



(12)发明专利申请

(10)申请公布号 CN 110938649 A
(43)申请公布日 2020.03.31

(21)申请号 201811113130.0

(22)申请日 2018.09.25

(71)申请人 康码(上海)生物科技有限公司
地址 201321 上海市浦东新区芙蓉花路500
弄8号楼3楼

(72)发明人 郭敏 丁晓辉 刘显成 杨宁
王绍杰 董颖颖 王静 代田纯
于雪

(51)Int.Cl.
C12N 15/81(2006.01)
C12N 15/67(2006.01)

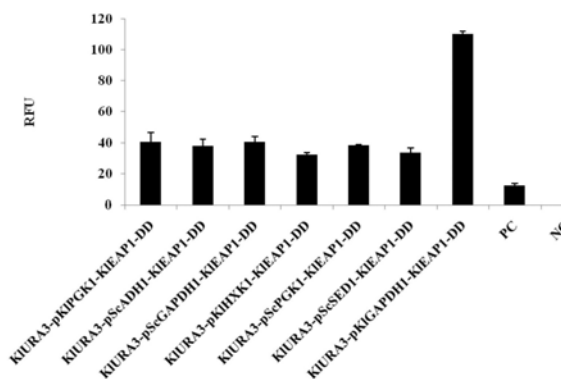
权利要求书1页 说明书20页
序列表8页 附图2页

(54)发明名称

一种提高外源蛋白表达量的蛋白合成体系及其应用方法

(57)摘要

本发明提供了一种提高外源蛋白表达量的蛋白合成体系及其应用方法,具体地,通过在编码eIF4E结合蛋白的核苷酸序列前插入强启动子序列构建的多核苷酸序列或载体,实现对Eap1p、p20等eIF4E结合蛋白过表达,并利用含有该改造的细胞裂解液进行外源蛋白合成。本发明提供了提高外源蛋白表达量的蛋白合成体系,本发明的合成体系可显著提高蛋白质翻译的效率。



1. 一种提高外源蛋白的表达量的体外蛋白合成体系,所述合成体系至少包括细胞裂解产物,所述细胞裂解产物来自于基因工程细胞,所述基因工程细胞的基因序列中整合有在编码eIF4E结合蛋白的核苷酸序列前插入强启动子序列构建的多核苷酸序列;且该合成体系中不添加或额外添加eIF4E结合蛋白。

2. 根据权利要求1所述的合成体系,其特征在于:所述强启动子序列选自GAPDH1、HXK1、PGK1、TEF1、TIF1、ADH1、SED1之一或其组合。

3. 根据权利要求2所述的合成体系,其特征在于:所述强启动子序列选自pScGAPDH1、pK1HXK1、pK1PGK1、pScPGK1、pScSED1、pScADH1、pK1GAPDH1之一或其组合。

4. 根据权利要求1-3的任意一项所述的合成体系,其特征在于:所述的eIF4E结合蛋白为Eap1p、p20、4E-BP1、4E-BP2、PHAS-I的一种或其组合。

5. 根据权利要求1-3的任意一项所述的合成体系,其特征在于:所述细胞为真核细胞。

6. 根据权利要求5所述的合成体系,其特征在于:所述真核细胞为哺乳动物细胞、植物细胞、酵母细胞、昆虫细胞之一或其任意组合。

7. 根据权利要求6所述的合成体系,其特征在于:酵母选自酿酒酵母、克鲁维酵母属酵母之一或其组合。

8. 根据权利要求7所述的合成体系,其特征在于:所述克鲁维酵母属酵母选自乳酸克鲁维酵母、马克斯克鲁维酵母、多布克鲁维酵母之一或其任意组合。

9. 一种体外合成蛋白的方法,其特征在于,所述方法包括以下步骤:

(i) 提供一蛋白合成体系,其中所述的合成体系为权利要求1-8的任一合成体系;和

(ii) 在适合表达蛋白的条件下,在编码所述外源蛋白的DNA或RNA模板存在下,孵育所述体外蛋白合成体系,从而表达所述的外源蛋白。

10. 根据权利要求9所述的体外蛋白合成方法,其特征在于,所述方法还包括:(iii) 分离或检测所述外源蛋白。

一种提高外源蛋白表达量的蛋白合成体系及其应用方法

技术领域

[0001] 本发明涉及生物工程技术领域,具体地,涉及一种提高外源蛋白表达量的蛋白合成体系及其应用方法。

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,其中受调控最多和限速的步骤是翻译的起始[1]。真核细胞的翻译起始可以分为两大类:“帽子结构”依赖的传统途径和“帽子结构”非依赖的途径。

[0003] 在细菌的翻译起始过程中,30S小亚基在3种起始因子的介导下直接在起始密码子的AUG附近形成起始复合体。真核生物的翻译起始则较为复杂,需要11种起始因子,并且核糖体在起始蛋白质合成时与mRNA结合的位点与原核生物不同[2]。真核生物的具有“帽子结构”(cap structure,m7GpppN)的mRNA翻译起始主要是通过依赖“帽子结构”的扫描机制进行的[1-3]。直接与帽子发生相互作用的是eIF4F,由eIF4G蛋白同时结合RNA解旋酶eIF4A和eIF4E而组成的翻译因子起始复合物。其中的eIF4E是帽子结合蛋白,为真核生物蛋白质合成途径中的关键调控靶点。在哺乳动物细胞中,与eIF4E发生相互作用的蛋白4E-BPs(eIF4E-binding proteins),通过结合eIF4E,使其不能和eIF4G结合,从而不能形成翻译起始复合物eIF4F,抑制依赖帽子结构的mRNA的翻译[4]。人源4E-BP1蛋白是哺乳动物雷帕霉素靶蛋白(mammalian target of rapamycin,mTOR)的底物之一,磷酸化水平受到mTOR信号通路的调控。当Thr37,Thr46和Ser65,Thr70残基被mTORC1先后磷酸化以后,4E-BP1将从eIF4E上解离,从而促进eIF4F起始复合物的形成和依赖帽子结构的翻译起始[5-6]。mTOR信号通路影响真核细胞中翻译起始及蛋白质合成,在细胞生长增殖过程中起重要作用。

[0004] mTOR是一类与细胞增殖紧密相关的丝/苏氨酸蛋白激酶,在进化上十分保守,从酵母到哺乳动物广泛存在,对细胞外包括生长因子,胰岛素,营养素,氨基酸,葡萄糖等多种刺激产生应答[7]。人类的编码TOR蛋白的基因和酵母中的同源性高达40%-60%[8]。在酵母体内没有和哺乳动物4E-BPs在结构上类似的蛋白,在功能上具有类似的结合eIF4E的蛋白有p20和EAP1p。这两种蛋白有一段高度相似的13个残基序列,其中包含了eIF4E的结合域YxxxxL Φ ,x为任意氨基酸残基, Φ 为疏水残基。除了这一段序列,这两种蛋白没有其他相似的地方,可能EAP1p还具有其他的功能会进一步影响TOR信号通路的下游[9]。对于其他真核生物的eIF4E结合蛋白,人源的4E-BP1、4E-BP2、鼠源的PHAS-I的蛋白序列都已有文献报道[10]。

[0005] 制备的酵母裂解液中含有大量的蛋白质合成翻译机器,可以通过外源添加模板而合成目标蛋白。但是酵母原来的mRNA也会继续使用资源合成非目标杂蛋白,影响目标蛋白

的合成效率和产量。本发明通过过表达Eap1p,与eIF4G竞争性结合eIF4E,减少翻译起始复合物eIF4F的形成,在体外抑制酵母裂解液中原来的mRNA帽子依赖的翻译过程,使更多的资源(如其他翻译因子,核糖体和氨基酸等)用于合成外源添加的不依赖帽子结构的体外蛋白合成体系模板,从而显著提高体外合成蛋白质的效率和产量。

[0006] 然而,本领域目前的体外生物合成系统的蛋白翻译合成的效率还较低,难以令人满意,因此,本领域亟需开发一种新的可提高外源蛋白翻译合成效率的体系和方法。

发明内容

[0007] 本发明的目的在于提供一种新的可提高外源蛋白翻译合成效率的合成体系和方法。同时,还提供经过改造的基因工程菌株及其应用,如基因工程菌株的细胞裂解液(细胞提取物)及其在无细胞蛋白合成中的应用。

[0008] 本发明第一方面提供了一种用于体外无细胞蛋白合成的基因工程菌株,其特征在于,所述基因工程菌株的基因组中整合有表达第一核酸构建物的第一外源基因表达盒,所述第一核酸构建物具有从5'至3'的式I结构:

[0009] A1-A2 (I);

[0010] 式中,

[0011] A1、A2分别为用于构成所述构建物的元件;

[0012] 各“-”独立地为键或核苷酸连接序列;

[0013] A1为启动子元件;

[0014] A2为eIF4E结合蛋白的编码序列;

[0015] 并且,在所述基因工程菌株中,eIF4E结合蛋白的表达或活性显著增强。

[0016] 在另一优选例中,所述eIF4E结合蛋白选自下组:EAP1p、p20、eIF4G1、eIF4G2、4E-BP1、4E-BP2、PHAS-I之一或其组合。

[0017] 在另一优选例中,所述eIF4E结合蛋白来源于酵母。

[0018] 在另一优选例中,所述酵母选自下组:酿酒酵母、克鲁维酵母属酵母、或其组合。

[0019] 在另一优选例中,所述克鲁维酵母属酵母选自下组:乳酸克鲁维酵母、马克斯克鲁维酵母、多布克鲁维酵母(*Kluyveromyces dobzhanskii*)、或其组合。

[0020] 在另一优选例中,所述eIF4E结合蛋白选自下组:EAP1p、p20、之一或其组合。

[0021] 在另一优选例中,所述eIF4E结合蛋白来源于人或鼠。

[0022] 在另一优选例中,所述eIF4E结合蛋白选自下组:4E-BP1、4E-BP2、PHAS-I、之一或其组合。

[0023] 在另一优选例中,所述启动子元件选自下组:GAPDH1、HXK1、PGK1、TEF1、TIF1、ADH1、SED1之一或其组合。

[0024] 在另一优选例中,所述启动子元件选自下组:pScGAPDH1、pK1HXK1、pK1PGK1、pScPGK1、pScSED1、pScADH1、pK1GAPDH1之一或其组合。

[0025] 在另一优选例中,所述eIF4E结合蛋白的表达或活性显著增强指eIF4E结合蛋白的表达或活性提高了 $\geq 30\%$,更佳的, $\geq 50\%$,更佳地, $\geq 70\%$,更佳地, $\geq 90\%$ 。

[0026] 在另一优选例中,所述eIF4E结合蛋白的表达或活性显著增强指与野生型菌株相比,所述工程菌株中的eIF4E结合蛋白的表达或活性(E1)/野生型菌株中的管家蛋白或参考

蛋白的表达或活性(E2) ≥ 1 , 较佳地, ≥ 2 , 更佳地, ≥ 3 。

[0027] 在另一优选例中, 所述菌株中的eIF4E结合蛋白的表达或活性的显著增强通过选自下组的方式实现: Crispr技术。

[0028] 在另一优选例中, 所述元件A2包括野生型和突变型的eIF4E结合蛋白的编码序列。

[0029] 在另一优选例中, 所述元件A2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所示核苷酸序列 $\geq 85\%$ 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序列相同活性的核苷酸。

