

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2017 年 7 月 6 日 (06.07.2017)



(10) 国际公布号
WO 2017/113786 A1

- (51) 国际专利分类号:
C12N 7/01 (2006.01) A61K 39/145 (2006.01)
C07K 14/11 (2006.01) A61K 35/76 (2015.01)
C12N 15/44 (2006.01) A61P 31/16 (2006.01)
C12N 15/63 (2006.01) C12N 5/10 (2006.01)
- (21) 国际申请号: PCT/CN2016/092778
- (22) 国际申请日: 2016 年 8 月 2 日 (02.08.2016)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201511029463.1 2015 年 12 月 31 日 (31.12.2015) CN
- (71) 申请人: 北京大学 (PEKING UNIVERSITY)
[CN/CN]; 中国北京市北京大学药学院天然药物及
仿生药物国家重点实验室, Beijing 100191 (CN)。
- (72) 发明人: 周德敏 (ZHOU, Demin); 中国北京市北京
大学药学院天然药物及仿生药物国家重点实验室,
Beijing 100191 (CN)。 司龙龙 (SI, Longlong); 中国
北京市北京大学药学院天然药物及仿生药物国家
重点实验室, Beijing 100191 (CN)。 周雪莹 (ZHOU,
Xueying); 中国北京市北京大学药学院天然药物及
仿生药物国家重点实验室, Beijing 100191 (CN)。

张子威 (ZHANG, Ziwei); 中国北京市北京大学药学院天然药物及仿生药物国家重点实验室, Beijing 100191 (CN)。 徐欢 (XU, Huan); 中国北京市北京大学药学院天然药物及仿生药物国家重点实验室, Beijing 100191 (CN)。 田振宇 (TIAN, Zhenyu); 中国北京市北京大学药学院天然药物及仿生药物国家重点实验室, Beijing 100191 (CN)。 张传领 (ZHANG, Chuanling); 中国北京市北京大学药学院天然药物及仿生药物国家重点实验室, Beijing 100191 (CN)。 肖苏龙 (XIAO, Sulong); 中国北京市北京大学药学院天然药物及仿生药物国家重点实验室, Beijing 100191 (CN)。 夏青 (XIA, Qing); 中国北京市北京大学药学院天然药物及仿生药物国家重点实验室, Beijing 100191 (CN)。 张礼和 (ZHANG, Lihe); 中国北京市北京大学药学院天然药物及仿生药物国家重点实验室, Beijing 100191 (CN)。

(74) 代理人: 北京骥驰知识产权代理有限公司
(BEIJING GEACH INTELLECTUAL PROPERTY
LAW OFFICE); 中国北京市西城区北礼士路甲 98
号阜成大厦 B 座 525 室, Beijing 100037 (CN)。

(81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG,

[见续页]

(54) Title: MUTANT VIRUS, PREPARATION METHOD THEREFOR AND APPLICATION THEREOF

(54) 发明名称: 突变的病毒、其制备方法和应用

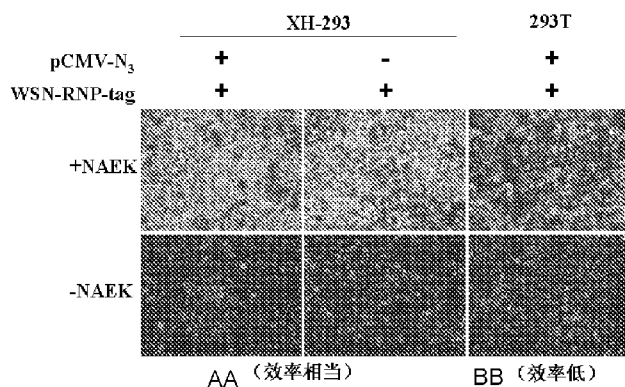


图 10

AA Equivalent efficiency
BB Low efficiency

(57) Abstract: The present invention provides a mutant virus, preferably an influenza virus, from a human or another animal, and a preparation method for the mutant virus. The method comprises of introducing a UAG codon upstream of an endogenous stop codon of one or more genes in a viral genome. The invention further provides applications of the mutant virus as an attenuated live vaccine or as a replicable and controllable safe virus model.

(57) 摘要: 提供一种人源的或其它动物的突变病毒及其制备方法, 优选流感病毒。该方法包括使用反向遗传技术在病毒基因组的一个或多个基因的自身终止密码子上游引入密码子 UAG。还提供了该突变病毒的应用, 如作为减毒活疫苗、复制可控的安全病毒模型等的用途。



BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

- (84) **指定国** (除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚

(AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

- 包括国际检索报告(条约第 21 条(3))。
- 包括说明书序列表部分(细则 5.2(a))。

突变的病毒、其制备方法和应用

技术领域

本发明属于生物制药领域，具体涉及突变的病毒，特别是在其一个或多个蛋白中包含非天然氨基酸的突变病毒，例如流感病毒。本发明还涉及制备突变病毒的方法。本发明进一步涉及突变病毒的应用，如作为病毒活疫苗、病毒强毒的安全模型等的用途。

背景技术

流感是由流感病毒所引起的一种急性呼吸道传播疾病，可季节性地感染禽类、哺乳动物和人类。流感病毒又可分为A、B、C三种类型，其中以A型（又称甲型）流感病毒的暴发最为频繁。甲型流感病毒属于正黏病毒科，其基因组共由8个独立的单链RNA片段组成，编码10种蛋白：血凝素蛋白（HA）；基质蛋白（M），包括M1和M2；神经氨酸酶（NA）；核壳蛋白（NP）；非结构蛋白（NS），包括NS1和NEP；PB1，PB2和PA三种聚合酶。这些蛋白对于流感病毒都具有重要的生物学功能。甲型流感病毒根据其主要表面抗原HA和NA的抗原性的差异分为不同亚型，目前已发现了18种亚型的HA蛋白和11种亚型的NA蛋白。A型流感病毒大规模流行可引起极高的发病率和死亡率，严重威胁着人类的健康(W.H.O. 2003; Coleman 2007)。A型流感病毒在二十世纪主要引起了三次大型流感，即1918年的H1N1，1957年的H2N2以及1968年的H3N2，共造成约5000万人死亡(Kilbourne 2006; Taubenberger, Hultin et al. 2007)。2009年甲型流感也是由H1N1流感病毒引起(Dawood, Jain et al. 2009; Zimmer and Burke 2009)，其传播之迅速，引起了世界的关注。据统计，全世界平均每年有30-50万人死于流感 (Fiore, Shay et al. 2007)。

自流感发现以来，科学家们一直致力于对于流感病毒的防控，接种疫苗目前是最有利于预防流感，控制流感传播的手段。20世纪30年代末流感病毒疫苗开始在人类中使用，目前的流感病毒疫苗分为以下几种：灭活病毒疫苗和减毒病毒疫苗，DNA疫苗，亚单位疫苗，重组病毒载体疫苗，类病毒颗粒疫苗。

目前应用的灭活流感疫苗为三价疫苗，包括H1N1、H3N2和B型流感病毒疫苗，多年的临床应用表明流感灭活疫苗具有很好的免疫效果和安全性，接种后可刺激机体产生相应的抗体，但缺点是不能刺激产生分泌型免疫球蛋白(sIgA)。另外，流感病毒在鸡胚上传代会发生抗原变异，且大部分鸡体内携带有多种病毒，疫苗有被这些病毒污染的可能。由于新流行的抗原变异株必须要能在鸡胚中高效复制，才有可能生产大量疫苗。随着制备疫苗所用的野生型病毒在鸡蛋中生长，其免疫原性在一定程度上被改变或降低。如果生产的疫苗株与当前流行株不匹配，就会失去免疫保护效果。近年来，一个重大进展就是利用哺乳动物细胞代替鸡胚培养，哺乳动物细胞主要是MDCK细胞和Vero细胞，具有无外源因子污染、易于规模化生产、抗原稳定等优点，哺乳动物细胞培养的流感疫苗具有较好的免疫原性，接种后不良反应轻，较安全。研究阶段取得了很好的实验结果。但还有一些问题未能解决，临床尚未见到使用。

减毒活疫苗主要包括温度敏感型疫苗、重配疫苗、冷适应减毒流感活疫苗、反向遗传技

术疫苗和复制缺陷流感疫苗5种类型。现在研制成功的是冷适应减毒流感活疫苗。该疫苗是一种降低了毒力并能在最佳适应温度下生长的流感病毒株。这种病毒只能在25℃左右复制，不能在37℃下传代，因此其感染只局限于上呼吸道，临床上无明显的流感症状。利用弱毒株与当前的流行毒株进行基因重组，得到带有弱毒株和流行株的HA和NA基因的重组病毒。多项研究证实，冷适应毒株生物学性质有很好的稳定性。弱毒流感疫苗与灭活疫苗的免疫后抗体阳转率都在50%~70%，能有效控制流感的流行。冷适应减毒活疫苗已在俄罗斯使用，美国也将被批准使用，其在接种途径、免疫效果方面比灭活疫苗有一定的优势，如可通过鼻内喷雾或者点滴方式来免疫。冷适应减毒活疫苗在上呼吸道复制，可诱导黏膜的sIgA和全身的体液及细胞免疫反应，产生比灭活疫苗更广泛更持久的保护。但冷适应减毒活疫苗可与其他流感病毒发生基因重配，产生有毒的重配株病毒，并且在二价或三价冷适应减毒活疫苗中可能会出现干扰现象。

理想的疫苗应具备下列条件：①免疫原性强，②毒性反应小，③遗传学上稳定，④在流行期能快速获得与流行株一致的抗原性。鉴于活疫苗相对于灭活疫苗的优势及活疫苗的安全考虑，使用哺乳动物细胞生产复制可控、遗传稳定、安全有效的新型流感病毒活疫苗，提高疫苗的安全性和有效性，将会促进流感疫苗快速发展。

遗传密码扩展技术

经过数年的研究，人们对原核生物核糖体的翻译机制已有较全面的理解，多种核糖体不同功能状态的晶体和电镜结构已得到解析，大多数氨 tRNA 合成酶的结构也已获得。基于这些研究成果，近年来发展起来了遗传密码扩展的技术——利用琥珀终止密码子 (TAG) 来编码多种非天然氨基酸并在生物活体内将其定点插入。到目前为止，这一技术已经将数种非天然氨基酸成功地在活细胞的目标蛋白质中进行了定点表达，为这些蛋白质赋予了新颖的物理、化学和生理性质。使用这一方法，可以将非天然氨基酸（包括亲和标记和光致异构化的氨基酸、羧基氨基酸和糖基化氨基酸）引入蛋白质中 (L. Wang 等人, (2001), *SCIENCE* 292:498-500; J.W. Chin 等, 2002, *Journal of the American Chemical Society* 124:9026-9027; J. W. Chin, & P. G. Schultz, 2002, *ChemBioChem* 11:1135-1137)。这些研究表明，有可能选择性地且常规地引入化学官能基团到蛋白质中，例如，羧基、炔基、和叠氮基团等特殊化学基团，这些基团一般能够有效且选择性地形成稳定的共价键，更加有利于蛋白质的定点特异修饰，改善蛋白质的性质。这一技术不仅可以用于蛋白质的定点修饰上，还被用在了活体生物体（如病毒、细菌等）的定点标记、定点修饰、复制的控制等方面。这种在基因组中引入了 TAG 终止密码子的生物体的增殖、蛋白表达依赖于相应的外源非天然氨基酸生物正交系统。

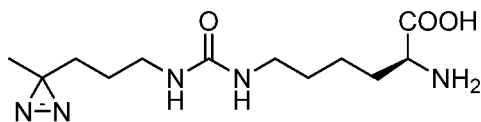
发明内容

发明人通过将琥珀终止密码子 (TAG) 引入到流感病毒的基因组中，并利用古甲烷球菌的 tRNA (tRNA^{Pyl}) 和吡咯赖氨酰-tRNA 合成酶 (tRNA^{Pyl}) 的蛋白质翻译系统使非天然氨基酸定点掺入到相应的蛋白的相应位点上，从而得到了定点突变的流感病毒。这类流感病毒的复制、蛋白表达等一系列的生命过程依赖于非天然氨基酸生物正交系统，即该类病毒可以在特别的非天然氨基酸生物正交系统中进行包装、繁殖、制备，而一旦脱离了该非天然氨基酸生物正交系统，病毒就不能进行复制、蛋白表达等生命活动。

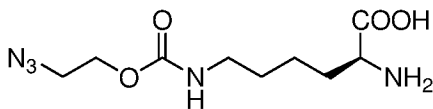
因此, 利用对非天然氨基酸依赖性制备的流感病毒, 具有很高的安全性。此外, 该定点突变的流感病毒可以进一步修饰, 进而得到定点修饰的流感病毒, 所述的定点修饰包括例如将一些免疫增强剂定点引入到流感病毒蛋白的特定氨基酸, 从而得到性能改善的流感病毒, 如得到免疫原性增强的修饰产物。

本发明最重要的突破是在流感病毒基因组中引入了琥珀终止密码子 (UAG), 使得流感病毒只有在非天然氨基酸生物正交系统中才可以复制。利用流感病毒对该非天然氨基酸生物正交系统的依赖性, 我们可以在该系统中进行流感病毒的大量制备, 由于动物和人体中并没有该生物正交系统, 因此制备出来的流感病毒在动物和人体中不能进行复制繁殖, 增加了病毒的安全性, 使得该流感病毒成了名副其实的流感病毒活疫苗。此外, 由于人工合成的非天然氨基酸可以修饰有不同功能的官能团, 可以为后续定点修饰的特异性提供保证, 例如 Click 反应仅仅能够发生在所述的非天然氨基酸上, 而不会在病毒蛋白的其他位点, 这对于定点修饰的流感病毒均质性及免疫原性的保留至关重要。

该突变系统的原理在于: 突变型的 $tRNA^{Pyl}$, $PylRS$ 满足下列关系: (1): $tRNA^{Pyl}$ 不能利用宿主细胞的赖氨酰 tRNA 酶, 只能被突变型的 $PylRS$ 酰化; (2): 突变型的 $PylRS$ 只能酰化 $tRNA^{Pyl}$, 不能酰化其它 tRNA, 因此, 突变性 $tRNA^{Pyl}$ 和 $PylRS$ 之间的关系是正交性的。这种正交性的酶能够、并且只有这种酶能够把非天然氨基酸酰化到这种正交的 tRNA 上, 并且只能酰化这种 tRNA, 而不能酰化其它的 tRNA。获得的正交赖氨酰 tRNA 合酶/tRNA 系统使非 20 种常见氨基酸的 Lys-diazirine 或 Lys-azido 等与琥珀密码子 TAG 相对应, 从而将非天然氨基酸定点引入到流感病毒的各个蛋白中。由于动物和人体中并没有该生物正交系统, 因此制备出来的流感病毒在动物和人体细胞中不能进行复制繁殖, 增加了病毒的安全性。因此, 本发明一方面涉及引入了非天然氨基酸的流感病毒, 其特征在于其中 PA、PB1、PB2、NP、NA、HA、NS 或 M 蛋白中至少一种蛋白的至少一个位点处的氨基酸被突变为非天然氨基酸; 优选地, 所述非天然氨基酸选自:



(I) 所示的 Lys-diazirine,



(II) 所示的 Lys- azido, 或

其它含有双吡丙啉、叠氮结构的非天然氨基酸中的至少 1 种。

本发明另一方面涉及一种制备引入了非天然氨基酸的流感病毒的方法, 其中包括:

(1) 构建含有编码流感病毒各个蛋白的基因的质粒, 其中编码所述蛋白的一个或多个基因在其终止密码子上游一个或多个位点处具有 TAG 突变; (2) 将 (1) 中所述质粒转染表达特异识别非天然氨基酸 Lys-diazirine 和 Lys-azido 的 tRNA 和 tRNA 合成酶的动物细胞中, 或者将 (1) 中所述质粒与表达特异识别非天然氨基酸例如 Lys-diazirine 和 Lys-azido 的 tRNA 和 tRNA 合成酶的质粒一起共转染动物细胞; 和 (3) 在添加了需要的非天然氨基酸的培养基中培养被转染的细胞, 获得在基因组中引入了 TAG 终止密码子、从而在对应的蛋白中引入了非天然氨基酸的流感病毒。所述动物细胞可以是例如哺乳动物细胞、鸟类细胞等, 如 293T 细胞。

在一个实施方案中,所述表达特异识别非天然氨基酸例如 Lys-diazirine 和 Lys-azido 的 tRNA 和 tRNA 合成酶的质粒是 pACYC-tRNA / PylRS 质粒,该质粒已保藏于中国普通微生物菌种保藏管理中心,保藏日为 2011 年 6 月 14 日。其在中国专利号为 201210214451.6 的专利中已经公开。保藏号为 CGMCC No: 4951 的分类命名为大肠埃希氏菌 (*Escherichia coli*) 的细菌细胞含有质粒 pACYC-tRNA / PylRS,该质粒可以表达特异识别非天然氨基酸 Lys-diazirine 和 Lys-azido 的 tRNA 和 tRNA 合成酶。

在另一个实施方案中,所述稳定表达 tRNA (tRNA^{Pyl}) 和吡咯赖氨酰-tRNA 合成酶 (tRNA^{Pyl}) 的动物细胞系是 HEK293-PYL,该细胞系已被保藏于中国普通微生物菌种保藏管理中心,菌种保藏地址:北京市朝阳区北辰西路 1 号院,中国科学院微生物研究所,保藏日为 2015 年 11 月 17 日,保藏号为 CGMCC No: 11592,其分类命名为人 HEK293T 细胞。用于本发明的动物细胞系还可以通过用表达特异识别非天然氨基酸例如 Lys-diazirine 和 Lys-azido 的 tRNA 和 tRNA 合成酶的质粒进行转染来制备。细胞转染的方法是本领域中熟知的。

本领域普通技术人员可通过多种方法确定流感病毒各个蛋白的编码基因中适于引入 TAG 突变的位点,例如通过生物信息学手段,分析流感病毒相应蛋白中氨基酸的保守性,再确定引入 TAG 的位点。

在本发明的一个具体的实施方案中,通过将带有琥珀密码子 TAG 的流感病毒基因插入到载体(例如 pHH21 质粒)中,将所述的质粒与 pACYC-tRNA / PylRS 一起转染宿主细胞例如 293T 细胞,即可通过在培养液中添加非天然氨基酸例如 Lys-diazirine 或 Lys-azido 获得定点突变的流感病毒。

在本发明的一个实施方案中,在流感病毒的 PA、PB1、PB2、NP 基因片段上分别引入 TAG,然后将 Lys-azido 定点修饰到流感病毒的 PA、PB1、PB2、NP 蛋白相应的位点,经中空纤维柱和凝胶层析方法或者蔗糖梯度密度离心方法,即可得到纯化后的突变型流感病毒。经体内实验初步证明,突变型的流感病毒具有非常高的安全性和遗传稳定性,并且与灭活病毒相比具有更好的免疫效果。

在本发明的一个实施方案中,为了提高流感病毒的产出效率,以及将来的工业化生产,发明人建立了可以稳定表达 tRNA (tRNA^{Pyl}) 和吡咯赖氨酰-tRNA 合成酶 (tRNA^{Pyl}) 的哺乳动物稳定细胞系 HEK293-PYL。该哺乳动物稳定细胞系还解决了传统使用鸡胚繁殖病毒易引起人体过敏等不良反应的缺点。本发明中提供了一种可在哺乳动物细胞中制备、在基因组和相应蛋白中分别引入琥珀终止密码子 UAG 和非天然氨基酸的突变流感病毒。该病毒也可以由稳定表达 tRNA (tRNA^{Pyl}) 和吡咯赖氨酰-tRNA 合成酶 (tRNA^{Pyl}) 的哺乳动物稳定细胞系(如 293T 细胞)来制备。制备的突变流感病毒具有很好的安全性和免疫活性。

国际上尚缺乏稳定表达 tRNA 的理想方法。为了提高突变型流感病毒的拯救效率,经过反复摸索,发明人建立了一套双慢病毒转导和质粒稳定转染结合的三轮筛选方法,建立了稳定表达正交 tRNA/氨酰 tRNA 合成酶的特殊细胞系,即可以稳定表达 tRNA (tRNA^{Pyl}) 和吡咯赖氨酰-tRNA 合成酶 (tRNA^{Pyl}) 的哺乳动物稳定细胞系 HEK293-PYL (图 11A)。发明人还构建了分别带有 puromycin 和 hygromycin 抗性的 2 个慢病毒过表达载体,两者分别携带氨酰 tRNA 合成酶和带有 39 位 TAG 突变的报告基因 GFP,通过两轮病毒转导 HEK-293T 细胞和 puromycin/hygromycin 筛选,得到稳定细胞株 pylRS/GFP^{39TAG}。之后,我们构建了携带 12 个拷贝数 tRNA 和 zeomycin 抗性的 bjmu-zeo-12tRNA 载体,质粒线性

化转染细胞株 pylRS/GFP^{39TAG}, 后经 UAA 存在下 zeomycin 筛选, 最后分离出 GFP 阳性细胞 (UAA 存在下细胞呈绿色, 去除 UAA 细胞呈无色), 从而得到了表达正交 tRNA/氨酰 tRNA 合成酶的稳定细胞系 HEK293-PYL。

在本发明的一个实施方案中, 利用反向遗传技术, 发明人可以将现用的流感病毒模型的任意基因替换成其他亚型或者毒株的基因, 从而制备出其他亚型或者毒株的流感病毒, 使得该方法可以应用于任意亚型或毒株的流感病毒, 而且制备出的病毒对非天然氨基酸生物正交系统具有严格的依赖性。此外, 利用反向遗传技术和基因密码子技术, 还可以制备多价流感病毒, 如含有 H1N1、H3N2、B 型流感病毒表面抗原的多价流感病毒。更重要的是, 制备出的突变型强毒和多价病毒具有非常高的安全性和有效性。

本发明提供了:

1. 一种突变的病毒, 其特征在于所述病毒的至少一种蛋白的编码核酸在终止密码子上游的一个或多个密码子处包含一个或多个 UAG 密码子。

所述“包含”是指在所述密码子处, UAG 密码子替换所述密码子、插入 UAG 密码子或同时用 UAG 密码子替换所述密码子和插入 UAG 密码子。

2. 项目 1 所述的突变的病毒, 其选自下述组中: 手足口病毒、柯萨奇病毒、丙肝病毒 HCV、乙肝病毒 HBV、甲肝病毒、丁型肝炎病毒、戊型肝炎病毒、EB 病毒、人乳头瘤病毒 HPV、单纯疱疹病毒 HSV、巨细胞病毒、水痘-带状疱疹病毒、水泡性口炎病毒、呼吸道合胞病毒 RSV、登革病毒、埃博拉病毒、寨卡病毒 Zika、SARS、中东呼吸综合征病毒、轮状病毒、狂犬病毒、麻疹病毒、腺病毒、艾滋病毒、脊髓灰质炎病毒、埃可病毒、乙型脑炎病毒、森林脑炎病毒、汉坦病毒、新型肠道病毒、风疹病毒、腮腺炎病毒、副流感病毒、蓝耳病毒、猪瘟疫病毒、口蹄疫病毒、细小病毒。

3. 项目 1 所述的突变的病毒, 其是流感病毒, 并且所述流感病毒中选自 PA、PB1、PB2、NP、NA、HA、NS 或 M 蛋白的至少一种蛋白的编码核酸在终止密码子上游的一个或多个密码子处包含一个或多个 UAG 密码子。

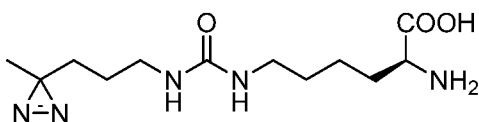
4. 项目 3 所述的突变的流感病毒, 其中选自 PA、PB1、PB2、NP、NA、HA、NS 或 M 蛋白的至少两种、三种或四种蛋白的编码核酸在终止密码子上游的一个或多个密码子处包含一个或多个 UAG 密码子。

5. 项目 3 所述的突变的流感病毒, 其中所述 PA、PB1、PB2 和 NP 蛋白中两种、三种或四种蛋白的编码核酸在终止密码子上游的一个或多个密码子处包含一个或多个 UAG 密码子。

