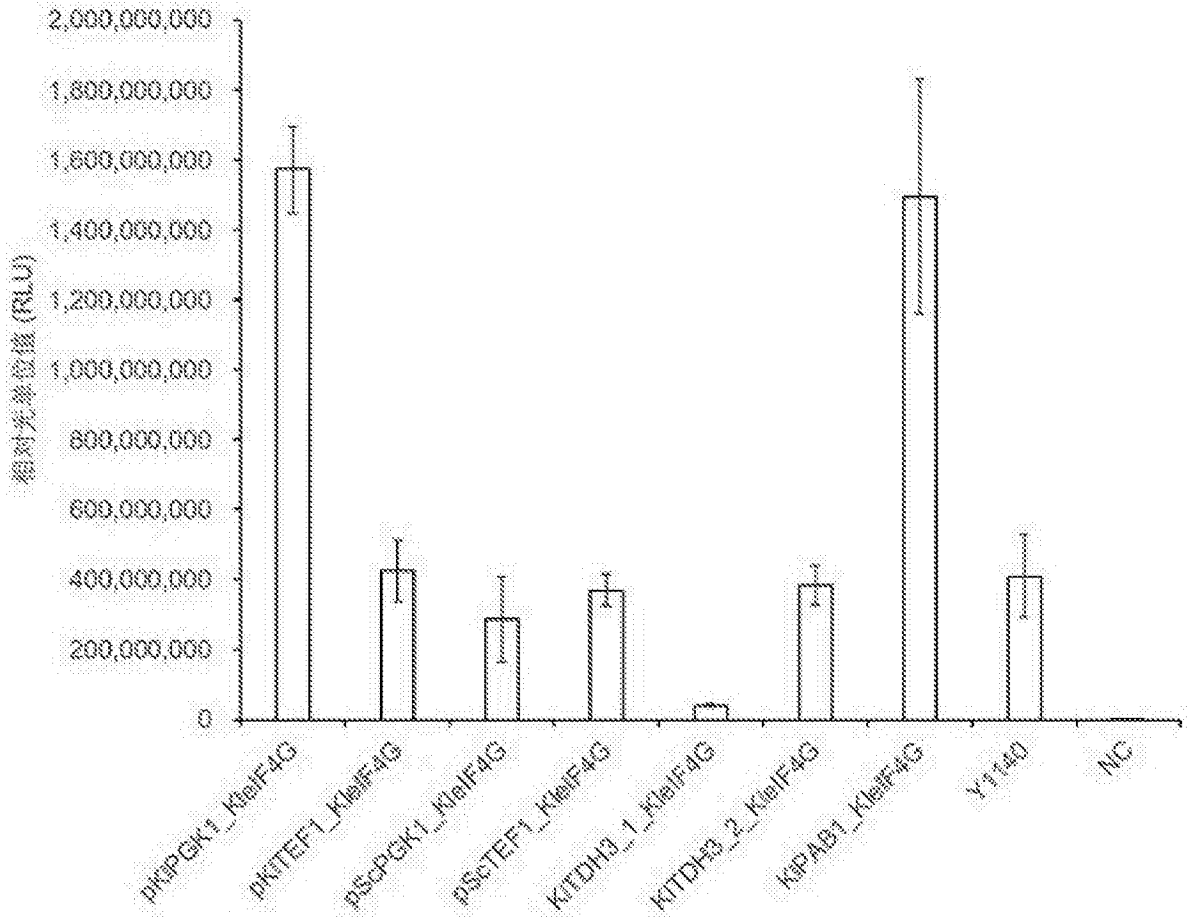


图12