[0030] 在另一优选例中, 所述的eIF4E结合蛋白为来自细胞(如酵母)的eIF4E结合蛋白。

[0031] 在另一优选例中, 所述酵母选自下组: 酿酒酵母、克鲁维酵母属酵母、或其组合。

[0032] 在另一优选例中, 所述克鲁维酵母属酵母选自下组: 乳酸克鲁维酵母、马克斯克鲁维酵母、多布克鲁维酵母(*Kluyveromyces dobzhanskii*)、或其组合。

[0033] 在另一优选例中, 所述的eIF4E结合蛋白为来自细胞(如人或鼠)的eIF4E结合蛋白。

[0034] 在另一优选例中, 所述细胞选自下组: 哺乳动物细胞(如HF9、HeLa、CHO、HEK293)、植物细胞、酵母细胞、昆虫细胞、或其组合。

[0035] 在另一优选例中, 所述细胞选自下组: HeLa、CHO、HF9、E μ Myc、HEK293、BY-2、酵母、小麦胚芽细胞、兔网织红细胞或其组合。

[0036] 在另一优选例中, 所述元件A2为衍生自细胞(如酵母)的eIF4E结合蛋白的编码序列。

[0037] 在另一优选例中, 所述A1、A2来源于细胞(如酵母)。

[0038] 在另一优选例中, 所述eIF4E结合蛋白来源于选自下组的一种或多种来源的酵母: 毕氏酵母、克鲁维酵母, 较佳地, 来源于克鲁维酵母。

[0039] 在另一优选例中, 所述克鲁维酵母包括马克斯克鲁维酵母、和/或乳酸克鲁维酵母。

[0040] 在另一优选例中, 所述编码eIF4E结合蛋白的核苷酸序列如SEQ ID NO.: 1所示。

[0041] 在另一优选例中, 所述eIF4E结合蛋白的蛋白序列具有SEQ ID NO.: 2所示的序列或其活性片段, 或者具有与SEQ ID NO.: 2所示多肽 $\geq 85\%$ 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。

[0042] 在另一优选例中, 所述eIF4E结合蛋白的蛋白序列如SEQ ID NO.: 2所示。

[0043] 在另一优选例中, 所述菌株选自下组: 克鲁维酵母菌株、毕赤酵母菌株、酿酒酵母菌株、裂殖酵母菌株、或其组合。

[0044] 本发明第二方面提供了一种本发明第一方面所述的基因工程菌株的用途, 用于提高体外蛋白合成效率。

[0045] 本发明第三方面提供了一种无细胞的细胞提取物, 所述细胞提取物来源于本发明第一方面所述的基因工程菌株。

[0046] 在另一优选例中, 所述细胞提取物为可溶性的细胞提取物。

[0047] 在另一优选例中, 所述细胞提取物来源选自下组的一种或多种细胞: 哺乳动物细

胞(如HF9、HeLa、CHO、HEK293)、酵母细胞、植物细胞、昆虫细胞或其组合。

[0048] 在另一优选例中,所述细胞提取物来源选自下组的一种或多种细胞:HeLa、CHO、HF9、EμMyc、HEK293、BY-2、酵母、小麦胚芽细胞、兔网织红细胞或其组合。

[0049] 在另一优选例中,所述细胞提取物包括酵母细胞提取物。

[0050] 在另一优选例中,所述酵母细胞选自下组的一种或多种来源的酵母:毕氏酵母、克鲁维酵母、或其组合;较佳地,所述的酵母细胞包括:克鲁维酵母,更佳地为马克斯克鲁维酵母、和/或乳酸克鲁维酵母。

[0051] 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。

[0052] 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。

[0053] 在另一优选例中,所述的酵母细胞提取物是用包括以下步骤的方法制备:

[0054] (i) 提供酵母细胞;

[0055] (ii) 对酵母细胞进行洗涤处理,获得经洗涤的酵母细胞;

[0056] (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;和

[0057] (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。

[0058] 在另一优选例中,所述的固液分离包括离心。

[0059] 在另一优选例中,在液态下进行离心。

[0060] 在另一优选例中,所述离心条件为 $5000-100000 \times g$,较佳地, $8000-30000 \times g$ 。

[0061] 在另一优选例中,所述离心时间为0.5-2h,较佳地,20min-50min。

[0062] 在另一优选例中,所述离心在 $1-10^{\circ}\text{C}$ 下进行,较佳地,在 $2-6^{\circ}\text{C}$ 下进行。

[0063] 在另一优选例中,所述的洗涤处理采用洗涤液在pH为7-8(较佳地,7.4)下进行处理。

[0064] 在另一优选例中,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。

[0065] 在另一优选例中,所述的破细胞处理包括高压破碎、冻融(如液氮低温)破碎。

[0066] 本发明第四方面提供了一种体外的无细胞的蛋白合成体系,所述无细胞的蛋白合成体系包括本发明第三方面所述的细胞提取物。

[0067] 在另一优选例中,所述蛋白合成体系由或基本由本发明第三方面所述的细胞提取物构成。

[0068] 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:

[0069] (a) 细胞提取物;

[0070] (b) 用于合成蛋白质的底物;

[0071] (c) 用于合成RNA的底物;

[0072] (d) 不含或含有RNA聚合酶。

[0073] 在另一优选例中,所述合成体系还包括选自下组的一种或多种组分:镁离子、钾离子、缓冲剂、能量再生系统、聚乙二醇、葡萄糖、磷酸盐、二硫苏糖醇(DTT)和任选的溶剂、蔗糖,所述溶剂为水或水性溶剂。

[0074] 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。

- [0075] 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。
- [0076] 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。
- [0077] 在另一优选例中,所述磷酸盐选自下组:磷酸钾、磷酸二氢钾、磷酸钠、磷酸铵、磷酸二氢钠、磷酸氢二钠、磷酸氢二钾、或其组合。
- [0078] 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。
- [0079] 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。
- [0080] 在另一优选例中,所述无细胞的蛋白合成体系还包括人工合成的tRNA。
- [0081] 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。
- [0082] 在另一优选例中,所述无细胞的蛋白合成体系还包括外源的用于指导蛋白质合成的DNA分子。
- [0083] 在另一优选例中,所述的DNA分子为线性的。
- [0084] 在另一优选例中,所述的DNA分子为环状的。
- [0085] 在另一优选例中,所述的DNA分子含有编码外源蛋白的序列。
- [0086] 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。
- [0087] 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。
- [0088] 在另一优选例中,所述无细胞的蛋白合成体系包括选自下组的成分:4-羟乙基哌嗪乙磺酸、醋酸钾、醋酸镁、核苷三磷酸、氨基酸、磷酸肌酸,二硫苏糖醇(DTT)、磷酸肌酸激酶、RNA聚合酶、或其组合。
- [0089] 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。
- [0090] 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,较佳地,分子量为3000-10000的聚乙二醇。
- [0091] 在另一优选例中,所述核苷三磷酸选自下组:腺嘌呤核苷三磷酸、鸟嘌呤核苷三磷酸、胞嘧啶核苷三磷酸、尿嘧啶核苷三磷酸、或其组合。
- [0092] 在另一优选例中,所述氨基酸为选自下组: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、组氨酸、或其组合。
- [0093] 在另一优选例中,所述氨基酸包括D型氨基酸和/或L型氨基酸。
- [0094] 在另一优选例中,所述蛋白合成体系中,所述醋酸钾的浓度为20-210mM,较佳地,30-210mM,更佳地,30-150mM,更佳地,30-60mM。
- [0095] 在另一优选例中,所述蛋白合成体系中,所述醋酸镁的浓度为1-10mM,较佳地,1-5mM,更佳地,2-4mM。
- [0096] 在另一优选例中,所述蛋白合成体系中,所述磷酸肌酸的浓度为10-50mM,较佳地,

20-30mM,更佳地,25mM。

[0097] 在另一优选例中,所述蛋白合成体系中,所述二硫苏糖醇(DTT)的浓度为0.2-15mM,较佳地,0.2-7mM,更佳地,1-2mM。

[0098] 在另一优选例中,所述蛋白合成体系中,所述磷酸肌酸激酶的浓度为0.1-1mg/mL,较佳地,0.2-0.5mg/mL,更佳地,0.27mg/mL。

[0099] 在另一优选例中,所述蛋白合成体系中,所述RNA聚合酶为T7RNA聚合酶,所述T7RNA聚合酶的浓度为0.01-0.3mg/mL,较佳地,0.02-0.1mg/mL,更佳地,0.027-0.054mg/mL。

[0100] 本发明第五方面提供了一种体外合成蛋白的方法,包括步骤:

[0101] (i) 提供本发明第四方面所述的体外的无细胞的蛋白合成体系,并加入外源的用于指导蛋白质合成的DNA分子;

[0102] (ii) 在适合条件下,孵育步骤(i)的蛋白合成体系一段时间,从而合成由所述外源DNA编码的蛋白质。

[0103] 在另一优选例中,所述的方法还包括:(iii) 任选地从所述蛋白合成体系中,分离或检测所述的由外源DNA编码的蛋白质。

[0104] 在另一优选例中,所述外源DNA来自原核生物、真核生物。

[0105] 在另一优选例中,所述外源DNA来自动物、植物、病原体。

[0106] 在另一优选例中,所述外源DNA来自哺乳动物,较佳地灵长动物,啮齿动物,包括人、小鼠、大鼠。

[0107] 在另一优选例中,所述的外源蛋白的编码序列编码选自下组的外源蛋白: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、萤光素酶突变体、 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。

[0108] 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。

[0109] 在另一优选例中,所述的外源DNA编码选自下组的外源蛋白: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、萤光素酶突变体、 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。

[0110] 在另一优选例中,所述外源DNA编码的蛋白质选自下组: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。

[0111] 但上述仅为外源蛋白的示范性介绍,并非仅包括上述外源蛋白;可以为任意能够外源表达的蛋白。

[0112] 在另一优选例中,所述步骤(ii)中,反应温度为20-37℃,较佳地,20-25℃。

[0113] 在另一优选例中,所述步骤(ii)中,反应时间为1-72h,较佳地,2-23h。

[0114] 本发明第六方面提供了一种体外蛋白合成体系,所述合成体系至少包括细胞裂解产物,所述细胞裂解产物来自于基因工程细胞,所述基因工程细胞的基因序列中整合有在编码eIF4E结合蛋白的核苷酸序列前插入强启动子序列构建的多核苷酸序列,用于在体外蛋白合成体系中提高外源蛋白的表达量;且该合成体系中不添加或额外添加eIF4E结合蛋白。

[0115] 与此等同的,提供一种提高外源蛋白表达量的体外蛋白合成体系,其包括:

[0116] (a) 细胞裂解液;

[0117] (b) 无或用于合成RNA的底物;

[0118] (c) 用于合成蛋白质的底物;

[0119] (d) 无或含有RNA聚合酶;

[0120] (e) 无或额外添加eIF4E结合蛋白;