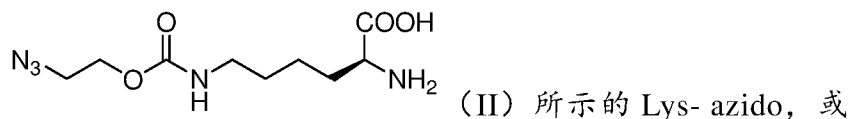
6. 项目 3 所述的突变的流感病毒, 其中所述流感病毒在选自说明书表 2a)、2b)、2c)、2d)、2e)、2f)、2g)、2h) 和 2i) 中所列举的一个或多个基因位点处包含 UAG 密码子。

7. 项目 3 所述的突变的流感病毒, 其中所述流感病毒在 PA 蛋白的 R266、PB1 蛋白的 R52、PB2 蛋白的 K33 和/或 NP 蛋白的 D101 的编码核酸密码子处含有 UAG 密码子。

8. 项目 1—7 中任一项所述的突变的病毒, 其特征在于, 所述病毒中, 所述蛋白在所述 UAG 密码子对应处包含非天然氨基酸; 优选地, 所述非天然氨基酸选自:

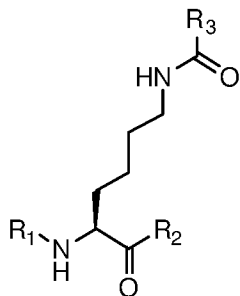


(I) 所示的 Lys-diazirine,

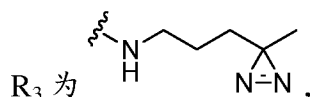


其它含有双吡丙啉、叠氮结构的非天然氨基酸中的至少 1 种。

9. 如项目 8 所述的突变的病毒, 其中所述的非天然氨基酸是位于第 n 位的 Lys-diazirine, 其在病毒蛋白中的连接方式如下式所示:

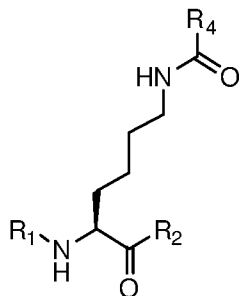


其中, 由 R_1 到 R_2 的方向为氨基酸序列的 N 末端到 C 末端方向, 第 n 位可以是流感病毒的 PA、PB1、PB2、NP、NA、HA、NS 或 M 蛋白中任何一种蛋白质氨基酸序列内的任意位置, 相应地, R_1 为第 1 至第 $n-1$ 位氨基酸残基, R_2 为第 $n+1$ 位至 C 末端的氨基酸残基,

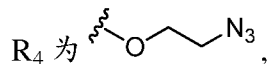


优选地, n 是该蛋白起始密码子下游第 $p-1$ 位到天然终止密码子上游第 q 位内的任意位置, 其中 p 和 q 独立地选自 1、2、3、4、5、6、7、8、9、10、15、20、30、40、50、60、70、80、90、100、125、150、200、250、300、350、400、450、500、600、700 或者直到其全长氨基酸序列长度减去 1。

10. 如项目 8 所述的突变的病毒, 其中所述的非天然氨基酸是位于第 n 位的 Lys-azido, 其在病毒蛋白中的连接方式如下式所示:



其中, 由 R_1 到 R_2 的方向为氨基酸序列的 N 末端到 C 末端方向, 第 n 位可以是病毒的 PA、PB1、PB2、NP、NA、HA、NS 或 M 蛋白中任何一种蛋白质的氨基酸序列内的任意位置, 相应地, R_1 为第 1 至第 $n-1$ 位氨基酸残基, R_2 为第 $n+1$ 位至 C 末端的氨基酸残基,



优选地, n 是该蛋白起始密码子下游第 $p-1$ 位到天然终止密码子上游第 q 位内的任意位置, 其中 p 和 q 独立地选自 1、2、3、4、5、6、7、8、9、10、15、20、30、40、50、60、70、80、90、100、125、150、200、250、300、350、400、450、500、600、700 或者直到其全长氨基酸序列长度减去 1。

11. 项目 8—10 中任一项的突变的病毒，其在所述蛋白质的多个位点处包含非天然氨基酸。

12. 如项目 11 所述的突变的病毒，其是流感病毒，并且在其中 PA 上的 R266、PB1 上的 R52、PB2 上的 K33、NP 上的 D101 处包含非天然氨基酸。

13. 如项目 1—12 中任一项所述的病毒，其中所述病毒的蛋白质编码序列中的正常终止密码子如果是 UAG 的话，已被突变成 UAA。

14. 如项目 3—13 中任一项所述的突变的病毒，其是流感病毒，并且，未经突变的 PB2 的氨基酸序列与 SEQ ID NO: 2 所示的核酸编码的氨基酸序列相同，

未经突变的 PB1 的氨基酸序列与 SEQ ID NO: 3 所示的核酸编码的氨基酸序列相同，

未经突变的 PA 的氨基酸序列与 SEQ ID NO: 4 所示的核酸编码的氨基酸序列相同，

未经突变的 HA 的氨基酸序列与 SEQ ID NO: 5 所示的核酸编码的氨基酸序列相同，

未经突变的 NP 的氨基酸序列与 SEQ ID NO: 6 所示的核酸编码的氨基酸序列相同，

未经突变的 NA 的氨基酸序列与 SEQ ID NO: 7 所示的核酸编码的氨基酸序列相同，

未经突变的 M 的氨基酸序列与 SEQ ID NO: 8 所示的核酸编码的氨基酸序列相同，或

未经突变的 NS 的氨基酸序列与 SEQ ID NO: 9 所示的核酸编码的氨基酸序列相同。

15. 根据项目 14 所述的突变的病毒，其是 WSN-RNP-TAG，在 PA 蛋白的 R266、PB1 蛋白的 R52、PB2 蛋白的 K33 和 NP 蛋白的 D101 的编码密码子处含有 UAG 密码子。

16. 项目 3 所述的突变的病毒，其是人源的或其它动物源的流感病毒，优选为 A、B 或 C 型流感病毒。

17. 项目 3 所述的突变的病毒，其中包含有说明书表 2a-表 2i 和表 4 中所列举的一个或多个突变。

18. 一种编码如项目 1-17 中任一项所述的突变的病毒或其经突变的蛋白质的核酸分子，其特征在于在天然终止密码子之外，还包含人为引入的密码子 UAG。

19. 制备项目 1—17 中任一项所述突变的病毒的方法，包括如下步骤：

(1) 将包含所述一个或多个被突变基因与合适的载体可操作地连接的突变核酸表达载体和包含其它未被突变基因的一个或多个核酸表达载体转染稳定表达 tRNA (tRNA^{Pyl}) 和吡咯赖氨酰-tRNA 合成酶 (tRNA^{Pyl}) 的动物细胞系，或者与表达 tRNA (tRNA^{Pyl}) 和吡咯赖氨酰-tRNA 合成酶 (tRNA^{Pyl}) 的质粒共同转染动物细胞系其它未被突变基因的一个或多个核酸表达载体；

(2) 将被转染的细胞在含有 Lys-diazirine 或者 Lys-azido 的培养基中培养；和

(3) 回收所述突变病毒。

20. 项目 19 所述的方法，其中步骤 (1) 中所述质粒是保藏日为 2011 年 6 月 14

日、保藏号为 CGMCC No: 4951 的大肠埃希氏菌 pACYC-tRNA / PylRS 中的质粒 pACYC-tRNA / PylRS。

21. 项目 19 所述的方法，其中所述动物细胞系选自哺乳动物细胞系、禽类动物细胞系、仓鼠细胞系等，优选选自 293 细胞、293T 细胞、Vero 细胞、A549 细胞、Hela 细胞、CHO 细胞、MDCK 细胞、sf9 细胞等。

22. 项目 19 所述的方法，其中所述稳定表达 tRNA (tRNA^{Pyl}) 和吡咯赖氨酰-tRNA 合成酶 (tRNA^{Pyl}) 的动物细胞系是以保藏号 CGMCC No: 11592 于 2015 年 11 月 17 日被保藏在 CGMCC 的 HEK293-PYL。

23. 筛选减毒病毒的方法，包括步骤：

(1) 基因突变：通过基因工程方法将病毒基因组中一个或多个基因的一个或多个选定的密码子突变为 TAG 密码子，得到一个或多个被突变的基因；

(2) 表达载体构建：将 (1) 得到的一个或多个被突变的基因与合适的载体可操作地连接，得到突变基因表达载体；

(3) 将步骤 (2) 中得到的突变基因表达载体和包含其它未被突变基因的一个或多个基因表达载体转染稳定表达 tRNA (tRNA^{Pyl}) 和吡咯赖氨酰-tRNA 合成酶 (tRNA^{Pyl}) 的动物细胞系，或者与表达 tRNA (tRNA^{Pyl}) 和吡咯赖氨酰-tRNA 合成酶 (tRNA^{Pyl}) 的质粒共同转染动物细胞系；

(4) 将被转染的细胞在含有 Lys-diazirine 或者 Lys-azido 的培养基中培养，收集培养基上清液，采用包含 Lys-diazirine 或者 Lys-azido 的培养板对上清液进行病毒克隆的非天然氨基酸依赖性的检测；和

(5) 将维持了对 Lys-diazirine 或者 Lys-azido 非天然氨基酸依赖性的突变体鉴定为减毒的病毒。

24. 项目 23 的方法，其在步骤 (5) 之后还包括：

(6) 将 (5) 中定点突变成功的候选物的突变位点进行组合，再次包装病毒，使得包装出的病毒在多个基因片段上均引入了 UAG，在对应的多个蛋白上均引入了非天然氨基酸，然后再次考察包装产物对非天然氨基酸的依赖性，保留经过长期传代仍旧维持着对非天然氨基酸依赖性的组合突变体设定为最优的定点突变成功候选物。

25. 项目 24 的方法，其中在步骤 (6) 之后还进一步包括：

(7) 对 (6) 中获得的突变型病毒进行安全性，鉴定出安全的病毒。

26. 项目 23 所述的方法，其中所述病毒选自下述组中：手足口病毒、柯萨奇病毒、丙肝病毒 HCV、乙肝病毒 HBV、甲肝病毒、丁型肝炎病毒、戊型肝炎病毒、EB 病毒、人乳头瘤病毒 HPV、单纯疱疹病毒 HSV、巨细胞病毒、水痘-带状疱疹病毒、水泡性口炎病毒、呼吸道合胞病毒 RSV、登革病毒、埃博拉病毒、寨卡病毒 Zika、SARS、中东呼吸综合征病毒、轮状病毒、狂犬病毒、麻疹病毒、腺病毒、艾滋病毒、脊髓灰质炎病毒、埃可病毒、乙型肝炎病毒、森林脑炎病毒、汉坦病毒、新型肠道病毒、风疹病毒、腮腺炎病毒、副流感病毒、蓝耳病毒、猪瘟病毒、口蹄疫病毒、细小病毒。

27. 项目 23 所述的方法，其中所述一个或多个选定的密码子独立地位于编码序列第 n 位，n 是该蛋白起始密码子下游第 p-1 位到天然终止密码子上游第 q 位内的任意位置，其中 p 和 q 独立地选自 1、2、3、4、5、6、7、8、9、10、15、20、30、40、50、60、70、80、90、100、125、150、200、250、300、350、400、450、500、600、700 或者直到其全长氨基酸序列长度减去 1。

28. 项目 23 所述的方法，其是流感病毒。
29. 项目 23 所述的方法，其中步骤 (3) 中所述质粒是保藏日为 2011 年 6 月 14 日、保藏号为 CGMCC No: 4951 的大肠埃希氏菌 pACYC-tRNA / PylRS 中的质粒 pACYC-tRNA / PylRS。
30. 项目 23 所述的方法，其中所述动物细胞系选自哺乳动物细胞系、禽类动物细胞系、仓鼠细胞系等，优选选自 293 细胞、293T 细胞、Vero 细胞、A549 细胞、Hela 细胞、CHO 细胞、MDCK 细胞、sf9 细胞等。
31. 项目 23 所述的方法，其中所述稳定表达 tRNA (tRNA^{Pyl}) 和吡咯赖氨酰-tRNA 合成酶 (tRNA^{Pyl}) 的动物细胞系是以保藏号 CGMCC No: 11592 于 2015 年 11 月 17 日被保藏在 CGMCC 的 HEK293-PYL。
32. 项目 23 所述的方法，其特征在于重复突变、包装和筛选步骤。
33. 项目 23—32 中任一项所述的方法，其中所述流感病毒是 A 型、B 型或 C 型流感病毒。
34. 含有有效量的项目 1-17 中任一项所述的突变的流感病毒的组合物。
35. 含有有效量的项目 1-17 中任一项所述的突变的流感病毒的疫苗。
36. 药物组合物，其中含有有效量的项目 1-17 中任一项所述的突变的流感病毒，以及药学上可以接受的赋形剂。
37. 项目 1-17 中任一项所述的突变的流感病毒在制备减毒活疫苗或者制备预防或治疗流感病毒感染相关药物中的用途。
38. 项目 1-17 中任一项所述的突变的流感病毒在预防和治疗流感病毒感染中的用途。
39. 可稳定表达 tRNA (tRNA^{Pyl}) 和吡咯赖氨酰-tRNA 合成酶 (tRNA^{Pyl}) 的哺乳动物稳定细胞系，其为 HEK293-PYL，保藏日为 2015 年 11 月 17 日，其保藏号为 CGMCC No: 11592。

上述的组合物、药物组合物和疫苗可以在本发明制备定点突变的流感病毒的基础上采用本领域常规技术制备；它们可以用于预防或治疗流感病毒感染，包括人和动物的流感病毒感染。

附图说明：

图 1. 用生物信息学工具 Consurf 计算流感病毒各个蛋白氨基酸序列保守性的工作流程。

图 2. 流感病毒各个蛋白上进行点突变后，拯救流感病毒的效果。位点仅作为代表，不局限于所示位点。

图 2A. 流感病毒 NP 蛋白点突变后，拯救流感病毒的效果。

图 2B. 流感病毒 PA 蛋白点突变后，拯救流感病毒的效果。

图 2C. 流感病毒 PB2 蛋白点突变后，拯救流感病毒的效果。

图 2D. 流感病毒 PB1 蛋白点突变后，拯救流感病毒的效果。

图 2E. 流感病毒 HA 蛋白点突变后，拯救流感病毒的效果。

图 2F. 流感病毒 NA 蛋白点突变后，拯救流感病毒的效果。

图 2G. 流感病毒 NS 蛋白点突变后，拯救流感病毒的效果。

图 2H. 流感病毒 M1 蛋白点突变后，拯救流感病毒的效果。

图 3.在流感病毒 NP、PA、PB1、PB2 上同时引入点突变后,拯救流感病毒的效果。即选择流感病毒的 NP-S3、PA-S1、PB1-S4、PB2-S4 这四个位点,进行组合,同时突变这四个位点,使用所得的质粒,拯救流感病毒。将所得到的流感病毒命名为 WSN-RNP-TAG。

图 4.对制备出的 WSN-RNP-TAG 的遗传稳定性进行考察,说明制备出的突变型流感病毒在多次传代后,仍旧是稳定的。在实验过程中,对该病毒连续传代时间长达 4 个月,制备的病毒一直维持着对非天然氨基酸的依赖性。

图 5.通过 qRT-PCR 说明,使用我们发明的方法制备的流感病毒的产量很高,接近野生型流感病毒的产量。

图 6.通过测定野生型和突变型 WSN-RNP-TAG 的病毒的抗原红细胞凝集效价,说明突变型流感病毒和野生型病毒而的抗原红细胞凝集效价是基本一致的,进一步说明,制备出的突变型流感病毒的产量很高。

图 7.为了检测突变型病毒 WSN-RNP-TAG 的安全性和免疫原性,设计了如下的动物实验方案。其中病毒的接种方式是鼻腔接种,取血方式是眼眶取血。

图 8. 突变型病毒 WSN-RNP-TAG 的动物水平安全性考察结果(每组小鼠数量为 10 只)。(A)当接种野生型流感病毒时,所有的小鼠体重均明显下降,当接种突变型 WSN-RNP-TAG 或者灭活的流感病毒 WSN-positive control 时,所有的小鼠体重均没有明显下降。(B)当接种野生型流感病毒时,小鼠在接种后第 10 天全部死亡。当接种突变型 WSN-RNP-TAG 或者灭活的流感病毒 WSN-positive control 时,观察期间所有的小鼠均存活。说明,突变型 WSN-RNP-TAG 在动物水平具有很好的安全性。

图 9 突变型 WSN-RNP-TAG 在动物水平的有效性考察。

图 9A.按照图 7 的实验方案,取各组小鼠肺组织,通过 qRT-PCR 的方式检测肺组织中病毒的相对含量。结果如图所示,当用灭活病毒或者突变型 WSN-RNP-TAG 对小鼠免疫后,小鼠肺组织中的病毒含量下降,并呈现免疫剂量依赖性。此外,突变型 WSN-RNP-TAG 的免疫效果优于灭活病毒的免疫效果。

图 9B.通过 Elisa 实验检测免疫前后的各组小鼠体内抗 HA 抗体的含量。结果说明免疫后 14 天和免疫后 21 天小鼠体内的抗体含量均增加,且 21 天的抗体含量高于 14 天时的抗体含量。

图 9C. 按照图 7 的实验方案,对各组小鼠分别进行免疫之后,用 100LD50 的 WSN 进行攻毒。并进行了小鼠体重的考察,发现我们制备的突变型病毒 WSN-RNP-tag 可以很好地保护动物不受病毒的感染。并且,1 倍剂量的突变型 WSN-RNP-tag 对小鼠的保护效果甚至和 10 倍剂量的灭活 WSN 的保护效果相当。这充分显示了我们制备的减毒流感疫苗相对于灭活疫苗具有更好的优势。

图 9D. 按照图 7 的实验方案,对各组小鼠分别进行免疫之后,用 100LD50 的 WSN 进行攻毒。并进行了小鼠死亡率的考察,发现我们制备的突变型病毒 WSN-RNP-tag 可以很好地保护动物不受病毒的感染。并且,1 倍剂量的突变型 WSN-RNP-tag 对小鼠的保护效果甚至和 10 倍剂量的灭活 WSN 的保护效果相当。这充分显示了我们制备的减毒流感疫苗相对于灭活疫苗具有更好的优势。

图 10.比较使用普通哺乳动物细胞 293T 和稳定表达 tRNA (tRNA^{Pyl}) 和吡咯赖氨酰-tRNA 合成酶 (tRNA^{Pyl}) 的哺乳动物稳定细胞系 HEK293-PYL 拯救突变型流感病毒的效率。结果说明,使用 HEK293-PYL 细胞系拯救突变型流感病毒的效率远远高于使用普通 293T 细胞。

图 11A. 正交 tRNA/氨酰 tRNA 合成酶的稳定细胞系 HEK293-PYL 筛选方法的建立。

图 11B. 双病毒过表达体系的构建。

图 11C. bjmu-12t-zeo 载体的构建。

图 12 接种流感病毒疫苗后的小鼠体内抗 NA 和 NP 蛋白的抗体含量测定，WSN-RNP-tag 可以诱导小鼠产生 NA 特异性抗体和 NP 的特异性抗体。

图 13 WSN-RNP-tag 对小鼠肺内病毒特异性抗体 IgA 的诱导作用。相对于灭活的 WSN，WSN-RNP-tag 可以诱导高水平的 IgA。

图 14 WSN-RNP-tag 免疫后的小鼠肺内 NP 特异性的 CD8+T 细胞的含量。灭活的 WSN 几乎不能影响 NP 特异性的 CD8+T 细胞的含量，而 WSN-RNP-tag 免疫后的小鼠肺内 NP 特异性的 CD8+T 细胞的含量明显增加。

图 15 显示了发明人制备的含 UAG 的突变型流感病毒可以抑制野生型流感病毒的复制和毒力。(A) 基因组中含有 4 个 UAG 的突变型病毒抑制了野生型病毒的噬斑形成。Wild type 为野生型病毒；CAIV 是冷适应减毒流感疫苗。(B) 含有 UAG 的突变型病毒抑制了野生型病毒的噬斑形成，抑制活性随着 UAG 数量的增加而增强。(C) 在动物水平上，含有 UAG 的突变型病毒降低了野生型病毒感染造成的动物死亡率。(D) 在动物水平上，含有 UAG 的突变型病毒降低了野生型病毒感染造成的动物体重丢失。(E) 在动物水平上，含有 UAG 的突变型病毒降低了野生型病毒感染造成的动物降低了动物肺组织中病毒的滴度。★, $P < 0.05$; ★★, $P < 0.01$; ★★★, $P < 0.001$ 。

为了更好地理解本发明，发明人用实施例对具体试验进行阐述和说明，其中所述实施例仅用于说明，并不限定本发明的保护范围。任何与本发明等价的变体或者实施方案都包括在本发明中。

实施例 1：包含定点突变的流感病毒 WSN 的基因载体的构建

(1)辅助质粒的获得

从保藏号为 CGMCC No: 4951 的大肠埃希氏菌 pACYC-tRNA / PylRS 的（中国普通微生物菌种保藏管理中心，北京市朝阳区北辰西路 1 号院，中国科学院微生物研究所，保藏日为 2011 年 6 月 14 日，分类命名为大肠埃希氏菌（*Escherichia coli*），其中含有质粒 pACYC-tRNA / PylRS）中获取质粒 pACYC-tRNA / PylRS（以下简称该质粒为辅助质粒），该质粒可以表达特异识别非天然氨基酸 Lys-diazirine 和 Lys-azido 的 tRNA 和 tRNA 合成酶。

(2)拯救野生型流感病毒 WSN 的质粒的获得

根据 pubmed 公布的流感病毒 A/WSN/1933 的基因序列，经全基因合成，获得该流感病毒各个基因片段的基因。流感病毒各基因序列分别如 SEQ ID NO: 2-9 所示。然后将其分别连接在 pHH21、pCDNA 3（neo）、pcAAGGS/MCS 载体上，获得拯救野生型流感病毒 WSN 的质粒。获得的质粒的命名及构成如表 1 所示。

简称	质粒名称	关键基 因	酶切位点	构建好的质粒结构	关键基因对应的 序列号
Ben1	pHH21	PB2	BsmBI	pPolI-WSN-PB2	SEQ ID NO: 2

Ben2	pHH21	PB1	BsmBI	pPolI-WSN-PB1	SEQ ID NO: 3
Ben3	pHH21	PA	BsmBI	pPolI-WSN-PA	SEQ ID NO: 4
Ben4	pHH21	HA	BsmBI	pPolI-WSN-HA	SEQ ID NO: 5
Ben5	pHH21	NP	BsmBI	pPolI-WSN-NP	SEQ ID NO: 6
Ben6	pHH21	NA	BsmBI	pPolI-WSN-NA	SEQ ID NO: 7
Ben7	pHH21	M	BsmBI	pPolI-WSN-M	SEQ ID NO: 8
Ben8	pHH21	NS	BsmBI	pPolI-WSN-NS	SEQ ID NO: 9
Ben9	pcDNA3 (neo)	PB2	EcoRI	pcDNA3(neo)-PB2	SEQ ID NO: 2
Ben10	pcDNA3 (neo)	PB1	EcoRI	pcDNA3(neo)-PB1	SEQ ID NO: 3
Ben11	pcDNA3 (neo)	PA	EcoRI	pcDNA3(neo)-PA	SEQ ID NO: 4
Ben13	pcAGGS/MCS	NP	EcoRI	pcAGGS/MCS-NP	SEQ ID NO: 6

(3) 定点突变位点的选择

通过生物信息学工具 Consurf 对流感病毒各个蛋白的氨基酸的保守性进行分析,并根据已被解析的流感病毒蛋白的晶体结构 (NP-PDB:2IQH; PA-PDB: 4IUJ; PB1-PDB: 3A1G、2ZNL; PB2-PDB: 4ENF; NA-PDB: 3TI6; HA-PDB: 1RVT; M-PDB: 4PUS、2RLF、3BKD; NS-PDB: 3L4Q) 选择保守、相对保守、相对不保守、不保守的氨基酸位点进行突变。在每个蛋白上选择的突变位点分别如表 2a)-表 2i) 所示。

表 2a) NP 蛋白上被选择的位点及保守性

名称	氨基酸位点	保守性(分值范围 1-9, 数值越大, 越保守)	名称	氨基酸位点	保守性(分值范围 1-9, 数值越大, 越保守)
NP-S1	K48	3	NP-S12	K90	8
NP-S2	L49	7	NP-S13	K113p	9
NP-S3	D101	1	NP-S14	A123	8
NP-S4	G102	1	NP-S15	R150	8
NP-S5	G126	6	NP-S16	T151	8
NP-S6	D127	1	NP-S17	M163	8
NP-S7	D128	2	NP-S18	L166	8
NP-S8	Y148	7	NP-S19	G169	9
NP-S9	S28	9	NP-S20	S170	9
NP-S10	M32	9	NP-S21	L172	8
NP-S11	K87	8	NP-S22	R175	8

