



(12)发明专利

(10)授权公告号 CN 106255758 B

(45)授权公告日 2020.06.23

(21)申请号 201480071836.9

(22)申请日 2014.11.03

(65)同一申请的已公布的文献号
申请公布号 CN 106255758 A

(43)申请公布日 2016.12.21

(30)优先权数据
2013904242 2013.11.01 AU
2014900370 2014.02.07 AU

(85)PCT国际申请进入国家阶段日
2016.06.30

(86)PCT国际申请的申请数据
PCT/AU2014/050330 2014.11.03

(87)PCT国际申请的公布数据
W02015/061858 EN 2015.05.07

(73)专利权人 赛门蒂斯有限公司
地址 澳大利亚维多利亚

(72)发明人 P·M·豪利 刘良

(74)专利代理机构 北京安信方达知识产权代理
有限公司 11262

代理人 郑霞

(51)Int.Cl.

C12N 15/863(2006.01)

C12N 5/071(2006.01)

C12N 5/10(2006.01)

C12N 7/00(2006.01)

C12N 15/33(2006.01)

(56)对比文件

CN 101541972 A,2009.09.23,全文.

YIFAN ZHANG, et al..Immature Viral Envelope Formation Is interrupted at the Same Stage by lac Operator-Mediated Repression of the Vaccinia Virus D13L Gene and by the Drug Rifampicin.《VIROLOGY》.1992,第187卷(第2期),第652页左栏第2段.

ANNA L. RAMSEY-EWING, et al..Complementation of a Vaccinia Virus Host-Range K1L Gene Deletion by the Nonhomologous CP77 Gene.《VIROLOGY》.1996,第222卷(第1期),第75页摘要,第78页左栏倒数第1段-右栏第2段.

审查员 陈皓

权利要求书1页 说明书42页
序列表22页

(54)发明名称

病毒载体制造

(57)摘要

本发明涉及经修饰的哺乳动物细胞,其中该细胞的基因组被修饰以包含在启动子的控制下表达CP77的序列,使得经修饰的细胞系维持在未经修饰的细胞中较不能繁殖或不能繁殖的痘病毒的繁殖。

1. 一种经修饰的中国仓鼠卵巢 (CHO) 细胞, 其中所述经修饰的CHO细胞的基因组被修饰以包含在组成型启动子的控制下编码CP77的序列, 使得所述经修饰的CHO细胞维持在未经修饰的CHO细胞中较不能繁殖或不能繁殖的痘苗病毒的繁殖, 并且其中所述经修饰的CHO细胞的基因组还包含在组成型启动子的控制下编码D13L的序列。

2. 根据权利要求1所述的经修饰的CHO细胞, 其中所述经修饰的CHO细胞的基因组包含在组成型启动子的控制下编码K1L的序列。

3. 根据权利要求1或2所述的经修饰的CHO细胞, 其中所述经修饰的CHO细胞是连续细胞系。

4. 根据权利要求1至3中任一项所述的经修饰的CHO细胞, 其中所述CP77基因的表达支持所述病毒的繁殖, 以产生与在允许细胞系中观察到的病毒产量等价的病毒产量。

5. 根据权利要求1至4中任一项所述的经修饰的CHO细胞, 其中所述CP77基因的表达支持大于500的病毒复制扩增率。

6. 根据权利要求1至5中任一项所述的经修饰的CHO细胞, 其中所述CP77由为了在哺乳动物细胞中表达而密码子优化的连续的核苷酸序列编码。

7. 一种用于繁殖在中国仓鼠卵巢 (CHO) 细胞中不繁殖的痘苗病毒的方法, 所述方法包括在CHO细胞中体外繁殖所述病毒, 其中所述CHO细胞被修饰以在组成型启动子的控制下编码并表达CP77并在组成型启动子的控制下表达D13L。

8. 根据权利要求7所述的方法, 其中所述CHO细胞被修饰以在组成型启动子的控制下编码并表达K1L。

病毒载体制造

[0001] 相关申请

[0002] 本申请与2013年11月1日提交的澳大利亚专利申请号2013904242和2014年2月7日提交的澳大利亚专利申请号2014900370相关并且要求其优先权,这些申请的每一个的完整内容通过引用并入本文。

发明领域

[0003] 本发明涉及适于繁殖并因此适于制造基于痘病毒的药物的细胞和细胞系的开发。特别地,本说明书涉及,为了用于繁殖这样的痘病毒以用于制造治疗剂或预防剂的重组的经修饰的细胞基底。

[0004] 背景

[0005] 本说明书的参考文献的著录细节在说明书的末尾列出。

[0006] 在本说明书中对任何在先出版物(或来自它的信息)或者任何已知的事物的参考不能且不应认为是承认或认可或者以任何形式暗示,该在先出版物(或来自它的信息)或者已知事物形成本说明书所属技术领域中的公知常识的一部分。

[0007] 本说明书中提及的所有出版物通过引用以其整体并入本文。

[0008] 痘病毒家族包括2个亚家族,脊索动物痘病毒亚科(Chordopoxvirinae)和昆虫痘病毒亚科(Entomopoxvirinae)。脊索动物痘病毒亚科包括8个属,包括包含感染人的种(例如,天花病毒(variola virus)、天花的致病因子、牛痘病毒(cowpox virus)(其形成詹纳在1796年报道的原始天花疫苗)、痘苗病毒(vaccinia virus)(被用作第二代天花疫苗)和猴痘病毒(monkeypox virus))的正痘病毒属(Orthopoxviridae)、和包含感染鸟的种的禽痘病毒属(Avipoxviridae)病毒,诸如鸡痘病毒(fowlpox virus)和金丝雀痘病毒(canarypox virus)。除了它们作为抗原在天花疫苗中的用途,对重组的基于痘苗病毒的病毒和禽痘病毒作为“骨架”载体的用途存在很大兴趣。作为细胞质内载体,正痘病毒属能够,除其他以外,递送外源抗原至宿主细胞质和将抗原加工成用于在细胞表面呈现的肽的抗原加工通路。这样的表达外源抗原的载体在针对疾病诸如AIDS、结核病、疟疾和癌症的疫苗的开发中使用,所述疾病已被证明难以用其他疫苗接种策略治疗。

[0009] 正痘病毒属具有大小范围从副痘病毒的130kb至禽痘病毒(avipoxvirus)的超过300kb的线性双链DNA基因组,并且其在宿主中的生命周期完全在宿主细胞质中度过。痘病毒大体上独立于其宿主细胞和宿主细胞分子工作,特别是对于参与早期mRNA合成的过程。但是,宿主分子表现为被用于中间和晚期病毒转录的起始或终止。痘病毒产生结构多样的“宿主范围因子(host range factor)”,其特异性地靶向并操控宿主信号通路以准许允许病毒复制的细胞条件。多数痘病毒可以结合并感染哺乳动物细胞,但继发感染是允许的(能够产生感染性的病毒粒)或非-允许的(大体上不能产生感染性的病毒粒)取决于参与的具体痘病毒和具体的细胞类型。目前,对分子水平的痘病毒-宿主相互作用存在相对少的了解,特别是宿主-范围基因,以及哪些因子是调节关系以有利于病毒和细胞繁殖二者必要的。对于宿主范围基因参考文献的综述,可以参考全文并入本文的Werden等人,2008。

[0010] 有关痘苗病毒毒株关于其作为天花病毒疫苗和随后作为病毒载体的观察,已经从1960年代早期发表至今。痘苗病毒的某些毒株,包括用作天花疫苗的毒株,能够在人细胞中繁殖并因此造成健康风险,诸如病毒性脑炎的发展。为了开发更安全的疫苗,来自安卡拉的痘苗病毒毒株(被称为“CVA”)在非-人细胞中传代超过500次。在此过程中痘苗病毒基因组改变,大体上包括与原始CVA基因组相比的至少6个主要缺失的发生。经修饰的病毒在人中致病性更低,但仍能够产生保护性免疫应答。此减毒的痘苗病毒被称为MVA(经修饰的痘苗病毒安卡拉)并且还由传代次数分类,因为发现具有不同传代次数的病毒在遗传上和表型上不同。但是,传代次数515MVA515被认为是遗传稳定的。在1990年代早期观察到,MVA毒株,诸如MVA572及其衍生物MVA F6,能够在非-允许细胞(其中病毒将不繁殖)中以高水平表达痘苗病毒蛋白和异源(重组的)蛋白,使得能够开发MVA作为感兴趣的异源分子的载体,所述感兴趣的异源分子诸如编码用于疫苗或疗法递送的抗原的那些。

[0011] 最近,已经尝试通过向CVA引入MVA的6个大的、已知的缺失来生产具有MVA性质的经修饰的痘苗病毒。有趣的是,这不产生具有减弱的MVA性质的病毒。有人提出,宿主范围基因的缺乏可能对观察的减毒负责,但是,这还未被证实(参见,例如Meyer等人, *Journal of General Virology* (1991) 72:1031-1038)。

[0012] 痘病毒组成特征在于大的、线性dsDNA基因组、用于繁殖的细胞质位点和复杂的病毒粒形态学的病毒的大家族。痘苗病毒是此组病毒的代表性病毒并且是关于病毒形态发生研究的最多的病毒。痘苗病毒病毒粒以“砖形”或以“卵形”与膜结合的颗粒出现,具有复杂的内部结构,该内部结构以两侧具有“侧体(lateral bodies)”的、被围住的、两头凹陷的核心为特征。病毒粒装配通路包括包含新月形的膜的制造,该膜发展成不成熟的病毒粒(IV),并且随后演化成成熟病毒粒(MV)。超过70种特定的基因产物被包含于痘苗病毒病毒粒内,其中已经描述了超过50种特定的基因中的突变对痘苗病毒装配的影响。

[0013] 痘苗病毒通过其表面膜与宿主细胞的细胞质膜的融合进入细胞,释放核心(和侧体)到细胞质并激活病毒的转录程序。病毒粒核心包含早期mRNA合成和修饰所需的全部病毒-编码酶。早期基因编码DNA增殖所需的基因,并且因此,随着早期基因表达达到峰值,病毒DNA增殖接着在被称为“工厂”的细胞质位点发生。早期基因还编码中间转录因子,并且中间基因反过来编码晚期转录因子,这样中间基因和晚期基因在病毒DNA增殖的前提条件的起始后连续表达。因此,全部病毒基因以时间串联转录,早期、中间和晚期类通过类别-特异性转录启动子和病毒编码的转录因子区分。此外,只有增殖的基因组是用于中间和晚期转录的合格的模板。这两类基因一起编码病毒粒结构蛋白、病毒粒酶和装配因子,并且是装配新的后代病毒颗粒所需的。

[0014] 病毒摄入和早期表达后不久,在细胞内形成感染-特异性细胞质结构域,该结构域的密度一致并且有时被大小随着时间增加的来自内质网(ER)的潴泡包围。这些结构域代表病毒DNA增殖的位点并且经常被称为“病毒工厂”。

[0015] 病毒装配在病毒工厂内随着刚性的新月形结构(三维的杯形器)的形成开始。在高分辨电子显微镜下,这些新月形结构的外层由规则隔开的被称为“针(spicule)”的突出物组成。新月表面上沿长度生长而保持相同的曲度,直到它们成为被称为不成熟的病毒粒(IV)的闭合的圆(三维的球)。IV被密度一致但具有比周围工厂可辨别的更大的电子密度的“病毒原质体”材料充满。随着IV的形成,还发生衣壳化的DNA的摄入;这些在电子显微照片

中作为在IV内的电子密集、圆形或卵形亚结构域观察到,被称为“拟核”。包含浓缩的DNA的拟核的IV经常被称为“IVN”。IVN向成熟病毒粒(MV)的形态发生要求数个病毒粒蛋白前体通过蛋白水解裂解的成熟。大多数成熟病毒粒在工厂外被发现并可以以簇存在于工厂周围或以从最近的工厂的显著距离明显地分开。

[0016] 痘病毒病毒粒以3种感染性形式存在:成熟病毒粒(MV)、包裹病毒粒(WV)和细胞外病毒粒(EV)。MV,病毒的最简单的形式,是包含两端凹陷的、包含DNA的核心的具膜颗粒,所述核心两侧具有充满核心的凹面的侧体。MV通常仅在细胞内发现并且仅由细胞裂解释放。WV由被2个源自反式-高尔基潴泡的另外的脂双层围绕的MV组成。WV,其外膜包含特征性的病毒蛋白,是EV的前体并也在细胞内发现。EV由已经通过最外层的WV膜与细胞质膜的融合外排的WV组成,将MV包裹于一个另外的膜中。发现一部分EV与细胞表面连接,而发现一些游离于细胞外基质中。认为EV对病毒在生物体内的扩散是重要的。

[0017] 从病毒取出必需基因和在宿主细胞中补足(complementation)是本领域已知的用于产生用于治疗 and 疗法的安全和有效的病毒载体的一般提议策略。对具有增强的安全性、表达和/或免疫原性的改善的减毒的正痘病毒载体存在需求,和对于安全地和经济地繁殖和制造这样的正痘病毒载体的方法存在相称的需求。

[0018] 出于研究目的,许多哺乳动物细胞系被用于重组蛋白质的制造。这些包括兔、仓鼠、灵长动物和人源细胞系诸如,不加限制地,和如本领域已知的,HEK293、293T、143B、CHO、HeLa、Vero和BHK、HepG2和3T3细胞。但是,显著地,大部分重组治疗蛋白的GMP生产已经在CHO细胞中进行,使此细胞系成为用于人治疗剂的研究最多的并被批准的细胞系。CHO细胞可以在悬浮培养中生长至非常高的密度,并且用于纯化产品的下游工艺是非常充分开发的。相关的是,痘苗病毒,诸如痘苗病毒-COP或痘苗病毒-WR,和其衍生物诸如MVA和NYVAC都将不在CHO细胞中增殖。

[0019] 在局部病毒启动子和痘病毒RNA聚合酶的控制下,来自痘病毒内的病毒宿主范围基因的表达式是本领域已知的。

[0020] 之前没有描述过来自哺乳动物细胞的基因组的某些病毒宿主范围基因的及时的表达以救援或允许病毒复制并允许宿主细胞存活。

[0021] 概述

[0022] 本说明书描述了被转化以表达痘病毒宿主范围因子的体外培养的哺乳动物细胞和用于繁殖病毒载体的方法,包括培养这样的细胞。

[0023] 在第一方面,本发明提供了一种经修饰的哺乳动物细胞,其中所述细胞的基因组被修饰以包含在启动子的控制下编码CP77的序列,使得所述经修饰的细胞系维持在未经修饰的细胞中较不能繁殖或不能繁殖的痘病毒的繁殖。

[0024] 在第二方面,本发明提供了用于繁殖在CHO细胞中不繁殖的正痘病毒的方法,该方法包括在哺乳动物细胞系中体外繁殖痘病毒,其中所述细胞系被修饰以在启动子的控制下编码并表达CP77。

[0025] 在不同的代替的实施方案中,经修饰的细胞被进一步修饰以包含在启动子的控制下编码D13L的序列和/或包含在启动子的控制下编码K1L的序列。

[0026] 本文中对“CP77”、“K1L”和“D13L”的指代包括功能性直向同源物和功能变体。

[0027] 特定实施方案的详细描述

[0028] 本发明不局限于特定程序或试剂、试剂的具体制剂和各种医疗方法,因为这些可以变化。本文使用的术语仅为了描述特定的实施方案的目的并且不意图是限制性的。除非另外定义,否则本文使用的所有技术和科学术语具有与本发明所属领域内普通技术人员通常的理解相同的意思。

[0029] 与本文描述的那些材料和方法类似的或等效的任何材料和方法可以被用于实践或测试本发明。实施人员尤其参考:Sambrook等人,(1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Press, Plainsview, N.Y.; Ausubel等人,(1999) *Current Protocols in Molecular Biology* (Supplement 47), John Wiley & Sons, New York; Murphy等人,(1995) *Virus Taxonomy* Springer Verlag: 79-87, Mahy Brian WJ和Kangro O Hillar (Eds): *Virology Methods Manual* 1996, Academic Press; 和 Davison AJ和Elliott RM (Eds): *Molecular Virology, A practical Approach* 1993, IRL Press at Oxford University Press; Perkus等人, *Virology* (1990) 179 (1): 276-86 或本领域的定义和术语和本领域技术人员已知的其他方法。