[0121] 其中,细胞裂解液为包含有在编码eIF4E结合蛋白的核苷酸序列前插入强启动子序列构建的多核苷酸序列或载体的细胞的裂解液。

[0122] 在另一优选例中,所述强启动子序列选自GAPDH1、HXK1、PGK1、TEF1、TIF1、ADH1、SED1之一或其组合。

[0123] 在另一优选例中,所述强启动子序列选自pScGAPDH1、pK1HXK1、pK1PGK1、pScPGK1、pScSED1、pScADH1、pK1GAPDH1之一或其组合。

[0124] 在另一优选例中,所述的eIF4E结合蛋白为来源于酵母的EAP1p、p20、人源的4E-BP1、4E-BP2、鼠源的PHAS-I的一种或其组合。

[0125] 在另一优选例中,所述细胞为真核细胞。

[0126] 在另一优选例中,所述细胞为哺乳动物细胞、植物细胞、酵母细胞、昆虫细胞之一或其任意组合。

[0127] 在另一优选例中,所述细胞为酵母、小麦胚芽细胞、兔网织红细胞之一或其任意组合。

[0128] 在另一优选例中,所述酵母选自酿酒酵母、克鲁维酵母属酵母之一或其组合。

[0129] 在另一优选例中,所述克鲁维酵母属酵母选自乳酸克鲁维酵母、马克斯克鲁维酵母、多布克鲁维酵母之一或其任意组合。

[0130] 在另一优选例中,所述细胞裂解产物为构成体外蛋白合成体系必不可少的一个组分,可以是液态的细胞裂解液,也可以是将细胞裂解液经冷冻干燥后得到的冻干粉,或其它形态。如果是冻干粉,使用时需要加入缓冲剂复溶,所述缓冲剂为体外合成体系中常用的缓冲剂(包括但不限于4-羟乙基哌嗪乙磺酸、三羟甲基氨基甲烷、或其组合)。细胞裂解产物中包含eIF4E结合蛋白,用于在体外蛋白合成体系中发挥提升外源蛋白表达的作用。

[0131] 在另一优选例中,所述细胞裂解产物可以本身含有细胞内源性表达的eIF4E结合蛋白,即细胞本身含有eIF4E结合蛋白的编码基因,通过在eIF4E结合蛋白的编码序列上游插入强启动子来增强eIF4E结合蛋白的表达。

[0132] 在另一优选例中,所述细胞裂解产物也可以本身不含有内源性eIF4E结合蛋白,即细胞本身不含有eIF4E结合蛋白的编码基因,而是经过改造,整合有在编码eIF4E结合蛋白的核苷酸序列前插入强启动子序列构建的多核苷酸序列或载体,来获得eIF4E结合蛋白的高表达。

[0133] 在另一优选例中,所述细胞裂解产物所包含的eIF4E结合蛋白可以是同源的,也可以是异源的。比如,可以直接在乳酸克鲁维酵母中EAP1p蛋白(K1EAP1p)的编码序列上游插入强启动子来增强eIF4E结合蛋白的表达;也可以在乳酸克鲁维酵母中整合有在编码酿酒酵母的EAP1p蛋白(ScEAP1p)核苷酸序列前插入强启动子序列构建的多核苷酸序列,来增强ScEAP1p蛋白的表达;也可以在酵母细胞中整合有在编码人源4E-BP1、4E-BP2蛋白的核苷酸序列前插入强启动子序列构建的多核苷酸序列,来增强4E-BP1、4E-BP2蛋白的表达。

[0134] 在另一优选例中,所述细胞裂解产物可以来自于单一细胞,也可以来源于两种或多种细胞的裂解产物。当来自于两种或多种细胞的裂解产物时,eIF4E结合蛋白可以在所有细胞内过表达,也可以在其中几种细胞内过表达,但至少在其中一种细胞内过表达。

[0135] 在另一优选例中,eIF4E结合蛋白可以额外添加到体外蛋白合成体系中,不论所述细胞是否经过改造使改细胞裂解产物中含有过表达的eIF4E结合蛋白。即便是使用未经过表达eIF4E结合蛋白改造的细胞,通过另行额外添加eIF4E结合蛋白同样可以获得相同的技术效果。额外添加的eIF4E结合蛋白可以通过在其它细胞内表达或无细胞体外合成后,经纯化或不经纯化获得。

[0136] 在另一优选例中,体外蛋白合成体系中可以含有一种或多种eIF4E结合蛋白,所述eIF4E结合蛋白通过细胞内源表达而存在于细胞裂解产物中和/或通过额外添加。

[0137] 本发明第七方面提供了一种体外合成蛋白的方法,所述方法包括以下步骤:

[0138] (i) 提供一蛋白合成体系,其中所述的合成体系为本发明第六方面所述的合成体系;和

[0139] (ii) 在适合表达蛋白的条件下,在编码所述外源蛋白的DNA或RNA模板存在下,孵育所述体外蛋白合成体系,从而表达所述的外源蛋白。

[0140] 在另一优选例中,所述方法还包括:(iii) 分离或检测所述外源蛋白。

[0141] 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

[0142] 图1显示了pKM-CAS1.0-K1URA3的质粒图谱。

[0143] 图2为体外翻译活性测定结果。

具体实施方式

[0144] 经过广泛而深入的研究,通过大量筛选和摸索,首次筛选得到了一种新的基因改造菌株,该菌株的基因组中整合有表达第一核酸构建物的第一外源基因表达盒,所述第一核酸构建物由启动子元件(如pScGAPDH1、pScHXX1、pScPGK1、pScTEF1、pScTIF1、pScADH1、pScTEF1等强启动子元件,其中p为promoter的简写,为启动子;Sc为酿酒酵母来源,K1为克

ATTCCCAGTTGGGATGCCACCACATATGATGGCCCCTCCTATGGGAATGCCTCCACAGCACTTCCAAGGTGGCTAT
GCTATGCCTCCTCCACCGCCACCTGGAATGGGACAGCCACGTATAATGAGCAACGGTAAAGGATCCCCGAACACT
TGGGTAACCCACAACAGCAGCCTCCTAGCGCTCCAAAACAGCCAGCAAGTAAAGTTGGTCCGCAAGGTCAAGTTCC
TTCACAAGGCCATCCTCAATCCCACCCACAACAACCTTATATGATGTCAGGTATGCCAATGAACTTCAATGGGCAA
AATGTTCCAATTCCAGTCAAGGTATCCCACCAAATGCATTCCCATATGGCCATCCAATGATGGCTCTACAGTTTC
AACAACAACAACAGCAGCAGCAGCAACAACAATACCCGCAGCAACAACAACACCCGCAGCAACCCGCGCCAAACAAA
TCCGCGCCAATAA;

[0155] 在一优选实施方式中,所述K1EAP1p的蛋白序列如SEQ ID NO.:2所示。

[0156] MSDTAEENGLTRFLMKVKENLDTSQSEAQTETEPIEFKYTFDYKPEFSSGKVVYSRNELLAIREQVAE
EDVTNLASELPNKKFWRLPVPGSNVGGRKGSNTRGNHDDKFGGKDRGAQSGSRNARNNRNSKRQGGKAGKESNEE
YIALEEQMESTGNPMADFENWRNKMELERQKRGLDSDAGKSDSPAGLPAQSFSSISDFFNLKPDDQKSAPLEEL
EPADSSSEVSKQTFEKQDRDSQDVQHGSQGGISKSNSRFSFFFQGGNSPDASDNRIAKPPSQSSEDSRPAAGS
RLLSLFNTDSPSSDSVVQHQQPEKPMVNNPPGLTQQSSSTLSVSSVASSNHSQPHSGPRAAKDNDVNGSVFLKSLMT
KGSENMMQQPLVSAPPGLSQNQSQIPNITQQQQQQQQQQHQQQQQQQQQQHIQRPHQQQQQPQQRKLQIHHQH
SQHAHPAQQQQHGKPPQNIASGAPMQAPPGFVGMPPHMMAPPMPMPQHFGQGYAMPPPPPPGMGQPRIMSNGK
GFPEHLGNPQQQPPSAPKQPASKVGPQGQVPSQGHPQSHPPQPYMMSGMPMNFNGQNPVIPAQGIPPNAFPYGHM
MALQFQQQQQQQQQQYPPQQQHPQQPRQTNPRQ。

[0157] 外源基因表达盒

[0158] 如本文所用,术语“外源基因表达盒”是指带有表达第一核酸构建物的第一外源基因表达盒。

[0159] 本申请的工程菌株是通过将“整合第一外源基因表达盒的重组菌”进行原生质体融合得到的。该工程菌株同时整合有第一外源基因表达盒。

[0160] 体外表达系统

[0161] 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。

[0162] 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。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,更好的大规模发酵特性、食品安全的级别、以及同时具有蛋白翻译后修饰的能力等。

[0163] 在本发明中,酵母体外表达系统不受特别限制,一种优选的酵母体外表达系统为克鲁维酵母表达系统(更佳地,乳酸克鲁维酵母表达系统)。

[0164] 体外表达系统不限于酵母,更具体地为乳酸克鲁维酵母表达系统,实施例仅利用该系统进行验证性说明,其他真核细胞体外表达系统同样适用。

[0165] 蛋白合成体系

[0166] 一种体外蛋白合成体系,所述合成体系至少包括细胞裂解产物,所述细胞裂解产物来自于基因工程细胞,所述基因工程细胞的基因序列中整合有在编码eIF4E结合蛋白的核苷酸序列前插入强启动子序列构建的多核苷酸序列,用于在体外蛋白合成体系中提高外源蛋白的表达量;且该合成体系中不添加或额外添加eIF4E结合蛋白。

[0167] 与此等同的,提供一种提高外源蛋白表达量的体外蛋白合成体系,其包括:

[0168] (a) 细胞裂解液;

[0169] (b) 无或用于合成RNA的底物;

[0170] (c) 用于合成蛋白质的底物;

[0171] (d) 无或含有RNA聚合酶;

[0172] (e) 无或额外添加eIF4E结合蛋白;

[0173] 其中,细胞裂解液为包含有在编码eIF4E结合蛋白的核苷酸序列前插入强启动子序列构建的多核苷酸序列或载体的细胞的裂解液。

[0174] 在另一优选例中,所述强启动子序列选自GAPDH1、HXK1、PGK1、TEF1、TIF1、ADH1、SED1之一或其组合。

[0175] 在另一优选例中,所述强启动子序列选自pScGAPDH1、pK1HXK1、pK1PGK1、pScPGK1、pScSED1、pScADH1、pK1GAPDH1之一或其组合。

[0176] 在另一优选例中,所述的eIF4E结合蛋白为来源于酵母的EAP1p、p20、人源的4E-BP1、4E-BP2、鼠源的PHAS-I的一种或其组合。

[0177] 在另一优选例中,所述细胞为真核细胞。

[0178] 在另一优选例中,所述细胞为哺乳动物细胞、植物细胞、酵母细胞、昆虫细胞之一或其任意组合。

[0179] 在另一优选例中,所述细胞为酵母、小麦胚芽细胞。

[0180] 在另一优选例中,所述酵母选自酿酒酵母、克鲁维酵母属酵母之一或其组合。

[0181] 在另一优选例中,所述克鲁维酵母属酵母选自乳酸克鲁维酵母、马克斯克鲁维酵母、多布克鲁维酵母之一或其任意组合。

[0182] 具体的,对于体外无细胞的蛋白合成体系的组成成分而言而言,所述合成体系包括:

[0183] (a) 细胞提取物;

[0184] (b) 用于合成蛋白质的底物;

[0185] (c) 用于合成RNA的底物;

[0186] (d) 不含或含有RNA聚合酶。

[0187] 在另一优选例中,所述合成体系还包括选自下组的一种或多种组分:镁离子、钾离子、缓冲剂、能量再生系统、聚乙二醇、葡萄糖、磷酸盐、二硫苏糖醇(DTT)和任选的溶剂,所述溶剂为水或水性溶剂。

[0188] 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,荧光素酶DNA,RNA聚合酶。

[0189] 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T7RNA聚合酶。

[0190] 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%,较佳地,30-60%,更佳地,40-50%。

[0191] 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。

[0192] 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。

[0193] 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:

[0194] (i) 提供酵母细胞;

[0195] (ii) 对酵母细胞进行洗涤处理,获得经洗涤的酵母细胞;

[0196] (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;

[0197] (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。

[0198] 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。

[0199] 在一优选实施方式中,所述离心在液态下进行。

[0200] 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。

[0201] 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。

[0202] 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。

[0203] 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。

[0204] 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。

[0205] 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。

[0206] 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、0.08mM。

[0207] 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)。

[0208] 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。

[0209] 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:22mM,pH为7.4的4-羟乙基哌嗪乙磺酸,30-150mM醋酸钾,1.0-5.0mM醋酸镁,1.5-4mM核苷三磷酸混合物,0.08-0.24mM的氨基酸混合物,25mM磷酸肌酸,1.7mM二硫苏糖醇,0.27mg/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,0.027-0.054mg/mL T7RNA聚合酶。

[0210] 本发明的主要优点包括:

[0211] (i) 本发明首次构建整合由强启动子元件,如pScGAPDH1、pScPGK1、pScTEF1、pScTIF1、pScADH1、pScTEF1、pK1ADH1、pK1GAPDH1、pK1HXX1、pK1PGK1、pK1TEF1、pK1TIF1和K1EAP1p蛋白的编码序列构成的核酸构建物;及其包含该核酸构建物的工程菌株。

[0212] (ii) 本发明首次意外的发现,来源于本发明的工程菌株的细胞提取物(如酵母细胞提取物)可显著提高体外蛋白质合成体系产生蛋白质的效率。

[0213] (iii) 本发明首次发现,本发明的菌株具有比野生型菌组和更强的体外蛋白质合成能力,其编码合成的eGFP蛋白放出的相对荧光单位值达到110.5,而野生型酵母菌株合成的eGFP蛋白的相对荧光单位值仅有12.3,提高约8.98倍。

[0214] (iv) 本发明可实现EAP1p蛋白的过表达,从而显著提高体外蛋白质合成体系产生蛋白质的效率。

[0215] 下面结合具体实施例,进一步阐述本发明。应理解,这些实施例仅用于说明本发明而不适用于限制本发明的范围。下列实施例中未注明具体条件的实验方法,通常按照常规条件,例如Sambrook等人,分子克隆:实验室手册(New York: Cold Spring Harbor Laboratory Press,1989)中所述的条件,或按照制造厂商所建议的条件。除非另外说明,否则百分比和份数是重量百分比和重量份数。

[0216] 如无特别说明,则本发明实施例中所用的材料和试剂均为市售产品。

[0217] 实施例1通过CRISPR-Cas9技术在K1EAP1基因前插入强启动子

[0218] 1.1K1URA3序列检索及CRISPR gRNA序列确定

[0219] URA3基因是酵母V号染色体上的一个基因,能进行阳性选择和阴性选择,是酵母遗传工程中最重要遗传标记物之一。实施例选择URA3作为筛选标记,当然也可选择其他常用的筛选标记(遗传标记物)。

[0220] 基于酿酒酵母(*Saccharomyces cerevisiae*)中URA3基因序列,在UniProt数据库中以URA3基因进行BLAST比对分析,确定乳酸克鲁维酵母中URA3同源基因序列,命名为K1URA3(位于染色体E的2034995..2035798)。在K1URA3基因终止密码子,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,为40%-60%;避免poly T结构的的存在。确定的K1URA3gRNA序列为GCGCTCCCCATTAATTATAC,位于染色体E的424927...424936位点。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。

[0221] 1.2CRISPR-Cas9介导的在K1EAP1前插入强启动子质粒构建

[0222] 为了实现K1EAP1p的过量表达,通过CRISPR-Cas技术在K1EAP1基因前分别插入pScADH1、pScGAPDH1、pScPGK1、pScSED1、pScTEF1、pScTIF1和pK1GAPDH1、pK1HXX1、pK1PGK1、

pK1TEF1、pK1TIF1启动子。质粒构建及转化方法如下：

[0223] 1.2.1CRISPR质粒构建

[0224] 使用引物PF1:CTTGAAACAAGGTGCGCAAGGTTTTAGAGCTAGAAATAGCAAGTTAAAAT,PR1:C TTGCGCACCTTGTTC AAGTGCATCGGCCGGAATCGAACCCGGGCCC,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 L10 \times digestion buffer,37 $^{\circ}$ C温浴3h。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30min,42 $^{\circ}$ C热激45s后,加入1mL LB液体培养基37 $^{\circ}$ C振荡培养1h,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-K1URA3(图1)。

[0225] 1.2.2供体DNA片段扩增

[0226] 为了快速扩增线性供体DNA片段,通过OVER-LAP PCR扩增得到线性供体DNA序列。

[0227] 在UniProt数据库中分别查找酿酒酵母ADH1、GAPDH1、PGK1、SED1、TEF1和TIF1基因,各以其起始密码子上游1000bp为ADH1、GAPDH1、PGK1、SED1、TEF1和TIF1启动子序列,分别命名为pScADH1、pScGAPDH1、pScPGK1、pScSED1、pScTEF1和pScTIF1。

[0228] 基于酿酒酵母中ADH1、GAPDH1、HXK1、PGK1、TEF1和TIF1基因序列,在UniProt数据库中分别以ADH1、GAPDH1、HXK1、PGK1、TEF1和TIF1基因进行BLAST比对分析,确定乳酸克鲁维酵母中ADH1、GAPDH1、HXK1、PGK1、TEF1和TIF1同源基因序列,各以其起始密码子上游1000bp为ADH1、GAPDH1、HXK1、PGK1、TEF1和TIF1启动子序列,分别命名为pK1ADH1、pK1GAPDH1、pK1HXK1、pK1PGK1、pK1TEF1、pK1TIF1,其碱基序列见序列表的SEQ ID NO.:3~8。

[0229] 基于酿酒酵母中EAP1基因序列,在UniProt数据库中以EAP1基因进行BLAST比对分析,确定乳酸克鲁维酵母中EAP1同源基因序列,命名为K1EAP1(位于染色体B的1231530..1233434)。

[0230] (1)以乳酸克鲁维酵母基因组DNA为模板扩增同源臂序列,引物分别为:F:ACACATTACTTGCTCGAGCAT;R:GTACACCCGAAACAACAAAAGGATTTAATGGGGAGCGCTGATTCTCTTT,和F: TGAGAAGGTTTTGGGACGCTCGAAGTTATACAGGAACTTAATAGAACAAATCAC;R: CCTAACGGGATTTTCGCTTCGTGA进行PCR扩增;

[0231] 以酿酒酵母基因组DNA为模板扩增ADH1启动子序列,以引物F:AAAGAGAATCAGCGCTCCCCATTAATCCTTTTGTGTTTCCGGGTGTAC;R:CATTCTCTTCTGCTGTGTCGGACATAGTTGATTGTATGCTTGGTATAGCT进行PCR扩增;

[0232] 以乳酸克鲁维酵母基因组DNA为模板扩增EAP1序列,以引物F:AGCTATACCAAGCATA CAATCAACTATGTCCGACACAGCAGAAGAGAATG;R:GTGATTTGTTCTATTAAGTTTCCTGTATAACTTCGAGCGTCCCAAAACCTTCTCA进行PCR扩增。

[0233] 然后将四次扩增产物都稀释50倍,各吸1.5 μ L混合作模板扩增供体序列,以引物GCCAGCGTCAATACACTCCC和TTACGACAATGCCTAGTTGAGTGC进行PCR扩增。所得产物经1%的琼脂糖凝胶电泳,回收纯化即得供体DNA,命名为K1URA3-pScADH1-K1EAP1-DD。

[0234] (2)以乳酸克鲁维酵母基因组DNA为模板扩增同源臂序列,引物分别为F:ACACATTACTTGCTCGAGCAT和R:TTTGAATGGCAGTATTGATAATGATTAATGGGGAGCGCTGATTCTCTTT,引物F: TGAGAAGGTTTTGGGACGCTCGAAGTTATACAGGAACTTAATAGAACAAATCAC和R:

CCTAACGGGATTTTCGCTTCGTGA进行PCR扩增；

[0235] 以酿酒酵母基因组DNA为模板扩增GPADH1启动子序列，以引物F:AAAGAGAATCAGCGCTCCCCATTAATCATTATCAATACTGCCATTTCAA和R:CATTCTTCTGCTGTGTCGGACATTTGTTTGTATGTGTGTTATTC进行PCR扩增；

[0236] 以乳酸克鲁维酵母基因组DNA为模板扩增EAP1序列，以引物F:GAATAAACACACATAACAAAACAAATGTCCGACACAGCAGAAGAGAATG和R:GTGATTTGTTCTATTAAGTTTCCTGTATAACTTCGAGCGTCCCAAACCTTCTCA进行PCR扩增。

[0237] 然后将四次扩增产物都稀释50倍，各吸1.5 μ L混合作模板扩增供体序列，以引物GCCAGCGTCAATACACTCCC和TTACGACAATGCCTAGTTGAGTGC进行PCR扩增。所得产物经1%的琼脂糖凝胶电泳，回收纯化即得供体DNA，命名为K1URA3-pScGPADH1-K1EAP1-DD。

[0238] (3) 以乳酸克鲁维酵母基因组DNA为模板扩增同源臂序列，引物分别为F:ACACATTACTTGCCTCGAGCAT和R:CACGAGTAATTCTTGCAAATGCCTATTAATGGGAGCGCTGATTCTCTTT，引物F: TGAGAAGGTTTTGGGACGCTCGAAGTTATACAGGAACTTAATAGAACAAATCAC和R: CCTAACGGGATTTTCGCTTCGTGA进行PCR扩增；

[0239] 以酿酒酵母基因组DNA为模板扩增PGK1启动子序列，以引物F:AAAGAGAATCAGCGCTCCCCATTAATAGGCATTTGCAAGAATTACTCGTG和R:CATTCTTCTGCTGTGTCGGACATGTTTTATATTTGTTGTA AAAAGTA进行PCR扩增；

[0240] 以乳酸克鲁维酵母基因组DNA为模板扩增EAP1序列，以引物F:TACTTTTTACAACAAATATAAAACAATGTCCGACACAGCAGAAGAGAATG和R:GTGATTTGTTCTATTAAGTTTCCTGTATAACTTCGAGCGTCCCAAACCTTCTCA进行PCR扩增。