表 2b) PB1 蛋白上被选择的位点及保守性:

名称	氨基酸位点	保守性(分值范围 1-9, 数值越大, 越保守)	名称	氨基酸位点	保守性(分值范围 1-9, 数值越大, 越保守)
PB1-S1	K11	4	PB1-S12	I241	9
PB1-S2	A14	1	PB1-S13	A242	9
PB1-S3	Q15	7	PB1-S14	I262	9

PB1-S4	R52	1	PB1-S15	Y30	7
PB1-S5	T105	1	PB1-S16	R126	9
PB1-S6	D685	7	PB1-S17	M227	9
PB1-S7	Y705	3	PB1-S18	K229	9
PB1-S8	K736	1	PB1-S19	D230	9
PB1-S9	I18	9	PB1-S20	G65	9
PB1-S10	S31	9	PB1-S21	P651	9
PB1-S11	A140	9	PB1-S22	S375	8

表 2c) PA 蛋白上被选择的位点及保守性:

名称	氨基酸位点	保守性(分值范围 1-9, 数值越大, 越保守)	名称	氨基酸位点	保守性(分值范围 1-9, 数值越大, 越保守)
PA-S1	R266	9	PA-S5	K281	7
PA-S2	L270	8	PA-S6	K289	6
PA-S3	D272	4	PA-S7	K318	5
PA-S4	P274	8	PA-S8	K328	5

表 2d) PB2 蛋白上被选择的位点及保守性:

名称	氨基酸位点	保守性(分值范围 1-9, 数值越大, 越保守)	名称	氨基酸位点	保守性(分值范围 1-9, 数值越大, 越保守)
PB2-S1	Q13	7	PB2-S5	T35	8
PB2-S2	T23	7	PB2-S6	S320	3
PB2-S3	T24	8	PB2-S7	D417	5
PB2-S4	K33	9	PB2-S8	A424	7

表 2e) HA 蛋白上被选择的位点及保守性:

名称	氨基酸位点	保守性(分值范围 1-9, 数值越大, 越保守)	名称	氨基酸位点	保守性(分值范围 1-9, 数值越大, 越保守)
HA-S1	F12	1	HA-S12	T35	9
HA-S2	T19	4	HA-S13	V43	9
HA-S3	A26	8	HA-S14	F160	8
HA-S4	V33	9	HA-S15	G300	8
HA-S5	K39	3	HA-S16	G317	9
HA-S6	N48	8	HA-S17	C319	9
HA-S7	K57	2	HA-S18	P320	9
HA-S8	K60	1	HA-S19	L328	9
HA-S9	D18	8	HA-S20	G333	9
HA-S10	C21	9	HA-S21	N336	9
HA-S11	G23	8	HA-S22	P338	9

表 2f) NA 蛋白上被选择的位点及保守性:

名称	氨基酸位点	保守性(分值范围 1-9, 数值越大, 越保守)	名称	氨基酸位点	保守性(分值范围 1-9, 数值越大, 越保守)
NA-S1	N2	9	NA-S12	A186	9
NA-S2	P3	8	NA-S13	Y192	9
NA-S3	N4	9	NA-S14	S202	9
NA-S4	Q5	8	NA-S15	I7	8
NA-S5	K6	6	NA-S16	I8	8
NA-S6	V16	1	NA-S17	G11	8
NA-S7	N28	7	NA-S18	C14	7
NA-S8	I29	5	NA-S19	C76	8
NA-S9	L118	9	NA-S20	K86	9
NA-S10	L142	9	NA-S21	K244	5
NA-S11	S144	9	NA-S22	K331	9

表 2g) NS1 蛋白上被选择的位点及保守性:

名称	氨基酸位点	保守性(分值范围 1-9, 数值越大, 越保守)	名称	氨基酸位点	保守性(分值范围 1-9, 数值越大, 越保守)
NS-S1	S8	9	NS-S12	K41	8
NS-S2	Q10	9	NS-S13	K110	9
NS-S3	S83	9	NS-S14	K131	9
NS-S4	A86	1	NS-S15	M1	1
NS-S5	H101	4	NS-S16	D2	8
NS-S6	F103	1	NS-S17	P3	6
NS-S7	A122	9	NS-S18	N4	7
NS-S8	K126	7	NS-S19	V6	3
NS-S9	T5	9	NS-S20	S7	7
NS-S10	F9	9	NS-S21	L43	9
NS-S11	R37	9	NS-S22	A132	9

表 2h) M1 蛋白上被选择的位点及保守性:

名称	氨基酸位点	保守性(分值范围 1-9, 数值越大, 越保守)	名称	氨基酸位点	保守性(分值范围 1-9, 数值越大, 越保守)
M-S1	P16	7	M-S8	L74	1
M-S2	S17	9	M-S9	S13	9
M-S3	G18	7	M-S10	V31	9
M-S4	K35	2	M-S11	S53	9
M-S5	T37	1	M-S12	A96	9
M-S6	I51	9	M-S13	V97	9
M-S7	R72	3	M-S14	I131	9

表 2 i) M2 蛋白上被选择的位点及保守性:

名称	氨基酸位点	保守性(分值范围 1-9, 数值越大, 越保守)	名称	氨基酸位点	保守性(分值范围 1-9, 数值越大, 越保守)
M2-S1	S22	9	M2-S5	W41	7
M2-S2	S23	9	M2-S6	K49	9
M2-S3	D24	9	M2-S7	K56	9
M2-S4	H37	9	M2-S8	K60	9

(4) 定点突变的引物设计以及突变载体构建

以 Ben1 pPolI-WSN-PB2; Ben2 pPolI-WSN-PB1; Ben3 pPolI-WSN-PA; Ben4 pPolI-WSN-HA; Ben5 pPolI-WSN-NP; Ben6 pPolI-WSN-NA; Ben7 pPolI-WSN-M; Ben8 pPolI-WSN-NS 作为模板质粒, 利用定点突变试剂盒 (QuikChange® Lightning Site-Directed Mutagenesis Kits, Catalog #210518), 根据常规的引物设计原则设计引物, 按说明书操作表 2a) -表 2i) 中列出的氨基酸位点对应的编码核苷酸突变为琥珀终止密码子 TAG, 经测序验证突变成功。

一些蛋白上被选择位点的突变采用如表 3 所示的序列来进行。

表 3

被突变蛋白	引物名称	引物序列 (5' 至 3')
NP	NP-S3 g326t_a327a_t328g_	For 5'-ggacctatatacaggagagtataggaaagtggaggagaga-3'
		Rev 5'-tctctctccactttccctatactctctctgtatataggtcc-3'
PA	PA-S1 c796t_g797a_a798g_	5'-gcccatccggaagtctaagtggctatgtgttgattcaaaaagggtca-3'
		5'-tgaacctttttgaaatcaacaccatagccacttagacttccggatgggc-3'
PB1	PB1-S4 a154t_g155a_g156g_	5'-gtttgtgtccatcttcctattctgagtactgatgtgtcct-3'
		5'-aggacacatcagtactcagaatagggaagatggacaacaac-3'
PB2	PB2-S4 a97t_a98a_g99g_	5'-ctgtctctctgatgtgtactacttgattatggccatag-3'
		5'-catatggccataatcaagtagtacacatcaggaagacag-3'

其他突变可以利用被突变位点上下游的核苷酸序列、根据常规知识设计的突变引物来引入。

(5) 可稳定表达 tRNA (tRNA^{Pyl}) 和吡咯赖氨酰-tRNA 合成酶 (tRNA^{Pyl}) 的哺乳动物稳定细胞系 HEK293-PYL 的建立。

构建了分别带有 puromycin 和 hygromycin 抗性的 2 个慢病毒过表达载体, 两者分别携带氨酰 tRNA 合成酶和在第 39 位编码密码子处携带 TAG 突变的报告基因 GFP, 通过两轮病毒转导 HEK-293T 细胞和 puromycin/hygromycin 筛选, 得到稳定细胞株 pylRS/GFP^{39TAG}。之后, 构建了携带 12 个拷贝数 tRNA 和 zeomycin 抗性的 bjmu-zeo-12tRNA 载体, 质粒线性化转染细胞株 pylRS/GFP^{39TAG}, 后经非天然氨基酸 (UAA) 存在下 zeomycin 筛选, 最后分离出 GFP 阳性细胞 (UAA 存在下细胞呈绿色, 去除 UAA 细胞呈无色), 从而得到了表达正交 tRNA/氨酰 tRNA 合成酶的稳定细胞系 HEK293-PYL (图 11A)。

a. 载体的构建

从 psd31 载体出发, 先通过 BamHI/xbaI 酶切位点将 sv40-puro^R 基因分别替换成

IRES-puro^R和 IRES-hygro^R基因,这样就得到了2个不同抗性的病毒载体 psd31-IRES-puro^R和 psd31-IRES-hygro^R。其中, IRES 为内部核糖体进入序列 (internal ribosome entry site), IRES 序列常用于多顺反子基因表达。例如,在目的基因之后插入 IRES 序列,后面是选择标记基因,这样转录出来的 mRNA 就可以同时表达两种蛋白。利用 IRES 系统过表达目的基因有2个优势:1.目的基因与标记基因共用一个启动子,避免了假阳性的出现;2. IRES 翻译效率低于传统翻译起始位点,使得目的基因表达量高于标记基因。所以,我们在 IRES 位点的前面通过 BamHI 酶切位点分别引入 CMV-pyIRS 序列和 CMV-GFP^{39TAG} 序列,就得到了能同时过表达两个目的蛋白的双病毒体系 psd31-CMV-pyIRS-IRES-puro^R/psd31-CMV-GFP^{39TAG}-IRES-hygro^R。

发明人用质粒稳定转染的方法过表达 tRNA。为了保证 tRNA 的表达量,发明人构建得到了载体 bjmu-12t-zeo, 其序列如 SEQ ID NO: 1 所示。(图 11C)。

b. 慢病毒的包装和转导

先包装 psd31-CMV-pyIRS-IRES-puro^R 病毒,转导 HEK293T 细胞, puromycin 筛选浓度为 0.6ug/ml, 得到稳定细胞系 1 号后, 再加入 psd31-CMV-GFP^{39TAG}-IRES-hygro^R 病毒, hygromycin 筛选浓度为 200ug/ml, 得到稳定细胞系 2 号。

c. 质粒的稳定转染

发明人通过质粒稳定转染,进行了第三轮筛选,最后得到了稳定表达正交 tRNA/氨酰 tRNA 合成酶的特殊细胞系,步骤如下:

A. 将 bjmu-12t-zeo 载体酶切线性化后,转染表达 pyIRS 和 GFP^{39TAG} 蛋白的稳定细胞系 2 号 (10cm 培养皿, 每皿 10ug 质粒, 转染时不能有抗生素的存在)。

B. 转染 6 小时后换液, 加入非天然氨基酸。

C. 转染 48 小时后, 观察绿色荧光, 换液, 加入 400ug/ml 的 zeomycin。

D. 每 3 天换液, 直到 blank 组全部死亡, 转染组形成克隆。

E. 分离纯化 GFP 阳性克隆, 继续用剂量减半的 zeomycin 扩大培养, 得到 12t-zeo 稳定细胞系 HEK293-PYL。

(6) 定点突变后的流感病毒的拯救

按照 Neumann, G.; et al. *Proc Natl Acad Sci U S A* **1999**, 96, 9345-50 中公开的正常的拯救流感病毒方法, 将拯救流感病毒所用的 12 个质粒共转染稳定细胞系, 只是用定点突变的质粒替换这 12 个质粒中相应的质粒。对应于六孔板的每个孔, 每种质粒加 0.1ug。转染后, 观察细胞的病变情况, 筛选出可以拯救出病毒并且对非天然氨基酸具有依赖性的突变位点。筛选出的位点根据蛋白及突变的位点进行命名。示例说明, 将 Ben3 pPolI-WSN-PA 质粒上进行突变, 突变成功后, 将该质粒与拯救流感病毒的其他质粒 Ben1 pPolI-WSN-PB2; Ben2 pPolI-WSN-PB1; Ben4 pPolI-WSN-HA; Ben5 pPolI-WSN-NP; Ben6 pPolI-WSN-NA; Ben7 pPolI-WSN-M; Ben8 pPolI-WSN-NS; Ben9 pcDNA 3 (neo)-PB2; Ben10 pcDNA 3 (neo)-PB1; Ben11 pcDNA 3 (neo)-PA; Ben13 pcAGGS/MCS-NP 共同转染 (5) 建立的稳定细胞系, 从而拯救出在流感病毒 PA 基因片段的 S1 位点上引入密码子 TAG 的突变型流感病毒, 命名为 PA-S1, 命名对应于表 2a)-表 2i) 中示出的突变位点。

依照同样的方法, 可获取其他位点引入密码子 TAG 的突变型流感病毒, 并根据突变

位点进行命名。

表 4 突变型病毒的包装效率和逃逸频率：

突变病毒 名称	蛋白	TAG 位点	相对包装效率 (%)	逃逸频率 (第一代)	逃逸频率 (第二十代)
NP-D101	NP	D101	~80%	$2.00\text{E-}09 \pm 1.20\text{E-}09$	$8.00\text{E-}09 \pm 7.10\text{E-}09$
NP-G102	NP	G102	~50%	$8.90\text{E-}08 \pm 5.89\text{E-}09$	$4.10\text{E-}07 \pm 9.40\text{E-}08$
NP-D128	NP	D128	~33%	$4.16\text{E-}07 \pm 2.13\text{E-}07$	$5.90\text{E-}07 \pm 7.26\text{E-}08$
NP-G126	NP	G126	~40%	$3.21\text{E-}08 \pm 6.50\text{E-}09$	$1.10\text{E-}07 \pm 5.20\text{E-}08$
NP-R150	NP	R150	~31%	未测	未测
NP-M163	NP	M163	67%	$7.00\text{E-}10 \pm 1.02\text{E-}10$	$2.00\text{E-}9 \pm 8.90\text{E-}10$
NP-G169	NP	G169	~25%	未测	未测
PB1-R52	PB1	R52	~67%	$7.10\text{E-}07 \pm 1.10\text{E-}07$	$1.21\text{E-}06 \pm 4.42\text{E-}07$
PB1-T105	PB1	T105	~57%	$6.24\text{E-}06 \pm 3.12\text{E-}06$	$7.35\text{E-}05 \pm 3.62\text{E-}05$
PB1-K736	PB1	K736	~29%	未测	未测
PB1-K11	PB1	K11	~25%	未测	未测
PB1-Y30	PB1	Y30	~25%	未测	未测
PB1-G65	PB1	G65	~31%	未测	未测
PB1-R126	PB1	R126	~25%	未测	未测
PB1-M227	PB1	M227	~31%	未测	未测
PB1-K229	PB1	K229	~25%	未测	未测
PB1-D230	PB1	D230	~25%	未测	未测
PB1-S375	PB1	S375	~57%	$3.20\text{E-}07 \pm 1.06\text{E-}07$	$5.10\text{E-}07 \pm 4.25\text{E-}07$
PB2-S320	PB2	S320	~36%	未测	未测
PB2-Q13	PB2	Q13	~67%	$5.80\text{E-}04 \pm 2.20\text{E-}05$	$8.90\text{E-}01 \pm 8.90\text{E-}02$
PB2-T24	PB2	T24	~50%	$6.00\text{E-}06 \pm 6.12\text{E-}07$	$1.30\text{E-}05 \pm 1.10\text{E-}06$
PB2-K33	PB2	K33	~67%	$3.00\text{E-}09 \pm 1.12\text{E-}09$	$7.00\text{E-}09 \pm 3.28\text{E-}09$
PB2-T35	PB2	T35	~67%	$3.50\text{E-}04 \pm 3.08\text{E-}04$	$9.20\text{E-}01 \pm 4.25\text{E-}02$
PA-D272	PA	D272	~80%	$9.30\text{E-}06 \pm 1.00\text{E-}06$	$5.91\text{E-}04 \pm 2.25\text{E-}04$
PA-K289	PA	K289	~25%	未测	未测
PA-K318	PA	K318	~33%	未测	未测
PA-K328	PA	K328	~67%	$1.12\text{E-}06 \pm 2.45\text{E-}07$	$2.10\text{E-}05 \pm 6.20\text{E-}06$
PA-R266	PA	R266	~80%	$1.00\text{E-}08 \pm 1.10\text{E-}08$	$6.80\text{E-}08 \pm 1.80\text{E-}08$
PA-L270	PA	L270	~80%	$6.70\text{E-}07 \pm 1.09\text{E-}07$	$3.90\text{E-}06 \pm 5.13\text{E-}07$

突变病毒 名称	蛋白	TAG 位点	相对包装效率 (%)	逃逸频率 (第一代)	逃逸频率 (第二十代)
NA-V16	NA	V16	~31%	未测	未测
NA-K6	NA	K6	~24%	未测	未测
NA-I29	NA	I29	~40%	$3.10\text{E-}06 \pm 5.34\text{E-}07$	$6.10\text{E-}06 \pm 1.11\text{E-}06$
NA-K244	NA	K244	~36%	未测	未测
NA-N2	NA	N2	~33%	未测	未测
NA-I7	NA	I7	~25%	未测	未测
NA-I8	NA	I8	~33%	未测	未测
NA-G11	NA	G11	~36%	未测	未测
NA-N28	NA	N28	~50%	$9.30\text{E-}05 \pm 8.56\text{E-}06$	$1.21\text{E-}03 \pm 2.12\text{E-}04$
NA-C76	NA	C76	~25%	未测	未测
HA-K57	HA	K57	~57%	$6.20\text{E-}04 \pm 1.30\text{E-}05$	$1.80\text{E-}01 \pm 1.50\text{E-}02$
HA-G317	HA	G317	~25%	未测	未测
HA-C319	HA	C319	~25%	未测	未测
HA-G333	HA	G333	~25%	未测	未测
HA-N336	HA	N336	~25%	未测	未测
NS-M1	NS	M1	~25%	未测	未测
NS-V6	NS	V6	~25%	未测	未测
NS-A86	NS	A86	~100%	$3.00\text{E-}06 \pm 6.15\text{E-}07$	$2.30\text{E-}04 \pm 8.29\text{E-}05$
NS-F103	NS	F103	~67%	$8.90\text{E-}03 \pm 1.12\text{E-}03$	$9.10\text{E-}01 \pm 3.30\text{E-}02$
NS-H101	NS	H101	~67%	$4.50\text{E-}06 \pm 7.30\text{E-}07$	$6.10\text{E-}06 \pm 1.08\text{E-}06$
NS-D2	NS	D2	~25%	未测	未测
NS-N4	NS	N4	~25%	未测	未测
NS-S7	NS	S7	~25%	未测	未测
NS-S8	NS	S8	~33%	$8.70\text{E-}07 \pm 4.50\text{E-}07$	$5.10\text{E-}05 \pm 8.55\text{E-}06$
NS-R37	NS	R37	~67%	$6.10\text{E-}07 \pm 7.31\text{E-}08$	$1.20\text{E-}06 \pm 1.17\text{E-}06$
NS-K41	NS	K41	~67%	$1.20\text{E-}06 \pm 8.22\text{E-}07$	$7.20\text{E-}05 \pm 3.96\text{E-}05$
NS-L43	NS	L43	~67%	$6.50\text{E-}07 \pm 3.27\text{E-}07$	$2.13\text{E-}06 \pm 7.24\text{E-}07$
NS-S83	NS	S83	~100%	$8.90\text{E-}06 \pm 6.60\text{E-}07$	$6.89\text{E-}03 \pm 1.27\text{E-}03$

NS-K110	NS	K110	~67%	$2.10\text{E-}06 \pm 2.13\text{E-}06$	$5.98\text{E-}05 \pm 1.02\text{E-}05$
NS-A122	NS	A122	~44%	$2.10\text{E-}07 \pm 1.13\text{E-}08$	$9.20\text{E-}06 \pm 2.54\text{E-}06$
NS-K126	NS	K126	~67%	$6.79\text{E-}05 \pm 1.27\text{E-}05$	$8.21\text{E-}04 \pm 8.56\text{E-}05$
NS-K131	NS	K131	~50%	$5.60\text{E-}07 \pm 3.55\text{E-}07$	$7.68\text{E-}06 \pm 6.05\text{E-}07$
NS-A132	NS	A132	~25%	未测	未测
M2-S23	M2	S23	~25%	未测	未测
M2-D24	M2	D24	~25%	未测	未测
M2-H37	M2	H37	~50%	$8.00\text{E-}06 \pm 7.75\text{E-}07$	$1.10\text{E-}05 \pm 3.21\text{E-}06$
M2-W41	M2	W41	~25%	未测	未测
M2-K49	M2	K49	~57%	$7.90\text{E-}03 \pm 1.11\text{E-}03$	$7.50\text{E-}01 \pm 2.25\text{E-}02$
M2-K60	M2	K60	~57%	$5.70\text{E-}03 \pm 3.29\text{E-}03$	$6.30\text{E-}01 \pm 2.13\text{E-}01$

突变病毒 名称	蛋白	TAG 位点	相对包装效率 (%)	逃逸频率 (第一代)	逃逸频率 (第二十代)
PTC-2	PA	R266	~67%	$1.20\text{E-}10 \pm 4.46\text{E-}11$	$3.10\text{E-}10 \pm 6.45\text{E-}11$
	PB2	K33			
PTC-3	PA	R266	~67%	<1.00E-11	<1.00E-11
	PB2	K33			
	PB1	R52			
PTC-4A	PA	R266	~67%	<1.00E-11	<1.00E-11
	PB2	K33			
	PB1	R52			
	NP	D101			
PTC-4B	PA	R266	~57%	<1.00E-11	<1.00E-11
	PB2	K33			
	PB1	S375			
	NP	M163			
PTC-5	PA	R266	~50%	<1.00E-11	<1.00E-11
	PB2	K33			
	PB1	R52			
	NP	D101			
	NS	K131			
PTC-6	PA	R266	~50%	<1.00E-11	<1.00E-11
	PB2	K33			
	PB1	R52			
	NP	D101			
	NS	K131			
	M2	H37			

PTC-7	PA	R266			
	PB2	K33			
	PB1	R52			
	NP	D101	~50%	<1.00E-11	<1.00E-11
	NS	K131			
	M2	H37			
	NA	N28			
PTC-8	PA	R266			
	PB2	K33			
	PB1	R52			
	NP	D101	~50%	<1.00E-11	<1.00E-11
	NS	K131			
	M2	H37			
	NA	N28			
	HA	K57			

发明人从上述突变位点中，选择了拯救流感病毒效率高且遗传稳定的位点 PA-S1、PB1-S4、PB2-S4、NP-S3 进行组合，即使用这四种质粒与拯救流感病毒的其他质粒 Ben4 pPolI-WSN-HA； Ben6 pPolI-WSN-NA； Ben7 pPolI-WSN-M； Ben8 pPolI-WSN-NS； Ben9 pcDNA 3 (neo)-PB2； Ben10 pcDNA 3 (neo)-PB1； Ben11 pcDNA 3 (neo)-PA； Ben13 pcAGGS/MCS-NP 共同转染（5）建立的稳定细胞系，拯救出在 PA-S1、PB1-S4、PB2-S4、NP-S3 位点同时引入 TAG 的突变型流感病毒，命名为 WSN-RNP-tag。

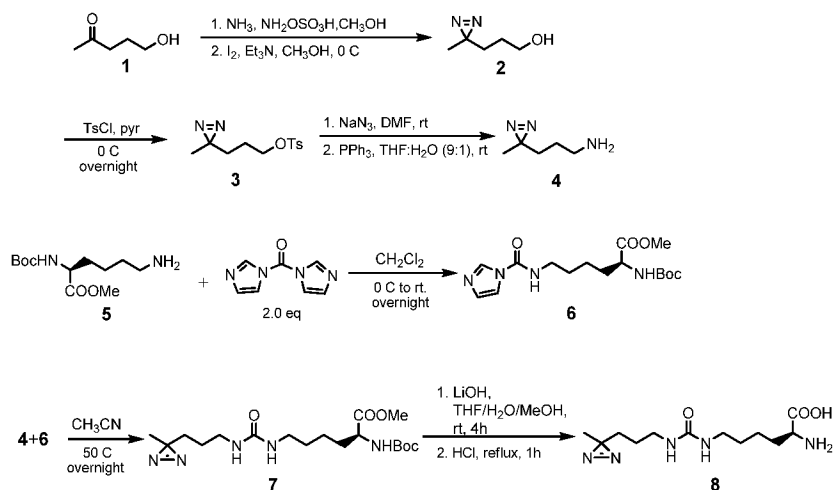
实施例 2：定点突变的流感病毒的表达和纯化

本发明中构建的引入了密码子 TAG 的流感病毒拯救质粒在可以稳定表达 tRNA (tRNA^{Pyl}) 和吡咯赖氨酰-tRNA 合成酶 (tRNA^{Pyl}) 的哺乳动物稳定细胞系中进行转录和表达后，细胞利用这套蛋白质翻译系统使非天然氨基酸 Lys-diazirine 或者与 Lys-diazirine 结构近似的叠氮非天然氨基酸 Lys-azido 掺入到流感病毒的相应蛋白中，从而在流感病毒相应蛋白的被突变位点造成定点突变。