[0030] 尽管与本文描述的那些方法和材料相似或等效的任何方法和材料都可用于实践或测试本发明,描述了优选的方法和材料。为了本发明的目的,以下定义了以下术语。

[0031] 遍及本说明书中,除非上下文另外要求,否则词语“包括(comprise)”、“包括(comprises)”和“包括(comprising)”将被理解为指包括陈述的步骤或要素或步骤或要素的组,但不排除任何其他步骤或要素或步骤或要素的组。因此,术语“包括(comprising)”等的使用表示,所列的要素是要求的或强制的,但其他要素是任选的且可以或可以不存在。在减毒的正痘病毒载体的上下文中,本发明的载体被修饰以通过包含必需的成熟或装配基因的缺失来减毒,但是,涵盖进一步的修饰诸如以运载抗原或其他蛋白。

[0032] “由……组成”意指包括并且局限于短语“由……组成”之中的任何事物。因此,短语“由……组成”表示,所列出的要素是要求的或强制的,并且不可以存在其他要素。“主要由……组成”意指包括在短语中列举的任何要素,并且局限于不影响或有助于本公开内容中对列举的要素具体说明的活性或作用的其他要素。因此,短语“主要由……组成”表示,所列出的要素是要求的或强制的,但其他要素是任选的并且根据它们是否影响所列举的要素的活性或作用可以存在或可以不存在。

[0033] 如本文使用的,除非上下文另外清楚指明,否则单数形式“一(a)”、“一(an)”和“该(the)”包括复数指代物。因此,例如,对“细胞”的指代包括单个细胞,以及2个或更多个细胞;对“生物体”的指代包括一个生物体,以及2个或更多个生物体;等等。在一些实施方案中,“一(an)”意指“一个或多于一个”。

[0034] 如本文使用的,“和/或”指并且涵盖一个或更多个相关的列出的项目的任何和全部组合,以及当用代替物(或)解释时没有组合。

[0035] 如本文使用的,“减毒(attenuation)”或“减毒的(attenuated)”意指病毒载体毒力的降低。毒力被定义为病毒在特定宿主中引起疾病的能力。不能产生感染性病毒的痘病毒载体可以初始地感染细胞但大体上不能在宿主内完全复制其本身或繁殖或引起病症。这是期望的,因为载体将其蛋白质或核酸递送到宿主细胞的细胞质,但不伤害受试者。

[0036] “控制元件”或“控制序列”意指特定痘病毒、载体、质粒或细胞中可操作地连接的编码序列的表达必需的核酸序列(例如,DNA)。适用于真核细胞的控制序列包括转录控制序

列,诸如启动子、多腺苷酸化信号、转录增强子、翻译控制序列诸如翻译增强子和内部核糖体结合位点(IRES)、调节mRNA稳定性的核酸序列、以及将由转录的多核苷酸编码的产物靶向到细胞内的细胞内隔室或细胞外环境的靶向序列。

[0037] 当提供序列时,涵盖相应的序列。“相应于(corresponds to)”、“相应(corresponding)”或“相应于(corresponding to)”意指对参照核酸序列表现出大体上序列同一性的核酸序列(例如,与全部或一部分的参照核酸序列至少约50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或甚至最高达100%序列同一性)或对参照氨基酸序列表现出大体上序列同一性的氨基酸序列(例如,与全部或一部分的参照氨基酸序列至少50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或甚至最高达100%序列相似性或同一性)。

[0038] 在治疗或预防病症或用于调节对靶抗原或生物体的免疫应答的上下文中,“有效量”意指对需要这样的治疗或预防的个体施用的试剂(例如,如本文描述的减毒的正痘病毒载体)或包含其的组合物的量,所述施用以单剂量形式或作为一系列的部分,所述量对于预防该病症的症状的发生、制止这样的症状和/或治疗现有症状有效或对调节对靶抗原或生物体的免疫应答有效。有效量将根据待被治疗的个体的健康和身体状况、待被治疗的个体的分类学类群、组合物的制剂、医学状况的评价和其他相关因素变化。预期该量将落入可以通过常规试验确定的相对宽的范围。

[0039] 如本文使用的,术语“编码(encode)”、“编码(encoding)”等指核酸提供另一种核酸或多肽的能力。例如,如果核酸序列可以被转录和/或翻译以产生多肽或如果它可以被加工成可以被转录和/或翻译以产生多肽的形式,则该核酸序列被称为能够“编码”多肽。这样的核酸序列可包括编码序列或编码序列和非-编码序列二者。因此,术语“编码(encode)”、“编码(encoding)”等包括由DNA分子的转录产生的RNA产物、由RNA分子的翻译产生的蛋白质、由DNA分子的转录以形成RNA产物和随后的RNA产物的翻译而形成的蛋白质、或由DNA分子的转录以提供RNA产物、加工RNA产物以提供经加工的RNA产物(例如,mRNA)和随后的经加工的RNA产物的翻译而形成的蛋白质。

[0040] 术语“内源的”指通常在宿主生物体内发现的基因或核酸序列或区段。

[0041] 术语“可表达的(expressible)”、“表达的(expressed)”及其变化形式指细胞将核苷酸序列转录成RNA并任选地翻译该mRNA以合成提供生物功能或生物化学功能的肽或多肽的能力。

[0042] 如本文使用的,术语“基因”包括能够被用于产生mRNA,任选地具有另外的元件以辅助此过程的核酸分子。基因可以或可以被用于产生功能蛋白。基因可以包括编码区域和非编码区域二者(例如,内含子、调节元件、启动子、增强子、终止序列和5'和3'非翻译区域)。

[0043] 术语“异源核酸序列”、“异源核苷酸序列”、“异源多核苷酸”、“外源多核苷酸

(foreign polynucleotide)”、“外源多核苷酸(exogenous polynucleotide)”等可互换使用以指通过实验操作被引入到生物体的基因组中的任何核酸(例如,包含IRES的核苷酸序列)并且可以包括在该生物体中发现的基因序列,只要该被引入的基因相对于修饰前的病毒基因组序列包含某种修饰(例如,点突变、缺失、至少1个核苷酸的取代或添加、内切核酸酶裂解位点的存在、loxP位点的存在等)。

[0044] 术语“异源多肽”、“外源多肽(foreign polypeptide)”和“外源多肽(exogenous polypeptide)”可互换使用以指由如以上定义的“异源核酸序列”、“异源核苷酸序列”、“异源多核苷酸”、“外源多核苷酸(foreign polynucleotide)”和“外源多核苷酸(exogenous polynucleotide)”编码的任何肽或多肽。

[0045] 术语“哺乳动物细胞”意指,出于繁殖痘病毒载体的目的,可以向其中引入包括本发明的减毒的正痘病毒载体的载体的细胞。在一个实施方案中,细胞是连续的细胞系。经修饰的细胞是能够连续分裂的细胞系是不太必要的。哺乳动物或更高级的真核细胞可以根据本发明被修饰并随后被转化或被永生化的(imortalized)以成为连续分裂的细胞系。但是,修饰前的细胞方便地是本领域已知的被良好表征和连续分裂的与生物技术相容的连续细胞系。这样的细胞可以从保藏机构诸如美国典型培养物保藏中心(ATCC)或欧洲细胞培养物保藏中心(ECACC)方便地获得。

[0046] 适当的哺乳动物细胞系包括但不限于可从例如ATCC获得的RK18、BHK、VERO、HBOC-143B、HaCat、HepG2、HeLa、HT1080、HEK-293、RD、COS-7、CHO、Jurkat、HUT、SUPT、C8166、MOLT4/克隆8、MT-2、MT-4、H9、PM1、CEM、骨髓瘤细胞(例如,SB20细胞)和CEMX174。

[0047] 可以使用用于产生表达异源基因的经修饰的细胞系的任何本领域公认的基因组工程方法。参考转座子技术的应用以通过使用piggyBac载体将基因插入到细胞基因组。但是,许多方法是公认的用于向细胞引入基因,所述方法包括,不加限制地,反转录病毒转导(例如,MoMLV等)、慢病毒转导、质粒转染和整合、用于转导细胞系的其他病毒系统诸如腺病毒、AAV(腺病毒伴随病毒)、EBV和用于通过与线性DNA同源重组进行位点特异性插入的基因组编辑技术、工程化大范围核酸酶(engineered meganuclease)、转录-激活物样效应核酸酶(TAL-核酸酶)、锌指核酸酶(ZFN)和CRISPR。

[0048] 在一个实施方案中,哺乳动物细胞是CHO细胞。不编码病毒宿主名称基因(viral host name gene)的现有技术的CHO细胞系不支持大体上不能在人在中繁殖的痘苗病毒或痘苗病毒衍生物的制造。如本领域技术人员已知的,中国仓鼠卵巢(CHO)细胞衍生自仓鼠(中国地鼠(*Cricetulus griseus*))的卵巢,是最常用于生物工业和重组蛋白治疗剂(包括抗体)的GMP生产的哺乳动物细胞。用于此目的的CHO细胞的普及性部分源自它们的快速生长和高蛋白质产量。因此,已经充分地表征了CHO细胞系。适当的CHO细胞系不限制地包括A2、A2H、XrS6、CHO-K1、CHO/dhfr、RR-CHO-K1、UT-I、P22、CHO-1C6、Lec1、Lec2、Lec8、Pro-5和CDKXB1系。时常使用以登录号ATCC CLL-61或ATCC CRL-9618保藏于ATCC的CHO-K1细胞系。CHO-K1细胞系作为由Puck T. (1957)的成年中国仓鼠的卵巢的生物活检起始的亲本CHO细胞系的亚克隆衍生。本说明书描述了经修饰的CHO细胞以外的经修饰的细胞。

[0049] 在一些实施方案中,哺乳动物细胞是人细胞、灵长动物细胞、仓鼠细胞或兔细胞。

[0050] 细胞可以是单细胞,或者可以在组织培养中作为液体培养物、单层或类似物生长。宿主细胞也可以直接或间接地衍生于组织或可以在包括动物的生物体中存在。

[0051] 应该理解,如本文预期的“诱导”免疫应答包括引发或刺激免疫应答和/或增强之前存在的免疫应答。

[0052] 如本文使用的,术语“内部核糖体进入位点”或“IRES”指允许在mRNA内的编码区域内部的位点或在mRNA的5'末端的3'位点起始该mRNA的翻译的病毒、细胞或合成的(例如,重组的)核苷酸序列,以提供位于该内部核糖体进入位点的下游(即,其3')的可操作地连接的编码区域的翻译。这使得翻译不依赖于5'帽结构,并且不依赖于mRNA的5'末端。IRES序列提供可操作地连接的编码区域的翻译的起始所必需的顺式作用序列。

[0053] 如本文使用的,术语“分离的”意指描述处于不同于该化合物天然存在的环境的环境下的细胞、感兴趣的化合物(例如,重组痘病毒、核酸分子诸如基因组、多肽等)。“分离的”意指包括样品内的化合物,该样品大体上富含感兴趣的化合物和/或其中感兴趣的化合物被部分或大体上纯化。

[0054] 如本文使用的,术语“可操作地连接的(operably connected)”或“可操作地连接(operably linked)”指一种并列方式,其中如此所述的元件呈允许其按照其预期的方式行使功能的关系。例如,与编码序列“可操作地连接的”转录控制序列指转录控制序列相对于编码序列的定位和/或定向允许编码序列在与该转录控制序列相容的条件下表达。在另一个实例中,与正痘病毒编码序列可操作地连接的IRES指IRES相对于正痘病毒编码序列的定位和/或定向允许正痘病毒编码序列的不依赖帽的翻译。

[0055] 如本文使用的,术语“开放阅读框”和“ORF”在本文中互换使用以指编码序列的翻译起始密码子和翻译终止密码子之间编码的氨基酸序列。术语“起始密码子”(例如,ATG)和“终止密码子”(例如,TGA、TAA、TAG)指编码序列中3个相邻核苷酸的单位(“密码子”),其分别明确蛋白质合成(mRNA翻译)的起始和链终止。

[0056] 如本文使用的,术语“亲代病毒”将被理解为指被修饰以掺入异源核酸序列以形成本发明的重组病毒的病毒。

[0057] 如本文使用的,术语“多核苷酸”、“多核苷酸序列”、“核苷酸序列”、“核酸”或“核酸序列”指mRNA、RNA、cRNA、cDNA或DNA。该术语通常指长度至少10个碱基的核苷酸的聚合物形式,为核糖核苷酸或脱氧核糖核苷酸或核苷酸的任一种类型的经修饰的形式。该术语包括RNA或DNA的单链和双链形式。

[0058] 本文中“多肽”、“肽”、“蛋白质”和“蛋白质分子”可互换使用以指包含氨基酸残基的聚合物的分子或由氨基酸残基的聚合物组成的分子,并且指其变体和合成的类似物。因此,这些术语适用于其中一个或多个氨基酸残基是合成的非天然存在的氨基酸(诸如相应天然存在的氨基酸的化学类似物)的氨基酸聚合物以及天然存在的氨基酸聚合物。

[0059] 如本文使用的,当用于“核酸分子”、“多核苷酸”等时,术语“重组的”被理解为意指可以在宿主细胞或本文描述的无细胞体系中转录和/或翻译的人工核酸结构(即,非复制的cDNA或RNA;或复制子、自复制的cDNA或RNA)。重组核酸分子或多核苷酸可以被插入到载体中。可以使用非病毒载体诸如质粒表达载体或病毒载体。载体的种类和插入根据本发明的核酸构建体的技术是技术人员已知的。在本发明描述的布置中,根据本发明的核酸分子或多核苷酸不存在于自然。换句话说,异源核苷酸序列不天然地与亲本病毒基因组的元件(例如,启动子、ORF、多腺苷酸化信号、核酶)组合。

[0060] 如本文使用的,术语“重组病毒”将被理解为指包含至少一种异源核酸序列的“亲

本病毒”。

[0061] 如本文使用的,术语“序列同一性”指比较窗口中的序列在逐个碱基或逐个氨基酸的基础上的相同程度。因此,“序列同一性百分比”通过以下方式计算:在比较窗口中比较两条最佳地对齐的序列,确定在两条序列上出现相同核酸碱基(例如,A、T、C、G、I)或者相同氨基酸残基(例如Ala、Pro、Ser、Thr、Gly、Val、Leu、Ile、Phe、Tyr、Trp、Lys、Arg、His、Asp、Glu、Asn、Gln、Cys和Met)的位置的数目以得到匹配的位置数目,用匹配的位置数目除以比较窗口中的位置总数目(即窗口大小),再将结果乘以100以得到序列同一性的百分比。出于本发明的目的,“序列同一性”将被理解为意指由DNASIS计算机程序(对于Windows,版本2.5;可从Hitachi Software engineering Co.,Ltd.,South San Francisco,California,USA获得)使用如软件随附的参考手册的标准默认值计算的“匹配百分比”。

[0062] 术语“信号序列”或“信号肽”指指导蛋白质从胞质溶胶向某些细胞器诸如例如细胞核、线粒体基质和内质网的翻译共转运或翻译后转运的短肽(约3至约60个氨基酸长)。对于具有ER靶向信号肽的蛋白质,信号肽通常在该蛋白质被转运到ER后通过信号肽酶从前体形式裂解,并且所得蛋白质沿着分泌途径移动到其细胞内位置(例如,高尔基体、细胞膜或细胞壁)或细胞外位置。如本文使用的,“ER靶向信号肽”包括通常在该蛋白质部分或全部插入通过ER膜进入ER腔后酶解除去的氨基末端疏水序列。因此,本领域已知,序列的信号前体形式可以作为蛋白质的前体形式的一部分存在,但通常不在该蛋白质的成熟形式出现。