[0241] 然后将四次扩增产物都稀释50倍，各吸1.5 μ L混合作模板扩增供体序列，以引物GCCAGCGTCAATACACTCCC和TTACGACAATGCCTAGTTGAGTGC进行PCR扩增。所得产物经1%的琼脂糖凝胶电泳，回收纯化即得供体DNA，命名为K1URA3-pScPGK1-K1EAP1-DD。

[0242] (4) 以乳酸克鲁维酵母基因组DNA为模板扩增同源臂序列，分别以引物F:ACACATTACTTGCCTCGAGCAT和R:TTGGGTGGAATGTTGTCGTTTTTCCTTAATGGGAGCGCTGATTCTCTTT，引物F: TGAGAAGGTTTTGGGACGCTCGAAGTTATACAGGAACTTAATAGAACAAATCAC和R: CCTAACGGGATTTTCGCTTCGTGA进行PCR扩增；

[0243] 以酿酒酵母基因组DNA为模板扩增SED1启动子序列，以引物F:AAAGAGAATCAGCGCTCCCCATTAAGGAAAAACGACAACATTCCACCCAA和R:CATTCTTCTGCTGTGTCGGACATCTTAATAGAGCGACGTATTTTATT进行PCR扩增；

[0244] 以乳酸克鲁维酵母基因组DNA为模板扩增EAP1序列，以引物F:AATAAAATACGTTTCGTCTATTAAGATGTCCGACACAGCAGAAGAGAATG和R:GTGATTTGTTCTATTAAGTTTCCTGTATAACTTCGAGCGTCCCAAACCTTCTCA进行PCR扩增。

[0245] 然后将四次扩增产物都稀释50倍，各吸1.5 μ L混合作模板扩增供体序列，以引物GCCAGCGTCAATACACTCCC和TTACGACAATGCCTAGTTGAGTGC进行PCR扩增。所得产物经1%的琼脂糖凝胶电泳，回收纯化即得供体DNA，命名为K1URA3-pScSED1-K1EAP1-DD。

[0246] (5) 以乳酸克鲁维酵母基因组DNA为模板扩增同源臂序列，引物分别为F:ACACATTACTTGCCTCGAGCAT和R:CAGACAAAATTCAATAAAGTTGCCTTTAATGGGAGCGCTGATTCTCTTT，引物F: TGAGAAGGTTTTGGGACGCTCGAAGTTATACAGGAACTTAATAGAACAAATCAC和R:

CCTAACGGGATTTTCGCTTCGTGA进行PCR扩增；

[0247] 以酿酒酵母基因组DNA为模板扩增TEF1启动子序列，以引物F:AAAGAGAATCAGCGCTCCCCATTAAGGCAACTTTATTGAATTTTGTCTG和R:CATTCTCTTCTGCTGTGTCGGACATCTTAGATTAGATTGCTATGCTTTCT进行PCR扩增；

[0248] 以乳酸克鲁维酵母基因组DNA为模板扩增EAP1序列，以引物F:AGAAAGCATAGCAATCTAATCTAAGATGTCCGACACAGCAGAAGAGAATG和R:GTGATTTGTTCTATTAAGTTTCCTGTATAACTTCGAGCGTCCCAAACCTTCTCA进行PCR扩增。

[0249] 然后将四次扩增产物都稀释50倍，各吸1.5μL混合作模板扩增供体序列，以引物GCCAGCGTCAATACACTCCC和TTACGACAATGCCTAGTTGAGTGC进行PCR扩增。所得产物经1%的琼脂糖凝胶电泳，回收纯化即得供体DNA，命名为K1URA3-pScTEF1-K1EAP1-DD。

[0250] (6) 以乳酸克鲁维酵母基因组DNA为模板扩增同源臂序列，引物分别为F:ACACATTACTTGCCTCGAGCAT和R:ACAGAATTTTCTGGAGCCAAATTGTTAATGGGAGCGCTGATTCTCTTT，引物F: TGAGAAGGTTTTGGGACGCTCGAAGTTATACAGGAACTTAATAGAACAATCAC和R: CCTAACGGGATTTTCGCTTCGTGA进行PCR扩增；

[0251] 以酿酒酵母基因组DNA为模板扩增TIF1启动子序列，以引物F:AAAGAGAATCAGCGCTCCCCATTAACAATTTGGCTCCAGAAAAATTCTGT和R:CATTCTCTTCTGCTGTGTCGGACATGATGAACCTTCTCTATATTACT进行PCR扩增；

[0252] 以乳酸克鲁维酵母基因组DNA为模板扩增EAP1序列，以引物F:AGTGTAATATAGAGCA AAGTTCATCATGTCCGACACAGCAGAAGAGAATG和R:GTGATTTGTTCTATTAAGTTTCCTGTATAACTTCGAGCGTCCCAAACCTTCTCA进行PCR扩增。

[0253] 然后将四次扩增产物都稀释50倍，各吸1.5μL混合作模板扩增供体序列，以引物GCCAGCGTCAATACACTCCC和TTACGACAATGCCTAGTTGAGTGC进行PCR扩增。所得产物经1%的琼脂糖凝胶电泳，回收纯化即得供体DNA，命名为K1URA3-pScTIF1-K1EAP1-DD。

[0254] (7) 以乳酸克鲁维酵母基因组DNA为模板扩增同源臂序列，引物分别为F:ACACATTACTTGCCTCGAGCAT和R:ATGGCACACTGGTACTGCTTCGACTTTAATGGGAGCGCTGATTCTCTTT，引物F: TGAGAAGGTTTTGGGACGCTCGAAGTTATACAGGAACTTAATAGAACA和R: CCTAACGGGATTTTCGCTTCGTGA进行PCR扩增；

[0255] 以乳酸克鲁维酵母基因组DNA为模板扩增ADH1启动子序列，以引物F:AAAGAGAATCAGCGCTCCCCATTAAGTTCGAAGCAGTACCAGTGTGCCAT和R:CATTCTCTTCTGCTGTGTCGGACATTTTATCTTTTTTAGTATAGAGTTT进行PCR扩增；

[0256] 以乳酸克鲁维酵母基因组DNA为模板扩增EAP1序列，以引物F:AACTCTATACTAAAA AAAGATAAAATGTCCGACACAGCAGAAGAGAATG和R:TGTTCTATTAAGTTTCCTGTATAACTTCGAGCGTCCCAAACCTTCTCA进行PCR扩增。

[0257] 然后将四次扩增产物都稀释50倍，各吸1.5μL混合作模板扩增供体序列，以引物GCCAGCGTCAATACACTCCC和TTACGACAATGCCTAGTTGAGTGC进行PCR扩增。所得产物经1%的琼脂糖凝胶电泳，回收纯化即得供体DNA，命名为K1URA3-pK1ADH1-K1EAP1-DD。

[0258] (8) 以乳酸克鲁维酵母基因组DNA为模板扩增同源臂序列，引物分别为F:ACACATTACTTGCCTCGAGCAT和R:TTCTGGCAGAAATGTGGTGTATGGGTTAATGGGAGCGCTGATTCTCTTT，引物F: TGAGAAGGTTTTGGGACGCTCGAAGTTATACAGGAACTTAATAGAACA和R:

CCTAACGGGATTTTCGCTTCGTGA进行PCR扩增；

[0259] 以乳酸克鲁维酵母基因组DNA为模板扩增GAPDH1启动子序列，以引物F:AAAGAGAA TCAGCGCTCCCCATTAACCCATACACCACATTTCTGCCAGAA和R:CATTCTCTTCTGCTGTGTCGGACATTTTAT CTTTTTTTAGTATAGAGTTT进行PCR扩增；

[0260] 以乳酸克鲁维酵母基因组DNA为模板扩增EAP1序列，以引物F:ACATCAAAACAACAAA TTAACAAAATGTCCGACACAGCAGAAGAGAATG和R:TGTTCTATTAAGTTTCCTGTATAACTTCGAGCGTCCCA AACCTTCTCA进行PCR扩增。

[0261] 然后将四次扩增产物都稀释50倍，各吸1.5μL混合作模板扩增供体序列，以引物 GCCAGCGTCAATACACTCCC和TTACGACAATGCCTAGTTGAGTGC进行PCR扩增。所得产物经1%的琼脂糖凝胶电泳，回收纯化即得供体DNA，命名为K1URA3-pK1GAPDH1-K1EAP1-DD。

[0262] (9) 以乳酸克鲁维酵母基因组DNA为模板扩增同源臂序列，引物分别为F: ACACATTACTTGCCTCGAGCAT和R:TTCCGGGCCCCGCTAGGTCTTTTTTAAATGGGAGCGCTGATTCTCTTT，引物F: TGAGAAGGTTTTTGGGACGCTCGAAGTTATACAGGAAACTTAATAGAACA和R: CCTAACGGGATTTTCGCTTCGTGA进行PCR扩增；

[0263] 以乳酸克鲁维酵母基因组DNA为模板扩增HXK1启动子序列，以引物F:AAAGAGAATC AGCGCTCCCCATTAATAAAGACCTAGCGGGGCCCCGAAA和R:CATTCTCTTCTGCTGTGTCGGACATTCTTGAT ATTTATGTAATGTAATCT进行PCR扩增；

[0264] 以乳酸克鲁维酵母基因组DNA为模板扩增EAP1序列，以引物F:AGATTACATTACATAA ATATCAAGAATGTCCGACACAGCAGAAGAGAATG和R:TGTTCTATTAAGTTTCCTGTATAACTTCGAGCGTCCCA AACCTTCTCA进行PCR扩增。

[0265] 然后将四次扩增产物都稀释50倍，各吸1.5μL混合作模板扩增供体序列，以引物 GCCAGCGTCAATACACTCCC和TTACGACAATGCCTAGTTGAGTGC进行PCR扩增。所得产物经1%的琼脂糖凝胶电泳，回收纯化即得供体DNA，命名为K1URA3-pK1HXK1-K1EAP1-DD。

[0266] (10) 以乳酸克鲁维酵母基因组DNA为模板扩增同源臂序列，引物分别为F: ACACATTACTTGCCTCGAGCAT和R:GTCTACGAGTGCTGAGGGCAGGCCCTTAATGGGAGCGCTGATTCTCTTT，引物F: TGAGAAGGTTTTTGGGACGCTCGAAGTTATACAGGAAACTTAATAGAACA和R: CCTAACGGGATTTTCGCTTCGTGA进行PCR扩增；

[0267] 以乳酸克鲁维酵母基因组DNA为模板扩增PGK1启动子序列，以引物F:AAAGAGAATC AGCGCTCCCCATTAAGGGCCTGCCCTCAGCACTCGTAGAC和R:CATTCTCTTCTGCTGTGTCGGACATTTTTATT AATTCTTGATCGATTTTT进行PCR扩增；

[0268] 以乳酸克鲁维酵母基因组DNA为模板扩增EAP1序列，以引物F:AAAAATCGATCAAGAA TTAATAAAAATGTCCGACACAGCAGAAGAGAATG和R:TGTTCTATTAAGTTTCCTGTATAACTTCGAGCGTCCCA AACCTTCTCA进行PCR扩增。