下面，发明人对 Lys-diazirine 和 Lys-azido 这两种非天然氨基酸的成功掺入和突变蛋白质的生产性能进行了检测。

(1) 非天然氨基酸 Lys-diazirine 的合成和鉴定

非天然氨基酸 Lys-diazirine 的化学合成反应式如下。



如上式所示，将原料 1 (5-羟基-2-戊酮) 15 mL 与液氨 40mL 在 -40°C 下搅拌反应 5h，之后降温至 -60°C ，缓慢滴加 $\text{NH}_2\text{OSO}_3\text{H}$ 的甲醇溶液，加毕升至室温，反应过夜。滤除沉淀，向上清液中加入三乙胺，冰浴条件下缓慢加入 I_2 ，至反应液颜色变深，不再产生气泡为止。反应完全后蒸除溶剂，经乙醚萃取后干燥。蒸除乙醚，剩余液体减压蒸馏获得 25.4g 无色粘稠液体产物 2。

将上述产物 2 用吡啶溶解， 0°C 搅拌下加入 11 g TsCl ，反应过夜。待反应完全后将反应液倒入浓盐酸与冰水的混合液中，乙醚萃取，醚层分别用 1N 盐酸和 1N NaOH 洗涤。有机相干燥柱分得到 11.8g 无色粘稠液体产物 3。

将上述产物 3 用 DMF 溶解，加入 NaN_3 室温反应隔夜至反应完全，加入大量水，乙醚萃取。蒸除乙醚，剩余产物用 THF :水(9:1)混溶，加入三苯基磷，室温反应。反应完后加 1N HCl 混匀，旋干 THF ，二氯甲烷把未反应的原料， PPh_3 和 O=PPh_3 洗掉，液相加 1N NaOH 调 pH 到 12，二氯甲烷萃取出 4.0 g 产物 4。

将 5.2 g 原料 5 (Boc-Lys-OMe) 与羰基二咪唑反应，制备出 5.9 g 化合物 6。之后化合物 6 与上述产物 4 (4.0 g) 偶联得到化合物 7，最后经过两步脱保护，将 Boc 和甲酯脱除，得到目标 4.5 g 产物 8，即 Lys-diazirine。经谱学验证，结果为：

$^1\text{H NMR}$ (400 MHz, D_2O): δ 3.10 (1H, t, $J = 6.3$ Hz), 2.96 (4H, m), 1.25 (10H, m), 0.90 (3H, s); $^{13}\text{C NMR}$ (100 MHz, D_2O): 183.63, 160.66, 56.00, 39.80, 39.30, 34.49, 30.84, 29.20, 26.75, 23.92, 22.43, 18.80; HREIMS m/z 308.16937 $[\text{M}+1]^+$ (calcd for $\text{C}_{12}\text{H}_{22}\text{N}_5\text{NaO}_3$, 308.16931), 证明所得到的 Lys-diazirine 结构正确。

(2) 突变流感病毒的非天然氨基酸 Lys-diazirine 掺入拯救

将实施例 1 的步骤 (6) 定点突变后的流感病毒的拯救中获得的突变型流感病毒包装质粒共转染步骤(5)中的稳定细胞系，6 小时后换成新的培养基，培养基中含有 1% 的 FBS、2 $\mu\text{g/ml}$ 的 TPCK-trypsin、1mM 的非天然氨基酸 Lys-diazirine，并以无非天然氨基酸 Lys-diazirine 的培养基作为对照。此拯救实验采用的阳性对照为野生型流感病毒 WSN，除了拯救病毒的质粒不同之外，其余条件均与突变型流感病毒的拯救条件相同。转染完成后，每天观察细胞的状态。被转染细胞在含非天然氨基酸培养基中培养时出现病变、而在不含非天然氨基酸的培养基中培养时不出现病变的流感病毒突变体被鉴定为阳性突变体。与之相比，被野生型流感病毒所感染的细胞在培养时，无论培养基中有没有非天然氨基酸，均会出现病变。

(3) Lys-diazirine 突变流感病毒的纯化

1) .当步骤(2)拯救突变型流感病毒的稳定系细胞完全病变时,收集细胞上清,于 5000g 离心 10min,取上清过 0.45um 的滤膜。

2) .使用蔗糖梯度密度离心的方法纯化流感病毒。具体步骤如下:将 1) 中的病毒液用 50ml 离心管(高速专用)于 10^5 g 离心 2h,沉淀用 1mlPBS 重悬。

3) .用 NTE Buffer (100 mM NaCl, 10 mM Tris-Cl, pH7.4, 1mM EDTA) 溶解蔗糖,配成 20%蔗糖溶液,过 0.45 um 滤膜。

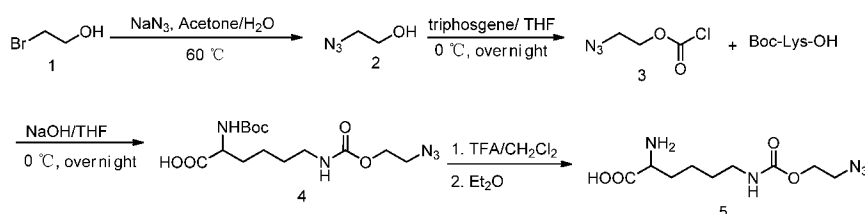
4) .将步骤 3) 的蔗糖加入到 50ml 或者 15ml 离心管中,将 2) 中的 PBS 重悬液滴在蔗糖溶液上。 11×10^4 g, 离心 2h。

5) .沉淀加约 15ml NTE buffer, 11×10^4 g, 继续离心 2h。

6) .将步骤 5) 中的沉淀用 PBS 重悬。

(4) 突变流感病毒的 Lys-azido 掺入表达及纯化

非天然氨基酸 Lys-azido 的化学合成反应式如下



如上式所述,将原料 1 (2-溴乙醇) 2.3mL 溶于 90mL 丙酮以及 15mL 水的混合溶液,加入 NaN₃ 3.12g, 60°C 油浴加热回流反应 20h。冷却至室温,旋蒸除去丙酮,无水乙醚萃取 (30mL×8),无水 Na₂SO₄ 干燥,旋蒸除去溶剂得 2.62g 无色液体产物 2。

将 2 (500mg, 5.74mmol) 加入到三光气 (1.70g, 5.74mmol) 的 THF(10ml) 溶液中。0°C 搅拌反应 8h, 溶剂蒸干。剩余物在真空下干燥 1h, 得到无色油状产物 3。

将 3 溶解在 1.5ml 的 THF 中并缓慢加入 Boc-Lys-OH (1.7g, 6.88mmol) 的 1M NaOH (20ml) /THF(5ml) 的溶液中。0°C 搅拌反应 12h 并逐渐升温到室温。重新将反应液冷却到 0°C 并用 0°C 的 1M 的盐酸溶液将反应液 pH 值调整至 2~3。反应液用 EtOAc 萃取 (30mL×5), 有机层用 2×100ml 的饱和食盐水洗涤。无水 Na₂SO₄ 干燥有机层、过滤、旋蒸除去溶剂得到 1.65g 无色粘稠液体产物 4 不用进一步纯化。

将 4 溶于 15mL CH₂Cl₂ 中, 搅拌下缓慢滴加 15mL TFA, 室温下反应 30min 后蒸出溶剂, 剩余液体产物用 5mL 甲醇溶解, 加入 100mL 乙醚, 析出大量白色固体沉淀, 过滤干燥得到 1.38g 白色固体终产物 5。¹H NMR (D₂O): δ = 1.22-1.45 (m, 4 H), 1.67-1.73 (m, 2H), 2.99 (m, 2 H), 3.38 (m, 2 H), 3.70 (m, 1 H), 4.09 (m, 2 H). ¹³C NMR (D₂O): δ = 21.4, 28.4, 29.6, 39.5, 53.4, 56.2, 57.8, 116.0 (TFA), 153.1, 162.3 (TFA), 172.9. HRMS: m/z calcd for C₉H₁₇N₅O₄ [M]⁺: 259.1281; found: 259.1283, 证明得到的 Lys-azido 结构正确。

除了用 Lys-azido 替换 Lys-diazirine 外, 其它条件与前述步骤 2—3 相同, 观察引入突变的流感病毒是否引起细胞病变, 即可确认突变是否成功, 从而得到在相应突变位点引入 UAG 的突变型流感病毒。

部分的突变型流感病毒的拯救情况作为示例, 可见图 2。所述结果表明, 在流感

病毒的 8 个蛋白质的编码基因片段上均可引入密码子 UAG，并且可以拯救出突变型流感病毒。其中，PA-S1、PB1-S4、PB2-S4、NP-S3 这四个位点拯救突变型流感病毒的效率高，且具有遗传稳定性，是较为理想的突变位点。将这四个位点进行组合，即使用这四种质粒与拯救流感病毒的其他质粒 Ben4 pPolI-WSN-HA； Ben6 pPolI-WSN-NA； Ben7 pPolI-WSN-M； Ben8 pPolI-WSN-NS； Ben9 pcDNA 3 (neo)-PB2； Ben10 pcDNA 3 (neo)-PB1； Ben11 pcDNA 3 (neo)-PA； Ben13 pcAGGS/MCS-NP 共同转染（5）建立的稳定细胞系，拯救出在 PA-S1、PB1-S4、PB2-S4、NP-S3 位点同时引入密码子 UAG 的突变型流感病毒，命名为 WSN-RNP-tag（PTC-4A）。突变型流感病毒 WSN-RNP-tag 的拯救效率很高，且产率高，而且具有很高的遗传稳定性。是发明人制备出来的最终突变产物。

5. 突变型流感病毒 WSN-RNP-tag 拯救效率的考察

发明人比较了使用普通哺乳动物细胞 293T 和稳定表达 tRNA (tRNA^{Pyl}) 与吡咯赖氨酰-tRNA 合成酶 (tRNA^{Pyl}) 的哺乳动物稳定细胞系 HEK293-PYL 拯救突变型流感病毒 WSN-RNP-tag 的效率。

考察实验分为三组：第一组，使用稳定细胞系 HEK293-PYL，同时转染 1.2ug 可表达 tRNA (tRNA^{Pyl}) 和吡咯赖氨酰-tRNA 合成酶 (tRNA^{Pyl}) 的质粒和 1.2ug 拯救突变型流感病毒的质粒；第二组，使用稳定细胞系，只转染 2.4ug 拯救突变型流感病毒的质粒；第三组，使用普通 293T 细胞，同时转染 1.2ug 可表达 tRNA (tRNA^{Pyl}) 和吡咯赖氨酰-tRNA 合成酶 (tRNA^{Pyl}) 的质粒和 1.2ug 拯救突变型流感病毒的质粒。结果如图 10 所示，使用稳定细胞系 HEK293-PYL 拯救流感病毒的效率远远高于使用普通 293T 细胞。（图 10 中 NAEK 存在情况的说明）

此外，发明人还考察了制备出的突变型流感病毒 WSN-RNP-tag 的产量。通过测定相同条件下制备的野生型和突变型病毒的 M 基因节段的 mRNA 水平，说明制备出的突变型流感病毒的产量与野生型基本一致（图 5）。另外通过测定野生型和突变型 WSN-RNP-TAG 的病毒的抗原红细胞凝集效价，说明突变型流感病毒和野生型病毒的抗原红细胞凝集效价是基本一致的（图 6），进一步说明，制备出的突变型流感病毒的产量很高。具体的实验操作可参考 qRT-PCR 的实验步骤和病毒的抗原红细胞凝集效价测定实验方法。

实施例 3. 突变型流感病毒 WSN-RNP-tag 在细胞水平的安全性考察

发明人通过对制备出的突变型流感病毒 WSN-RNP-tag 进行长期的传代培养，来考察该突变型病毒在突变位点处的 UAG 密码子的稳定性。具体实验如下：将新制备出的突变型流感病毒 WSN-RNP-tag 按照 MOI=0.01 接种在新的培养基中并感染稳定系细胞，培养基中含有 1%FBS、2ug/ml 的 TPCK-trypsin、1mM 非天然氨基酸 NAEK，并用无非天然氨基酸 NAEK 的培养基作为对照。待含有 1mM NAEK 培养基中的细胞完全病变后，取出上清，过 0.45 um 的滤膜，再按照 1/1000 的比例接种在新的培养基中并感染稳定系细胞，同样以无非天然氨基酸 NAEK 的培养基作为对照。如此重复，进行长期的病毒传代。

由图 4 可知，经过长期传代，突变型 WSN-RNP-tag 一直维持着对非天然氨基酸的依赖性，即只能在含有非天然氨基酸的培养基中才能大量复制，使细胞发生病变。这说明在流感病毒基因中引入的 UAG 密码子是一直稳定存在的，进而说明所选择的位点在遗传上是稳定的。

实施例 4. 突变型流感病毒 WSN-RNP-tag 在动物水平的安全性和有效性考察

实验流程可参考图 7，具体实施方案如下：

- 1) 将 80 只小鼠分成 8 组，每组 10 只。
- 2) 饲养一天，让小鼠适应环境。
- 3) 第二天，称量小鼠的体重，并从每组小鼠中选择 5 只，取它们的血清（采血体积在 20-40ul 即可，防止小鼠失血死亡。采集的血清在 -80℃ 冷冻保存）。
- 4) 饲养两天后（让被采血的动物回复正常状态），根据组别和病毒液的管号，接种下面相应的病毒液。

接种病毒的方法：麻醉，以滴鼻的方式向小鼠的鼻腔内接种 50ul 病毒液。半数致死剂量 LD₅₀ 是 10000 个病毒颗粒/50ul。一倍致死剂量是 10*LD₅₀，即相当于 10⁵ 个病毒颗粒/50ul。10 倍致死剂量是 100*LD₅₀，即相当于 10⁶ 个病毒颗粒/50ul。

第一组接种 1 号管的病毒液（成分：PBS）；

第二组接种 2 号管的病毒液（成分：1 倍致死剂量的 WSN-wild type）；

第三组接种 3 号管的病毒液（成分：1 倍致死剂量的 WSN-RNP-tag）；

第四组接种 4 号管的病毒液（成分：5 倍致死剂量的 WSN-RNP-tag）；

第五组接种 5 号管的病毒液（成分：10 倍致死剂量的 WSN-RNP-tag）；

第六组接种 6 号管的病毒液（成分：1 倍致死剂量的灭活 WSN）；

第七组接种 7 号管的病毒液（成分：5 倍致死剂量的灭活 WSN）；

第八组接种 8 号管的病毒液（成分：10 倍致死剂量的灭活 WSN）；

每组之间防止交叉感染。

- 5) 接种以后的每一天，称量并记录小鼠的体重，并记录小鼠的死亡情况。

- 6) 分别在接种病毒液后的第 14 天和第 21 天，再次取那 5 只小鼠的血清（采血体积在 20-40ul 即可，采集的血清在 -80℃ 冷冻保存）。

- 7) 在接种病毒液后的第 21 天，每只小鼠鼻腔内感染新的 50ul 病毒液，感染剂量为 100LD₅₀。

然后继续观察 2-3 周内，称量并记录小鼠的体重，并记录小鼠的死亡情况。此外，在感染病毒液后的第三天，从每组小鼠中取 3 只小鼠，取它们的肺组织，研磨匀浆后，收集上清（可置于 -80℃ 冷冻保存）。剩下的小鼠继续观察。

由图 8 可知，当接种野生型流感病毒时，所有的小鼠体重均明显下降，并在接种后第 10 天全部死亡。当接种突变型 WSN-RNP-TAG 或者灭活的流感病毒 WSN-positive control 时，所有的小鼠体重均没有明显下降，且观察期间所有的小鼠均存活。说明，突变型 WSN-RNP-TAG 在动物水平具有很好的安全性。

由图 9A 可知，按照图 7 的实验方案，取各组小鼠肺组织，通过 qRT-PCR 的方式检测肺组织中病毒的相对含量。结果说明，当用灭活病毒或者突变型 WSN-RNP-TAG 对小鼠免疫后，小鼠肺组织中的病毒含量下降，并呈现免疫剂量依赖性。此外，突变型 WSN-RNP-TAG 的免疫效果优于灭活病毒的免疫效果。

由图 9B 可知，通过 Elisa 实验检测免疫前后的各组小鼠体内抗 HA 抗体的含量。结果说明免疫后 14 天和免疫后 21 天小鼠体内的抗体含量均增加，且 21 天的抗体含量高于 14 天时的抗体含量。

由图 9C 可知，按照图 7 的实验方案，对各组小鼠分别进行免疫之后，用 100×LD₅₀

的 WSN 进行攻毒。并进行了小鼠体重和死亡率的考察，发现制备的突变型病毒 WSN-RNP-tag 可以很好地保护动物不受病毒的感染。并且，1 倍剂量的突变型 WSN-RNP-tag 对小鼠的保护效果甚至和 10 倍剂量的灭活 WSN 的保护效果相当。这充分显示了发明人制备的减毒流感疫苗相对于灭活疫苗具有更好的免疫原性优势。

实施例 5 接种流感病毒疫苗后的小鼠体内抗 NA 和 NP 蛋白的抗体含量测定

按照与实施例 4 接种病毒方法完全一致的方法，用 WSN-RNP-tag 和灭活的 WSN 接种小鼠。

使用 Elisa 方法检测小鼠体内抗 NA 和 NP 蛋白的抗体含量。具体来说，将蛋白 NA 或 NP 用包被液稀释，30ng/100ul。Elisa 专用 96 孔板用上述稀释液于 4℃包被过夜。用 Elisa 洗涤液洗涤后，用含 3%BSA 的洗涤液于 37℃阻断 1h，然后将小鼠血清用 0.5%BSA 的洗涤液稀释后加入相应的空中，37℃孵育 1h，弃上清。用洗涤液洗 5 次。然后加入 HRP 标记的 goat anti-mouse IgG，于 37℃孵育 1h，弃上清。用洗涤液洗 5 次。用 TMB 显色液显色 5-10min，并用 Elisa 终止液终止反应。于 450nm 测 OD 值。

结果如图 12 所示，WSN-RNP-tag 可以诱导小鼠产生 NA 特异性抗体，其抗体产生水平高于灭活疫苗，说明发明人所制备的疫苗具有更好的免疫原性。更为重要的是，发明人发明的 WSN-RNP-tag 还可以诱导小鼠产生针对流感病毒内部蛋白 NP 的抗体，这有利于为小鼠提供交叉免疫保护。

实施例 6 WSN-RNP-tag 对小鼠肺内病毒特异性抗体 IgA 的诱导作用

具体实施方案如下：

- 1) 将 15 只小鼠分成 3 组，每组 5 只。
- 2) 饲养一天，让小鼠适应环境。
- 3) 第二天，根据组别和病毒液的管号，接种下面相应的病毒液。

接种病毒的方法：麻醉，以滴鼻的方式向小鼠的鼻腔内接种 50ul 病毒液。与实施例 4 中定义相同，半数致死剂量 LD50 是 10000 个病毒颗粒/50ul。一倍致死剂量是 10*LD50，10 倍致死剂量是 100*LD50。

第一组接种 1 号管的病毒液（成分：PBS）；

第二组接种 2 号管的病毒液（成分：10 倍致死剂量的 WSN-RNP-tag）；

第三组接种 3 号管的病毒液（成分：10 倍致死剂量的灭活 WSN）；

每组之间防止交叉感染。

4) 分别在接种病毒液后第 21 天，取每组小鼠的肺组织，用 PBS 洗涤，收集洗涤液。收集的洗涤液可在-80℃冷冻保存。

使用 Elisa 的方法检测 IgA 的产生情况。具体来说，将纯化的 WSN 病毒用包被液稀释，0.5ug/100ul。Elisa 专用 96 孔板用上述稀释液于 4℃包被过夜。用 Elisa 洗涤液洗涤后，用含 3%BSA 的洗涤液于 37℃阻断 1h，然后将小鼠血清用 0.5%BSA 的洗涤液稀释后加入相应的空中，37℃孵育 1h，弃上清。用洗涤液洗 5 次。然后加入 HRP 标记的 goat anti-mouse IgA，于 37℃孵育 1h，弃上清。用洗涤液洗 5 次。用 TMB 显色液显色 5-10min，并用 Elisa 终止液终止反应。于 450nm 测 OD 值。

结果如图 13 所示，说明，相对于灭活的 WSN，WSN-RNP-tag 可以诱导高水平的 IgA，

这有利于为小鼠提供交叉免疫保护。

实施例7 WSN-RNP-tag 免疫后的小鼠肺内 NP 特异性的 CD8+T 细胞的含量变化

试验方法：小鼠免疫三周后，取小鼠肺组织，分离其中的 T 淋巴细胞。将 T 淋巴细胞用 anti-mouse CD8a-APC 抗体和流感 NP366-374-tetramer-PE 染色。通过细胞流式实验测流感特异性 CD8+ T 细胞的数量（比例），具体操作可参考文献 Budimir, N. et al. Heterosubtypic cross-protection induced by whole inactivated influenza virus vaccine in mice: influence of the route of vaccine administration. Influenza Other Respir Viruses 7, 1202-1209 (2013)。

结果如图 14 所示，说明，灭活的 WSN 几乎不能影响 NP 特异性的 CD8+T 细胞的含量，而 WSN-RNP-tag 免疫后的小鼠肺内 NP 特异性的 CD8+T 细胞的含量明显增加，这有利于为小鼠提供交叉免疫保护。

实施例8 发明人制备的含有 UAG 密码子的病毒疫苗对野生型病毒的抑制作用考察。

首先在细胞水平考察了含有 UAG 密码子的病毒对野生型病毒复制的影响。具体实施方案如下：

1. 用 MDCK 细胞铺 6 孔板，每孔 5×10^5 个细胞。
2. 24 小时后，用野生病毒（MOI=0.01）或者野生型病毒与突变病毒（MOI=0.1 或者 1）的混合液共同感染细胞。
3. 24 小时后，取 200 μ l 细胞上清，并且每间隔 12 小时取一次，连续至 72 小时。
4. 测定上清中野生型病毒的滴度。

结果如图 15A 和 B 所示，突变型病毒降低了野生型病毒的子代病毒的滴度，且升高突变型病毒的滴度或者增加突变型病毒基因组中 UAG 密码子的数量会增强突变型病毒对野生型病毒复制的抑制作用。

然后，在动物水平考察了含有 UAG 密码子的病毒对野生型病毒复制的影响。具体实施方案如下：

1. 每组 15 只 Balb/C 小鼠。用野生型病毒（ 2×10^4 PFU）或者野生型病毒与突变病毒（ 2×10^6 PFU）的混合液通过鼻腔接种的方式感染 Balb/C 小鼠；或者先用野生病毒（ 2×10^4 PFU）感染小鼠，24 小时后再用突变病毒（ 2×10^6 PFU）感染小鼠。
2. 3 天后，每组随机取出 5 只小鼠处死，取出肺组织，测定肺组织中野生型病毒的滴度。
3. 剩下的 10 只小鼠被观察存活率和体重，持续两周。

结果如图 15C, D 和 E 所示。当单独接种野生型病毒时，所有的小鼠体重下降，并最终死亡；当同时接种野生型病毒和突变型病毒的混合液时，小鼠体重下降不明显，最终小鼠的存活率为 90%，同时肺组织中的病毒明显降低；当接种野生型病毒 24 小时后再接种突变型病毒，小鼠体重有所下降，小鼠的存活率为 40%。这说明突变型病毒可以抑制野生型病毒在小鼠体内的复制，从而为小鼠提供保护。这些结果也说明，本专利提供的技术可以制备具有治疗作用的疫苗。