[0063] “相似性”指相同的或构成如以下表A中定义的保守取代的氨基酸的百分比数目。相似性可以通过使用序列比对程序诸如GAP确定(Deveraux等人,1984,Nucleic Acids Research12:387-395)。以这种方式,与本文列举的序列相似或大体上不同长度的序列可以通过在比对中插入缺口来比较,这样的缺口通过例如由GAP使用的比较算法确定。

[0064] 表A:示例性保守氨基酸取代

	原始残基	示例性取代
[0065]	Ala	Ser
	Arg	Lys
	Asn	Gln、 His
	Asp	Glu
	Cys	Ser
	Gln	Asn
	Glu	Asp
	Gly	Pro
	His	Asn、 Gln
	Ile	Leu、 Val
	Leu	Ile、 Val
	Lys	Arg、 Gln、 Glu
	Met	Leu、 Ile,

原始残基	示例性取代
Phe	Met、Leu、Tyr
Ser	Thr
Thr	Ser
Trp	Tyr
Tyr	Trp、Phe
Val	Ile、Leu

[0067] 用于对齐比较窗口的序列的最佳对齐可以通过计算机化的算法实现 (Wisconsin Genetics Software Package Release 7.0 中的 GAP、BESTFIT、FASTA 和 TFASTA, Genetics Computer Group, 575 Science Drive Madison, WI, USA) 进行或通过检查进行, 并且用所选不同方法中的任一种产生最佳的对齐 (即, 在比较窗口产生最高同源性百分比)。也可以参考 BLAST 家族的程序如, 例如 Altschul 等人, 1997, Nucl. Acids Res. 25:3389 公开的。序列分析的详细讨论可以在 Ausubel 等人, “Current Protocols in Molecular Biology”, John Wiley & Sons Inc, 1994-1998, Chapter 15 的 Unit 19.3 找到。

[0068] 术语“受试者”、“患者”、“宿主”或“个体”在本文中可互换使用, 指期望疗法或预防的任何受试者, 特别是脊椎动物受试者, 并且甚至更特别是哺乳动物受试者。落入本发明的范围的适当的脊椎动物包括但不限于脊索动物 (Chordata) 亚门的任何成员, 包括灵长动物 (例如, 人、猴子和猿, 并且包括猴子的种, 诸如来自猕猴属 (*Macaca*) (例如, 猕猴 (*cynomolgus monkey*) 诸如食蟹猴 (*Macaca fascicularis*) 和/或恒河猴 (*Macaca mulatta*)) 和狒狒 (*Papio ursinus*), 以及狨猴 (来自狨属 (*Callithrix*) 的种)、鼠猴 (来自松鼠猴属 (*Saimiri*) 的种) 和绢毛猴 (来自怪柳猴属 (*Saguinus*) 的种) 以及猿的种, 诸如黑猩猩 (*Pan troglodytes*))、啮齿动物 (例如, 小鼠、大鼠、豚鼠)、兔形目 (例如, 家兔、野兔)、牛科动物 (例如, 牛)、羊 (例如, 绵羊)、山羊 (*caprines*) (例如, 山羊 (*goats*))、猪 (*porcines*) (例如, 猪 (*pigs*))、马科动物 (例如, 马)、犬科动物 (例如, 狗)、猫科动物 (例如, 猫)、鸟类 (例如, 鸡、火鸡、鸭、鹅、宠物鸟诸如金丝雀、虎皮鹦鹉等)、海洋动物 (例如, 海豚、鲸鱼)、爬行动物 (蛇、青蛙、蜥蜴等) 和鱼。优选的受试者是需要病症的治疗或预防的人。但是, 应理解, 上述术语不暗示出现症状。

[0069] 本文使用术语“转基因”来描述已经或将要被人工引入到宿主生物体的基因组并且被传递给该宿主的后代的遗传材料。在一些实施方案中, 它对其被引入至其中的哺乳动物细胞或正痘病毒载体赋予期望的特性, 或以其他方式产生期望的治疗或诊断结果。

[0070] 如本文使用的, 术语“治疗 (treatment)”、“治疗 (treating)”等指获得期望的药理作用和/或生理学作用。该作用可以在完全或部分预防疾病或其症状方面是预防性的, 和/或可以在部分或完全治愈疾病和/或可归因于该疾病的不利作用方面是治疗性的。如本文使用的, “治疗 (treatment)”包括哺乳动物, 特别是人的疾病的任何治疗, 并且包括: (a) 在可能倾向于患病但尚未被诊断为患病的受试者中预防疾病的发生; (b) 阻止疾病, 即, 阻止其发展; 和 (c) 减轻疾病, 即, 引起疾病的消退。

[0071] 有关生物体、多肽或核酸序列的术语“野生型”、“天然的 (natural)”、“天然的 (native)”等意指该生物体、多肽或核酸序列是天然存在的或可以在未改变的、突变的或以其他方式由人操作的至少一种天然存在的生物体中获得。

[0072] 变体包括与参照分子或其互补形式的全部或部分充分相似的核酸分子,使得在中严格度或高严格度条件下可以实现选择性杂交,或在包含至少约15个核苷酸的比较窗口中,与定义参照痘病毒宿主范围因子的核苷酸序列具有约60%至90%或90%至98%的序列同一性。优选地,杂交区域为约12至约18个核苷碱基或更大的长度。优选地,特定核苷酸序列和参照序列之间的同一性百分比为至少约80%或85%,或更优选地约90%相似或更大,诸如约95%、96%、97%、98%、99%或更大。涵盖80%和100%之间的同一性百分比。核苷酸序列的长度取决于其提议的功能。涵盖同源物。术语“同源物(homolog)”、“同源基因”或“同源物(homologs)”广泛地指功能上和结构上相关的分子,包括来自其他种的那些。同源物和直向同源物(ortholog)是变体的实例。

[0073] 核酸序列同一性可以用以下方式确定。本发明的核酸序列被用于使用程序BLASTM版本2.1搜索核酸序列数据库,诸如GenBank数据库(可以从网址<http://www.ncbi.nlm.nih.gov/blast/>进入)(基于Altschul等人,(1997)Nucleic Acids Research 25:3389-3402)。以无缺口模式使用程序。使用默认过滤以除去由于低复杂度的区域的序列同源性。使用BLASTM的默认参数。

[0074] 氨基酸序列同一性可以用以下方式确定。本发明的多肽序列被用于使用BLASTP程序搜索多肽序列数据库,诸如GenBank数据库(可以从网址<http://www.ncbi.nlm.nih.gov/blast/>进入)。以无缺口模式使用程序。使用默认过滤以除去由于低复杂度的区域的序列同源性。使用BLASTP的默认参数。可以使用SEG程序过滤低复杂度的序列。

[0075] 优选的序列将在严格条件下与参照序列或其互补物杂交。术语“在严格条件下杂交”和其语法同义词,指核酸分子与靶核酸分子(诸如在DNA或RNA印迹,诸如Southern印迹或Northern印迹上固定的靶核酸分子)在规定的温度和盐浓度的条件下杂交的能力。对于长度大于约100碱基的核酸分子,典型的严格杂交条件是不超过低于天然双螺旋的解链温度(T_m)的25℃至30℃(例如,10℃)(参见,一般的,Sambrook等人,(以上);Ausubel等人,(1999))。对于大于约100个碱基的核酸分子的 T_m ,可以用公式 $T_m = 81.5 + 0.41\% (G+C - \log(Na^+))$ 计算。对于具有长度小于100个碱基的核酸分子,示例性的严格杂交条件是低于 T_m 5℃至10℃。

[0076] 在本文中,术语“缺失”意指除去靶基因的编码区域的全部或部分。该术语还包括消除靶基因的基因表达或消除或大体上下调编码的蛋白质的水平或活性的突变或转化的任何形式。

[0077] 对“基因”的指代包括对应于外显子或基因的开放阅读框的DNA。本文中对“基因”的指代还包括:由转录和/或翻译调节序列和/或编码区域和/或非翻译序列(即,内含子、5'-和3'-非翻译序列)组成的经典基因组基因;或对应于编码区域(即外显子)和基因的5'-和3'-非翻译序列的mRNA或cDNA。

[0078] “调节元件”或“调节序列”意指特定宿主细胞中可操作地连接的编码序列的表达必需的核酸序列(例如,DNA)。适用于原核细胞的调节序列,例如,包括启动子和任选地顺式作用序列,诸如操纵子序列和核糖体结合位点。适用于真核细胞的控制序列包括启动子、多腺苷酸化信号、转录增强子、翻译增强子、调节mRNA稳定性的前导序列或尾随序列,以及将由转录的多核苷酸编码的产物靶向到细胞内的细胞内区室或靶向到细胞外环境的靶向序列。

[0079] 适用于实现本发明的经修饰的哺乳动物细胞的嵌合构建体包含编码正痘病毒宿主范围因子的核酸序列,该核酸序列可操作地连接到调节序列。调节序列适当地包含转录和/或翻译控制序列,其将与细胞中的表达相容。典型地,转录和翻译调节控制序列包括但不限于,启动子序列、5'非编码区、顺式调节区诸如转录调节蛋白或翻译调节蛋白的功能结合位点、上游开放阅读框、核糖体结合序列、转录起始位点、翻译起始位点和/或编码前导序列的核苷酸序列、终止密码子、翻译终止位点和3'非翻译区。构思了本领域已知的组成型或诱导型启动子。启动子可以是天然存在的启动子,或组合多于一个启动子的元件的杂合启动子(hybrid promoters)。

[0080] 预期的启动子序列可以是哺乳动物细胞天然的或可以衍生自其中该区域在选择的生物体中有功能的代替性来源。启动子的选择将根据计划的宿主细胞不同。例如,可以用于在哺乳动物细胞中表达的启动子包括金属硫蛋白启动子,其可以响应于重金属诸如镉被诱导, β -肌动蛋白启动子以及病毒启动子诸如SV40大T抗原启动子、人巨细胞病毒(CMV)即刻早期(IE)启动子、劳氏肉瘤病毒LTR启动子、小鼠乳房瘤病毒LTR启动子、腺病毒主要晚期启动子(Ad MLP)、简单疱疹病毒启动子和HPV启动子,特别是HPV上游调节区域(URR),除了其他以外。所有这些启动子都被详细描述并且在本领域可容易地获得。

[0081] 本文中也可以使用增强子元件来提高哺乳动物构建体的表达水平。实例包括SV40早期基因增强子,如例如在Dijkema等人,(1985)EMBO J.4:761中描述的,衍生自劳氏肉瘤病毒的长末端重复(LTR)的增强子/启动子,如例如Gorman等人,(1982)Proc.Natl.Acad.Sci.USA 79:6777中描述的,和衍生自人CMV的元件,如例如Boshart等人,(1985)Cell 41:521中描述的,诸如包括在CMV内含子A序列中的元件。

[0082] 嵌合构建体还可以包含3'非-翻译序列。3'非-翻译序列指基因的包含含有多腺苷酸化信号和能够实现mRNA加工或基因表达的任何其他调节信号的DNA区段的那部分。多腺苷酸化信号的特征在于实现向mRNA前体的3'末端添加多腺苷酸串(polyadenylic acid tract)。多腺苷酸化信号通常由与规范形式5' AATAAA-3'的同源性的出现识别,尽管变异并非不常见的。3'非翻译调节DNA序列优选地包括从约50至1,000nt并且除多腺苷酸化信号和能够实现mRNA加工或基因表达的任何其他调节信号以外可以包含转录和翻译终止序列。

[0083] 在一些实施方案中,嵌合构建体还包含可选择的标记基因以允许选择包含该构建体的细胞。选择基因是本领域公知的并且将与感兴趣的细胞中的表达相容。

[0084] 在一个实施方案中,正痘病毒结构基因或装配基因的表达是在启动子的控制下。在一个非限制性的实施方案中,启动子是指导在缺乏对宿主细胞的显著的毒性作用下维持病毒繁殖的充分水平的CP77的表达的细胞组成型启动子,诸如人EF1 α (人延伸因子1 α 基因启动子)、DHFR(二氢叶酸还原酶基因启动子)或PGK(磷酸甘油酸激酶基因启动子)。启动子也可以是诱导型的,诸如细胞诱导型启动子,MTH(来自金属硫蛋白基因)。病毒启动子,诸如CMV、RSV、SV-40和MoU3,也在哺乳动物细胞中使用。

[0085] 在第一方面,本发明提供了经修饰的哺乳动物细胞,其中细胞的基因组被修饰以包含在启动子的控制下编码CP77的序列,使得经修饰的细胞系维持在未经修饰的细胞中较不能繁殖或不能繁殖的痘病毒的繁殖。

[0086] 在第二方面,本发明提供了使在CHO细胞中不繁殖的正痘病毒繁殖的方法,该方法包括在哺乳动物细胞系中体外繁殖痘病毒,其中该细胞系被修饰以在启动子的控制下编码

并表达CP77。

[0087] 在一个实施方案中,细胞的基因组还包含在启动子的控制下编码D13L的序列和/或还包含在启动子的控制下编码K1L的序列。

[0088] 在另一个实施方案中,细胞是连续的细胞系,优选地CHO细胞。

[0089] 在其他实施方案中,细胞可以是人细胞、灵长动物细胞、仓鼠细胞或兔细胞。

[0090] 在另一个实施方案中,CP77基因、D13L基因和/或K1L基因是在哺乳动物启动子的控制下。

[0091] 在另一个实施方案中,CP77基因的表达支持病毒的繁殖以产生与在允许细胞系中观察到的病毒产量相当的病毒产量,并且优选地支持大于500的病毒复制扩增率。

[0092] 在另一个实施方案中,CP77、D13L和/或K1L由为了在哺乳动物细胞中表达而密码子优化的核苷酸的连续序列编码。

[0093] 示例K1L序列在SEQ ID NO:6和SEQID NO:7中提供。

[0094] 本说明书广泛地涉及经修饰的(重组的或经转化的)细胞和用于在培养的高级真核宿主细胞中体外生产病毒载体的方法,其中该细胞被以增强或有利于病毒载体在体外细胞的群体内或通过体外细胞的群体的繁殖的方式遗传修饰。在一个非限制性实施方案中,说明书提供了基于痘苗病毒的痘病毒在经修饰的中国仓鼠卵巢(CHO)细胞中的繁殖。

[0095] 如本领域技术人员已知的,中国仓鼠卵巢(CHO)细胞衍生自仓鼠(中国地鼠)的卵巢,是最常用于生物工业和重组蛋白治疗剂(包括抗体)的GMP生产的哺乳动物细胞。用于此目的的CHO细胞的普及性部分源自它们的快速生长和高蛋白质产量。因此,已经充分地表征了CHO细胞系。适当的CHO细胞系不限制地包括A2、A2H、XrS6、CHO-K1、CHO/dhfr、RR-CHO-K1、UT-I、P22、CHO-1C6、Lec1、Lec2、Lec8、Pro-5和CDKXB1系。时常使用以登录号ATCC CLL-61或ATCC CRL-9618保藏于ATCC的CHO-K1细胞系。CHO-K1细胞系作为由Puck T. (1957)的成年中国仓鼠的卵巢的生物活检起始的亲本CHO细胞系的亚克隆衍生。本说明书描述了经修饰的CHO细胞以外的经修饰的细胞。

[0096] 为了使产生的病毒的产量最大化,可以使用标准技术使细胞系诸如CHO细胞适应于悬浮培养。对重组体或经修饰的细胞的指代包括其子代。细胞可以以任何形式销售,包括冷冻的或液体悬浮液形式。细胞可以以被痘病毒载体感染的形式销售。

[0097] 本文中CP77的指代意指Spheener等人, (1988) 和Hsiao等人, (2006) 中提到的牛痘宿主范围基因和如本文确定的当在哺乳动物宿主细胞的基因组中作为异源基因表达时能够支持痘病毒生长的其功能直向同源物和经修饰的形式。对“经修饰的形式”的指代包括来自野生型或参照序列的变异(诸如缺失、取代或插入)。参照核苷酸序列以SEQ ID NO:1提供。经修饰的形式在编码区域或其包含至少200个连续碱基对的一个或多个部分中共有至少80%序列同一性。经修饰的形式包括如本文所述的、为了在哺乳动物或其他高级真核细胞(包括CHO细胞)中表达而优化的密码子优化形式。对CP77的指代包括直向同源物,即,存在于其他种或以不同名称鉴定的具有相同功能的基因。牛痘宿主范围因子CP77也被称为VHR1、CHOhr和CPXV025。痘苗病毒Western Reserve (WR) 毒株中的CHOhr/CP77基因是大体上碎片化的并且不能产生功能因子。CP77不在MVA和哥本哈根毒株中存在。