[0269] 然后将四次扩增产物都稀释50倍，各吸1.5μL混合作模板扩增供体序列，以引物 GCCAGCGTCAATACACTCCC和TTACGACAATGCCTAGTTGAGTGC进行PCR扩增。所得产物经1%的琼脂糖凝胶电泳，回收纯化即得供体DNA，命名为K1URA3-pK1PGK1-K1EAP1-DD。

[0270] (11) 以乳酸克鲁维酵母基因组DNA为模板扩增同源臂序列，引物分别为F: ACACATTACTTGCCTCGAGCAT和R:TAAAGTGGTTTCATCGTGAACCGTTAATGGGAGCGCTGATTCTCTTT，引物F: TGAGAAGGTTTTTGGGACGCTCGAAGTTATACAGGAAACTTAATAGAACA和R:

CCTAACGGGATTTTCGCTTCGTGA进行PCR扩增；

[0271] 以乳酸克鲁维酵母基因组DNA为模板扩增TEF1启动子序列，以引物F:AAAGAGAATCAGCGCTCCCCATTAAACGGTTTCACGATGAAACCACTTTA和R:CATTCTCTTCTGCTGTGTCGGACATTTTAAATGTTACTTCTCTGCAGTT进行PCR扩增；

[0272] 以乳酸克鲁维酵母基因组DNA为模板扩增EAP1序列，以引物F:AACTGCAAGAGAAGTAACATTA AAAATGTCCGACACAGCAGAAGAGAATG和R:TGTTCTATTAAGTTTCCTGTATAACTTCGAGCGTCCCAAAACCTTCTCA进行PCR扩增。

[0273] 然后将四次扩增产物都稀释50倍，各吸1.5μL混合作模板扩增供体序列，以引物GCCAGCGTCAATACACTCCC和TTACGACAATGCCTAGTTGAGTGC进行PCR扩增。所得产物经1%的琼脂糖凝胶电泳，回收纯化即得供体DNA，命名为K1URA3-pK1TEF1-K1EAP1-DD。

[0274] (12) 以乳酸克鲁维酵母基因组DNA为模板扩增同源臂序列，引物分别为F:ACACATTACTTGCCTCGAGCAT和R:AGATACGTCTTCAACAATGTTGAACCTAATGGGAGCGCTGATTCTCTTT，引物F:TGAGAAGGTTTTGGGACGCTCGAAGTTATACAGGAACTTAATAGAACA和PR2:CCTAACGGGATTTTCGCTTCGTGA进行PCR扩增；

[0275] 以乳酸克鲁维酵母基因组DNA为模板扩增TIF1启动子序列，以引物F:AAAGAGAATCAGCGCTCCCCATTAAAGTTCAACATTGTTGAAGACGTATCT和R:CATTCTCTTCTGCTGTGTCGGACATCTTTACAGTTATGGATTTTCTAGTT进行PCR扩增；

[0276] 以乳酸克鲁维酵母基因组DNA为模板扩增EAP1序列，以引物F:AACTAGAAAATCCATAACTGTAAAGATGTCCGACACAGCAGAAGAGAATG和R:TGTTCTATTAAGTTTCCTGTATAACTTCGAGCGTCCCAAAACCTTCTCA进行PCR扩增。

[0277] 然后将四次扩增产物都稀释50倍，各吸1.5μL混合作模板扩增供体序列，以引物GCCAGCGTCAATACACTCCC和TTACGACAATGCCTAGTTGAGTGC进行PCR扩增。所得产物经1%的琼脂糖凝胶电泳，回收纯化即得供体DNA，命名为K1URA3-pK1TIF1-K1EAP1-DD。

[0278] 1.2.3 乳酸克鲁维酵母转化及阳性鉴定

[0279] 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25mL 2×YPD液体培养基中振荡培养过夜，取2mL菌液于50mL液体2×YPD培养基中继续振荡培养2-8h。20℃条件下3000g离心5min收集酵母细胞，加入500μL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500μL该溶液中。分装50μL至1.5mL离心管中，-80℃保存。

[0280] 从-80℃冰箱取出100μL感受态细胞，冰上融化，加入200ng gRNA&Cas9质粒(或gRNA/cas9片段)和1000ng供体DNA片段，混匀，全部转入电击杯中，冰浴2min；1.5kV，200Ω，25μF电击，立即加入700μL YPD，30℃，200rpm摇床孵育1~3h。吸取200μL涂布于固体YPD(200μg/mL G418)培养基，30℃培养2~3天至单菌落出现。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1mL YPD(200μg/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。

[0281] 1.3 改造菌株体外翻译活性测定

[0282] 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，并加入绿色荧光蛋白基因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℃的环境

中,静置孵育约2-6h。反应结束后,立即放置于Envision 2120多功能酶标仪(Perkin Elmer),读数,检测eGFP信号强弱,相对荧光单位值(Relative Fluorescence Unit,RFU)作为活性单位。

[0283] 作为对照,PC为未经改造的野生酵母菌株,将其制备成体外蛋白质合成体系,按照同样的方法进行测定蛋白翻译能力;NC是在制备的体外蛋白合成体系中不加入萤火虫荧光素酶基因DNA模板,而加入相应体积的水。

[0284] 1.4结果

[0285] 结果显示,在12个改造结构当中,有7个改造表现出比野生型酵母细胞的蛋白质合成能力有明显的、显著的提高,分别是pK1PGK1::K1EAP1、pScADH1::K1EAP1、pK1GAPDH1::K1EAP1、pK1HXK1::K1EAP1、pScPGK1::K1EAP1、pScSED1::K1EAP1、pScGAPDH1::K1EAP1。而其他5个改造的蛋白合成能力未提升或者是提高水平不显著,数值范围在12.5-15.6之间。

[0286] 其中,K1EAP1前插入启动子pK1GAPDH1的结构pK1GAPDH1::K1EAP1,表现出比野生型酵母菌株更强的体外蛋白质合成能力。其编码合成的eGFP蛋白放出的相对荧光单位值达到110.5,而野生型酵母菌株合成的eGFP蛋白的相对荧光单位值仅有12.3,提高约8.98倍,这表明对K1EAP1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。

[0287] 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。

[0288] 参考文献

[0289] 1.Sonenberg&Hinnebusch.,Regulation of translation initiation in eukaryotes:mechanisms and biological targets.Cell,2009.136(4):731-45.

[0290] 2.Dever et al.,Mechanism and Regulation of Protein Synthesis in Saccharomyces cerevisiae.Genetics,2016.203(1):65-107.

[0291] 3.M.,D.J.R.G.,Nucleic Acid.Encyclopedia of Cell Biology.Elsevier,2015.

[0292] 4.Pause et al.,Insulin-dependent stimulation of protein synthesis by phosphorylation of a regulator of 5'-cap function.Nature,1994.371:762-767.

[0293] 5.Gingras et al.,Regulation of 4E-BP1 phosphorylation:a novel two-step mechanism.Genes Dev,1999a.13:1422-1437.

[0294] 6.Gingras et al.,Hierarchical phosphorylation of the translation inhibitor 4E-BP1.Genes Dev,2001.15:2852-2864.

[0295] 7.Roux,P.P and Topisirovic,I.Regulation of mRNA translation by signaling pathways.Cold Spring Harb Perspect Biol 2012.4:a012252

[0296] 8.Adami,A.,et al.,Structure of TOR and Its Complex with KOG1.Molecular Cell,2007.27:p.509-516.

[0297] 9.Cosentino,G.P.,et al.,Eaplp,a novel eukaryotic translation initiation factor 4E-associated protein in Saccharomyces cerevisiae.2000.20:p.4604-4613.

[0298] 10.Pause et al.,Insulin-dependent stimulation of protein synthesis by phosphorylation of a regulator of 5'-cap function.Nature,1994.371:762-767.

序列表

<110> 康码(上海)生物科技有限公司

<120> 一种提高外源蛋白表达量的蛋白合成体系及其应用方法

<130> 2018

<141> 2018-09-25

<160> 8

<170> SIPOSequenceListing 1.0

<210> 1

<211> 1905

<212> DNA

<213> 乳酸克鲁维酵母 (*Kluyveromyces lactis*)

<400> 1

```

atgtccgaca cagcagaaga gaatggcctt acgaggtttc ttatgaaggt gaaggaaaat 60
ctggatacca gccagtcaga agctcaaact gagaccgaac ccattgaatt taagtacact 120
tttgattata agcctgaatt cagcagtggc aaggtggtct attcacggaa cgaattgctt 180
gctattcgcg agcaagttgc agaggaagac gtgaccaatc tagcttctga attgcctaac 240
aagaagtttt ggagattacc cgttccagga tccaacgtcg gtggtaggaa gggatccaac 300
accagaggta accatgatga taagtttggc ggaaaggaca gaggtgcca gagtggtagc 360
agaaatgcca ggaacaatag aaatagcaag cgtcaaggtg gtaaaaaggc tgggaaggag 420
agcaacgaag aatacattgc tttggaagaa cagatggaat ctacaggaa tccaatggca 480
gattttgaga attggagaaa caaatgaag gaactggagc gtcaaaaaag agggttggac 540
tctgatgcgg gcaaagattc ggattctcct gcaggattgc cagctcaaag tttcagctcc 600
atatccgatt tcttcaattt aaagccggat gacaaaaga gtgctccatt ggaagagctt 660
gagccagcgg actcgtcggg agacgtttcg aaacaacat ttgaaaaca agatagagat 720
tccaagatg ttcaacacgg cagtcaaggc caaggcataa gcaagagcaa ttcttcaga 780
ttctcctcgt tcttccaagg tgtaattct ccggatgcca gtgataaccg tcctatcgcg 840
aagcctccaa gtcaatcttc ggaagatagc agaccagctg cgggatccag gctactctct 900
ctattcaata ctgactcacc atcttcagat tcggtagttc aacaccaaca acctgaaaag 960
cccatggtaa acaatccacc tggcttact caacagtett ctaccacatc attgtcgtcg 1020
gtcgttctt cgaatcattc ccaacctcac tcaggctctc gtgctgcaa agataatgat 1080
gtcaatggca gcgtgttttt gaagagtta atgactaaag gcagtgaaa tatgatgcaa 1140
cagcccttag tctctgccc tccgggctt tccaaaate aatctcagat acctaacatc 1200
acacagcagc agcagcagca gcagcagcag cagcagcatc agcaacagca acaacagcaa 1260
caacagcaac acattcagcg cccacatcaa cagcagcagc agcaacctca acaacgcaaa 1320
cttcaacaaa ttcacaaaca acacagccag catgctcctc ctgctcagca gcagcaacat 1380
ggtaaaccac cacaaaacat tgetcaatca ggtgctccaa tgcaggcacc tccaggattc 1440
ccagttggga tgccaccaca tatgatggcc cctcctatgg gaatgcctcc acagcacttc 1500
caaggtggct atgctatgcc tcctccaccg ccacctggaa tgggacagcc acgtataatg 1560

```

agcaacggta aaggattccc cgaacacttg ggtaaccac aacagcagcc tcctagcgct 1620
 ccaaaacagc cagcaagtaa agttgggtccg caaggtcaag ttccttcaca aggccatcct 1680
 caatcccacc cacaacaacc ttatatgatg tcaggtatgc caatgaactt caatgggcaa 1740
 aatgttccaa ttccagctca aggtatccca ccaaatgcat tcccatatgg ccatccaatg 1800
 atggctctac agtttcaaca acaacaacag cagcagcagc aacaacaata cccgcagcaa 1860
 caacaacacc cgcagcaacc gcgccaaca aatccgcgcc aataa 1905