以上所述制备流感病毒减毒活疫苗的方法，也可以适用于其他任何种类的病毒，优选地，手足口病毒、柯萨奇病毒、丙肝病毒 HCV、乙肝病毒 HBV、甲肝病毒、丁型肝炎病毒、戊型肝炎病毒、EB 病毒、人乳头瘤病毒 HPV、单纯疱疹病毒 HSV、巨细胞病毒、水痘-带状疱疹病毒、水泡性口炎病毒、呼吸道合胞病毒 RSV、登革病毒、埃博拉病毒、

寨卡病毒 Zika、SARS、中东呼吸综合征病毒、轮状病毒、狂犬病毒、麻疹病毒、腺病毒、艾滋病毒、脊髓灰质炎病毒、埃可病毒、乙型脑炎病毒、森林脑炎病毒、汉坦病毒、新型肠道病毒、风疹病毒、腮腺炎病毒、副流感病毒、蓝耳病毒、猪瘟病毒、口蹄疫病毒、细小病毒。

以上所述的仅是本发明的一些实施方式。对于本领域的普通技术人员来说，在不脱离本发明创造构思的前提下，还可以做出若干变形和改进，这些都属于本发明的保护范围。

序列表

<110> 北京大学

<120> 突变的病毒、其制备方法和应用

<130> 2

<160> 9

<170> PatentIn version 3.3

<210> 1

<211> 8482

<212> DNA

<213> 人工序列

<400> 1

```

gacgaaaggg cctcgtgata cgcctatattt tataggttaa tgtcatgata ataatggttt      60
cttagacgtc aggtggcact tttcggggaa atgtgcgcgg aaccctattt tgtttatattt      120
tctaaataca ttcaaatatg tatccgctca tgagacaata accctgataa atgcttcaat      180
aatattgaaa aaggaagagt atgagtattc aacatttccg tgtcgccctt attccctttt      240
ttgcggcatt ttgccttcct gtttttgctc acccagaaac gctggtgaaa gtaaaaagatg      300
ctgaagatca gttgggtgca cgagtgggtt acatcgaact ggatctcaac agcggtaaga      360
tccttgagag ttttcgcccc gaagaacgtt ttccaatgat gaggcacttt aaagtctctgc      420
tatgtggcgc ggtattatcc cgtattgacg cggggcaaga gcaactcggc cgccgcatac      480
actattctca gaatgacttg gttgagtact caccagtcac agaaaagcat cttacggatg      540
gcatgacagt aagagaatta tgcagtgcct ccataacatc gaggatgata actgcggcca      600
acttacttct gacaacgata ggaggaccga aggagctaac cgtttttttg cacaacatgg      660
gggatcatgt aactcgcctt gatcgttggg aaccggagct gaatgaagcc ataccaaacg      720
acgagcgtga caccacgatg cctgtagcaa tggcaacaac gttgcgcaa ctattaactg      780
gcgaactact tactctagct tcccggcaac aattaataga ctggatggag gcggataaag      840
ttgcaggacc acttctgcgc tcggcccttc cggtcggctg gtttattgct gataaatctg      900
gagccggtga gcgtgggtct cgcggtatca ttgcagcact ggggccagat ggtaagccct      960
cccgtatcgt agttatctac acgacgggga gtcaggcaac tatggatgaa cgaaatagac      1020
agatcgtcga gatagggtgcc tcatgatta agcattggta actgtcagac caagtttact      1080
catatatact ttagattgat ttaaaacttc atttttaatt taaaaggatc taggtgaaga      1140
tcctttttga taatctcatg accaaaatcc cttaacgtga gttttcgttc cactgagcgt      1200
cagaccccgat agaaaagatc aaaggatcct cttgagatcc tttttttctg cgcgtaatct      1260
gctgcttgca aacaaaaaaa ccaccgtac cagcgggtgg ttgtttgccg gatcaagagc      1320
taccaactct ttttccgaag gtaactggct tcagcagagc gcagatacca aatactgtcc      1380
ttctagtgtg gccgtagtta ggccaccact tcaagaactc tgtagcaccg cctacatacc      1440
tcgctctgct aatcctgtta ccagtggctg ctgccagtgg cgataagtcg tgtcttaccg      1500
ggttggactc aagacgatag ttaccggata aggcgcagcg gtcgggctga acgggggggt      1560
cgtgcacaca gccagccttg gacgcaacga cctacaccga actgagatac ctacagcgtg      1620

```

agctatgaga	aagcgccacg	cttcccgaag	ggagaaaggc	ggacaggtat	ccggtaaagcg	1680
gcagggtcgg	aacaggagag	cgcacgaggg	agcttccagg	gggaaacgcc	tggtatcttt	1740
atagtcctgt	cgggtttcgc	cacctctgac	ttagcgctcg	atttttgtga	tgctcgtcag	1800
gggggcggag	cctatggaaa	aacgccagca	acgcggcctt	tttacggttc	ctggcctttt	1860
gctggccttt	tgctcacatg	ttctttcctg	cgttatcccc	tgattctgtg	gataaccgta	1920
ttaccgcctt	tgagttagct	gataccgctc	gccgcagccg	aacgaccgag	cgcagcgagt	1980
cagttagcga	ggaagcggaa	gagcgcccaa	tacgcaaacc	gcctctcccc	gcgcgttggc	2040
cgattcatta	atgcagctgg	cacgacaggt	ttcccgactg	gaaagcgggc	agtgagcgca	2100
acgcaattaa	tgtgagttag	ctcactcatt	aggcacccca	ggctttacac	tttatgcttc	2160
cggctcgtat	gttgtgtgga	attgtgagcg	gataacaatt	tcacacagga	aacagctatg	2220
accatgatta	cgccaagctt	gcatgcctgc	aggtcgacga	acgtgacgt	catcaacccg	2280
ctccaaggaa	tcgcgggccc	agtgtcacta	ggcgggaaca	cccagcgcgc	gtgcgccctg	2340
gcaggaagat	ggctgtgagg	gacaggggag	tggcgccctg	caatatattg	atgtcgctat	2400
gtgttctggg	aaatcaccat	aaacgtgaaa	tgtcttttga	tttgggaatc	ttataagttc	2460
tgtatgagac	cacagatccc	cggaaacctg	atcatgtaga	tcgaatggac	tctaaatccg	2520
ttcagccggg	ttagattccc	ggggtttccg	ccatttttct	cgcgcgcgc	gccatctcta	2580
ggccccgcgc	ggccccctcg	cacagacttg	tgggagaagc	tcggctactc	ccctgccccg	2640
gttaattttg	atataatatt	tcctagtaac	tatagagget	taatgtgcga	taaaagacag	2700
ataatctgtt	ctttttaata	ctagctacat	tttcatgat	aggcttggat	ttctataaga	2760
gatacaaata	ctaaattatt	attttaaaaa	acagcacaaa	aggaaactca	ccctaactgt	2820
aaagtaattg	tgtgttttga	gactataaat	atcccttggg	gaaaagcctt	gtttggaaac	2880
ctgatcatgt	agatcgaatg	gactctaaat	ccgttcagcc	gggttagatt	cccgggggtt	2940
ccgccatttt	tctcgacaag	gtcgggcagg	aagagggcct	atttcccatg	attccttcat	3000
atttgcatat	acatacaagg	ctgttagaga	gataattaga	attaatttga	ctgtaaacac	3060
aaagatatta	gtacaaaata	cgtgacgtag	aaagtaataa	tttcttgggt	agtttgcagt	3120
tttaaaatta	tgttttaaaa	tggactatca	tatgcttacc	gtaacttgaa	agtatttcga	3180
tttcttggct	ttatatatct	tgtggaaagg	acgaaacacc	ggaaacctga	tcatgtagat	3240
cgaatggact	ctaaatccgt	tcagccgggt	tagattcccc	gggtttccgc	catttttctc	3300
gacgaacgct	gacgtcatca	acccgctcca	aggaatcgcg	ggcccagtgt	cactaggcgg	3360
gaacacccag	cgcgcgtgcg	ccctggcagg	aagatggctg	tgagggacag	gggagtggcg	3420
ccctgcaata	tttgcatgtc	gctatgtgtt	ctgggaaatc	accataaacg	tgaaatgtct	3480
ttggattttg	gaatcttata	agttctgtat	gagaccacag	atccccggaa	acctgatcat	3540
gtagatcgaa	tggactctaa	atccgttcag	ccgggttaga	ttcccggggt	ttccgccatt	3600
tttctcgacg	acgcgcctat	ctctaggccc	gcgcgcggcc	cctcgcacag	acttgtggga	3660
gaagctcggc	tactccccctg	ccccggttaa	tttgcatata	atatttccta	gtaactatag	3720
aggcttaatg	tgcgataaaa	gacagataat	ctgttctttt	taatactagc	tacattttac	3780
atgataggct	tggattttcta	taagagatac	aaatactaaa	ttattatttt	aaaaaacagc	3840
acaaaaggaa	actcacccta	actgtaaagt	aattgtgtgt	tttgagacta	taaatatccc	3900
ttggagaaaa	gccttgtttg	gaaacctgat	catgtagatc	gaatggactc	taaatccgtt	3960
cagccgggtt	agattccccg	ggtttccgcc	atttttctcg	acaaggctcg	gcaggaagag	4020
ggcctatttc	ccatgattcc	ttcatatttg	catatacgat	acaaggctgt	tagagagata	4080
attagaatta	atttgactgt	aaacacaaag	atatttagtac	aaaatacgtg	acgtagaaag	4140
taataatttc	ttgggttagtt	tgcagtttta	aaattatgtt	ttaaaatgga	ctatcatatg	4200
cttaccgtaa	cttgaaagta	tttcgatttc	ttggctttat	atatcttgtg	gaaaggacga	4260
aacaccggaa	acctgatcat	gtagatcgaa	tggactctaa	atccgttcag	ccgggttaga	4320

ttccccgggt	ttccgccatt	tttctcgact	ctagaggatc	cctgcagtat	ttagcatgcc	4380
ccacccatct	gcaaggcatt	ctggatagtg	tcaaaacagc	cggaaatcaa	gtccgtttat	4440
ctcaaacttt	agcatttttg	gaataaatga	tatttgctat	gctggttaaa	ttagatttta	4500
gttaaatttc	ctgctgaagc	tctagtacga	taagtaactt	gacctaatg	taaagttgag	4560
atttccttca	ggtttatata	gcttgtgcgc	cgcctgggta	cctcggaaac	ctgatcatgt	4620
agatcgaatg	gactctaaat	ccgttcagcc	gggttagatt	cccggggttt	ccgccatttt	4680
tggatctaag	gtcgggcagg	aagagggcct	atttcccatg	attccttcat	atttgcatat	4740
acgatacaag	gctgttagag	agataattag	aattaatttg	actgtaaaca	caaagatatt	4800
agtacaaaat	acgtgacgta	gaaagtaata	atttcttggg	tagtttgcag	ttttaaaatt	4860
atgttttaaa	atggactatc	atatgcttac	cgtaacttga	aagtatttcg	atttcttggc	4920
tttatataat	ttgtggaaag	gacgaaacac	cggaaacctg	atcatgtaga	tcgaatggac	4980
tctaaatccg	ttcagccggg	ttagattccc	gggggtttccg	ccattttttg	atctgaacgc	5040
tgacgtcatc	aacccgctcc	aaggaatcgc	gggccagtg	tcactaggcg	ggaacaccca	5100
gcgcgcgtgc	gccctggcag	gaagatggct	gtgagggaca	ggggagtggc	gccctgcaat	5160
atttgcatgt	cgctatgtgt	tctgggaaat	caccataaac	gtgaaatgtc	tttggatttg	5220
ggaatcttat	aagtcttgta	tgagaccaca	gatccccgga	aacctgatca	tgtagatcga	5280
atggactcta	aatccgttca	gccgggttag	attcccgggg	tttccgccat	ttttggatct	5340
ctgcagtatt	tagcatgccc	cacccatctg	caaggcattc	tggatagtgt	caaaacagcc	5400
ggaaatcaag	tccgtttatc	tcaaacttta	gcattttggg	aataaatgat	atttgctatg	5460
ctggttaaat	tagatttttag	ttaaatttcc	tgctgaagct	ctagtacgat	aagtaacttg	5520
acctaagtgt	aaagttgaga	tttcccttcag	gtttatatag	cttgtgcgcc	gcctgggtac	5580
ctcggaaacc	tgatcatgta	gatcgaatgg	actctaaatc	cgttcagccg	ggttagattc	5640
ccgggggttt	cgccattttt	ggatctaagg	tccggcagga	agagggccta	tttcccatga	5700
ttccttcata	tttgcatata	cgatacaagg	ctgttagaga	gataattaga	attaatttga	5760
ctgtaaacac	aaagatatta	gtacaaaata	cgtgacgtag	aaagtaataa	tttcttgggt	5820
agtttgcagt	tttaaaatta	tgttttaaaa	tggactatca	tatgcttacc	gtaacttgaa	5880
agtatttcga	tttcttggct	ttatatatct	tgtggaaagg	acgaaacacc	ggaaacctga	5940
tcatgtagat	cgaatggact	ctaaatccgt	tcagccgggt	tagattcccc	gggtttccgc	6000
cattttttgga	tctgaacgct	gacgtcatca	acccgctcca	aggaatcgcg	ggcccagtg	6060
cactaggcgg	gaacaccag	cgcgcgtgcg	ccctggcagg	aagatggctg	tgagggacag	6120
gggagtggcg	ccctgcaata	tttgcatgtc	gctatgtgtt	ctgggaaatc	accataaacg	6180
tgaatgtct	ttggatttgg	gaatcttata	agttctgtat	gagaccacag	atccccggaa	6240
acctgatcat	gtagatcgaa	tggactctaa	atccgttcag	ccgggttaga	ttcccggggt	6300
ttccgccatt	tttgcatctc	cgggtaccct	gtgccttcta	gttgccagcc	atctgttgtt	6360
tgccccctcc	ccgtgccttc	cttgaccctg	gaaggtgcca	ctcccactgt	cctttcctaa	6420
taaaatgagg	aaattgcac	gcattgtctg	agtaggtgtc	attctattct	gggggggtgg	6480
gtggggcagg	acagcaaggg	ggaggattgg	gaagacaata	gcaggcatgc	tggggatgcg	6540
gtgggctcta	tggcttctga	ggcggaagaa	accagctggg	gctctagggg	gtatccccac	6600
gcgccctgta	gcggcgcat	aagcgcggcg	ggtgtgggtg	ttacgcgcag	cgtgaccgct	6660
acacttgcca	gcgccctagc	gcccgtcctt	ttcgctttct	tcccttcctt	tctcgcacg	6720
ttcgccggct	ttccccgtca	agctctaaat	cggggcatcc	ctttagggtt	ccgatttagt	6780
gctttacggc	acctcgaccc	caaaaaactt	gattagggtg	atggttcacg	tagtgggcca	6840
tcgccctgat	agacggtttt	tcgccctttg	acgttggagt	ccacgttctt	taatagtgga	6900
ctcttgttcc	aaactggaac	aacactcaac	cctatctcgg	tctattcttt	tgatttataa	6960
gggatttttg	ggatttcggc	ctatttggtta	aaaaatgagc	tgatttaaca	aaaatttaac	7020

```

gcgaattaat tctgtggaat gtgtgtcagt taggggtgtgg aaagtcccca ggctccccag 7080
gcaggcagaa gtatgcaaag catgcatctc aattagtcag caaccagggtg tggaaagtcc 7140
ccaggctccc cagcaggcag aagtatgcaa agcatgcata tcaattagtc agcaaccata 7200
gtcccccccc taactccgcc catccccccc ctaactccgc ccagttccgc ccattctccg 7260
ccccatggct gactaatttt ttttatttat gcagaggccg aggccgcctc tgcctctgag 7320
ctattccaga agtagtgagg aggccttttt ggaggcctag gcttttgcaa aaagctcccc 7380
ggagcttgta tatccatttt cggaattcat ggccaagttg accagtgccg ttccgggtgct 7440
caccgcgcgc gacgtcgccg gacggtcgga gttctggacc gaccggctcg ggttctcccc 7500
ggacttcgtg gaggacgact tcgccggtgt ggtccgggac gacgtgaccc tgttcacag 7560
cgcggtccag gaccagggtg tgccggacaa caccctggcc tgggtgtggg tgcgcggcct 7620
ggacgagctg tacgccgagt ggtcggaggt cgtgtccacg aacttccggg acgcctccgg 7680
gccggccatg accgagatcg gcgagcagcc gtggggggcg gagttcgccc tgcgcgaccc 7740
ggccggcaac tgcgtgcaact tcgtggccga ggagcaggac tgagcgggac tctgggggtc 7800
gaaatgaccg accaagcgac gcccacctg ccatcacgag atttcgattc caccgccgcc 7860
ttctatgaaa ggttgggctt cggaatcggt ttccgggacg ccggttgat gatcctccag 7920
cgcggggata tcatgctgga gttcttcgcc caccccaact tgtttattgc agcttataat 7980
ggttacaaat aaagcaatag catcacaat ttacaaaata aagcattttt ttcactgcat 8040
tctagtgtgt gtttgtccaa actcatcaat gtatcttata atgtctgact ggccgtcggt 8100
ttacaacgtc gtgactggga aaaccctggc gttaccaaac ttaatgcct tgcagcacat 8160
ccccctttcg ccagctggcg taatagcgaa gagggccgca ccgacgccc ttcccaacag 8220
ttgcgcagcc tgaatggcga atggcgccgt atgcgggtatt ttctccttac gcatctgtgc 8280
ggtatttcac accgcatatg gtgcactctc agtacaatct gctctgatgc cgcatagtta 8340
agccagcccc gacacccgcc aacacccgct gacgcgccct gacgggcttg tctgctcccc 8400
gcatccgctt acagacaagc tgtgaccgtc tccgggagct gcatgtgtca gaggttttca 8460
ccgtcatcac cgaaacgcgc ga 8482

```

<210> 2

<211> 2290

<212> DNA

<213> Influenza A virus

<400> 2

```

atggaagaa taaaagaact aaggaatcta atgtcgcagt ctgcactcg cgagatactc 60
acaaaaacca ccgtggacca tatggccata atcaagaagt acacatcagg aagacaggag 120
aagaaccag cacttaggat gaaatggatg atggcaatga aatatccaat tacagcagac 180
aagaggataa cggaatgat tcctgagaga aatgagcagg gacaaacttt atggagtaaa 240
atgaatgacg ccggtacaga ccgagtgatg gtatcacctc tggctgtgac atggtggaat 300
aggaatggac cagtgacaag tacagttcat tatecaaaaa tctacaaaac ttattttgaa 360
aaagtcgaaa ggttaaaaca tggaaccttt ggccctgtcc attttagaaa ccaagtcaaa 420
atacgtcgaa gagttgacat aaatcctggt catgcagatc tcagtccaa agaggcacag 480
gatgtaatca tggaagttgt tttccctaac gaagtgggag ccaggatact aacatcgga 540
tcgcaactaa cgacaaccaa agagaagaaa gaagaactcc agggttgcaa aatttctcct 600
ctgatggtgg catacatgtt ggagagagaa ctggtccgca aaacgagatt cctcccagtg 660
gctggtggaa caagcagtggt gtacattgaa gtgttgcat tggaccaagg aacatgctgg 720

```

gaacagatgt	acactccagg	aggggaggcg	aggaatgatg	atgttgatca	aagcttaatt	780
attgctgcta	gaaacatagt	aagaagagcc	acagtatcag	cagatccact	agcatcttta	840
ttggagatgt	gccacagcac	gcagattggg	ggagtaagga	tggtaaacad	ccttaggcag	900
aacccaacag	aagagcaagc	cgtggatatt	tgcaaggctg	caatgggact	gagaattagc	960
tcatccttca	gttttgggtg	attcacatth	aagagaacaa	gcggatcatc	agtcaagaga	1020
gaggaagagg	tgcttacggg	caatcttcag	acattgaaga	taagagtgc	tgagggatat	1080
gaagagttca	caatggttgg	gagaagagca	acagctatac	tcagaaaagc	aaccaggaga	1140
ttgattcagc	tgatagttag	tgggagagac	gaacagtcga	ttgccgaagc	aataattgtg	1200
gccatggtat	tttcacaaga	ggattgtatg	ataaaagcag	ttagagggtg	cctgaatttc	1260
gtcaataggg	cgaatcagcg	attgaatccc	atgcaccaac	ttttgagaca	ttttcagaag	1320
gatgcaaagg	tgctctttca	aaattgggga	attgaatcca	tcgacaatgt	gatgggaatg	1380
atcgggatat	tgcctgacat	gactccaagc	accgagatgt	caatgagagg	agtgagaatc	1440
agcaaaatgg	gggtagatga	gtattccagc	gcggagaaga	tagtggtgag	cattgaccgt	1500
tttttgagag	ttagggacca	acgtgggaat	gtactactgt	ctcccaggga	ggtcagtga	1560
acacagggaa	cagagaaact	gacaataact	tactcatcgt	caatgatgtg	ggagattaat	1620
ggctctgaat	cagtgtttgg	caatacctat	cagtggatca	tcagaaactg	ggaaactgtt	1680
aaaattcagt	ggctccagaa	tcctacaatg	ctgtacaata	aaatggaatt	tgagccattt	1740
cagtcttttag	ttccaaaggc	cgtttagaggc	caatacagtg	ggtttgtgag	aactctgttc	1800
caacaaatga	gggatgtgct	tgggacattt	gataccgctc	agataataaa	acttcttccc	1860
ttcgcagccg	ctccaccaa	gcaaagtgg	atgcagttct	cctcattgac	tataaatgtg	1920
aggggatcag	gaatgagaat	acttgtaagg	ggcaattctc	cagtattcaa	ctacaacaag	1980
accactaaaa	gactcacagt	tctcggaag	gatgctggcc	ctttaactga	agaccagat	2040
gaaggcacag	ctggagttag	gtccgcagtt	ctgagaggat	tcctcattct	gggcaaagaa	2100
gacaggagat	atggaccagc	attaagcata	aatgaactga	gcaaccttgc	gaaaggagag	2160
aaggctaatt	tgctaattgg	gcaaggagac	gtggtgttgg	taatgaaacg	gaaacggaac	2220
tctagcatac	ttactgacag	ccagacagcg	accaaagaa	ttcggatggc	catcaattag	2280
tgctgaatag						2290