[0098] 本文中对“CP77”、“D13L”和“K1L”的指代包括功能直向同源物和功能变体。“功能”指本文描述的由培养的细胞的哺乳动物基因组指导的支持表达细胞细胞质内的痘病毒繁

殖的表达(即转录和翻译)质量。术语“繁殖”包括细胞内繁殖和细胞间繁殖并且涵盖成熟病毒颗粒的产生。

[0099] 在一个实施方案中,经修饰的细胞的群体维持在未经修饰的对照细胞中较不能繁殖或不能繁殖的痘病毒的繁殖。不能繁殖指,通常不发生细胞到细胞或受试者到受试者传递。

[0100] 对“未经修饰的对照细胞”的指代包括经修饰的细胞在被修饰以在启动子的控制下从其基因组编码并表达至少一种病毒宿主范围因子之前存在的经修饰的细胞以及本领域已知的其他适当的对照细胞,所述病毒宿主范围因子选自由CP77、K1L和SPI-1组成的组。

[0101] 例如,痘病毒诸如MVA和痘苗病毒不能在CHO细胞中繁殖。但是,如本文意外地确定的,被重组修饰以在启动子的控制下从其基因组表达CP77的CHO细胞不仅能够支持痘苗病毒的繁殖,还能支持痘苗病毒衍生物诸如MVA的繁殖。在本发明的经修饰的CHO-CP77细胞中繁殖后,MVA和痘苗病毒仍然不能在未经修饰的CHO细胞或其他适当的对照中繁殖。

[0102] 本文中对“痘病毒”的指代包括编码感兴趣的异源分子(诸如感兴趣的抗原)的作为药物、预防剂、诊断剂或治疗剂的重组痘病毒载体。这样的重组痘病毒载体通常用作抗非-痘病毒诱导的疾病或病症的疫苗。并且,术语痘病毒包括提议用作抗痘病毒感染(诸如天花(variola)或天花感染)的治疗疫苗或预防疫苗的分离的痘病毒及其衍生物。

[0103] 本文中对“基于痘苗病毒的”的指代包括痘苗病毒和衍生物以及痘苗病毒的经修饰的形式。基于痘苗病毒的衍生物包括但绝不限于MVA和NYVAC。对MVA和NYVAC的指代包括本领域已知的这些痘病毒的毒株或衍生物。经修饰的形式可以在1至10个或更多个基因中具有修饰。术语“修饰”意图意指从野生型或参照序列的变异(诸如缺失、取代或插入)。对“减毒的”的指代包括与相同病毒的非减毒的形式相比,在相关受试者中不繁殖或大体上以更小的程度繁殖的痘病毒。该术语还包括在受试者中非病原性的病毒。

[0104] 在一个实施方案中,细胞是连续的细胞系。经修饰的细胞是能够连续分裂的细胞系是不太必要的。哺乳动物或更高级的真核细胞可以根据本发明被修饰并随后被转化或被永生化以成为连续分裂的细胞系。但是,修饰前的细胞方便地是本领域已知的被良好表征和连续分裂的与生物技术相容的连续细胞系。这样的细胞可以从保藏机构诸如美国典型培养物保藏中心(ATCC)或欧洲细胞培养物保藏中心(ECACC)方便地获得。

[0105] 适当的哺乳动物细胞系包括但不限于可从例如ATCC获得的RK18、BHK、VERO、HBOC-143B、HaCat、HepG2、HeLa、HT1080、HEK-293、RD、COS-7、CHO、Jurkat、HUT、SUPT、C8166、MOLT4/克隆8、MT-2、MT-4、H9、PM1、CEM、骨髓瘤细胞(例如,SB20细胞)和CEMX174。

[0106] 在一个实施方案中,细胞是CHO细胞。现有技术的CHO细胞系不编码病毒宿主名称基因,不支持大体上不能在人在中复制的痘苗病毒或痘苗病毒衍生物的制造。

[0107] 在一些实施方案中,细胞是人细胞、灵长动物细胞、仓鼠细胞或兔细胞。

[0108] 适用于实现本发明的经修饰的哺乳动物细胞的嵌合构建体包含编码痘病毒宿主范围因子的核酸序列,该核酸序列可操作地连接到调节序列。调节序列适当地包含转录和/或翻译控制序列,所述转录和/或翻译控制序列将与细胞中的表达相容。典型地,转录和翻译调节控制序列包括但不限于,启动子序列、5'非编码区、顺式调节区诸如转录调节蛋白或翻译调节蛋白的功能结合位点、上游开放阅读框、核糖体结合序列、转录起始位点、翻译起始位点和/或编码前导序列的核苷酸序列、终止密码子、翻译终止位点和3'非翻译区。构思

了本领域已知的组成型或诱导型启动子。启动子可以是天然存在的启动子,或组合多于1个启动子的元件的杂合启动子。

[0109] 预期的启动子序列可以是哺乳动物细胞天然的或可以衍生自其中该区域在选择的生物体中有功能的代替性来源。启动子的选择将根据计划的宿主细胞不同。例如,可以用于在哺乳动物细胞中表达的启动子除了其他以外包括金属硫蛋白启动子,其可以响应于重金属诸如镉被诱导, β -肌动蛋白启动子以及病毒启动子诸如SV40大T抗原启动子、人巨细胞病毒(CMV)即刻早期(IE)启动子、劳氏肉瘤病毒LTR启动子、小鼠乳房瘤病毒LTR启动子、腺病毒主要晚期启动子(Ad MLP)、简单疱疹病毒启动子和HPV启动子,特别是HPV上游调节区域(URR)。所有这些启动子已经在本领域中很好地描述并可以容易地获得。

[0110] 本文中也可以使用增强子元件来提高哺乳动物构建体的表达水平。实例包括SV40早期基因增强子,如例如在Dijkema等人,(1985)EMBO J.4:761中描述的,衍生自劳氏肉瘤病毒的长末端重复(LTR)的增强子/启动子,如例如Gorman等人,(1982)Proc.Natl.Acad.Sci.USA 79:6777中描述的,和衍生自人CMV的元件,如例如Boshart等人,(1985)Cell 41:521中描述的,诸如CMV内含子A序列中包括的元件。

[0111] 嵌合构建体还可以包含3'非翻译序列。3'非翻译序列指基因的包含含有多腺苷酸化信号和任何其他能够实现mRNA加工或基因表达的调节信号的DNA区段的那部分。多腺苷酸化信号的特征在于实现向mRNA前体的3'末端添加多腺苷酸串。多腺苷酸化信号通常由与规范形式5' AATAAA-3'的同源性的出现识别,尽管变异并非不常见的。3'非翻译调节DNA序列优选地包括从约50至1,000nt并且除多腺苷酸化信号和能够实现mRNA加工或基因表达的任何其他调节信号以外可以包含转录和翻译终止序列。

[0112] 在一些实施方案中,嵌合构建体还包含可选择的标记基因以允许选择包含该构建体的细胞。选择基因是本领域公知的并且将与在感兴趣的细胞中的表达相容。

[0113] 在一个实施方案中,病毒宿主范围基因的表达是在启动子的控制下。在一个非限制性的实施方案中,启动子是指导在缺乏对宿主细胞的显著的毒性作用下维持病毒繁殖充分水平的CP77的表达的细胞组成型启动子,诸如人EF1 α (人延伸因子1 α 基因启动子)、DHFR(二氢叶酸还原酶基因启动子)或PGK(磷酸甘油酸激酶基因启动子)。启动子也可以是诱导型的,诸如细胞诱型启动子,MTH(来自金属硫蛋白基因)。病毒启动子,诸如CMV、RSV、SV40和MoU3,也在哺乳动物细胞中使用。

[0114] 方便地,在一个实施方案中,病毒宿主范围基因的表达支持病毒的繁殖以产生与在允许细胞系中观察到的病毒产量等价的病毒产量。例如,病毒宿主范围基因的表达支持大于500的病毒复制扩增率。病毒宿主范围基因的表达支持在缺乏明显的宿主细胞毒性下的病毒繁殖。明显的宿主细胞毒性指由于不成熟的宿主细胞死亡或不能分裂而降低病毒产量的病毒宿主范围因子表达的水平。技术人员熟悉定性地或定量地评价宿主细胞参数和病毒参数的方法,所述宿主细胞参数诸如宿主细胞存活率和繁殖,所述病毒参数诸如病毒宿主范围基因表达、病毒复制和病毒产量。

[0115] 在一个实施方案中,痘病毒是除编码功能CP77或CP77直向同源物的正痘病毒以外的脊索动物痘病毒(chordopox virus)。牛痘病毒编码CP77并且因此不被包括于此方面。正痘病毒包括,水牛痘病毒(buffalopox virus)、牛痘病毒(cowpox virus)、骆驼痘病毒(camelpox virus)、鼠痘病毒(ectromelia virus)、猴痘病毒(monkeypox virus)、兔痘病

毒(rabbitpox virus)、浣熊痘病毒(racoonpox virus)、tetrapox virus、痘苗病毒、田鼠痘病毒(volepox virus)、臭鼬痘病毒(skunkpox virus)和瓦森伊修病毒(Uasin Gishu disease virus)。其他属包括副痘病毒属(parapoxvirus)、禽痘病毒属(avipoxvirus)、羊痘病毒属(capripoxvirus)、野兔痘病毒属(leporipoxvirus)、猪痘病毒(swinepoxvirus)、软疣痘病毒属(molluscipoxvirus)和亚塔痘病毒属(yatapoxvirus)。

[0116] 在一个实施方案中,痘病毒是MVA或大体上不能在人/受试者中复制的MVA衍生物。

[0117] 在一个实施方案中,痘病毒是痘苗病毒或大体上在人中在体内非复制性的痘苗病毒衍生物。

[0118] 在另一个实施方案中,痘病毒适合用作痘病毒疫苗。

[0119] 在又一个实施方案中,痘病毒是编码并表达感兴趣的异源分子的重组痘病毒载体,所述感兴趣的异源分子诸如具有医疗益处的抗原,其中该重组痘病毒载体在受试者中用作诊断剂、治疗剂或预防剂。

[0120] 本文中对K1L的指代意指Shisler和Jin (2004) 描述的基因及其直向同源物或经修饰的形式。

[0121] 本文中对SPI-1的指代意指Brookes等人, (1995) 描述的宿主范围基因或其直向同源物和经修饰的形式。

[0122] 在一些实施方案中,并且为了避免疑问,本发明的增强的病毒繁殖过程不需要向痘病毒基因组加入基因。当然,这不排除出于其他目的的对病毒载体的修饰,诸如,不加限制地,以编码异源分子作为用于疫苗目的的感兴趣的抗原或以在受试者中产生免疫应答。

[0123] 痘病毒宿主范围基因从宿主细胞核内的转录和编码的产物的翻译在被感染的细胞中发生并且对于在宿主细胞细胞质中的痘病毒繁殖是足够的。不限于任何特定作用方式,认为CP77是病毒保护剂。

[0124] 在一个说明性实施方案中,由被感染的细胞系表达的痘病毒宿主范围基因是CP77。如实施例4中所示,当CHO细胞核被修饰以编码和表达CP77时,它能够如同它是允许细胞系(诸如143B)一样维持病毒扩增。典型地,在感染后2天内观察到融合的噬斑。

[0125] 在另一个实施方案中,在经修饰的细胞中的病毒繁殖水平提供至少10至5000的扩增率。在一些实施方案中,扩增率在500和3000之间,或在1000和4000之间。

[0126] 在另一个实施方案中,驱动异源痘病毒宿主范围基因的表达的启动子提供了细胞中痘病毒宿主范围异源基因表达的水平。CP77表达的水平可以与由牛痘病毒在允许细胞中产生的表达的水平相似,或超过牛痘病毒在允许细胞中产生的表达的水平。

[0127] 在另一个实施方案中,驱动异源痘病毒宿主范围基因的表达的启动子提供了足以允许病毒繁殖至少达到在允许细胞中病毒繁殖的水平的细胞中异源基因表达的水平。

[0128] 在一个实施方案中,CHO细胞系中的痘病毒生产与阳性对照细胞中痘病毒产生的水平相等或超过阳性对照细胞中痘病毒产生的水平。

[0129] 在一个实施方案中,CHO细胞中MVA病毒产生的水平与CEF细胞中MVA病毒产生的水平大体相等或超过CEF细胞中MVA病毒产生的水平。

[0130] 在一个实施方案中,CP77、K1L和/或SPI-1由为了在哺乳动物细胞中表达而密码子优化的核苷酸的连续序列编码。

[0131] 如本文进一步描述的,编码CP77的密码子优化的核酸序列可以与编码Brighton

Red毒株的牛痘CP77蛋白的序列(UniprotKB/Swiss-Prot:P12932.1)具有少于80%或少于70%的核苷酸序列同一性。在一些实施方案中,密码子优化的序列具有SEQ ID NO:1列出的序列,或者密码子优化的序列是包含与SEQ ID NO:1列出的序列具有至少70%序列同一性的核酸序列的功能变体。在一些实施方案中,CP77病毒宿主范围因子具有SEQ ID NO:2列出的氨基酸序列,或者与SEQ ID NO:2列出的氨基酸序列具有至少70%氨基酸序列同一性。

[0132] 在一些实施方案中,构思了试剂盒,所述试剂盒包括如本文描述的从其基因组表达CP77的经修饰的哺乳动物细胞的群体,诸如克隆群体,或主要由从其基因组表达CP77的经修饰的哺乳动物细胞的群体,诸如克隆群体组成。在一些实施方案中,经修饰的细胞不包含痘苗病毒。

[0133] 本说明书还描述了用于制造在CHO细胞中不繁殖的痘病毒的工艺或方法,该工艺包括在哺乳动物细胞系中体外繁殖痘病毒,其中该细胞系被修饰以在启动子的控制下编码并表达CP77。此工艺可以还包括分离病毒颗粒。

[0134] 方便地,细胞系是本领域技术人员已知的适用于药物或治疗剂、诊断剂或预防剂的制造的哺乳动物细胞系。

[0135] 说明书描述了一种经修饰的CHO细胞,其中该CHO细胞被修饰以在启动子的控制下从其基因组编码CP77并表达CP77。

[0136] 在一个实施方案中,经修饰的CHO细胞系维持在未经修饰的对照CHO细胞中较不能繁殖或不能繁殖的病毒的繁殖,所述未经修饰的对照CHO细胞是不表达CP77的CHO细胞。

[0137] 在一个实施方案中,病毒是除编码CP77的正痘病毒以外的正痘病毒。如本领域已知的,牛痘病毒是编码CP77的痘病毒。

[0138] 在一些实施方案中,病毒是痘苗病毒或大体上在人/受试者中在体内非复制性的痘苗病毒衍生物。

[0139] 在一些实施方案中,病毒是MVA。

[0140] 在另一个实施方案中,说明书提供了繁殖在人中大体上非-复制性的痘病毒的方法,该方法包括:(i)培养已经被转化以表达CP77的CHO细胞;和(ii)用该在人中大体上非-复制性的痘病毒感染来自(i)的培养的CHO细胞。

[0141] 在另一个实施方案中,说明书提供了繁殖在人中大体上非-复制性的痘病毒的方法,该方法包括:(i)培养已经被转化以表达CP77和D13L和/或K1L的CHO细胞;和(ii)用该在人中大体上非-复制性的痘病毒感染来自(i)的培养的CHO细胞。

[0142] 在另一个实施方案中,说明书提供了繁殖MVA的方法,该方法包括:(i)培养已经被转化以表达CP77和D13L和/或K1L的CHO细胞;和(ii)用MVA感染来自(i)的培养的CHO细胞。