<210> 2

<211> 634

<212> PRT

<213> 乳酸克鲁维酵母 (*Kluyveromyces lactis*)

<400> 2

Met	Ser	Asp	Thr	Ala	Glu	Glu	Asn	Gly	Leu	Thr	Arg	Phe	Leu	Met	Lys	1	5	10	15
Val	Lys	Glu	Asn	Leu	Asp	Thr	Ser	Gln	Ser	Glu	Ala	Gln	Thr	Glu	Thr	20	25	30	
Glu	Pro	Ile	Glu	Phe	Lys	Tyr	Thr	Phe	Asp	Tyr	Lys	Pro	Glu	Phe	Ser	35	40	45	
Ser	Gly	Lys	Val	Val	Tyr	Ser	Arg	Asn	Glu	Leu	Leu	Ala	Ile	Arg	Glu	50	55	60	
Gln	Val	Ala	Glu	Glu	Asp	Val	Thr	Asn	Leu	Ala	Ser	Glu	Leu	Pro	Asn	65	70	75	80
Lys	Lys	Phe	Trp	Arg	Leu	Pro	Val	Pro	Gly	Ser	Asn	Val	Gly	Gly	Arg	85	90	95	
Lys	Gly	Ser	Asn	Thr	Arg	Gly	Asn	His	Asp	Asp	Lys	Phe	Gly	Gly	Lys	100	105	110	
Asp	Arg	Gly	Ala	Gln	Ser	Gly	Ser	Arg	Asn	Ala	Arg	Asn	Asn	Arg	Asn	115	120	125	
Ser	Lys	Arg	Gln	Gly	Gly	Lys	Lys	Ala	Gly	Lys	Glu	Ser	Asn	Glu	Glu	130	135	140	
Tyr	Ile	Ala	Leu	Glu	Glu	Gln	Met	Glu	Ser	Thr	Gly	Asn	Pro	Met	Ala	145	150	155	160
Asp	Phe	Glu	Asn	Trp	Arg	Asn	Lys	Met	Lys	Glu	Leu	Glu	Arg	Gln	Lys	165	170	175	
Arg	Gly	Leu	Asp	Ser	Asp	Ala	Gly	Lys	Asp	Ser	Asp	Ser	Pro	Ala	Gly	180	185	190	
Leu	Pro	Ala	Gln	Ser	Phe	Ser	Ser	Ile	Ser	Asp	Phe	Phe	Asn	Leu	Lys	195	200	205	
Pro	Asp	Asp	Gln	Lys	Ser	Ala	Pro	Leu	Glu	Glu	Leu	Glu	Pro	Ala	Asp	210	215	220	

Ser Ser Glu Asp Val Ser Lys Gln Thr Phe Glu Lys Gln Asp Arg Asp
 225 230 235 240
 Ser Gln Asp Val Gln His Gly Ser Gln Gly Gln Gly Ile Ser Lys Ser
 245 250 255
 Asn Ser Ser Arg Phe Ser Ser Phe Phe Gln Gly Gly Asn Ser Pro Asp
 260 265 270
 Ala Ser Asp Asn Arg Pro Ile Ala Lys Pro Pro Ser Gln Ser Ser Glu
 275 280 285
 Asp Ser Arg Pro Ala Ala Gly Ser Arg Leu Leu Ser Leu Phe Asn Thr
 290 295 300
 Asp Ser Pro Ser Ser Asp Ser Val Val Gln His Gln Gln Pro Glu Lys
 305 310 315 320
 Pro Met Val Asn Asn Pro Pro Gly Leu Thr Gln Gln Ser Ser Thr Thr
 325 330 335
 Ser Leu Ser Ser Val Ala Ser Ser Asn His Ser Gln Pro His Ser Gly
 340 345 350
 Pro Arg Ala Ala Lys Asp Asn Asp Val Asn Gly Ser Val Phe Leu Lys
 355 360 365
 Ser Leu Met Thr Lys Gly Ser Glu Asn Met Met Gln Gln Pro Leu Val
 370 375 380
 Ser Ala Pro Pro Gly Leu Ser Gln Asn Gln Ser Gln Ile Pro Asn Ile
 385 390 395 400
 Thr Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln His Gln Gln Gln
 405 410 415
 Gln Gln Gln Gln Gln Gln Gln His Ile Gln Arg Pro His Gln Gln Gln
 420 425 430
 Gln Gln Gln Pro Gln Gln Arg Lys Leu Gln Gln Ile His Gln Gln His
 435 440 445
 Ser Gln His Ala His Pro Ala Gln Gln Gln Gln His Gly Lys Pro Pro
 450 455 460
 Gln Asn Ile Ala Gln Ser Gly Ala Pro Met Gln Ala Pro Pro Gly Phe
 465 470 475 480
 Pro Val Gly Met Pro Pro His Met Met Ala Pro Pro Met Gly Met Pro
 485 490 495
 Pro Gln His Phe Gln Gly Gly Tyr Ala Met Pro Pro Pro Pro Pro Pro
 500 505 510
 Gly Met Gly Gln Pro Arg Ile Met Ser Asn Gly Lys Gly Phe Pro Glu
 515 520 525
 His Leu Gly Asn Pro Gln Gln Gln Pro Pro Ser Ala Pro Lys Gln Pro

<211> 1200

<212> DNA

<213> 人工序列 (artificial sequence)

<400> 4

```

cccatacacc acatttctgc cagaatttca tgattaccgc gccatgaatt ttccagcttc 60
cagttgtttt tcttctggaa atttcggctt ccggttaaaa aactaaaata actcaacatg 120
gaaagaaatt ggaatctgag gtttagtgga cgctgggatt ccactagagt cacaagctcc 180
ctagacatcc gaggacaact gacgaaactc tggcgttatt ttttccaaaa tagatttatac 240
cgtttgtaaa gtggctgctc gttcgtggtt gcaaacggaa aaaaagaacc aaaaatccat 300
ttttccgaaa cttcatagat tcttccgaaa tattcagagg tgaaaaagct cgagagagtt 360
cagacgcaca aaacatgget gaaccaggca cgaagttcca ctataccatc gaatatgatg 420
gatttgaaaag cagatggtaa agcaaagaga gtgacggggt cattcaacga gtaatgggtt 480
gagcaagtga ttgcctagag gatgaaggag gtggtacttc tgtttgtcac tagcaggata 540
gaaaaaatta tcattatctc tcagaaacgt aatagaagcg ttttcacata gagctacggg 600
tttgcataata cttctgtagt ttttgttcac acctgctagt tgttgacatc acgcaataag 660
atattttttg ataccggttg aagatggttc tctcttatg tcggtcagct gtcgtcgtt 720
taggtataat tctgtgctac gattacgtac agagtagata ttagagacca tgggtataatt 780
caattgtata ttaaaactat gttgagaaat actggaagac agtaatcagt tgattaattc 840
gagaatcccc gtgtaccagc caagtagtgt gtaagtgtaa ttgcgtgtca cttctttttt 900
ttctctactg cttagtacct ttctttcttt tctctctgta gggttgggaa tccagtattg 960
tgggctaatag aactgagtca cataatgtgg ttatgttcca atataggtac cacctttggt 1020
caagatttag ttttctaatt gaataaaat acagaggtta tttcaaccta attgagatta 1080
aggagagact tattttacta tagtatatat ttatttataa ttacttattg ttactccaat 1140
ccccaagtag attagattta atcaatcaca cacaacatc aaaacaaca attaacaaaa 1200

```

<210> 5

<211> 1000

<212> DNA

<213> 人工序列 (artificial sequence)

<400> 5

```

aaaaagacct agcggggccc ggaaaaatac ggtggtaaaa gtttttcgct atgattttcc 60
agagttttcac ttcagaagag ccagtggaaa ctcaaaaca atctctccag attttttcgg 120
aataacatca tacattcaac acccaaatag aagatacgaa agcacgagat ataacagaaa 180
aggatgtgca gaacctctga cttgcttac ctgagagca ataaccaatt catatagctg 240
ataggtagtt acgtttacta cttaactact tttccccaa ccaaatctga ttctaattcc 300
ccgctgaaa aaaaaaaaaa cacaaaagaa gagaacacag aataagatac aaccacagag 360
agtacgtgcg tacgtgctg catctcgcca ctccacttc gatttatctg cccgcataa 420
actagagatt gaaaatgtgc tcttgctga gctcctccgc gacttctcgg taaaacctg 480
cttctctat tattccgta gcgagcctgg aaaaatttca acaacaactt ttttttttca 540
gttcagcttc cacaaagttt cctcttcac aaagttcctt agaaacgaaa aacttcatat 600

```

agacggatcc catttgaatg tggaggcagg caggaattca gatgaatttg accttgaat 660
 gtaaagcgca aaccactatt tggttccgac ttgacgttg gagtggcagt ttctggatca 720
 tttgcgtata taaatatata aatatgttgt tctatcgtag agttctttcc ttttcagttg 780
 taggttttgt ccaactgtgaa ttgttatacc atcatttctt aattacagac agccccaagt 840
 tttaaagtat acaagtctca ttactactta ctactactac tactactgct gtgattaatt 900
 gctattttac tagaaatact actagttata gcatcataag aagagtatta gagctaacgc 960
 aaaagctaaa cttttagatt acattacata aatatcaaga 1000

<210> 6

<211> 1200

<212> DNA

<213> 人工序列 (artificial sequence)

<400> 6

gggcctgccc tcagcactcg tagacacgag taacgtcttg agacctctcg tacaggaag 60
 cgacatatcg ttcaatagac tatggaacaa agtgtacacc gcagcgatat ctttgcattt 120
 gcaaaacgat tgaataagt acgtcgatgc taaatcctgg ataagtacgc tggatcgtg 180
 taagcccatg agaacgacac gttcctcatc actagaagcc gaactgttgt cttcagtggg 240
 gattggttcg acattttgcc aattgctgtc gatgtacct ttcaaagcca tgtaccttaa 300
 atcttcatcc ttggcaagta gattcatcgg gtgtgtttga agtaagaata tttgcttggt 360
 tttatggtat caaaggtata tgtttagaa gacaatttcc ggtaatcaa ttgtctgtct 420
 gctcagttta gcacatgtat agtacgttg acatagtcta caatattcag cattcagcat 480
 tcagtataca gcatatggct aaatgatcac aaatgtgatt gatgattga cacgactaga 540
 aaagagaacg aaaaaggaa attccatgtc acgtgcgttg gcacgtgaca tggaatatcg 600
 aagaaagaaa aaaaaaacga tctcgtccta gtggaagccc agagtctggt cccccggag 660
 tcttcccaaa acaagaagct gacacatggt gacacagaac accccacagc aaatgcacca 720
 cgctacgtag atcaggaagc ttaactctag cgacctgtcg ctgccccac agaacctcac 780
 ccgagaacca cacattacac gccgccagct cccactatac tcactttgct tcccttaagc 840
 gttctcacga ttcgttcgct gcccttctc aagagtcttc tgattctaatt tctcattcga 900
 aatcctctac agttaatgaa ttgcttgaca tgacattcat tgtctcatgg ttttggcttt 960
 ttggcttttg tcttttaaag ctatatcaac ttacatata aatatacgtc aaaaggggat 1020
 tcattaatta gaaaattctc tttttcaata gttgtattc attatcaatc tattcaactc 1080
 aattggttat tattttcact tttttgctat cctaaacct caacaatatt taaatatatc 1140
 tgttgctaca ttaagagtta cttcagaaat acaaaaaaa tcgatcaaga attaataaaa 1200