<210> 3

<211> 2303

<212> DNA

<213> Influenza A virus

<400> 3

atttgaatgg	atgtcaatcc	gactttactt	ttcttaaaag	tgccagcaca	aaatgctata	60
agcacaactt	tcctttatac	tggagaccct	ccttacagcc	atgggacagg	aacaggatac	120
accatggata	ctgtcaacag	gacacatcag	tactcagaaa	ggggaagatg	gacaacaaac	180
accgaaactg	gagcaccgca	actcaaccgc	attgatgggc	cactgccaga	agacaatgaa	240
ccaagtgggt	atgcccaaac	agattgtgta	ttggaagcaa	tggccttcct	tgaggaatcc	300
catcctggta	tctttgagac	ctcgtgtctt	gaaacgatgg	aggttgttca	gcaaacacga	360
gtggacaagc	tgacacaagg	ccgacagacc	tatgactgga	ctctaaatag	gaaccagcct	420
gctgcaacag	cattggccaa	cacaatagaa	gtgttcagat	caaatggcct	cacggccaat	480
gaatctggaa	ggctcataga	cttccttaag	gatgtaatgg	agtcaatgaa	caaagaagaa	540
atggagatca	caactcattt	tcagagaaag	agacgagtga	gagacaatat	gactaagaaa	600

atggtgacac	agagaacaat	aggtaaaagg	aagcagagat	tgaacaaaag	gagttatcta	660
attagggcat	taaccctgaa	cacaatgacc	aaagatgctg	agagagggaa	gctaaaacgg	720
agagcaattg	caaccccagg	gatgcaaata	agggggtttg	tatactttgt	tgagacacta	780
gcaaggagta	tatgtgagaa	acttgaacaa	tcaggattgc	cagttggagg	caatgagaag	840
aaagcaaagt	tggcaaagt	tgtaaggaag	atgatgacca	attctcagga	cactgaaatt	900
tctttcacca	tacttgga	taacaccaa	tggaacgaaa	atcagaaccc	tcggatgttt	960
ttggccatga	tcacatatat	aaccagaaat	cagcccgaat	ggttcagaaa	tgttctaagt	1020
attgctccaa	taatgttctc	aaacaaaatg	gcgagactgg	gaaaggggta	catgtttgag	1080
agcaagagta	tgaacttag	aactcaaata	cctgcagaaa	tgctagcaag	catcgatttg	1140
aaatacttca	atgattcaac	tagaaagaag	attgaaaaaa	tccggccgct	cttaatagat	1200
gggactgcat	cattgagccc	tggaatgatg	atgggcatgt	tcaatatgtt	aagtactgta	1260
ttaggcgtct	ccatcctgaa	tcttgacaa	aagagacaca	ccaagactac	ttactgggtg	1320
gatggtcttc	aatcttctga	tgattttgct	ctgatttgta	atgcacccaa	tcatgaaggg	1380
attcaagccg	gagtcaacag	gttttatcga	acctgtaagc	tacttggaat	taatatgagc	1440
aagaaaaagt	cttacataaa	cagaacaggt	acatttgaat	tcacaagttt	tttctatcgt	1500
tatgggtttg	ttgccaat	cagcatggag	cttcccagct	ttgggggtgc	tgggatcaac	1560
gagtctgcgg	acatgagtat	tggagttact	gtcatcaaaa	acaatatgat	aaacaatgat	1620
cttgggtccag	caaccgctca	aatggccctt	cagctgttca	tcaaagatta	caggtacacg	1680
taccggtgcc	atagaggtga	cacacaaata	caaaccgaa	gatcatttga	aataaagaaa	1740
ctgtgggagc	aaaccattc	caaagctgga	ctgctggctc	ccgacggagg	cccaaattta	1800
tacaacatta	gaaatctcca	cattcctgaa	gtctgcttga	aatgggaatt	aatggatgag	1860
gattaccagg	ggcgtttatg	caaccactg	aaccatttgg	tcaaccataa	agacattgaa	1920
tcagtgaaca	atgcagtgat	aatgccagca	catgggtccag	ccaaaaacat	ggagtatgat	1980
gctgttgcaa	caacacactc	ctggatcccc	aaaagaaatc	gatccatctt	gaatacaagc	2040
caaagaggaa	tacttgaaga	tgaacaaatg	taccaaaggt	gctgcaactt	atttgaaaaa	2100
ttcttcccca	gcagttcata	cagaagacca	gtcgggatat	ccagtatggg	ggaggctatg	2160
gtttccagag	cccgaattga	tgacagaatt	gatttcgaat	ctggaaggat	aaagaaagag	2220
gagttcactg	agatcatgaa	gatctgttcc	accattgaag	agctcagacg	gcaaaaaatag	2280
tgaatttagc	ttgtccttca	tga				2303

<210> 4

<211> 2233

<212> DNA

<213> Influenza A virus

<400> 4

agcgaaagca	ggtactgatt	caaaatggaa	gattttgtgc	gacaatgctt	caatccgatg	60
attgtcgagc	ttgcggaaaa	ggcaatgaaa	gagtatggag	aggacctgaa	aatcgaaaca	120
aacaaatttg	cagcaatatg	cactcacttg	gaagtgtgct	tcatgtattc	agattttcac	180
ttcatcgatg	agcaaggcga	gtcaatagtc	gtagaacttg	gcgatccaaa	tgcacttttg	240
aagcacagat	ttgaaataat	cgagggaaga	gatcgcacaa	tagcctggac	agtaataaac	300
agtatttgca	acactacagg	ggctgagaaa	ccaaagtttc	taccagattt	gtatgattac	360
aagaagaata	gattcatcga	aattggagta	acaaggagag	aagttcacat	atactatctg	420
gaaaaggcca	ataaaattaa	atctgagaag	acacacatcc	acattttctc	attcactggg	480

gaggaaatgg ccacaaaggc cgactacact ctcgatgaag aaagcagggc taggatcaaa 540
 accaggctat tcaccataag acaagaaatg gctagcagag gcctctggga ttcctttcgt 600
 cagtccgaga gaggcgaaga gacaattgaa gaaagatttg aaatcacagg aacaatgcgc 660
 aagcttgccg accaaagtct cccgccaaac ttctccagcc ttgaaaattt tagagcctat 720
 gtggatggat tcgaaccgaa cggctacatt gagggcaagc ttctcctaat gtccaaagaa 780
 gtaaattgcta gaattgaacc ttttttgaaa tcaacaccac gaccacttag acttccggat 840
 gggcctccct gttctcagcg gtccaaattc ctgctgatgg atgccttaaa attaagcatt 900
 gaggacccaa gtcattgaggg agaggggata ccgctatatg atgcaatcaa atgcatgaga 960
 acattctttg gatggaagga acccaatgtt gttaaaccac acgaaaaggg aataaatcca 1020
 aattatcttc tgtcatggaa gcaagtactg gcagaactgc aggacattga gaatgaggag 1080
 aaaattccaa ggactaaaaa tatgaagaaa acgagtcagt taaagtgggc acttggtgag 1140
 aacatggcac cagaaaaggt agactttgac gattgtaaag atgtaggcga tttgaagcaa 1200
 tatgatagtg atgaaccaga attgaggtcg cttgcaagtt ggattcagaa tgagttcaac 1260
 aaggcatgtg aactgaccga ttcaagctgg atagagctcg atgagattgg agaagatgcg 1320
 gctccaattg aacacattgc aagcatgaga aggaattatt tcacagcaga ggtgtctcat 1380
 tgcagagcca cagaatacat aatgaagggg gtgtacatca atactgcctt gcttaatgca 1440
 tcctgtgcag caatggatga tttccaatta attccaatga taagcaagtg tagaactaag 1500
 gagggaaggg gaaagaccaa tttgtacggg ttcattcataa aaggaagatc ccacttaagg 1560
 aatgacaccg atgtgtgtaa ctttgtgagc atggagtttt ccctcactga cccaagactt 1620
 gaaccacaca aatgggagaa gtactgtgtt cttgaggtag gagatatgct tctaagaagt 1680
 gccataggcc atgtgtcaag gcctatgttc ttgtatgtga ggacaaatgg aacctcaaaa 1740
 attaaaatga aatgggggat ggaaatgagg cgttgccctc ttcagtcact tcaacaaatc 1800
 gagagtatga ttgaagctga gtctctgttc aaggagaaag acatgaccaa agagttcttt 1860
 gaaaacaaat cagaaacatg gcccgttgga gagtcccca aaggagtgga ggaaggttcc 1920
 attgggaagg tctgcagaac tttattggca aagtcggtat tcaacagctt gtatgcatct 1980
 ccacaactgg aaggattttc agctgaatca agaaaactgc ttcttatcgt tcaggctctt 2040
 agggacaacc tggaacctgg gacctttgat cttggggggc tatatgaagc aattgaggag 2100
 tgcctgatta atgatccctg ggttttgctt aatgcttctt ggttcaactc cttcctcaca 2160
 catgcattga gatagtgttg gcaatgctac tatttgctat ccatactgtc caaaaaagta 2220
 ccttgtttct act 2233

<210> 5

<211> 1698

<212> DNA

<213> Influenza A virus

<400> 5

atgaaggctt ttgtactagt cctgttatat gcatttgtag ctacagatgc agacacaata 60
 tgtataggct accatgcgaa caactcaacc gacactgttg acacaatatt cgagaagaat 120
 gtggcagtga cacattctgt taacctgctc gaagacagac acaacgggaa actatgtaaa 180
 ttaaaaggaa tagccccact acaattgggg aaatgtaaca tcaccggatg gctcttggga 240
 aatccagaat ggcactcact gcttcagcg agatcatggt cctacattgt agaaacacca 300
 aactctgaga atggagcatg ttatccagga gatttcacg actatgagga actgaggag 360
 caattgagct cagtatcatc attagaaaga ttcgaaatat ttcccaagga aagttcatgg 420

```

cccaaccaca cattcaacgg agtaacagta tcatgctccc ataggggaaa aagcagtttt 480
tacagaaatt tgctatggct gacgaagaag ggggattcat acccaaagct gaccaattcc 540
tatgtgaaca ataaaggga agaagtcctt gtactatggg gtgttcatca cccgtctagc 600
agtgatgagc aacagagtct ctatagtaat ggaaatgctt atgtctctgt agcgtcttca 660
aattataaca ggagattcac cccggaaata gctgcaaggc ccaaagtaaa agatcaacat 720
gggaggatga actattactg gaccttgcta gaacccggag acacaataat atttgaggca 780
actggtaatc taatagcacc atggtatgct ttcgcaactga gtagagggtt tgagtcgggc 840
atcatcacct caaacgcgtc aatgcatgag tgtaacacga agtgtcaaac accccaggga 900
tctataaaca gcaatctccc ttccagaat atacaccag tcacaatagg agagtgccca 960
aaatatgtca ggagtaccaa attgaggatg gttacaggac taagaaacat cccatccatt 1020
caatacagag gtctatttgg agccattgct ggttttattg aggggggatg gactggaatg 1080
atagatggat ggtatggtta tcatcatcag aatgaacagg gatcaggcta tgcagcggat 1140
caaaaaagca cacagaatgc cattaacagg attacaaaca aggtgaactc tgttatcgag 1200
aaaatgaaca ctcaattcac agctgtgggt aaagaattca acaacttaga aaaaaggatg 1260
gaaaatttaa ataaaaaagt tgatgatggg ttcttgaca ttggacata taatgcagaa 1320
ttgttagttc tactggaaaa tgaaagaact ttggatttcc atgacttaaa tgtgaagaat 1380
ctgtacgaga aagtaaaaag ccaattaaag aataatgcca aagaaatcgg aaatgggtgt 1440
tttgagtict accacaagtg tgacaatgaa tgcattgaaa gtgtaagaaa tgggacttat 1500
gattatccaa aatattcaga agaatacaag ttgaacaggg aaaagataga tggagtgaag 1560
ttggaatcaa tgggggtgta tcagattctg gcgactact caactgtcgc cagttcactg 1620
gtgcttttgg tctccctggg ggcaatcagt ttctggatgt gttctaattg gtctttgcag 1680
tgcagaatat gcacttga 1698

```

<210> 6

<211> 1526

<212> DNA

<213> Influenza A virus

<400> 6

```

tcactcacag agtgacatcg aaatcatggc gaccaaaggc accaaacgat cttacgaaca 60
gatggagact gatggagaac gccagaatgc cactgaaatc agagcatctg tcggaaaaat 120
gattgatgga attggacgat tctacatcca aatgtgcacc gaacttaaac tcagtgatta 180
tgagggacgg ctgattcaga acagcttaac aatagagaga atggtgctct ctgcttttga 240
cgagaggagg aataaatatc tagaagaaca tcccagtgcg gggaaagatc ctaagaaaac 300
tgaggacct atatacagga gagtagatgg aaagtggagg agagaactca tcctttatga 360
caaagaagaa ataagacgaa tctggcgcca agctaataat ggtgacgatg caacggctgg 420
tctgactcac atgatgatct ggcaatccaa ttggaatgat gcaacttacc agaggacaag 480
agctcttgtt cgcacaggaa tggatcccag gatgtgctca ctgatgcagg gttcaaccct 540
ccctaggagg tctggggccg caggtgctgc agtcaaagga gttggaacaa tggatgatga 600
attgatcaga atgatcaaac gtgggatcaa tgatcggaac ttctggaggg gtgagaatgg 660
acggagaaca aggattgctt atgaaagaat gtgcaacatt ctcaaaggga aatttcaaac 720
agctgcacaa agaacaatgg tggatcaagt gagagagagc cggaatccag gaaatgctga 780
gttcgaagat ctcatctttt tagcacggtc tgcactcata ttgagagggt cagttgctca 840
caagtcctgc ctgcctgcct gtgtgtatgg atctgccgta gccagtggat acgactttga 900

```

aagagaggga	tactctctag	tcggaataga	ccctttcaga	ctgcttcaaa	acagccaagt	960
atacagccta	atcagaccaa	atgagaatcc	agcacacaag	agtcaactgg	tgtggatggc	1020
atgccattct	gctgcatttg	aagatctaag	agtatcaagc	ttcatcagag	ggacgaaaagt	1080
ggtcccaaga	gggaagcttt	ccactagagg	agttcaaatt	gcttccaatg	aaaacatgga	1140
gactatggaa	tcaagtaccc	ttgaactgag	aagcagatac	tgggccataa	ggaccagaag	1200
tggagggaac	accaatcaac	agagggcctt	ctcgggcca	atcagcatac	aacctacgtt	1260
ctcagtacag	agaaatctcc	cttttgacag	accaaccatt	atggcagcat	tcactgggaa	1320
tacagagggg	agaacatctg	acatgagaac	cgaatcata	aggctgatgg	aaagtgcaag	1380
accagaagat	gtgtctttcc	aggggcgggg	agtcttcgag	ctctcggacg	aaaaggcaac	1440
gagcccgatc	gtgccctcct	ttgacatgag	taatgaagga	tcttattttct	tcggagacaa	1500
tgcagaggag	tacgacaatt	aaagaa				1526

<210> 7

<211> 1410

<212> DNA

<213> Influenza A virus

<400> 7

ggtcgacctc	cgaagttggg	gggagcgaaa	gcaggagttt	aaatgaatcc	aaaccagaaa	60
ataataacca	ttgggtcaat	ctgtatggta	gtcgggaataa	ttagccta	attgcaaata	120
ggaaatataa	tctcaatatg	gattagccat	tcaattcaaa	ccggaaatca	aaaccatact	180
ggaatatgca	accaaggcag	cattacctat	aaagttgttg	ctgggcagga	ctcaacttca	240
gtgatattaa	ccggcaattc	atctctttgt	cccatccgtg	ggtgggctat	acacagcaaa	300
gacaatggca	taagaattgg	ttccaaagga	gacgtttttg	tcataagaga	gccttttatt	360
tcatgttctc	acttggaaatg	caggaccttt	ttcttgactc	aaggcgctt	actgaatgac	420
aagcattcaa	gggggacctt	taaggacaga	agcccttata	gggccttaat	gagctgccct	480
gtcggatgaag	ctcgtccccc	gtacaattca	aggtttgaat	cggttgcttg	gtcagcaagt	540
gcatgtcatg	atggaatggg	ctggctaaca	atcggaattt	ctgggtccaga	tgatggagca	600
gtggctgtat	taaaatacaa	cggcataata	actgaaacca	taaaaagttg	gaggaagaat	660
atattagaaa	cacaagagtc	tgaatgtacc	tgtgtaaatg	gttcatgttt	taccataatg	720
accgatggcc	caagtgatgg	gctggcctcg	tacaaaattt	tcaagatcga	gaaggggaag	780
gttactaaat	caatagagtt	gaatgcacct	aattctcact	acgaggaatg	ttcctgttac	840
cctgataccg	gcaaagtgat	gtgtgtgtgc	agagacaatt	ggcacggttc	gaaccgacca	900
tgggtgtcct	tcgacaaaaa	cctagattat	aaaataggat	acatctgcag	tggggttttc	960
ggtgacaacc	cgcgtcccaa	agatggaaca	ggcagctgtg	gcccagtgtc	tgctgatgga	1020
gcaaacggag	taaagggatt	ttcatataag	tatggtaatg	gtgtttggat	aggaaggact	1080
aaaagtgaca	gttccagaca	tgggtttgag	atgatttggg	atcctaattg	atggacagag	1140
actgatagta	ggttctctat	gagacaagat	gttgtggcaa	tgactgatcg	gtcagggtac	1200
agcggaagtt	tcgttcaaca	tcctgagcta	acagggctag	actgtatgag	gccttgcttc	1260
tgggttgaat	taatcagggg	gctacctgag	gagaacgcaa	tctggactag	tgggagcatc	1320
atttcttttt	gtgggtgtgaa	tagtgatact	gtagattggg	cttggccaga	cgggtgctgag	1380
ttgccgttca	ccattgacaa	gtagtttgtt				1410

<210> 8
 <211> 900
 <212> DNA
 <213> Influenza A virus

<400> 8
 agcaaaagca ggtagatatt gaaagatgag tcttctaacc gaggtcgaaa cgtacgttct 60
 ctctatcgtc ccgtcaggcc ccctcaaagc cgagatcgca cagagacttg aagatgtctt 120
 tgcagggaag aacaccgatc ttgaggttct catggaatgg ctaaagacaa gaccaatcct 180
 gtcacctctg actaagggga ttttaggatt tgtgttcacg ctcaccgtgc ccagtgagcg 240
 gggactgcag cgtagacgct ttgtccaaaa tgctcttaat gggaacgaag atccaaataa 300
 catggacaaa gcagttaaac tgtgtaggaa gcttaagagg gagataacat tccatggggc 360
 caaagaaata gcactcagtt attctgctgg tgcacttgcc agttgtatgg gcctcatata 420
 caacaggata ggggctgtga ccactgaagt ggcatttggc ctggtatgcg caacctgtga 480
 acagattgct gactcccagc atcgggtctca taggcaaatg gtgacaacaa ccaatccact 540
 aatcagacat gagaacagaa tggttctagc cagcactaca gctaaggcta tggagcaaat 600
 ggctggatcg agtgagcaag cagcagaggc catggatatt gctagtcagg ccaggcaaat 660
 ggtgcaggcg atgagaacca ttgggactca tcctagctcc agtgctggtc taaaagatga 720
 tcttcttgaa aatttgcagg cctatcagaa acgaatgggg gtgcagatgc aacgattcaa 780
 gtgatcctct cgtcattgca gcaaatatca ttggaatctt gcacttgata ttgtggattc 840
 ttgategtct ttttttcaaa tgcatttata gtcgctttaa atacggtttg aaaagagggc 900

<210> 9
 <211> 890
 <212> DNA
 <213> Influenza A virus

<400> 9
 agcaaaagca gggtgacaaa gacataatgg atccaaacac tgtgtcaagc tttcaggtag 60
 attgctttct ttggcatgtc cgcaaaagag ttgcagacca agaactaggt gatgccccat 120
 tccttgatcg gcttcgccga gatcagaagt ccctaagagg aagaggcagc actctcggtc 180
 tggacatcga aacagccacc cgtgctggaa agcaaatagt ggagcggatt ctgaaggaag 240
 aatctgatga ggcactcaaa atgacatgg cctctgtacc tgcacgcgc tacctaactg 300
 acatgactct tgaggaaatg tcaaggcact ggttcatgct catgcccag cagaaagtgg 360
 caggccctct ttgtatcaga atggaccagg cgatcatgga taagaacatc atactgaaag 420
 cgaacttcag tgtgattttt gaccggctgg agactctaatt attactaagg gccttcaccg 480
 aagaggggac aattgttggc gaaatttcac cactgcctc tcttccagga catactgatg 540
 aggatgtcaa aaatgcagtt ggggtcctca tcggaggact tgaatggaat aataacacag 600
 ttcgagtctc tgaaactcta cagagattcg cttggagaag cagtaatgag aatgggagac 660
 ctccactcac tccaaaacag aaacggaaaa tggcgggaac aattaggtca gaagtttgaa 720
 gaaataagat ggttgattga agaagtgaga cacagactga agataacaga gaatagtttt 780
 gagcaataaa catttatgca agccttaca ctattgcttg aagtggagca agagataaga 840
 actttctcgt ttcagcttat ttaataataa aaaacaccct tgtttctact 890

PCT

打印件(原件为电子形式)

0-1	PCT/RO/134表(SAFE)有关保藏的微生物或其他生物材料的说明(PCT细则第13条之二)	
0-1-1	软件版本	CEPCT 版本 1.01.00 MT/FOP 20140331/0.20.5.21
0-2	国际申请号	PCT/CN2016/092778
0-3	申请人或代理人的档案号	IC16086PCN

1	下面的说明与本申请说明书中此处提到的保藏的微生物或其他生物材料相关:	
1-1	页码	8
1-2	行号:	6-8
1-3	保藏事项	
1-3-1	保藏单位名称	中国微生物菌种保藏管理委员会普通微生物中心
1-3-2	保藏单位地址	中国微生物菌种保藏委员会, 中国北京市2714信箱, 邮政编码:100080, Beijing (CN)。
1-3-3	保藏日期	2015年 11月 17日 (17.11.2015)
1-3-4	保藏号	CGMCC 11592
1-4	补充说明	
1-5	本说明是对下列指定国	所有指定国
1-6	单独提交的说明 这些说明将随后提交给国际局	

由受理局填写

0-4	本表格与国际申请一起收到: (是或否)	
0-4-1	授权官员	

由国际局填写

0-5	国际局收到本表格日期:	
0-5-1	授权官员	

权 利 要 求 书

1. 一种突变的病毒，其特征在于所述病毒的至少一种蛋白的编码核酸在终止密码子上游的一个或多个密码子处包含一个或多个 UAG 密码子。

2. 权利要求 1 所述的突变的病毒，其选自下述组中：手足口病毒、柯萨奇病毒、丙肝病毒 HCV、乙肝病毒 HBV、甲肝病毒、丁型肝炎病毒、戊型肝炎病毒、EB 病毒、人乳头瘤病毒 HPV、单纯疱疹病毒 HSV、巨细胞病毒、水痘-带状疱疹病毒、水泡性口炎病毒、呼吸道合胞病毒 RSV、登革病毒、埃博拉病毒、寨卡病毒 Zika、SARS、中东呼吸综合征病毒、轮状病毒、狂犬病毒、麻疹病毒、腺病毒、艾滋病毒、脊髓灰质炎病毒、埃可病毒、乙型脑炎病毒、森林脑炎病毒、汉坦病毒、新型肠道病毒、风疹病毒、腮腺炎病毒、副流感病毒、蓝耳病毒、猪瘟病毒、口蹄疫病毒、细小病毒。

3. 权利要求 1 所述的突变的病毒，其是流感病毒，并且所述流感病毒中选自 PA、PB1、PB2、NP、NA、HA、NS 或 M 蛋白的至少一种蛋白的编码核酸在终止密码子上游的一个或多个密码子处包含一个或多个 UAG 密码子。

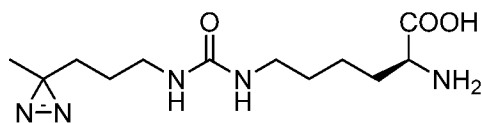
4. 权利要求 3 所述的突变的流感病毒，其中选自 PA、PB1、PB2、NP、NA、HA、NS 或 M 蛋白的至少两种、三种或四种蛋白的编码核酸在终止密码子上游的一个或多个密码子处包含一个或多个 UAG 密码子。

5. 权利要求 3 所述的突变的流感病毒，其中所述 PA、PB1、PB2 和 NP 蛋白中两种、三种或四种蛋白的编码核酸在终止密码子上游的一个或多个密码子处包含一个或多个 UAG 密码子。

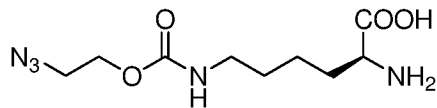
6. 权利要求 3 所述的突变的流感病毒，其中所述流感病毒在选自说明书表 2a)、2b)、2c)、2d)、2e)、2f)、2g)、2h) 和 2i) 中所列举的一个或多个基因位点处包含 UAG 密码子。

7. 权利要求 3 所述的突变的流感病毒，其中所述流感病毒在 PA 蛋白的 R266、PB1 蛋白的 R52、PB2 蛋白的 K33 和/或 NP 蛋白的 D101 的编码核酸密码子处含有 UAG 密码子。

8. 权利要求 1—7 中任一项所述的突变的病毒，其特征在于，所述病毒中，所述蛋白在所述 UAG 密码子对应处包含非天然氨基酸；优选地，所述非天然氨基酸选自：



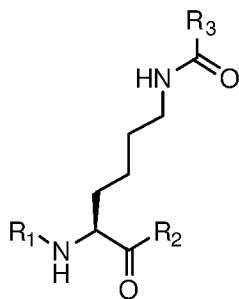
(I) 所示的 Lys-diazirine,



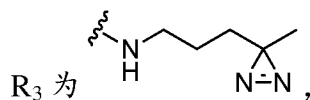
(II) 所示的 Lys- azido, 或

其它含有双吡丙啶、叠氮结构的非天然氨基酸中的至少 1 种。

9. 如权利要求 8 所述的突变的病毒，其中所述的非天然氨基酸是位于第 n 位的 Lys-diazirine，其在病毒蛋白中的连接方式如下式所示：

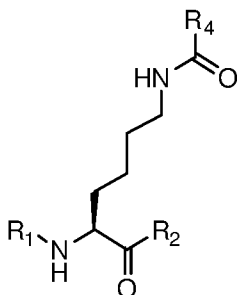


其中, 由 R_1 到 R_2 的方向为氨基酸序列的 N 末端到 C 末端方向, 第 n 位可以是流感病毒的 PA、PB1、PB2、NP、NA、HA、NS 或 M 蛋白中任何一种蛋白质氨基酸序列内的任意位置, 相应地, R_1 为第 1 至第 $n-1$ 位氨基酸残基, R_2 为第 $n+1$ 位至 C 末端的氨基酸残基,