[0143] 在另一个实施方案中,说明书提供了繁殖在人中大体上非-复制性的痘苗病毒衍生物的方法,该方法包括:(i)培养已经被转化以表达CP77和D13L和/或K1L的CHO细胞;和(ii)用该痘苗病毒衍生物感染来自(i)的培养的CHO细胞。

[0144] 在另一个实施方案中,说明书提供了繁殖编码异源蛋白的MVA的方法,该方法包括:(i)培养已经被转化以表达CP77和D13L和/或K1L的CHO细胞;和(ii)用该MVA感染来自(i)的培养的CHO细胞。

[0145] 在另一个实施方案中,说明书提供了繁殖在人中大体上非-复制性的、编码异源蛋白的痘苗病毒衍生物的方法,该方法包括:(i)培养已经被转化以表达CP77和D13L和/或K1L

的CHO细胞;和(ii)用该痘苗病毒衍生物感染来自(i)的培养的CHO细胞。

[0146] 在另一个实施方案中,本说明书提供了包含病毒宿主范围基因的核酸序列的人工制造的载体、多核苷酸或质粒,所述病毒宿主范围基因的核酸序列可操作地连接到用于在哺乳动物细胞系中表达的调节元件诸如启动子。在一些实施方案中,病毒基因是包含牛痘锚蛋白重复结构域的蛋白CP77基因(UniProtKB Swiss-Prot P12932.1[025LBR CP77protein]。在一个实施方案中,为了在哺乳动物细胞系中表达,对病毒宿主范围基因诸如CP77进行密码子优化。适当的载体和质粒是本领域已知的。

[0147] 在一个实施方案中,多核苷酸编码CP77。在一些实施方案中,多核苷酸包含在SEQ ID NO:1中列出的核苷酸序列(为了在哺乳动物细胞诸如CHO中表达,被密码子优化)。在一些实施方案中,分离的多核苷酸包含在SEQ ID NO:1中列出的核苷酸序列或编码在SEQ ID NO:2中列出的氨基酸序列的其变体。

[0148] 在另一个实施方案中,本说明书描述了用于将病毒宿主范围基因稳定插入到哺乳动物细胞的转座递送载体。

[0149] 本说明书还提供了将大体上对病毒非允许的哺乳动物或更高级的真核培养细胞转化成对该病毒允许的细胞的方法,该方法包括转化该细胞以表达CP77。

[0150] 在一些实施方案中,该方法包括用能够指导在哺乳动物启动子的控制下表达编码的CP77的载体转染细胞。

[0151] 在一些实施方案中,该载体是在哺乳动物启动子的控制下编码CP77的转座递送载体。

[0152] 在一个非限制性的实施方案中,启动子是指导在缺乏对宿主细胞的显著的毒性作用下维持病毒繁殖的充分水平的CP77的表达的细胞组成型启动子,诸如人EF1 α (人延伸因子1 α 基因启动子)、DHFR(二氢叶酸还原酶基因启动子)或PGK(磷酸甘油酸激酶基因启动子)。启动子也可以是诱导型的,诸如细胞诱导型启动子,MTH(来自金属硫蛋白基因)。病毒启动子,诸如CMV、RSV、SV40和MoU3,也在哺乳动物细胞中使用。

[0153] 在一些实施方案中,细胞是CHO细胞。

[0154] 在一些实施方案中,病毒是痘病毒。

[0155] 在一些实施方案中,痘病毒是在人中非-病原性或非-复制性的痘苗病毒衍生物。

[0156] “分离的”意指大体上或基本上不含通常在该材料天然状态下伴随的组分材料。例如,如本文使用的,“分离的多核苷酸”或“分离的多肽”等指多核苷酸或多肽分子从其天然细胞环境和从与细胞其他组分的关联的体外分离和/或纯化。不加限制地,分离的组合物、复合物、多核苷酸、肽或多肽可以指通过纯化分离的天然序列或指由重组或合成方法产生的序列。

[0157] 变体包括与参照分子或其互补形式的全部或部分充分相似的核酸分子,使得在中严格度或高严格度条件下可以实现选择性杂交,或在包含至少约15个核苷酸的比较窗口中,与定义参照痘病毒宿主范围因子的核苷酸序列具有约60%至90%或90%至98%的序列同一性。优选地,杂交区域为约12至约18个核苷碱基或更大的长度。优选地,特定核苷酸序列和参照序列之间的同一性百分比为至少约80%或85%,或更优选地约90%相似或更大,诸如约95%、96%、97%、98%、99%或更大。涵盖80%和100%之间的同一性百分比。核苷酸序列的长度取决于其提议的功能。涵盖同源物。术语“同源物(homolog)”、“同源基因”或“同

源物(homologs)”广泛地指功能上和结构上相关的分子,包括来自其他种的那些。同源物和直向同源物是变体的实例。

[0158] 核酸序列同一性可以用以下方式确定。本发明的核酸序列被用于使用程序BLASTM版本2.1搜索核酸序列数据库,诸如GenBank数据库(可以从网址<http://www.ncbi.nlm.nih.gov/blast/>进入)(基于Altschul等人,(1997)Nucleic Acids Research 25:3389-3402)。以无缺口模式使用程序。使用默认过滤以除去由于低复杂度的区域的序列同源性。使用BLASTM的默认参数。

[0159] 氨基酸序列同一性可以用以下方式确定。本发明的多肽序列被用于使用BLASTP程序搜索多肽序列数据库,诸如GenBank数据库(可以从网址<http://www.ncbi.nlm.nih.gov/blast/>进入)。以无缺口模式使用程序。使用默认过滤以除去由于低复杂度的区域的序列同源性。使用BLASTP的默认参数。可以使用SEG程序过滤低复杂度的序列。

[0160] 术语“在严格条件下杂交”和其语法同义词,指核酸分子与靶核酸分子(诸如在DNA或RNA印迹,诸如Southern印迹或Northern印迹上固定的靶核酸分子)在规定的温度和盐浓度的条件下杂交的能力。对于长度大于约100碱基的核酸分子,典型的严格杂交条件是不超过低于天然双螺旋的解链温度(T_m)的 25°C 至 30°C (例如, 10°C)(参见,一般的,Sambrook等人,(以上);Ausubel等人,(1999))。对于大于约100个碱基的核酸分子的 T_m 可以用公式 $T_m = 81.5 + 0.41\% (G+C - \log(\text{Na}^+))$ 计算。对于具有长度小于100个碱基的核酸分子,示例性的严格杂交条件是低于 T_m 5°C 至 10°C 。

[0161] “载体”意指衍生自例如质粒、噬菌体、酵母、病毒、哺乳动物、鸟类、爬行动物或鱼的多核苷酸分子,适当地DNA分子,多核苷酸可被插入或克隆至其中。载体优选地包含一个或更多个独特的限制位点并且能够在定义的宿主细胞包括靶细胞或组织或其祖细胞或组织中自主复制,或可以与定义的宿主的基因组整合,使得克隆的序列是可以复制的。因此,载体可以是自主复制的载体,即,作为染色体外实体存在的载体,其复制独立于染色体复制,例如,线性或闭合环状质粒、染色体外元件、微型染色体或人工染色体。载体可以包含用于保证自体复制的任何工具。可选地,载体可以是,当被引入到宿主细胞时,被整合到基因组并且与其被整合进入的染色体一起复制的载体。载体系统可以包括单个载体或质粒、两个或更多个载体或质粒,其一起包含待被引入到宿主细胞的基因组的全部DNA,或转座子。载体的选择通常将取决于载体与该载体将被引入的宿主细胞的相容性。载体还可以包括可以用于选择适当的转化体的选择标记诸如抗生素抗性基因。这样的抗性基因的实例是本领域的技术人员已知的。

[0162] 对“基因”的指代包括对应于基因的外显子的cDNA。本文中对“基因”的指代还包括:由转录和/或翻译调节序列和/或编码区域和/或非翻译序列(即,内含子、5'-和3'-非翻译序列)组成的经典基因组基因;或对应于编码区域(即外显子)和基因的5'-和3'-非翻译序列的mRNA或cDNA。

[0163] “调节元件”或“调节序列”意指特定宿主细胞中可操作地连接的编码序列的表达必需的核酸序列(例如,DNA)。适用于原核细胞的调节序列,例如,包括启动子和任选地顺式作用序列,诸如操纵子序列和核糖体结合位点。适用于真核细胞的控制序列包括启动子、多腺苷酸化信号、转录增强子、翻译增强子、调节mRNA稳定性的前导序列或尾随序列,以及将由转录的多核苷酸编码的产物靶向到细胞内的细胞内区室或靶向到细胞外环境的靶向序

列。

[0164] 涵盖这些序列的互补序列和部分。当与核酸分子相关使用时,术语“互补”和“互补的”指如通过沃森-克里克碱基配对确定的互补的核酸序列。例如,核酸序列5' CCATG3' 的互补序列是5' CATGG3' 。

[0165] 短语“与……特异性杂交”等指一个分子在严格条件下只与特定核苷酸序列结合、形成双链或杂交,当该特定核苷酸序列存在于复杂混合物(例如,总细胞)DNA或RNA中时。

[0166] 术语“受试者”或“个体”或“患者”在本文中可互换使用,指期望疗法或预防的任何受试者,特别是脊椎动物受试者,并且甚至更特别是哺乳动物受试者。落入本发明的范围的适当的脊椎动物包括但不限于,灵长动物、鸟类、家畜动物(例如,绵羊、牛、马、驴、猪)、实验室试验动物(例如,兔子、小鼠、大鼠、豚鼠、仓鼠)、伴侣动物(例如,猫、狗)和圈养的野生动物(例如,狐狸、鹿、澳洲野犬)。优选的受试者是需要治疗或预防的灵长动物诸如人。但是,应理解,上述术语不暗示出现症状。

[0167] 通过以下的非限制性实施例进一步描述本文可行的各种实施方案。

[0168] 实施例1

[0169] VACV-COP不能在CHO细胞中生长

[0170] 材料和试剂

[0171] -VACV-COP、VSS02、SEM120213,滴度: 1.6×10^8 pfu/mL

[0172] -Vero:WHO-VERO-MCB passage No 141,08/08/2005,Virax Holdings Limited

[0173] -CHO:SA-Pathology 7.05.2004

[0174] -生长培养基:RPMI,10%FBS,Pen/Strep

[0175] -维持培养基:RPMI,2%FBS,Pen/Strep

[0176] 以每细胞系一个6孔板(6-WP)将CHO和Vero细胞培养至融合。在室温下,用 4×10^4 pfu VACV-COP VSS02感染每孔45分钟,并且随后在37°C/5%CO₂孵育。对于每种细胞系,在感染后24h、48h、78h收获并合并2孔的内容物。通过冻融3次产生病毒提取物并且在-80°C储存直到准备好滴定。如实施例4中描述的方案描述的,使用Vero细胞滴定每种提取物。

[0177] 冻融后,可以使用匀质探头来破坏这些大的不溶的团块。待被收获的每孔包含2mL MM。对于每个时间点,合并2孔以产生每时间点的4mL的总体积。可以加入TE缓冲液以得到每时间点的6mL的总体积。

[0178] 滴定结果、病毒产量结果和生产收率在表1中制成表格。结果示出了,VACV-COP不能在CHO细胞中从低moi繁殖,不像Vero细胞,在Vero细胞中病毒生产随着时间增加。VACV-COP在CHO细胞中是非-允许的。

[0179] 实施例2

[0180] VACV+PH22[CP77]在CHO中对Vero中的多步生长——CP77在VACV中活跃

[0181] 在多步生长研究中与Vero细胞比较,来确定表达编码CP77的牛痘病毒BR025L基因的重组VACV-COP(VACV-PH22[CP77])在CHO的繁殖潜力。

[0182] VACV-PH22是表达来自编码CHO宿主范围蛋白CP77的牛痘病毒Brighton毒株的天然025L ORF的重组痘苗病毒哥本哈根毒株。痘苗病毒中此ORF也在BR025L天然启动子的控制下。克隆天然BR025L基因(天然启动子和ORF)以产生还编码红色荧光蛋白DsRed-Express2的pPH22。pPH22是将BR025L基因和DsRed基因插入到VACV-COP的B19R ORF的整合

载体, B19R ORF编码可溶的和细胞表面IFN α / β 受体蛋白。将BR025L和DsRed插入到VACV-COP通过同源重组实现, 所述同源重组是用低复制性的VACV-COP感染CHO和用pPH22转染的结果。只有包含BR025L基因的病毒将继续在CHO细胞中扩增, 这可以通过被红色荧光感染的细胞或噬斑的存在目视确认。

[0183] 将同源重组后3天提取的病毒在CHO中扩增3次, 并且随后在Vero细胞中滴定, 之后在CHO细胞中的此多步生长研究中使用。

[0184] 材料和试剂

[0185] -VACV-COP、VSS02、SEM120213, 滴度: 1.6×10^8 pfu/mL

[0186] -VACV-PH22[CP77], 滴度: 8×10^6 pfu/mL

[0187] -Vero: WHO-VERO-MCB passage No 141, 08/08/2005

[0188] -CHO: SA-Pathology 7.05.2004

[0189] -生长培养基: RPMI, 10% FBS, Pen/Strep用于CHO和Vero

[0190] -维持培养基: RPMI, 2% FBS, Pen/Strep用于CHO和Vero

[0191] 以每细胞系2x T25烧瓶将CHO和Vero细胞培养至融合。在室温下, 用 1×10^5 pfu VACV-COP VSS02和 1×10^5 pfu VACV-PH22感染每种细胞系的1个烧瓶45分钟, 并随后在37°C/5% CO₂孵育。对于每种细胞系, 在感染后96h收获每个烧瓶的内容物。通过冻融3次产生病毒提取物并且在-80°C储存直到准备好滴定。如实施例4中详细描述的方案描述的, 在24孔板格式下使用Vero细胞滴定每种提取物。

[0192] 在T25烧瓶中进行感染, 每细胞系2个烧瓶, 其中用VACV-COP感染烧瓶1, 并且用VACV-PH22感染烧瓶2。仅在感染后96h进行收获。结果在表2中制成表格。从表2可以看出: CHO细胞不能支持VACV-COP病毒后代产生。这证实了实施例1描述的结果。此外, 当感染本研究中的允许细胞系即Vero细胞时VACV-COP有活力, 将接种物水平扩增至超过200倍, 因此CHO中观察到的结果是由于宿主细胞限制。但是, 当CP77被重组痘苗病毒VACV-PH22表达时, 接种物的扩增达到约700倍。CP77的表达不将痘苗病毒宿主范围限制于仅CHO细胞, 因为用VACV-PH22感染vero细胞也扩增接种物。但是, 扩增水平可能不如无CP77的痘苗病毒好。由于滴定结果的标准差太大, 以至于产量的区别可能是不显著的。

[0193] 这些结果示出了, 通过产生与从允许细胞基底诸如Vero预期的产量相似的产量, 从病毒基因组表达的CP77足够使VACV-COP在非-允许CHO细胞中扩增。

[0194] CHO细胞的痘苗病毒感染期间CP77蛋白的可能的功能已经由Hsiao等人, (2006) 报道。他们提出, CP77与HMG20A结合并且从位于病毒工厂的新合成的痘苗病毒基因组除去HMG20A, 从而使痘苗病毒生命周期在CHO细胞中继续。本文假设, 由于HMG20a与基因组结合并且“锁定”基因组, CP77使新合成的基因组可用于在CHO细胞中原本不可得的包装。由于在允许细胞系诸如Vero中, 病毒扩增不要求CP77的功能, 本文提出, 存在具有此功能的代替的等效蛋白, 甚至细胞蛋白, 所述等效蛋白可能在至少CHO细胞中失活或不存在, 但在允许细胞系中有活性或存在。该代替的蛋白可能导致不需要CP77, 并且因此在痘苗病毒的进化中, 它可以随后被缺失或重排, 因为它的功能缺失对广泛宿主范围不重要。