<210> 7

<211> 1200

<212> DNA

<213> 人工序列 (artificial sequence)

<400> 7

acggtttcac gatgaaacca ctttagcgct gaagttgta agattcagaa ccagaactac 60
 attcagtgct agaagtgtat gcaggcgctc tctggtttcg ctatgtact ggtggagtgg 120

atgtgaatgg cctcagtctc gcttttagag agagtccac taagcagtcc aaagaaagct 180
 cccactggaa caggggaaag gagcctgtcc aagcaaagtc cttctcataa atggtgccaa 240
 agacccgcaa gcccaaagca attaccccc aaaaagaaat gatatagtgc aagatacgta 300
 tatgaccatg acttgactag gtgaaacagt gcagaaacag ccgcacaaaa gcagccctaa 360
 ccctcagagt cgattttact ctttcaggta ataaagcctc gacatcaatt ttagacagaa 420
 gccaggctgg cctcgagatt atagccatag gcaagcaaga ggagagaagg ggaggcccc 480
 catggggggc ctccccccc ctgtcaaggt ttggcagaac ctagcttcat taggccacta 540
 gcccagccta aaacgtcaac gggcaggagg aacactccca caagacggcg tagtattctc 600
 gattcataac cattttctca atcgaattac acagaacaca ccgtacaaac ctctctatca 660
 taactactta atagtcacac acgtactcgt ctaaatacac atcatcgtcc tacaagttca 720
 tcaaagtgtt ggacagacaa ctataccagc atggatctct tgtatcgggt cttttctccc 780
 gctctctcgc aataacaatg aacactgggt caatcatagc ctacacagggt gaacagagta 840
 gcgtttatac agggtttata cgggtgattc tacggcaaaa atttttcatt tctaaaaaaa 900
 aaaagaaaaa tttttcttc caacgctaga aggaaaagaa aatctaat aaattgattt 960
 ggtgattttc tgagagtcc ctttttcata tatcgaattt tgaatataaa aggagatcga 1020
 aaaaattttt ctattcaatc tgtttctgg ttttattga tagtttttt gtgtattatt 1080
 attatggatt agtactgggt tatatgggtt tttctgtata acttctttt attttagttt 1140
 gtttaatctt attttgagtt acattatagt tcctaactg caagagaagt aacattaaaa 1200

<210> 8

<211> 1000

<212> DNA

<213> 人工序列 (artificial sequence)

<400> 8

gttcaacatt gttgaagacg tatcttgaa agatgatgtc atcaaagttt tcgaagacaa 60
 gatggaaaga ctaccgggt cttactgcaa gatcggagat tcgatggatg gattccatac 120
 agaaaatgca gaagaccgtg atcgtgtgca aggtgttatt ggcaagcaa tgacacatat 180
 caacacctg ttttcagaca aggtatcca tgacctggt aataaaaaca ttgtgtttgt 240
 gcaagaagtc ggattagcat tgaaacgtt gcaattctg gtgagctatt ataattctgt 300
 tgatgacatt tcatgttcat ctattctcc agtggcgcaa ccaactctg ctgaaacgtc 360
 cgcaccccc ctctatctc caatatctc aaagaacggg gctggtttct tccacgttgc 420
 tgccaagagg gaaaatagag acccaattgg attcctaac atcactgggt ccacgtctcc 480
 cgttattgaa ccattgttc aatacgtcaa tgaactcac aaggaaggaa agcttcagta 540
 cgggtactct gtagtacatg gtgatacgtc ttcaacatac gctaaagagc acattcaggg 600
 tcttaacgag ctattttcga tgetacaaaa gttgttctt cccaactcat cctaaaagag 660
 atagaacaaa tctgatcctt ccataatat aaataaagag aatctgtata gtttacgata 720
 atgtgataaa gatgctgggt taagatatgt gaggcataag caacgatca gtcaccggtg 780
 cggacataaa tgcggtatcc tatgacatca cgtgatataa tcaactctgga cgagttgaaa 840
 aatttttaggt ttcagaccaa acgcctacaa taagcaatga ctttaaagaa ctgatgagat 900
 gaatagacta gattacttga ggttttaacg ttcatattgt tgttgacta gtcttttagc 960

attctgcagt gatagaacta gaaaatccat aactgtaaag 1000

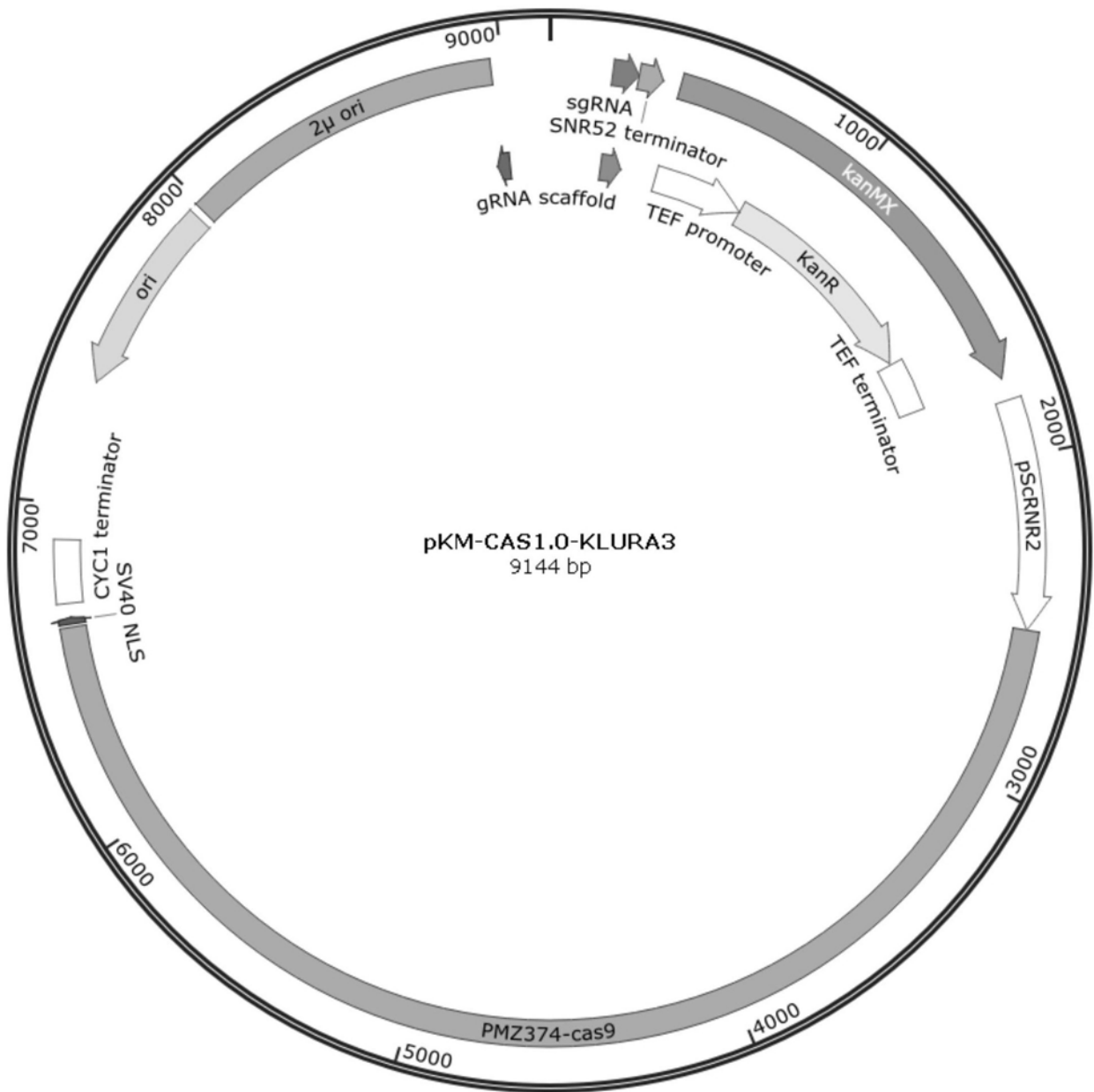


图1

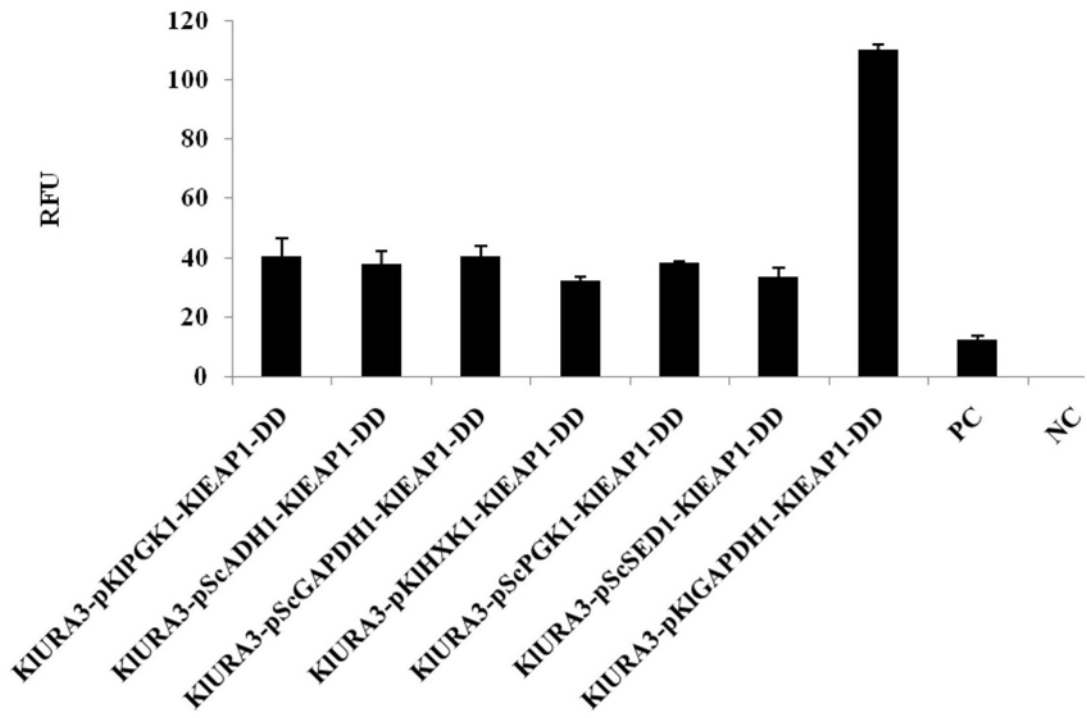


图2