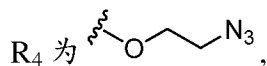


优选地, n 是该蛋白起始密码子下游第 $p-1$ 位到天然终止密码子上游第 q 位内的任意位置, 其中 p 和 q 独立地选自 1、2、3、4、5、6、7、8、9、10、15、20、30、40、50、60、70、80、90、100、125、150、200、250、300、350、400、450、500、600、700 或者直到其全长氨基酸序列长度减去 1。

10. 如权利要求 8 所述的突变的病毒, 其中所述的非天然氨基酸是位于第 n 位的 Lys- azido, 其在病毒蛋白中的连接方式如下式所示:



其中, 由 R_1 到 R_2 的方向为氨基酸序列的 N 末端到 C 末端方向, 第 n 位可以是病毒的 PA、PB1、PB2、NP、NA、HA、NS 或 M 蛋白中任何一种蛋白质的氨基酸序列内的任意位置, 相应地, R_1 为第 1 至第 $n-1$ 位氨基酸残基, R_2 为第 $n+1$ 位至 C 末端的氨基酸残基,



优选地, n 是该蛋白起始密码子下游第 $p-1$ 位到天然终止密码子上游第 q 位内的任意位置, 其中 p 和 q 独立地选自 1、2、3、4、5、6、7、8、9、10、15、20、30、40、50、60、70、80、90、100、125、150、200、250、300、350、400、450、500、600、700 或者直到其全长氨基酸序列长度减去 1。

11. 权利要求 8—10 中任一项的突变的病毒, 其在所述蛋白质的多个位点处包含非天然氨基酸。

12. 如权利要求 11 所述的突变的病毒, 其是流感病毒, 并且在其中 PA 上的 R266、PB1 上的 R52、PB2 上的 K33、NP 上的 D101 处包含非天然氨基酸。

13. 如权利要求 1—12 中任一项所述的病毒, 其中所述病毒的蛋白质编码序列中的正常终止密码子如果是 UAG 的话, 已被突变成 UAA。

14. 如权利要求 3—13 中任一项所述的突变的病毒，其是流感病毒，并且，
未经突变的 PB2 的氨基酸序列与 SEQ ID NO: 2 所示的核酸编码的氨基酸序列相同，
未经突变的 PB1 的氨基酸序列与 SEQ ID NO: 3 所示的核酸编码的氨基酸序列相同，
未经突变的 PA 的氨基酸序列与 SEQ ID NO: 4 所示的核酸编码的氨基酸序列相同，
未经突变的 HA 的氨基酸序列与 SEQ ID NO: 5 所示的核酸编码的氨基酸序列相同，
未经突变的 NP 的氨基酸序列与 SEQ ID NO: 6 所示的核酸编码的氨基酸序列相同，
未经突变的 NA 的氨基酸序列与 SEQ ID NO: 7 所示的核酸编码的氨基酸序列相同，
未经突变的 M 的氨基酸序列与 SEQ ID NO: 8 所示的核酸编码的氨基酸序列相同，或
未经突变的 NS 的氨基酸序列与 SEQ ID NO: 9 所示的核酸编码的氨基酸序列相同。

15. 根据权利要求 14 所述的突变的病毒，其是 WSN-RNP-TAG，在 PA 蛋白的 R266、PB1 蛋白的 R52、PB2 蛋白的 K33 和 NP 蛋白的 D101 的编码密码子处含有 UAG 密码子。

16. 权利要求 3 所述的突变的病毒，其是人源的或其它动物源的流感病毒，优选为 A、B 或 C 型流感病毒。

17. 权利要求 3 所述的突变的病毒，其中包含有说明书表 2a-表 2i 和表 4 中所列举的一个或多个突变。

18. 一种编码如权利要求 1-17 中任一项所述的突变的病毒或其经突变的蛋白的核酸分子，其特征在于在天然终止密码子之外，还包含人为引入的密码子 UAG。

19. 制备权利要求 1—17 中任一项所述突变的病毒的方法，包括如下步骤：

(1) 将包含所述一个或多个被突变基因与合适的载体可操作地连接的突变核酸表达载体和包含其它未被突变基因的一个或多个核酸表达载体转染稳定表达 tRNA (tRNA^{Pyl}) 和吡咯赖氨酰-tRNA 合成酶 (tRNA^{Pyl}) 的动物细胞系，或者与表达 tRNA (tRNA^{Pyl}) 和吡咯赖氨酰-tRNA 合成酶 (tRNA^{Pyl}) 的质粒共同转染动物细胞系其它未被突变基因的一个或多个核酸表达载体；

(2) 将被转染的细胞在含有 Lys-diazirine 或者 Lys-azido 的培养基中培养；和

(3) 回收所述突变病毒。

20. 权利要求 19 所述的方法，其中步骤 (1) 中所述质粒是保藏日为 2011 年 6 月 14 日、保藏号为 CGMCC No: 4951 的大肠埃希氏菌 pACYC-tRNA / PylRS 中的质粒 pACYC-tRNA / PylRS。

21. 权利要求 19 所述的方法，其中所述动物细胞系选自哺乳动物细胞系、禽类动物细胞系、仓鼠细胞系等，优选选自 293 细胞、293T 细胞、Vero 细胞、A549 细胞、Hela 细胞、CHO 细胞、MDCK 细胞、sf9 细胞。

22. 权利要求 19 所述的方法，其中所述稳定表达 tRNA (tRNA^{Pyl}) 和吡咯赖氨酰-tRNA

合成酶 (tRNA^{Pyl}) 的动物细胞系是以保藏号 CGMCC No: 11592 于 2015 年 11 月 17 日被保藏在 CGMCC 的 HEK293-PYL。

23. 筛选减毒病毒的方法, 包括步骤:

(1) 基因突变: 通过基因工程方法将病毒基因组中一个或多个基因的一个或多个选定的密码子突变为 TAG 密码子, 得到一个或多个被突变的基因;

(2) 表达载体构建: 将 (1) 得到的一个或多个被突变的基因与合适的载体可操作地连接, 得到突变基因表达载体;

(3) 将步骤 (2) 中得到的突变基因表达载体和包含其它未被突变基因的一个或多个基因表达载体转染稳定表达 tRNA (tRNA^{Pyl}) 和吡咯赖氨酰-tRNA 合成酶 (tRNA^{Pyl}) 的动物细胞系, 或者与表达 tRNA (tRNA^{Pyl}) 和吡咯赖氨酰-tRNA 合成酶 (tRNA^{Pyl}) 的质粒共同转染动物细胞系;

(4) 将被转染的细胞在含有 Lys-diazirine 或者 Lys-azido 的培养基中培养, 收集培养基上清液, 采用包含 Lys-diazirine 或者 Lys-azido 的培养板对上清液进行病毒克隆的非天然氨基酸依赖性的检测; 和

(5) 将维持了对 Lys-diazirine 或者 Lys-azido 非天然氨基酸依赖性的突变体鉴定为减毒的病毒。

24. 权利要求 23 的方法, 其在步骤 (5) 之后还包括:

(6) 将 (5) 中定点突变成功的候选物的突变位点进行组合, 再次包装病毒, 使得包装出的病毒在多个基因片段上均引入了 UAG, 在对应的多个蛋白上均引入了非天然氨基酸, 然后再次考察包装产物对非天然氨基酸的依赖性, 保留经过长期传代仍旧维持着对非天然氨基酸依赖性的组合突变体设定为最优的定点突变成功候选物。

25. 权利要求 24 的方法, 其中在步骤 (6) 之后还进一步包括:

(7) 对 (6) 中获得的突变型病毒进行安全性, 鉴定出安全的病毒。

26. 权利要求 23 所述的方法, 其中所述病毒选自下述组中: 手足口病毒、柯萨奇病毒、丙肝病毒 HCV、乙肝病毒 HBV、甲肝病毒、丁型肝炎病毒、戊型肝炎病毒、EB 病毒、人乳头瘤病毒 HPV、单纯疱疹病毒 HSV、巨细胞病毒、水痘-带状疱疹病毒、水泡性口炎病毒、呼吸道合胞病毒 RSV、登革病毒、埃博拉病毒、寨卡病毒 Zika、SARS、中东呼吸综合征病毒、轮状病毒、狂犬病毒、麻疹病毒、腺病毒、艾滋病毒、脊髓灰质炎病毒、埃可病毒、乙型脑炎病毒、森林脑炎病毒、汉坦病毒、新型肠道病毒、风疹病毒、腮腺炎病毒、副流感病毒、蓝耳病毒、猪瘟病毒、口蹄疫病毒、细小病毒。

27. 权利要求 23 所述的方法, 其中所述一个或多个选定的密码子独立地位于编码序列第 n 位, n 是该蛋白起始密码子下游第 p-1 位到天然终止密码子上游第 q 位内的任意位置, 其中 p 和 q 独立地选自 1、2、3、4、5、6、7、8、9、10、15、20、30、40、50、60、70、80、90、100、125、150、200、250、300、350、400、450、500、600、700 或者直到其全长氨基酸序列长度减去 1。

28. 权利要求 23 所述的方法, 其是流感病毒。

29. 权利要求 23 所述的方法, 其中步骤 (3) 中所述质粒是保藏日为 2011 年 6 月 14 日、保藏号为 CGMCC No: 4951 的大肠埃希氏菌 pACYC-tRNA / PylRS 中的质粒 pACYC-tRNA / PylRS。

30. 权利要求 23 所述的方法, 其中所述动物细胞系选自细胞系哺乳动物、禽类动物细胞系、仓鼠细胞系等, 优选选自 293 细胞、293T 细胞、Vero 细胞、A549 细胞、Hela 细胞、CHO

细胞、MDCK 细胞、sf9 细胞。。

31. 权利要求 23 所述的方法,其中所述稳定表达 tRNA(tRNA^{Pyl})和吡咯赖氨酰-tRNA 合成酶 (tRNA^{Pyl}) 的动物细胞系是以保藏号 CGMCC No: 11592 于 2015 年 11 月 17 日被保藏在 CGMCC 的 HEK293-PYL。

32. 权利要求 23 所述的方法,其特征在于重复突变、包装和筛选步骤。

33. 权利要求 23—32 中任一项所述的方法,其中所述流感病毒是 A 型、B 型或 C 型流感病毒。

34. 含有有效量的权利要求 1-17 中任一项所述的突变的流感病毒的组合物。

35. 含有有效量的权利要求 1-17 中任一项所述的突变的流感病毒的疫苗。

36. 药物组合物,其中含有有效量的权利要求 1-17 中任一项所述的突变的流感病毒,以及药学上可以接受的赋形剂。

37. 权利要求 1-17 中任一项所述的突变的流感病毒在制备减毒活疫苗或者制备预防或治疗流感病毒感染相关药物中的用途。

38. 权利要求 1-17 中任一项所述的突变的流感病毒在预防和治疗流感病毒感染中的用途。

39. 可稳定表达 tRNA (tRNA^{Pyl}) 和吡咯赖氨酰-tRNA 合成酶 (tRNA^{Pyl}) 的哺乳动物稳定细胞系,其为 HEK293-PYL,保藏日为 2015 年 11 月 17 日,其保藏号为 CGMCC No: 11592。

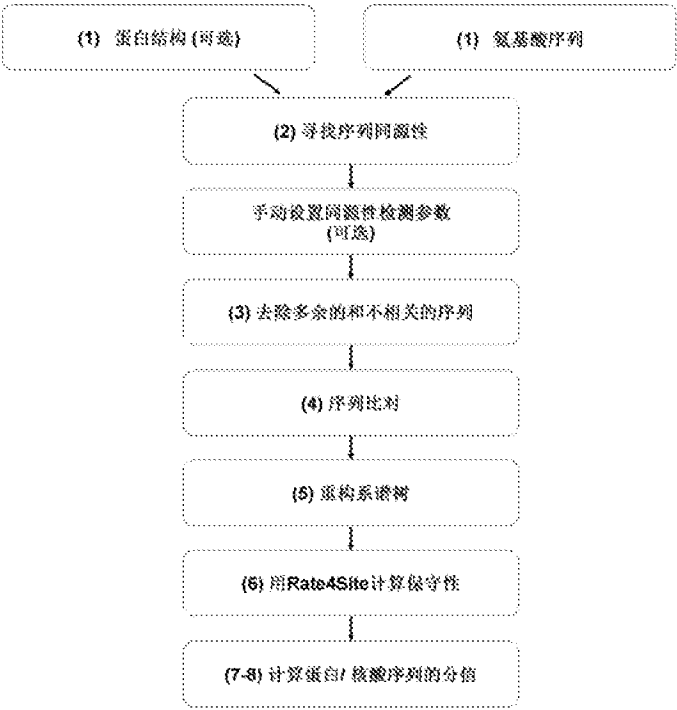


图 1

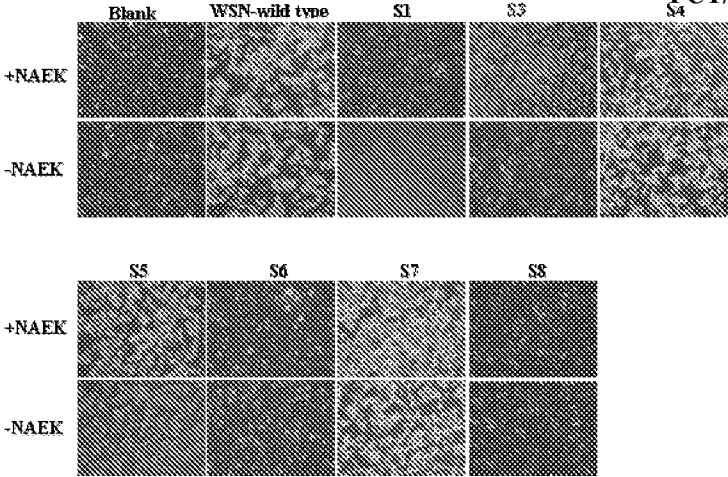


图 2A

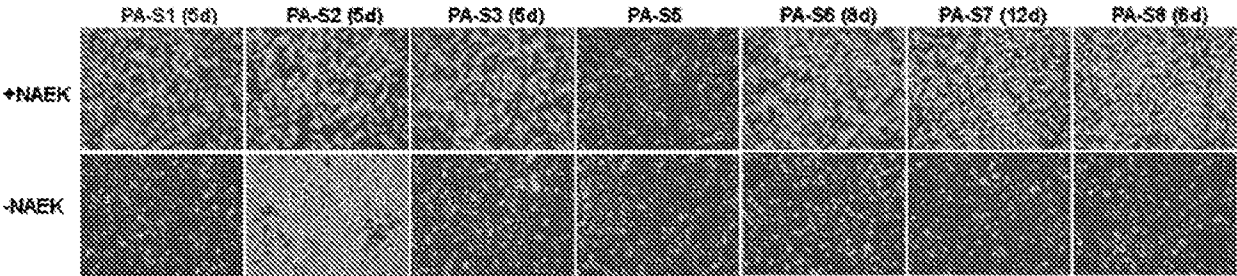


图 2B

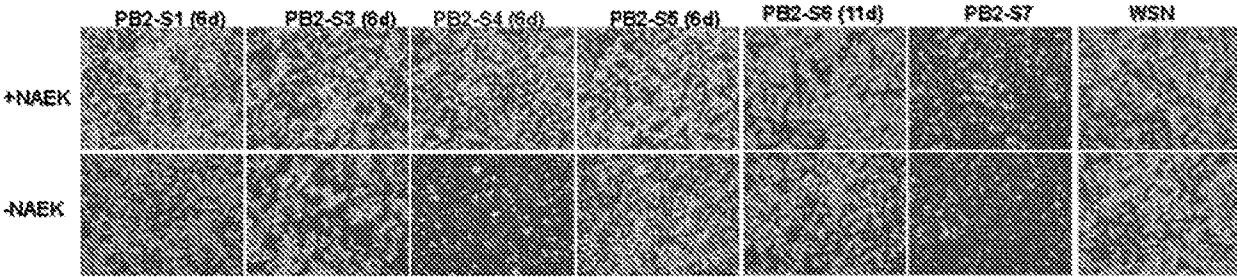


图 2C

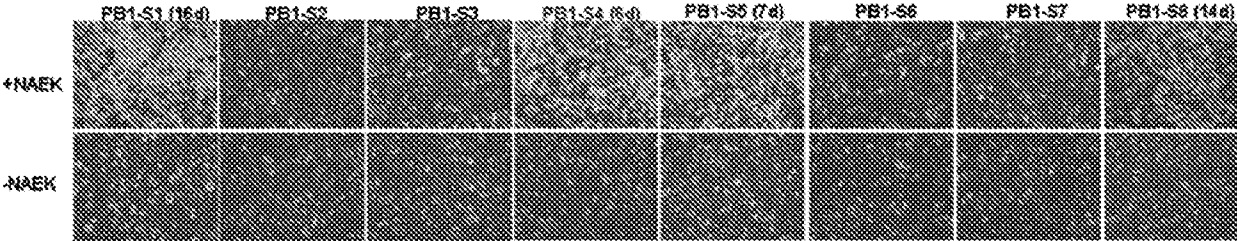


图 2D

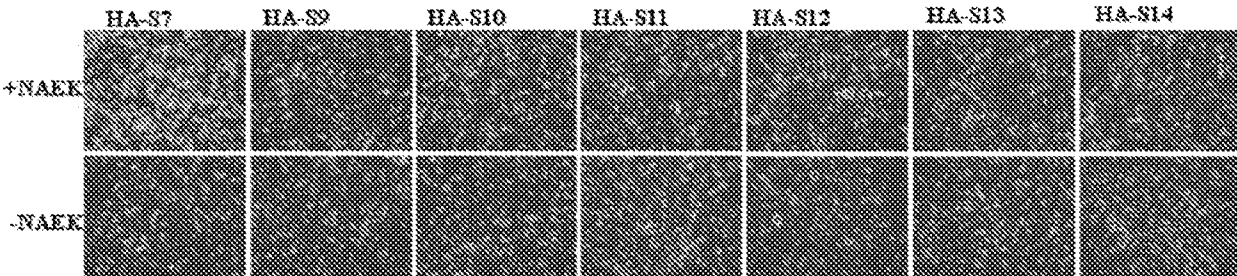


图 2E

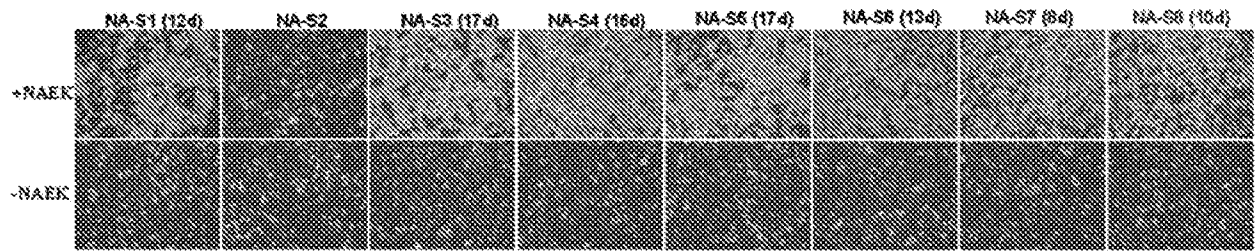


图 2F

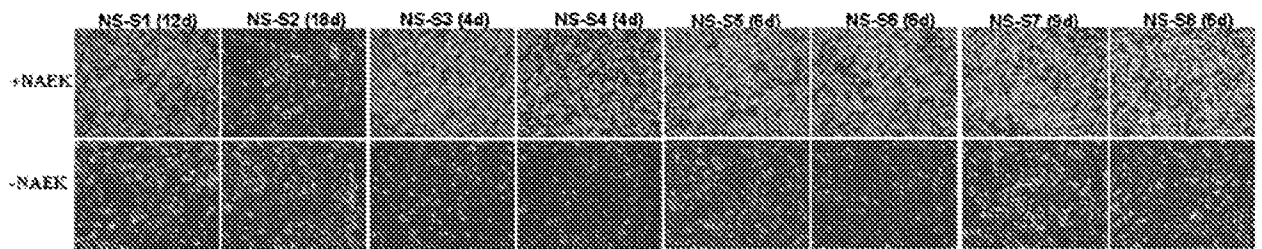


图 2G

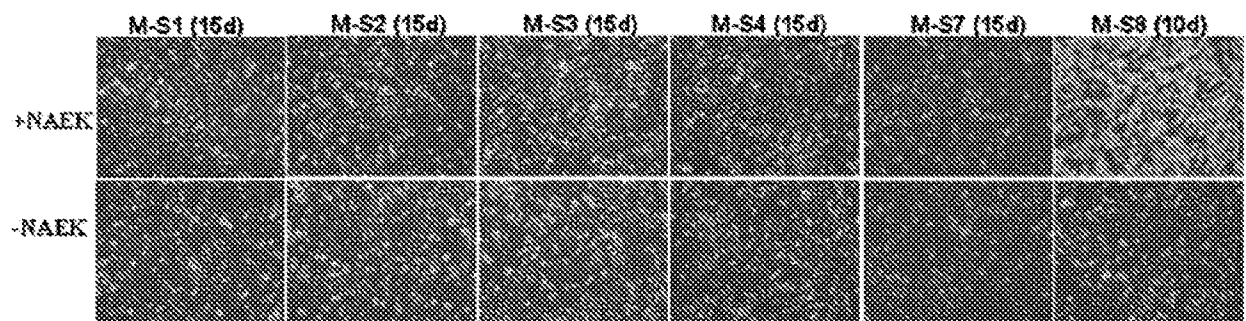


图 2H

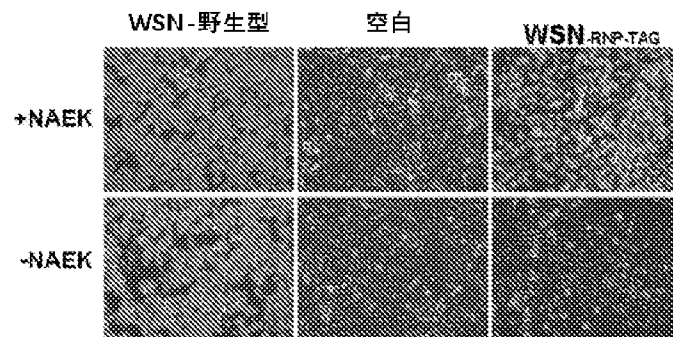


图 3

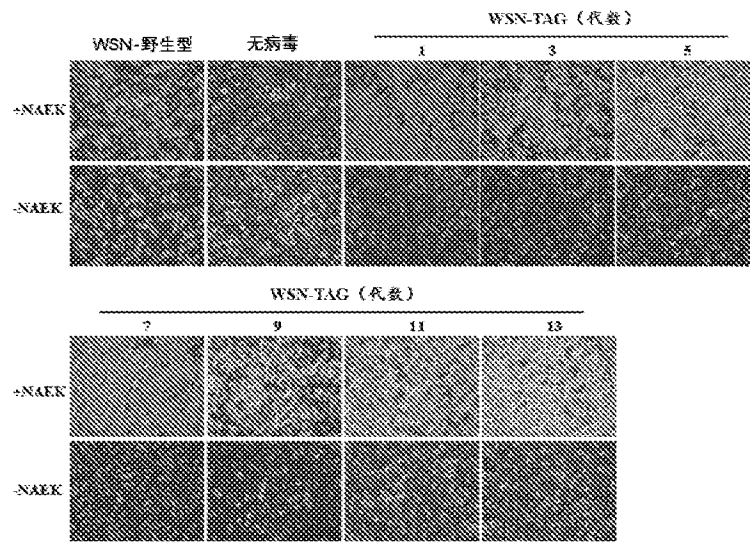


图 4

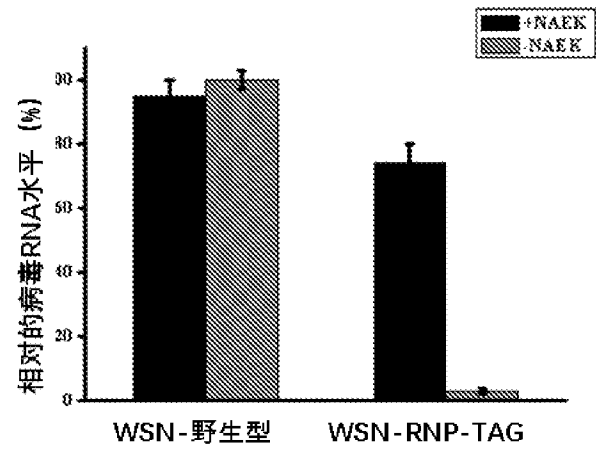


图 5

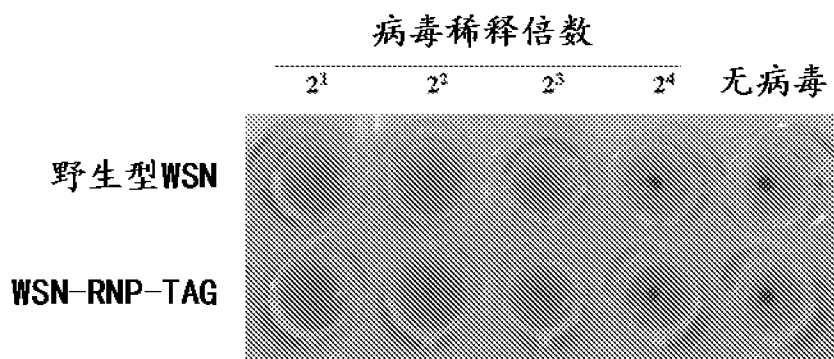
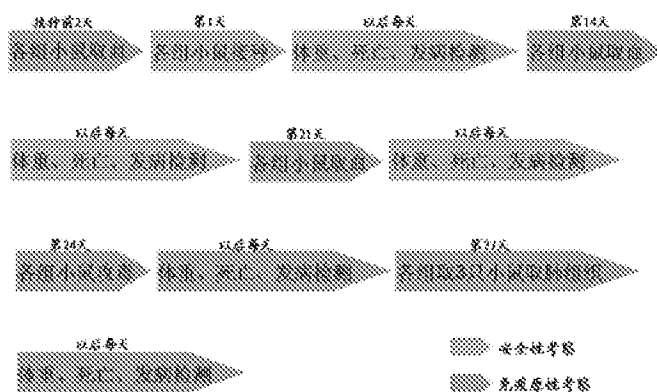