[0195] 实施例3

[0196] p-LL07-CHO (表达CP77的多克隆细胞系) 的构建

[0197] 构建CHO细胞系以表达CP77蛋白。表达绿色荧光蛋白 (EGFP) 的VACV-COP重组病毒

形成噬斑,其在感染的数天内发展成融合感染。

[0198] 痘苗病毒-COP (SCV401C) 是已经插入A39R ORF表达盒的哥本哈根毒株的重组痘苗病毒 (VACV-COP), 该A39R ORF表达盒由可操作地连接到增强的绿色荧光蛋白 (EGFP) 的蛋白编码序列的强痘苗病毒早期/晚期启动子组成并由痘病毒早期转录终止序列终止。在非允许细胞和允许细胞的感染时,EGFP将在被感染的细胞中表达,所述EGFP可以使用荧光显微镜显现。在允许细胞中,可以观察到随着病毒在细胞群体中扩散,绿色荧光从细胞到细胞扩散。

[0199] 由GeneArt GmbH (Germany) 通过从由牛痘病毒Brighton Red毒株UniProtKB/Swiss-Prot:P12932.1的025L ORF编码的CP77的氨基酸重建(回译(back translation)或逆向翻译(reverse translation))DNA序列合成地制造CP77蛋白编码序列,并且为了在哺乳动物细胞(CHO)细胞中表达,进行密码子优化。为了CP77蛋白编码序列的表达,参见SEQ ID NO:1的密码子优化的核苷酸序列。

[0200] 使用PCR引物将来自pPH51(含有密码子优化的CP77蛋白编码序列的克隆质粒)的密码子优化的CP77蛋白编码序列PCR扩增,其中5'引物被设计为在起始密码子周围加入kozak序列,并且3'引物被设计为在终止密码子之前加入flag-标签序列。将扩增的PCR产物通过Bsa I克隆位点亚克隆到从DNA2.0 Inc (USA) 购买的转座子piggybac载体pJ507-2 (Hyg⁺) 中。克隆到BsaI中有效地用CP77蛋白编码序列代替CometGPF编码序列以产生pLL07。现在,CP77成为在人组成型延伸因子1 α 启动子(EF1 α)的控制下并且在CP77与潮霉素抗性基因已经被稳定地整合到转染的细胞的基因组中之后,CP77与潮霉素抗性基因二者共表达。

[0201] 通过转座子辅助的CP77和潮霉素抗性表达盒的稳定插入转导CHO:将CHO细胞接种到6孔板的孔中,这样过夜孵育后它们为约50%融合度。使用来自Qiagen的Effectene转染试剂,按照生产商的说明,将1 μ g pLL07转染到1个孔的50%融合的CHO细胞。随后将转染细胞在生长培养基中孵育过夜。第二天,使用包含500 μ g/mL潮霉素B的生长培养基更换培养基用于选择转导的细胞。每2-3天更换选择培养基并在转导的细胞开始大量生长时,使用TrypLE Select (Life Technologies)将其回收并接种在T25烧瓶中用于进一步细胞扩增。

[0202] 通过蛋白质印迹法验证CP77表达(加Flag-标签的)

[0203] 产生兔抗-CP77血清:用连接到代表CP77蛋白的短内部氨基酸序列的KHL蛋白的15个氨基酸肽注射兔。此氨基酸序列为:SGSDVNIRSNNGYTC——UniProtKB Swiss-Prot P12932.1[025LBR Cp77蛋白]SEQ ID NO:2的氨基酸位置481至495。

[0204] 进行总计三次注射,间隔1个月,以产生对此KHL缀合的氨基酸序列的抗体。用蛋白质印迹法检验来自被注射的兔的血液的细菌表达的和Ni-NTA纯化的N-末端加His-标签的CP77蛋白。显示出此兔抗CP77血清清楚地识别重组CP77蛋白。

[0205] 检验p-LL07-CHO的CP77表达:用CHO和p-LL07-CHO接种2个T25烧瓶并且培养直到每个烧瓶中的细胞单层达到100%融合。收获来自每个烧瓶的细胞、用PBS洗涤并且随后在200 μ L中重悬。为此,向每管重悬细胞加入50 μ L 5X SDS-PAGE加样缓冲液,并且随后在98 $^{\circ}$ C孵育5min。将15 μ L每种细胞蛋白提取物加载到2个10%SDS-PAGE凝胶、电泳并随后通过电-印迹布点到Hybond ECL硝酸纤维素。

[0206] 随后在室温下用溶解于包含吐温20的Tris缓冲盐水(TBST)中的5%脱脂奶粉处理电-印迹的膜1小时以阻断膜上所有可用的非特异性抗体结合位点。随后,在用抗体探测前,

将膜在TBST中洗涤数次。在4℃,用HRP缀合的抗-DDDDK标签抗体[M2](ab49763,Sapphire Bioscience)的1:5000稀释液探测膜1过夜。在4℃,用抗-CP77抗血清的1:100稀释液探测膜2过夜,用TBST洗涤3次,随后用二抗,HRP缀合的抗-兔抗体(GE Healthcare)的1:5000稀释液探测2小时。随后,将2个膜在TBST中洗涤3次,并且用ECL蛋白质印迹检测试剂(GE Healthcare)处理并暴露于X-射线底片,如使用手册指导的。

[0207] CHO和p-LL7-CHO中的噬斑试验:将CHO和p-LL7-CHO细胞接种到多个6孔板中并且培养直到细胞单层达到100%融合。

[0208] 使用包含在早期/晚期痘苗病毒启动子控制下的EGFP表达盒(绿色荧光蛋白)的重组痘苗病毒(哥本哈根毒株)(SCV401C)感染细胞。由于SCV401C的未知的滴度,先在1ml MM培养基中稀释10 μ l储备病毒(Dil1:1:100稀释),并且随后用500 μ l病毒稀释液感染每孔。感染后1天,注意到高moi感染,其中大多数细胞是荧光绿。决定进行Dil 1的进一步的1:20稀释,通过将100 μ l Dil 1加入到2ml MM培养基中(Dil 2)进行。随后用500 μ l Dil 2将其用于感染每孔。

[0209] 在带有GFP滤光器(Cat#U-MGFPHQ,Olympus)的荧光显微镜(Olympus IX51)下观察病毒感染。使用cellSens数字成像软件(Olympus)捕获图像。

[0210] 从结果看出,如预期地,表达绿色荧光蛋白的VACV-COP(SCV401C)在CHO细胞中不繁殖。单细胞荧光绿是病毒进入细胞、表达其包括EGFP的基因但不能产生新的感染病毒颗粒并因此不能将感染扩散到邻近细胞的结果。但是,在表达CP77的CHO细胞系中,截止感染后第1天,痘苗病毒能够产生扩散到邻近细胞以形成感染焦点的新的感染性病毒。截止感染后接下来2天,这些感染焦点变成融合感染,截止到第3天,整个细胞单层被SCV401C感染。与亲本(原始)CHO细胞不同,表达宿主范围蛋白CP77的CHO细胞允许痘苗病毒感染。

[0211] HMG20A属于包含HMG盒结构域的蛋白质家族。HMG蛋白是识别扭曲的DNA结构,诸如十字形(cruiciform)的染色体重塑蛋白。它们还可以通过与DNA的小沟结合来诱导DNA弯曲。因此,包含HMG盒的蛋白质被认为在DNA复制、重组或修复期间对染色体重塑重要。此外,某些包含HMG盒的蛋白质可以通过与转录因子在启动子位点相互作用而影响基因转录。

[0212] Hsiao等人,2006发表的研究表明,在CHO-K1细胞中CP77与HMG20A结合。此宿主细胞蛋白,HMG20A,表现为在被痘苗病毒感染的CHO细胞中的病毒工厂与病毒DNA结合,并且假设该宿主细胞蛋白“锁定”病毒工厂中的DNA并且随后阻止痘苗病毒生命周期的下一个阶段并从而防止后代感染病毒的产生。CP77被牛痘病毒的表达表现为从病毒DNA除去了宿主HMG20A并且允许病毒生命周期重新开始,最终产生后代感染性病毒。

[0213] 但是,随着在缺乏病毒感染下CHO中CP77的表达,人们将预期此蛋白质在存在于胞质的新合成的HMG20A移位到细胞核之前就将其隔离。如果是这样的情况,细胞核中HMG20A的功能将会丢失,并且由于它在DNA复制、重组和修复期间起重要作用以及其在基因转录期间的功能,人们将预期CP77的表达将在细胞增殖和维持期间损害CHO细胞的完整性。出人意料地,情况表现为不是这样,因为随着连续培养,表达CP77的CHO细胞系容易地维持,与亲本CHO细胞系相比,没有对其经许多世代复制的能力有明显影响。

[0214] 实施例4

[0215] 多步生长研究

[0216] 在p-LL07-CHO中进行多步生长动力学研究:评价表达宿主范围基因CP77的CHO细

胞系对痘苗病毒感染的允许性质并比较病毒产生的水平和在天然允许人细胞系-143B中获得的产生水平。

[0217] 目的是比较p-LL07-CHO、CHO和143B细胞系中VACV-COP特征性的繁殖。在此研究中,通过检查与VACV-COP在CHO(非-允许的)和143B(允许的)细胞中特征性的扩增相比,VACV-COP在p-LL07-CHO细胞系中特征性的扩增来检验由p-LL07-CHO表达的CP77的功能。

[0218] 材料和方法

[0219] 细胞系设置

[0220] CHO设置:用CHO细胞接种一个6-孔板(6WP)并培养直到在生长培养基(RPMI+10%FBS)中达到融合。

[0221] 143B设置:用143B细胞接种一个6-孔板(6WP)并培养直到在生长培养基(RPMI+10%FBS)中达到融合。

[0222] p-LL07-CHO设置:用p-LL07-CHO细胞接种一个6-孔板(6WP)并培养直到在生长培养基(RPMI+10%FBS+500ug/ml潮霉素B)中达到融合。

[0223] VACV-COP的稀释:用0.01pfu于总体积500uL感染每个6WP的每个孔。假设在100%融合,每孔的细胞计数为 4×10^6 细胞。对于0.01pfu/孔的感染率,要求 4×10^4 pfu每孔,并且因此,将储备病毒在维持培养基(RPMI/2%FBS)中稀释到 8×10^4 pfu/mL。

[0224] 感染:从每板的每孔除去培养基,加入500uL稀释的病毒,并且静置在室温下孵育1小时以使细胞吸附病毒。随后除去病毒接种物,并且用1mL无菌PBS洗涤每个感染的孔一次,并且随后在37°C/5%CO₂在每孔2mL MM中孵育。

[0225] 收获:在感染后以下时间点:24h、48h和72h收获来自每板的2组孔。在收获当天,将细胞刮到培养基中,其中合并来自每种细胞系的两组孔并且在1000g离心5分钟使细胞沉淀。将每种细胞沉淀在1mL 10mM TrisHCl pH8中重悬。然后,将重悬的细胞沉淀进行3个循环的冻融并且在-80°C储存直到准备好滴定。

[0226] 在于24孔板中培养至100%融合的Vero和143B细胞中进行滴定。

[0227] 稀释:从冰柜取出病毒提取物,解冻并超声处理以使任何可见的团块匀质化。将病毒提取物在维持培养基中系列稀释至 10^{-8} 。

[0228] 感染:从每孔除去生长培养基,并在室温下用每种稀释液的1mL(每种稀释液4个孔,每种病毒1个平板,从 10^{-2} 稀释开始)感染1h。孵育后,将每个板移动到培养箱并在37°C/5%CO₂孵育3天,以发展噬斑。

[0229] 滴度的计算:包含20至50个噬斑的稀释液,计算噬斑并然后平均。用稀释的倒数乘以平均计数,并且,因为使用1mL感染,所得数字将为以pfu/mL计的滴度。

[0230] 标准误差的计算:使用以下公式用组成平均数的4个滴定值计算95%置信的标准误差(SE): $1.95 \times (Sd / \sqrt{n})$,其中:SD是来自小样本的标准差,n是滴定重复的数目(此情况下为4)。

[0231] 产量的计算:这是被滴定的病毒提取物中的病毒的总量:平均滴定值(pfu/mL) X 病毒提取物的总体积=pfu

[0232] 扩增率的计算:此数值代表扩增倍数相对于用作接种物的量:以pfu计的产量/以pfu计的接种量。

[0233] 在6孔板上进行多步生长动力学研究-每时间点每细胞系两个孔。用 4×10^4 pfu

VACV-COP感染每孔,并且对于收获,收获并合并每个细胞系和时间点的2个孔,从其制备病毒提取物。然后滴定总体积为1mL、代表由接种量为 8×10^4 pfu产生的2个合并的感染(4×10^4 pfu X 2)的此病毒提取物。使用两个指示细胞系143B和Vero进行滴定。

[0234] 分别在表3和表4中将滴定和病毒产量结果制成表格。VACV-COP在非允许CHO细胞中不扩增,即,它产生比输入接种物中使用的量更少的病毒。但是,在表达宿主范围基因CP77的CHO细胞系中的病毒扩增为大于输入接种物约2000倍(基于使用143B细胞作为指示细胞系的滴定结果)。来自允许细胞系的扩增为大于输入接种物约3000倍。由于标准误差与p-LL07-CHO和143B细胞之间的扩增的差异重叠,该差异不是统计上显著的。

[0235] 如果使用来自Vero指示细胞的滴定结果比较两个细胞系之间的病毒扩增,p-LL07-CHO中的扩增将似乎比在143B细胞中的扩增稍多,但是这不是统计上显著的。

[0236] 在此研究中还证实,痘苗病毒在CHO细胞中是不允许的,因为病毒生产的水平比用作接种物的病毒的量低得多。但是如果CHO表达编码CP77蛋白的牛痘病毒宿主范围基因,它现在变得对痘苗病毒是允许的并且支持病毒生产达到在允许细胞系中见到的相同水平。还值得注意的是,只要求表达一种宿主范围基因CP77来将此细胞系转化成用于痘苗病毒的生产的“可用的”允许细胞基底。非常期望从CHO生产痘苗病毒,因为CHO是一种生物技术友好的细胞系,因为它与细菌生长得一样快,它可以在限定的合成培养基中培养而不需要生物添加剂诸如胎牛血清,并且可以作为细胞悬液在生物反应器中培养。CHO也具有生产生物药物产品的历史,已被充分表征并且被生物医药管理机构诸如FDA和EMA所知和喜爱。

[0237] 在组成型细胞启动子的控制下表达CP77蛋白的转基因CHO细胞系允许痘苗病毒复制和繁殖达到在产生高水平子代病毒的天然允许细胞系中观察到的相同水平。

[0238] 实施例5

[0239] 在CHO-CP77细胞中的MVA繁殖

[0240] 在CHO、143B和p-LL07-CHO中的MVA+GFP繁殖

[0241] 材料和方法

[0242] 生长培养基 (GM): RPMI-1640, 补充有10% FBS、2mM L-谷氨酰胺、青霉素和庆大霉素、Hepes

[0243] 维持培养基 (MM): RPMI-1640, 补充有2% FBS、2mM L-谷氨酰胺、青霉素和庆大霉素、Hepes

[0244] 培养p-LL07-CHO的注意事项: 在GM加上500ug/mL潮霉素B中进行p-LL07-CHO细胞系的一般繁殖和维持,但是,对于在感染前铺板 (plating out) 和培养达到100%融合,在无潮霉素B的GM中培养细胞。

[0245] 细胞设置: 将143B、CHO和P-LL07-CHO细胞接种到多个6孔板中并且在生长培养基 (GM) 中、37°C/5% CO₂下培养直到细胞单层达到100%融合。每细胞系培养1个板。

[0246] 感染

[0247] ●使用包含GFP表达盒(绿色荧光蛋白)的重组MVA感染培养于6-孔板中的细胞。由于不知道MVA-GFP的滴度,根据以下将病毒在维持培养基 (MM) 中系列稀释:

[0248] ○Dil 1: 将20ul储备病毒在2ml MM培养基中稀释(1:100稀释)并通过剧烈漩涡混合。

[0249] ○Dil 2: 将500ul Dil 1加入到4.5ml MM中(1:10³稀释)并通过剧烈漩涡混合。

[0250] ○Di1 3:将500ul Di1 2加入到4.5ml MM中(1:10⁴稀释)并通过剧烈漩涡混合。

[0251] ○Di1 4:将500ul Di1 3加入到4.5ml MM中(1:10⁵稀释)并通过剧烈漩涡混合。

[0252] ●用500ul的以下病毒稀释液Di1 2、Di1 3和Di1 4感染每板的一个孔。

[0253] ●将每板孵育经过5天时间段并检查荧光细胞的焦点的发展和扩散。在此研究中, Di1 4经过3天时间段产生可辨别的荧光细胞的焦点。每板的未感染的孔是自体荧光的对照。

[0254] 显微镜观察

[0255] 在带有GFP滤光器(Cat#U-MGFPHQ,Olympus)的荧光显微镜(Olympus IX51)下观察病毒感染。使用cellSens数字成像软件(Olympus)捕获图像。

[0256] 结果

[0257] 结果示出了,如预期地,表达绿色荧光蛋白(GFP)的MVA在CHO和143B细胞中不繁殖。单细胞荧光绿是病毒进入细胞、表达其包括GFP的基因但不能产生新的感染病毒颗粒并因此不能将感染扩散到邻近细胞的结果。但是,在表达CP77的CHO细胞系中,截止感染后第1天并进一步发展到第3天,MVA能够产生扩散到邻近细胞以形成感染焦点的新的感染性病毒。

[0258] 实施例6

[0259] 从p-LL07-CHO感染收获的MVA的宿主范围限制的确认

[0260] 细胞设置

[0261] 将143B、CHO和BHK-21细胞接种到多个6孔板中并且在生长培养基(GM)中、37℃/5%CO₂下培养直到细胞单层达到100%融合。每细胞系培养1个板。

[0262] 来自实施例5的病毒收获

[0263] 如下,在感染后第5天从实施例5的Di1 3感染的P-LL07-CHO孔收获MVA+GFP:

[0264] ●从孔收集上清液和细胞并在1000g离心5分钟以沉淀被感染的细胞。

[0265] ●在500ul 100mM Tris-HCl pH8缓冲液中重悬细胞沉淀。

[0266] ●将重悬的细胞沉淀冻融至少3次以从被感染的细胞释放病毒。

[0267] 感染

[0268] ●由于不知道粗制病毒提取物的滴度,用与实施例5相同的方法将病毒在MM培养物中系列稀释:

[0269] ●用500ul的以下病毒稀释液Di1 2、Di1 3和Di1 4感染每板的一个孔。

[0270] ●将每板孵育经过5天时间段并检查荧光细胞的焦点的发展和扩散。在此实验中, Di1 3经过3天时间段产生可辨别的荧光细胞的焦点。

[0271] 显微镜观察

[0272] 在带有GFP滤光器(Cat#U-MGFPHQ,Olympus)的荧光显微镜(Olympus IX51)下观察病毒感染。使用cellSens数字成像软件(Olympus)捕获图像。

[0273] 结果

[0274] 从表达CP77的CHO细胞系收获的MVA仍通过不能在非允许细胞系CHO(仓鼠)和143B细胞(人)中繁殖保持其受限的宿主范围。在CHO和143B细胞感染后经过3天时间段缺乏绿色荧光焦点证明了,在这些细胞系中不产生感染性后代病毒。但是,此MVA仍保持其对BHK21细胞(仓鼠)的宿主范围,因为截止感染后1天可以看到感染的绿色荧光焦点,该焦点经过之后

3天其大小增加。

[0275] 结论

[0276] ●与亲本(原始)CHO细胞不同,表达宿主范围蛋白CP77的CHO细胞允许MVA感染。

[0277] ●在表达CP77的CHO细胞系中繁殖的MVA不增加其宿主范围至非允许细胞系诸如CHO和143B(人)。

[0278] 实施例7

[0279] pLL07的构建

[0280] 背景信息:将来自pPH51DNA模板的CP77(CHO密码子优化的)CDS PCR产物通过Clontech's InFusion克隆系统克隆到pJ507-2(Hyg+)PiggyBac系统中以产生pLL07。将加flag-标签的CP77序列插入到pJ507-2的BsaI,从而除去CometGFP序列(SEQ ID NO:3)。

[0281] 用于将来自pPH51质粒DNA的CP77-CHO基因PCR扩增的PCR引物对:AACACGTCTCGGGGgcccgccaccATGTTCTGACTACCTGGAAAATGAGGAAGTG(SEQ ID NO:4)和CAGGAAGACGCTTTTtca**CTTGTCATCGTCATCCTTGTAATC**CTGCTGCTCGAAGATCTTGTA(TACT)(SEQ ID NO:5)。在Inf-LL07-CP77-Rv引物中以黑体示出了Flag标签序列。

[0282] 质粒插入/盒配置(Plasmid INSERT/CASSETTE Configuration)如下。插入/盒图。将pLL7克隆#3送去测序。

[0283] 将15ABI测序文件和pLL07参考文件“pLL07_ref.sbd”一起输入到Lasergene's DNASTar Seqman计算机程序并组装成一个一致重叠群(consensus contig)。整理对齐序列以匹配参照序列的开始和结束,之后从对齐中删除参照序列以对一致序列的建立不产生影响。以名称为“pLL07_#3consensus.seq”的DNA序列文件保存重叠群的一致序列。确定重叠群的读取方向,发现其代表与参照序列相同的读取方向。使用Megalign(DNASTar)将“pLL07_#3consensus”与“pLL07_ref.sbd”参照序列人工对齐以帮助鉴定参照序列和pLL07_#3中的序列之间的差异。使用Seqman(默认设置)对来自AGRF-测序服务的pLL07Clone#3装配15ABI文件,产生一个覆盖LL07插入参照序列全长的一致重叠群。在pLL07_#3的序列和参照序列之间不存在差异。pLL07_#3重叠群一致中的插入序列与参照序列相同。

[0284] 实施例8

[0285] G1L和I7L在哺乳动物细胞中的表达

[0286] 构建编码加flag-标签的-G1L或加flag-标签的I7L的表达质粒并将其用于产生表达这些蛋白的转基因143B细胞系。尽管这些细胞已经在遗传霉素(Geneticin)的存在下扩增以阳性选择经转导的细胞,并且PCR分析已经确认了这些细胞系的基因组中这些蛋白编码序列的存在,当用抗DDDDK抗体探测时,蛋白质印迹分析未能检测表达的加Flag-标签的-蛋白质的存在。

[0287] C-末端加flag-标签的COP-G1L和COP-I7L的蛋白编码序列由GeneArt(Life Technologies)合成并且亚克隆到从DNA2.0Inc(USA)购买的pJ503-2(piggyBac-新霉素抗生素选择)的Bsa I位点。此克隆过程交换CometGFP为加Flag-标签的G1L或加Flag-标签的I7L蛋白编码序列。对于G1L piggyBac载体,所得克隆被命名为pLL08,并且对于I7L piggyBac载体,所得克隆被命名为pLL10。

[0288] 这两个质粒载体的关键特征是:加flag-标签的蛋白编码序列在组成型人启动子

EF1 α 的控制下,这些表达盒与NPT II表达盒(新霉素抗性基因)一起两侧是左转座子边界和右转座子边界以形成人工转座子元件。在人工转座子元件外但包含于相同质粒内的是介导转座元件不可逆地整合到宿主基因组中的转座子酶表达盒。通过将G418(遗传霉素)并入到细胞生长培养基中可以阳性选择携带这些转座子元件的细胞。

[0289] 质粒载体pLL08(加Flag-标签的-G1L)和pLL10(加Flag-标签的-I7L)被转染到143B细胞中以产生2种经转导的转基因细胞系:G1L-143B(包含G1L表达盒)和I7L-143B(包含I7L表达盒)。通过在生长培养基中包括遗传霉素将具有成功的转座子整合的细胞扩增至可工作的量。为了证明成功的整合,依照试剂盒的使用手册的指导,使用来自Qiagen的DNeasy DNA提取试剂盒从这些转基因细胞提取总细胞DNA。然后将提取的DNA用作模板以使用对G1L和I7L的PCR扩增特异性的PCR引物对的PCR扩增反应。对于其基因组中G1L和I7L DNA序列的存在,两种细胞系都是阳性。

[0290] 为了检测这些细胞系的G1L和I7L表达,通过使用缀合至HRP的抗-flag标签抗体[M2](抗-DDDDK抗体,Abcam#ab49763)检测每种加flag-标签的蛋白的存在来进行蛋白质印迹分析。使用从G1L-143B和I7L-143B细胞系提取的总蛋白的这些蛋白质印迹分析的结果显示出,如预期的,抗-flag-标签抗体不识别从143B提取的任何蛋白,并且如果它可以识别flag-标签-CP77转基因CHO细胞系(CP77-CHO)中表达的flag-标签蛋白,证明抗-flag-标签抗体可以识别加flag-标签的蛋白。但是,在来自G1L-143B样品的蛋白提取物中未能检测到加flag-标签的G1L蛋白。细菌表达的加flag-标签的I7L蛋白可以用抗-flag-标签抗体检测到,但该抗体未能检测I7L-143B细胞系的加flag-标签的I7L表达。

[0291] 由于在其相应细胞系中可以检测到G1L和I7L表达盒,结论可以是在缺乏痘苗病毒感染下,表达的蛋白质在合成后被快速降解,即G1和I7蛋白二者都是病毒特异性酶并且可以在缺乏其痘苗病毒特异性酶底物时不稳定,或者驱动这两个表达盒的启动子有缺陷。另一个假设,在缺乏痘苗病毒感染下,这些蛋白的表达对细胞“有毒”,并且在遗传霉素选择方法期间,具有沉默的转基因G1L和I7L表达盒而非NPT II表达盒的细胞使不表达G1L或I7L蛋白的遗传霉素抗性细胞扩增。由于在我们的实验室成功使用了pJ503-2piggyBac质粒进行COP-D13L在143B细胞中的表达,提出EF1 α 启动子是功能性的并且这些蛋白质在缺乏痘苗病毒感染下不稳定,或者驱动这些蛋白质的表达的启动子被选择性沉默以阻止在遗传霉素存在下的细胞扩增期间表达“有毒的”蛋白质。

[0292] 实施例9

[0293] 作为重要的结构成熟或装配蛋白的实例,D13L由哺乳动物细胞表达并救援具有D13L基因缺失的痘苗病毒,并且D13L缺失的痘苗病毒在受感染的细胞中表达蛋白质

[0294] D13-救援细胞系的构建——作为通过阻断装配/成熟过程使痘苗病毒减毒的实例,把哥本哈根毒株的D13L ORF作为用于缺失的靶。这样做时,首先必须构建表达此蛋白的细胞系,这样可以繁殖COP-D13L-缺失的病毒。对于救援细胞系的构建,选择中国仓鼠卵巢细胞系,通常被称为CHO,因为此细胞系是“生物技术”友好的。为了用COP-D13L缺失救援感染性的痘苗病毒,此细胞系必须使用细胞的转录机器而非痘苗病毒的转录机器从细胞的核基因组表达D13-蛋白。细胞表达的D13-蛋白的蛋白质氨基酸序列包含C-末端加标签的DYKDDDDK(Flag-标签,Hopp等人,1988)的氨基酸序列并且为了产生相应的核苷酸序列进行CHO密码子优化。将由哺乳动物启动子和哺乳动物多腺苷酸化信号序列组成的此加标签的

D13L-CHO密码子优化的表达盒稳定整合到核DNA通过Urschitz等人,2010和Matasci等人,2011报道的类型的转座子整合技术实现。CHO的转导通过使用从DNA2.0Inc (USA) 购买的piggy Bac载体系统实现。

[0295] D13L蛋白编码序列的构建——D13L蛋白编码序列由Life Technologies的GeneArt通过从痘苗病毒哥本哈根毒株的D13L ORF编码的D13-氨基酸重建DNA序列并为了在CHO细胞中表达进行密码子优化而合成地制造。VACV-COP D13L ORF的蛋白编码序列在序列表中示出。

[0296] D13L细胞转导载体的构建——将来自pLL7的密码子优化的D13-蛋白编码序列(D13LchoTagged) PCR扩增并且通过Bsa I克隆位点亚克隆到从DNA2.0Inc购买的转座子piggyBac载体pJ503-2 (具有CometGFP和Neo+的pHULK piggyBac哺乳动物表达载体) (Cat# pJ503-2) 以产生pLL19。通过In-Fusion克隆(Clontech:无连接酶克隆)在BsaI之间克隆除去了CometGFP编码序列并通过体外同源重组用加标签的D13-蛋白编码序列代替它。现在,D13LchoTagged蛋白编码序列在人延伸因子1 α 启动子(EF1a)的控制下并且将与新霉素抗性基因共表达,一旦二者已经被稳定地整合到转染的细胞的基因组中。通过由piggyBac载体的左转座子边界和右转座子边界结合的DNA序列的转座子整合介导向宿主基因组的稳定整合。

[0297] 使用以下引物对进行D13LchoTagged序列的PCR扩增:

[0298] 正向引物序列:

[0299] Inf-LL19-D13LC-Fw:5' -**AACACGTCTCGGGGgcccgcacc**ATGAACAACACCATCATCAA-3'

[0300] 大写的、粗体和加下划线的文字的序列代表与In-Fusion克隆必需的pJ503-2 (Neo+) 中的comet GFP上游的Bsa I位点同源的序列。小写的红色和加下划线的文字的序列是经修饰的Kozak序列。正常的大写的文字的序列与pLL17中的D13LchoTagged序列的5' 末端同源。

[0301] 反向引物序列

[0302] Inf-LL19-D13LC-Rv:5' -**CAGGAAGACGCTTTT**TCACTTGTCGTCGTCGTCCTTGTAG-3'

[0303] 大写的、粗体和加下划线的文字的序列代表与In-Fusion克隆必需的pJ503-2 (Neo+) 中的comet GFP下游的Bsa I位点同源的序列。正常的大写的文字的序列与pLL17中的D13LchoTagged序列的3' 末端同源。按照制造商的说明,通过InFusion克隆(Clontech)将预期的1719bp的PCR产物克隆到BsaI切割的pJ503-2 (Neo+) 中以产生pLL19。

[0304] 表达D13-蛋白的CHO细胞系的构建:将p-LL19-CHO-CHO细胞接种到6孔板的孔中,这样过夜孵育后它们为约50%融合度。使用来自Qiagen的Effectene转染试剂(Cat# 301425),按照生产商的说明,将1 μ g pLL19转染到1个孔的50%融合的CHO细胞。然后,在生长培养基(RPMI 1640/10%FBS/2mM Glutamax/Pen-Strep)中将转染的细胞孵育过夜。第二天,使用包含1000 μ g/mL遗传霉素的生长培养基更换培养基用于选择转导的细胞。每2至3天更换选择培养基。当转导的细胞生长至超过90至100%融合度时,使用TrypLE Select (Gibco-Invitrogen Corp,Cat#12563-029)将它们回收并接种到T25烧瓶用于进一步细胞扩增。

[0305] 通过蛋白质印迹验证D13表达——兔抗-D13抗血清生产：用连接到代表天然D13-蛋白的C-末端氨基酸的KLH蛋白的16个氨基酸肽注射兔子。此氨基酸序列为：CYDQGVSTIKIMGDNN。进行总计三次注射，间隔1个月，以产生对此氨基酸序列的抗体。通过对以下细胞提取物的蛋白质印迹分析检验来自经注射的兔的血清：143B全细胞提取物、来自表达加flag-标签的-D13L的143B转基因细胞系(p-LL06-143b)的全细胞提取物和经VACV-COP感染的143B细胞提取物。结果清楚地证明，兔抗-D13血清可以清楚地特异性识别D13-蛋白。

[0306] CHO-转导细胞的制备——如普通CHO细胞一样，在T25烧瓶中将D13LchoTagged转导的CHO多克隆扩增细胞系(p-LL19-CHO)培养到100%融合度。用TrypLE Select (Gibco-Invitrogen Corp, Cat#12563-029)收获来自每个烧瓶的细胞、通过低速离心(300g离心5分钟)沉淀、用PBS洗涤并在200uL PBS中重悬。