图 6

动物安全性和免疫原性实验流程



取血用于检测抗体；取肺组织用于检测肺内病毒滴度。

图 7

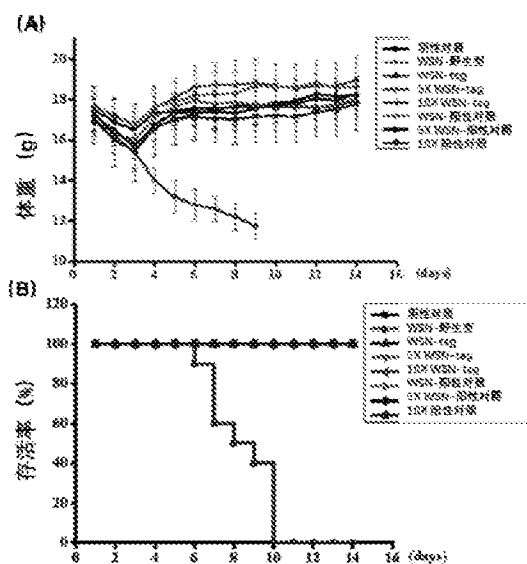


图 8

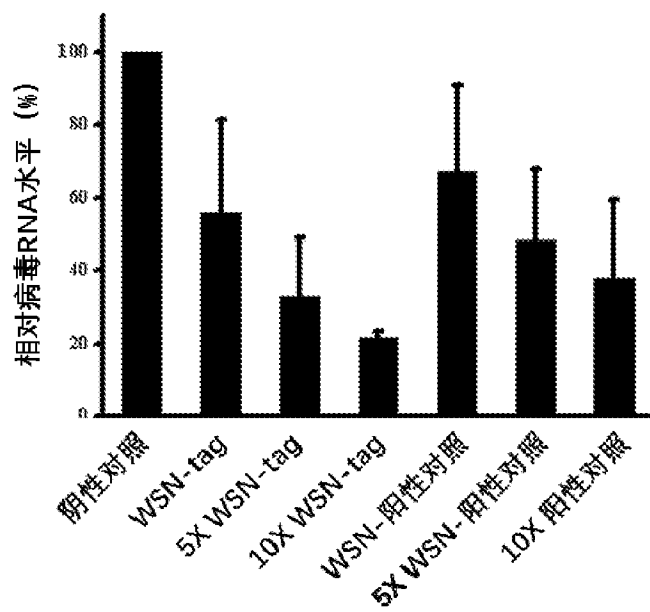


图 9A

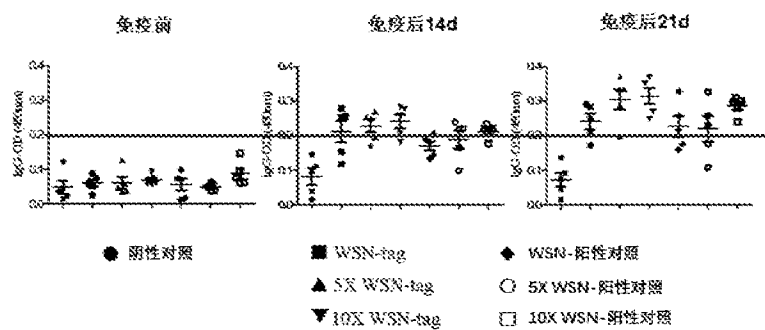


图 9B

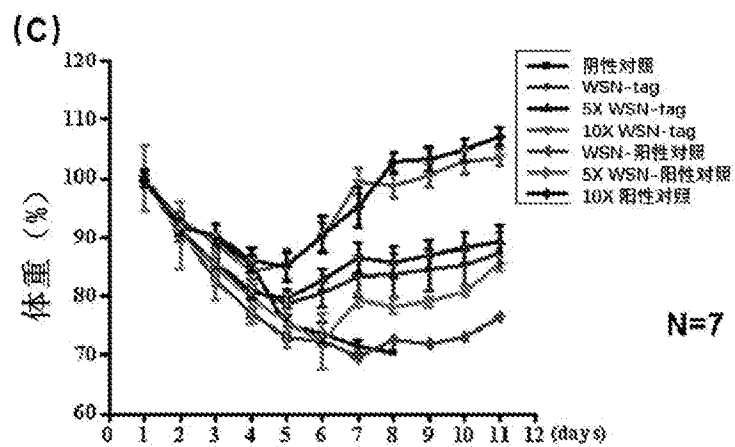


图 9C

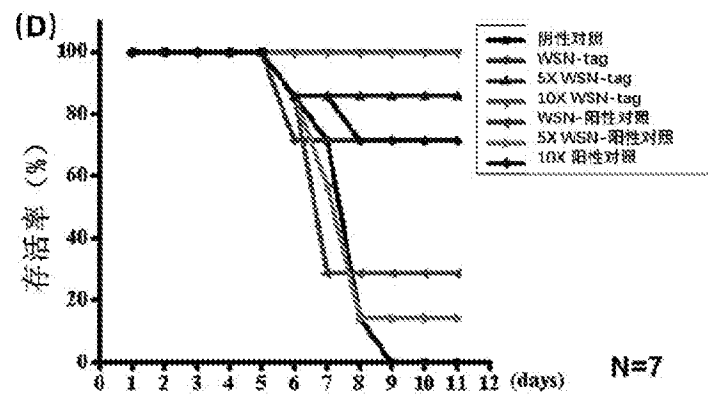


图 9D

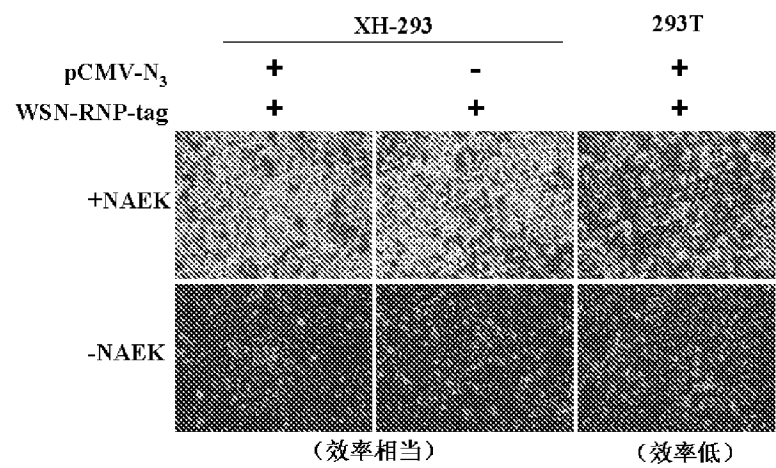


图 10

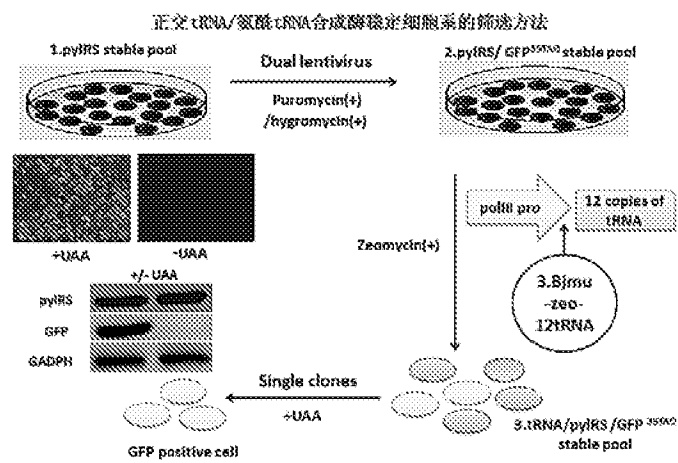


图 11A.

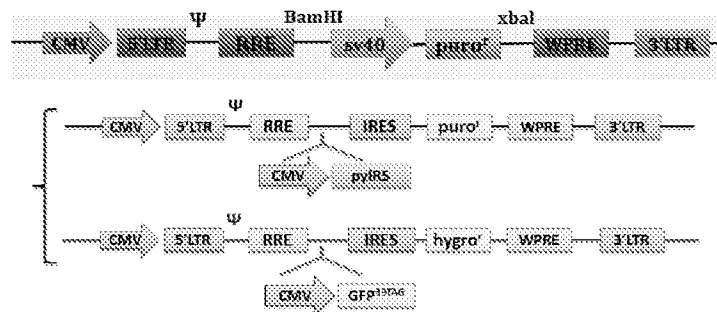


图 11B.

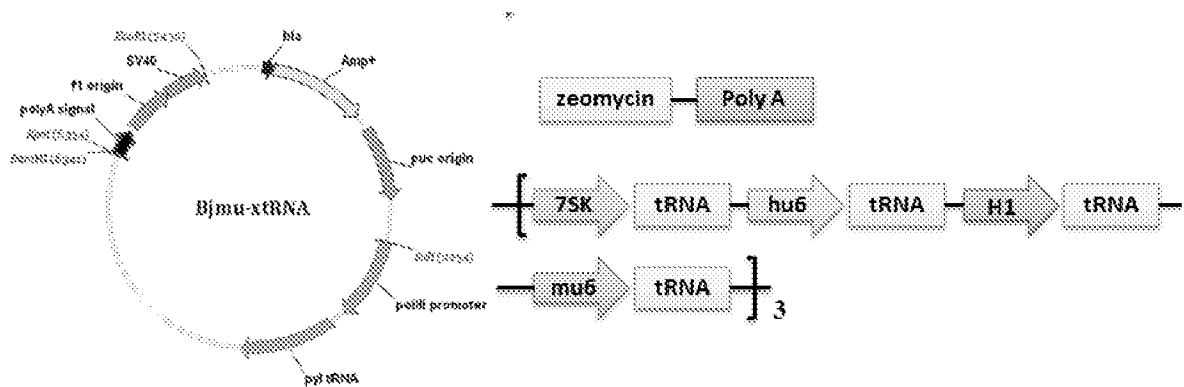


图 11C

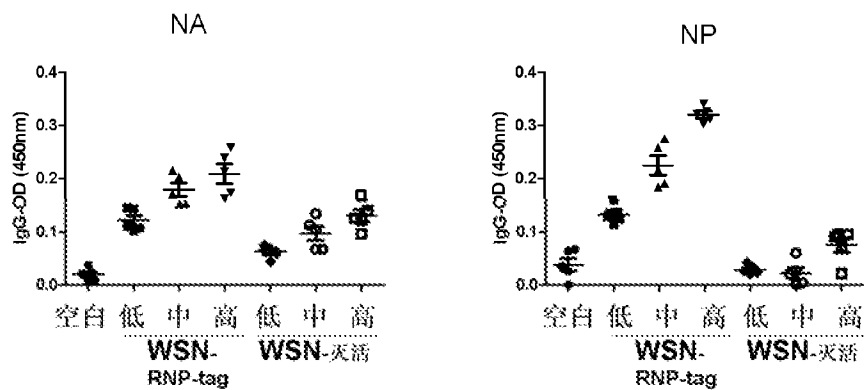


图 12

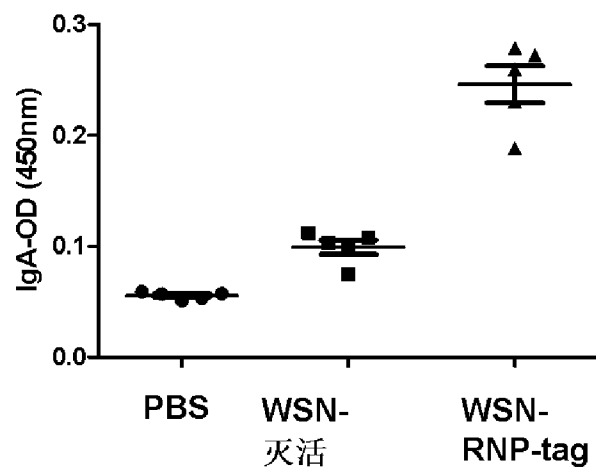


图 13

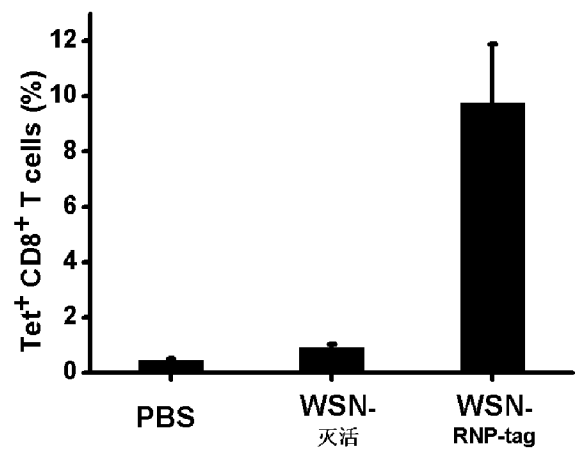


图 14

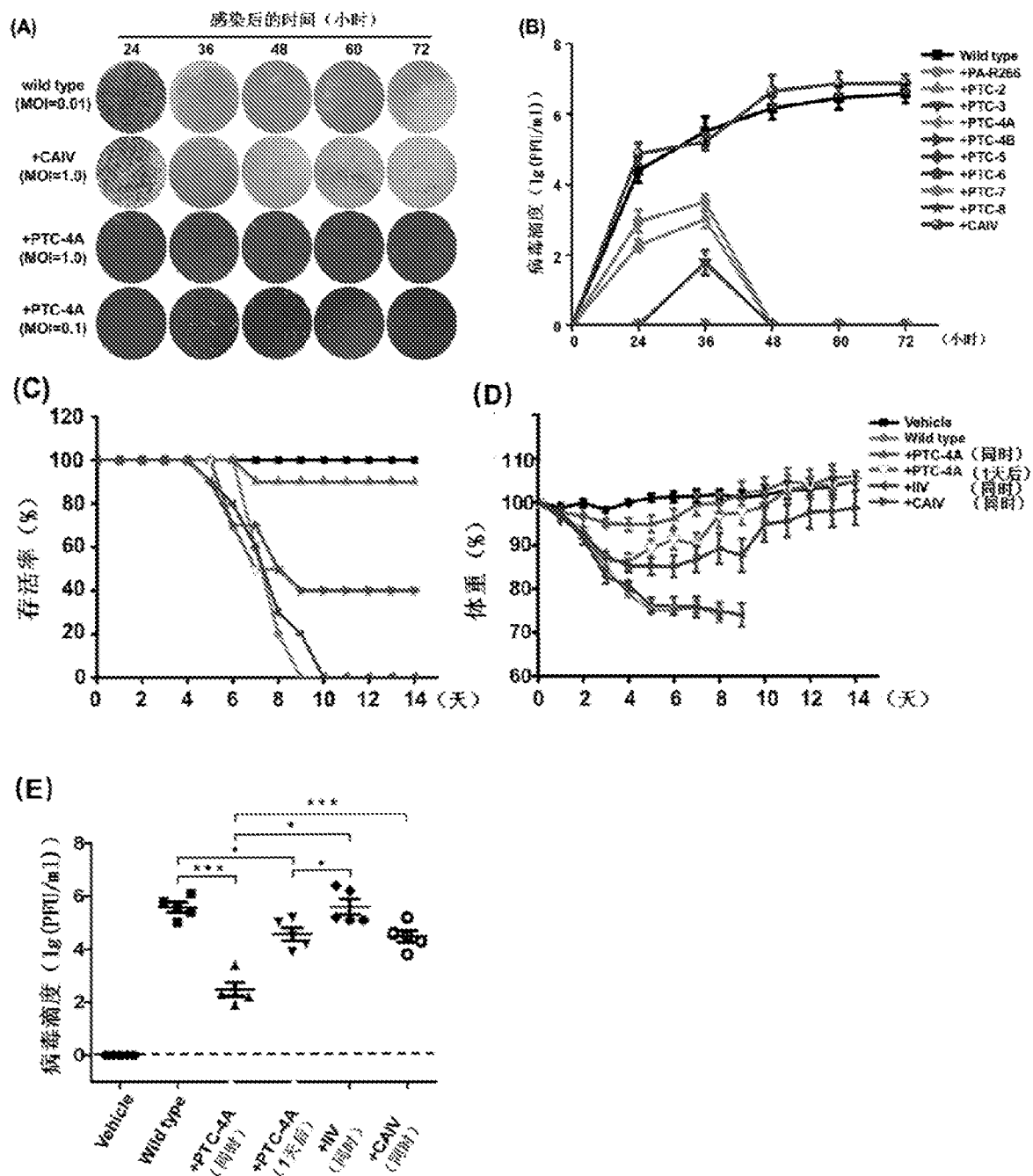


图 15

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2016/092778

A. CLASSIFICATION OF SUBJECT MATTER

C12N 7/01 (2006.01) i; C07K 14/11 (2006.01) i; C12N 15/44 (2006.01) i; C12N 15/63 (2006.01) i; A61K 39/145 (2006.01) i; A61K 35/76 (2015.01) i; A61P 31/16 (2006.01) i; C12N 5/10 (2006.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

C12N; C07K; A61K; A61P

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNABS, CPRSABS, DWPI, SIPOABS, CPEA, CNKI, ISI Web of Science: virus, mutat+, codon, unnatural amino acid, non-natural amino acid, UAG, TAG

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 104592364 A (PEKING UNIVERSITY), 06 May 2015 (06.05.2015), see description, paragraphs 6-82	1, 2, 8, 11, 18
A	YANG, Xia et al. "Genetic Code Expanding Technology", CHEMISTRY OF LIFE, vol.24, no.6, 15 December 2004 (15.12.2004), ISSN: 1000-1336, see pages 512-513	1-37, 39
A	TAKAHIRO, H. et al. "Incorporation of Non-Natural Amino Acids into Proteins", CURR OPIN CHEM BIOL., vol.6, no.6, 01 December 2002 (01.12.2002), ISSN: 1367-5931, see pages 809-815	1-37, 39

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier application or patent but published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 25 August 2016 (25.08.2016)	Date of mailing of the international search report 07 September 2016 (07.09.2016)
Name and mailing address of the ISA/CN: State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No.: (86-10) 62019451	Authorized officer XU, Yijun Telephone No.: (86-10) 62411083

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2016/092778

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing filed or furnished:
 - a. (means)
 - ☐ on paper
 - ☒ in electronic form
 - b. (time)
 - ☐ in the international application as filed
 - ☒ together with the international application in electronic form
 - ☐ subsequently to this Authority for the purposes of search
2. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.
3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2016/092778

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.: 38
because they relate to subject matter not required to be searched by this Authority, namely:
[1] claim 38 relate to a method for treating a human or animal body.
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/CN2016/092778

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 104592364 A	06 May 2015	WO 2015062516 A1	07 May 2015

A. 主题的分类		
C12N 7/01(2006.01)i; C07K 14/11(2006.01)i; C12N 15/44(2006.01)i; C12N 15/63(2006.01)i; A61K 39/145(2006.01)i; A61K 35/76(2015.01)i; A61P 31/16(2006.01)i; C12N 5/10(2006.01)i		
按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类		
B. 检索领域		
检索的最低限度文献(标明分类系统和分类号)		
C12N; C07K; A61K; A61P		
包含在检索领域中的除最低限度文献以外的检索文献		
在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))		
CNABS, CPRSABS, DWPI, SIPOABS, CPEA, CNKI, ISI Web of Science: 病毒, 突变, 密码子, 非天然氨基酸, virus, mutat+, codon, unnatural amino acid, non-natural amino acid, UAG, TAG		
C. 相关文件		
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
X	CN 104592364 A (北京大学) 2015年 5月 6日 (2015 - 05 - 06) 参见说明书第6-82段	1, 2, 8, 11, 18
A	杨侠等. "遗传密码扩充技术" 生命的化学, 第24卷, 第6期, 2004年 12月 15日 (2004 - 12 - 15), ISSN: 1000-1336, 参见第512-513页	1-37, 39
A	Takahiro Hohsaka等. "Incorporation of non-natural amino acids into proteins" Curr Opin Chem Biol., 第6卷, 第6期, 2002年 12月 1日 (2002 - 12 - 01), ISSN: 1367-5931, 参见第809-815页	1-37, 39
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。		
* 引用文件的具体类型: "A" 认为不特别相关的表示了现有技术一般状态的文件 "E" 在国际申请日的当天或之后公布的在先申请或专利 "L" 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的) "O" 涉及口头公开、使用、展览或其他方式公开的文件 "P" 公布日先于国际申请日但迟于所要求的优先权日的文件 "T" 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 "X" 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 "Y" 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 "&" 同族专利的文件		
国际检索实际完成的日期		国际检索报告邮寄日期
2016年 8月 25日		2016年 9月 7日
ISA/CN的名称和邮寄地址		受权官员
中华人民共和国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088		徐益君
传真号 (86-10)62019451		电话号码 (86-10)62411083

第I栏

核苷酸和/或氨基酸序列(续第1页第1.c项)

1. 关于国际申请中所公开的任何对要求保护的发明必要的核苷酸和/或氨基酸序列，国际检索是在下列基础上进行的：

a. (提交提供)

☐

纸件形式

☒

电子形式

b. (提交时间)

☐

含在申请提交时的国际申请中

☒

以电子形式与国际申请一起提交

☐

为检索之用随后提交本单位

2. ☐ 另外，在提交/提供了多个版本或副本的序列表的情况下，提供了关于随后提交的或附加的副本中的信息与申请时提交的申请中的信息相同或未超出申请时提交的申请中的信息范围（如适用）的所需声明。

3. 补充意见：

第II栏

某些权利要求被认为是不能检索的意见(续第1页第2项)

根据条约第17条(2)(a)，对某些权利要求未做国际检索报告的理由如下：

1. ☒ 权利要求： 38
因为它们涉及不要求本单位进行检索的主题，即：
[1] 权利要求38涉及人体或动物体的治疗方法。
2. ☐ 权利要求：
因为它们涉及国际申请中不符合规定的要求的部分，以致不能进行任何有意义的国际检索， 具体地说：
3. ☐ 权利要求：
因为它们是从属权利要求，并且没有按照细则6.4(a)第2句和第3句的要求撰写。

国际申请号
PCT/CN2016/092778

检索报告引用的专利文件	公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN 104592364 A	2015年 5月 6日	WO 2015062516 A1	2015年 5月 7日