[0307] 蛋白质印迹分析——向每种细胞悬液添加50uL 5x SDS-PAGE加样缓冲液，并且随后在98℃孵育5分钟。将15uL每种细胞蛋白提取物加载到2个10%SDS-PAGE凝胶、电泳并随后通过电-印迹布点到Hybond ECL硝酸纤维素上。随后在室温下用溶解于包含吐温20的Tris缓冲盐水(TBST)中的5%脱脂奶粉处理电-印迹的膜1小时以阻断膜上所有可用的非特异性抗体结合位点。在用抗体探测前，将膜在TBST中洗涤数次。在4℃，用HRP中缀合抗-DDDDK标签抗体[M2] (Abcam, Cat#ab49763)的1:5000稀释液探测膜1过夜。在4℃，用兔D13-抗血清的1:2000稀释液探测膜2过夜，用TBST洗涤3次，并且随后用二抗，HRP缀合的抗-兔抗体 (Abcam, Cat#ab97069)的1:5000稀释液探测2小时。随后，将2个膜在TBST中洗涤3次，并且用ECL蛋白质印迹检测试剂 (GE Healthcare) 处理并暴露于X-射线底片，如使用手册指导的。

[0308] 实施例10

[0309] 通过D13L ORF缺失使VACV-COP减毒

[0310] 为了使痘苗病毒减毒，将保守的晚期启动子序列与COP-D13L ORF蛋白编码序列的大部分一起被标记用于通过同源重组缺失，并且在它的位置，插入选择/报告盒这样可以在表达D13-蛋白的CHO细胞系中选择通过同源重组的成功的缺失并且其中感染可以用红色荧光显现。选择/报告盒由表达CP77的CHO宿主范围基因和在痘苗病毒启动子的控制下的DsRed-Express2序列组成。

[0311] COP-D13L缺失同源重组载体的构建——为了从VACV-COP通过同源重组缺失D13L ORF，在COP-D13L ORF每一侧侧翼设计两个同源重组臂。但是，由于COP-D13L ORF的启动子序列可能位于COP-D13L ORF的3'端，COP-D13L ORF的3'端的大约200bp是完整的。

[0312] 同源重组臂F1和F2的构建——使用以下示出的引物对从VACV-COP基因组DNA PCR扩增同源重组臂。粗体和加下划线的文字代表用于连接F1和F2臂与选择/报告盒和在来自Clontech的In-Fusion试剂盒中提供的线性化pUCk19的In-Fusion臂。In-Fusion连接将产生被命名为pLL09的环状的质粒。

[0313] 用于从VACV-COP DNA PCR扩增D13L-F1臂的PCR引物对：

[0314] Inf-PCR-D13L-F1-Fw: 5' **CGGTACCCGGGGATC**ACGAAAAATAATAGTAACCA-3'。粗体和加下划线的文字(15bp)代表与Clontech In-Fusion试剂盒中提供的线性化pUC19质粒的左端同源的序列。

[0315] Inf-PCR-D13L-F1-Rv:5' -**AATTTAGTGTGCGCG**TGGAAAAAGCTTACAATA AACTC-3'。粗体和加下划线的文字(15bp)代表与选择/报告盒的5'端同源的序列。预期的PCR产物大小:657bp

[0316] 用于从VACV-COP DNA PCR扩增D13L-F2臂的PCR引物对:

[0317] Inf-PCR-D13L-F2-Fw:5' -**ATATTAAATGCGCG**CAATAATGGAACAAGAACCCT-3'。粗体和加下划线的文字(15bp)代表与选择/报告盒的3'端同源的序列。

[0318] Inf-PCR-D13L-F2-Rv:5' -**CGACTCTAGAGGATC**GCGCTGAGGTCGGCAACTACG-3'。粗体和加下划线的文字(15bp)代表与Clontech In-Fusion试剂盒中提供的线性化pUC19质粒的右端同源的序列。预期的PCR产物大小:621p

[0319] 选择/报告盒的构建——选择/报告表达盒由来自牛痘病毒025L ORF (Brighton Red毒株)的CP77CHO宿主范围基因和DsRed-Express2的红色荧光蛋白编码序列组成。此表达盒由Life Technologies GeneArt合成地制造,这样CP77蛋白编码序列在其天然启动子(CP77ATG起始密码子上游的100bp序列)的控制下并且由痘病毒早期转录终止序列(T₅NT)终止。DsRed蛋白编码序列在痘苗病毒早期/晚期启动子的控制下,并且也由痘病毒早期转录终止序列终止。

[0320] pLL09的装配——按照生产商的说明,通过In-Fusion克隆将D13L-F1和D13L-F2PCR产物与选择/报告盒一起装配到来自ClonTech的InFusion试剂盒提供的pUC19中,以产生pLL09。

[0321] 通过同源重组缺失COP-D13L以及噬斑纯化以产生SCV104——通过缺失蛋白编码序列的大部分和保守的晚期启动子元件“TAAAT”使COP-D13L ORF失活。COP-D13L ORF在痘苗病毒晚期启动子的控制下,其中保守的晚期启动子元件对于启动子活性是重要的(Moss, 2007)——其缺失将导致COP-D13L ORF变得沉默。蛋白编码序列和保守的启动子元件的缺失通过VACV-COP和pLL09之间的同源重组实现,其中同源重组的结果是将CP77/DsRed表达盒插入到缺失的区域。此表达盒的插入使得重组病毒,现在被称为SCV104,能够在CHO细胞中繁殖,但仅在表达功能D13-蛋白的细胞诸如p-LL19-CHO中繁殖。同源重组之后,通过连续几轮的噬斑纯化清除污染性的“遗留(carryover)”亲代病毒(VACV-COP)。通过对插入位点的PCR分析监测污染性的亲代病毒的存在。

[0322] 同源重组——用p-LL19-CHO接种3个包含生长培养基(RPMI-1640/10%FCS/2mM Glutamax/Pen-Strep和1000ug/mL遗传霉素)的T25烧瓶并且在37°C/5%CO₂培养直到100%融合。在感染当天,用VACV-COP以moi 0.01pfu/细胞感染2个烧瓶,其中另一个烧瓶不被感染(未被感染的对照)。在室温下,感染烧瓶1和2 45分钟后,除去病毒接种物,并用PBS洗涤细胞单层2次。洗涤后,向包括烧瓶3的每个烧瓶加入4ml维持培养基(MM:RPMI-1640/2%FCS/2mM Glutamax/Pen-Strep),该烧瓶3也经历了相同的洗涤步骤。

[0323] 使用Effectene转染试剂(Qiagen,Cat No 301425)并且按照生产商的说明进行转染。简要地说,将16uL Enhancer加入到150uL Ec缓冲液的2ug线性化pLL09中,并在充分混合后在室温下放置5分钟。向此加入25uL Effectene转染试剂并在室温下放置10分钟。最后,向混合物加入1ml MM(RPMI-1640/2%FCS/2mM Glutamax/Pen-Strep)并一起缓缓彻底混合。然后,将此转染混合物添加到之前已经用VACV-COP感染的烧瓶1。

[0324] 在37°C/5%CO₂将烧瓶1(同源重组)、烧瓶2(仅感染的对照)和烧瓶3(未感染的对

照) 孵育过夜, 其中第二天用新鲜MM更换每个烧瓶的培养基——每烧瓶5mL并在37℃/5% CO₂继续培养直到总CPE仅可在烧瓶1中见到。在烧瓶2中应该不存在感染的迹象, 因为VACV-COP对CHO细胞是非感染性的并且烧瓶3中单层应该看起来健康。

[0325] 过夜感染后, 在用倒置荧光显微镜检查时, 可以清楚地辨认红色荧光细胞。决定通过将细胞刮入培养基来收获烧瓶1中的细胞, 然后通过低速离心(在室温下500g离心5分钟)沉淀并将细胞沉淀重悬于1mL 10mM Tris-HCl pH8。通过多个冻融循环制备病毒提取物并随后在-80℃保存以备用于噬斑纯化期。病毒构建体被称为SCV104。

[0326] 噬斑纯化方法——基本原理是系列稀释同源重组提取物并使用每种稀释液感染在48孔板中培养的p-LL19-CHO细胞的一行。目的是将病毒稀释至每孔1pfu感染, 并且然后在收获前、在感染后大约30小时后寻找仅包含1个荧光噬斑的孔。

[0327] 将p-LL19-CHO细胞接种到48孔板的每孔并在37℃/5%CO₂、在包含1000ug/mL遗传霉素的生长培养基(RPMI-1640/10%FBS/2mM Glutamax/pen-strep)中培养至100%。

[0328] 对于感染, 将同源重组提取物(SCV104)解冻并短暂地超声处理以打破团块和聚集体。使用以1mL体积的MM(RPMI/2%FBS/Glutamax/PenStrep)进行至10⁻⁵病毒提取物的10-倍系列稀释。对于每种稀释液, 从每孔除去生长培养基并用PBS洗涤一次后, 用500uL稀释的病毒接种48孔板的一行。将该48孔板在室温放置45分钟以使病毒吸附发生。病毒吸附后, 仔细地将病毒接种物从每孔除去, 其中通过由每孔500uL PBS组成的洗涤步骤除去残余的接种物。洗涤后, 向每孔加入500uL MM(RPMI/2%FBS/Glutamax/PenStrep)并且然后在37℃/CO₂孵育直到在荧光显微镜下可以清楚看到感染的荧光红色焦点。

[0329] 对于收获, 仅选择在可能的最低稀释度包含单个荧光焦点的孔。从选择的孔小心地除去培养基并加入100uL 10mM TrisHCl pH8。将板冻融3次并然后回收选择的孔的内容物。对于每个回收的噬斑, 将噬斑纯化方法再重复5次。

[0330] 然后通过100%融合度感染包含p-LL19-CHO细胞的6孔板的1个孔进一步扩增选择的克隆, 通过从该孔除去生长培养基并加入于PBS中稀释到500uL的10uL病毒提取物。在室温下45分钟后, 向孔加入2mL MM并在37℃/5%CO₂进一步孵育3天直到大部分细胞在荧光显微镜下发红色荧光。将感染的孔中的细胞刮到培养基中并然后在500g沉淀5分钟。将沉淀的细胞重悬于500uL 10mM TrisHCl pH8并短暂地超声处理以制造病毒提取物。

[0331] CP77/DsRed表达盒插入到SCV104的D13L ORF的PCR验证——进行PCR分析以确定同源重组和噬斑纯化后D13L ORF是否事实上被CP77/DsRed表达盒取代。设计PCR引物对以结合两个侧翼的同源重组臂的区域的外侧, 这样它们将与“原始”DNA序列而非“引入的”DNA结合。以这样的方式设计引物对意味着, 此引物对可以使用VACV-COP和SCV104作为用于PCR扩增的DNA模板, 并且PCR产物的大小将指示D13L ORF中的表达盒的插入或不具有插入的病毒, 即VACV-COP的检测。此PCR测定不仅指示了插入的存在, 也可以鉴定在多轮噬斑纯化后是否仍然存在残余的污染性亲本病毒(VACV-COP)。不想要的VACV-COP的痕量污染的存在是非常不期望的, 因为此污染可以对SCV104提供反式D13帮助并从而降低SCV104的减毒。

[0332] 使用QIAGEN DNeasy Tissue试剂盒(Cat#69504)的DNA提取——按照制造商的说明, 使用DNeasy Tissue试剂盒从200uL的上述SCV104扩增的病毒提取病毒DNA。简单地说, 这通过将20uL蛋白酶K加入到200uL病毒提取物并充分混合进行。向其加入200uL缓冲液AL, 然后充分混合, 之后在56℃(热阻断)孵育10分钟。孵育后, 加入200uL 100%乙醇并充分混

合,并且然后将总体积加到离心柱。如生产商使用者手册说明的,通过离心使液体通过离心柱,之后用AW1和AW2缓冲液洗涤离心柱。用两次100uL AE缓冲液洗脱与离心柱结合的DNA。将洗脱的DNA合并到单个管并准备用于PCR分析,或在4℃保存直到准备进行PCR分析。

[0333] PCR扩增——使用从SCV104提取的DNA(参见以上4.2.3.1节)作为用于PCR扩增的模板以确定插入是否在D13L ORF内发生。以下描述的PCR引物与D13L ORF外的侧翼的序列结合,这样从SCV104DNA的扩增应该扩增4360bp的PCR产物,而非从亲本病毒VACV-COP扩增的2881bp的PCR产物。PCR分析示出了,SCV104的第6个噬斑纯化克隆仅扩增对应于大于4000bp、小于5000bp的产物,表明SCV104在D13L ORF中包含CP77/DsRed盒。缺乏约3000bp的PCR产物意味着,在SCV104扩增的储备物中不存在可检测的水平亲本病毒污染。

[0334] 引物对细节

[0335] ID_D12L_LL04_Fw:5'-TACAAAATCAAATAATGGTCGAAAC-3'

[0336] ID_A2L_LL04_Rv:

[0337] 5'-TGCCAAGAAAACACTCCTTCTAAGACAAT-3'

[0338] 实施例11

[0339] 通过噬斑感染测定检验SCV104的减毒

[0340] 由于由D13L ORF编码的蛋白对病毒装配重要,人们将期望,SCV104救援病毒的细胞入侵将不能在允许痘苗病毒的正常细胞中产生感染的后代病毒粒并且因此,初始感染不能扩散到相邻的细胞以形成感染的不断扩散的焦点。为了证实情况是否是这样,将本发明描述的D13L缺失的病毒(SCV104)系列稀释至这样的点,此时可以使用低数量的噬斑形成单位来感染培养于6孔板的细胞,这样可以在几天的时间段监测感染病毒的细胞向细胞的扩散。在此研究中,研究3种细胞系:允许痘苗病毒的143B细胞,不允许痘苗病毒但如果痘苗病毒表达CP77蛋白则允许痘苗病毒的CHO细胞和表达D13L蛋白的重组细胞系p-LL19-CHO。已经具有插入到D13L ORF的CP77/DsRed表达盒从而无功能D13-蛋白可以被表达的本发明描述的病毒应该仅在p-LL19-CHO细胞系中形成感染性的感染焦点。143B细胞单层中的感染的焦点的存在表明在病毒群体混合物中存在污染性的痘苗病毒,因为污染物将提供反式D13L帮助。CHO单层中的感染的焦点的存在意味着,CP77/DsRed盒的插入未使D13L ORF失活。

[0341] 细胞设置——将143B、CHO和p-LL19-CHO细胞接种到多个6孔板中并且在生长培养基(RPMI-1640/10%FBS/2mM Glutamax/pen-strep,并且仅对于p-LL19-CHO的1000ug/mL遗传霉素)中在37℃/5%CO₂培养直到细胞单层为100%融合。

[0342] 感染——使用扩增的3rd噬斑纯化的病毒提取物来感染细胞。由于不知道此病毒的滴度,如下将储备病毒10倍梯度稀释至10⁻⁴:首先,在6ml MM培养基中稀释60uL储备病毒(Dil 1;1:100稀释),然后通过向7.2ml MM培养基中加入800uL稀释1来制备进一步的稀释液(稀释2;10⁻³稀释),并且然后重复以制成10⁻⁴稀释液。对于感染,将500uL每种稀释液加至6孔板的每孔——每细胞系使用1个板。在室温下进行45分钟病毒吸附,其中,从每孔除去病毒接种物并且用PBS洗涤每孔后,之后每孔加入2mL MM。将板在37℃/5%CO₂培养并每日在荧光显微镜下观察。10⁻²稀释液产生最佳的用于观察的病毒感染率。

[0343] 显微镜观察——在带有DsRed滤光器(Cat#U-MRFPHQ,Olympus)的荧光显微镜(Olympus IX51)下观察病毒感染。使用cellSens数字成像软件(Olympus)捕获图像。

[0344] 结果——10⁻²稀释液感染产生用于观察初始单细胞感染和感染向相邻细胞的扩散

的最佳结果。对143B细胞的检查示出了在第1天的单细胞感染的存在,其中截止第2天和第3天病毒不能向相邻细胞扩散,表明已经发生病毒进入单个细胞,但截止第2天和第3天没有产生感染性病毒后代。对于CHO单个感染细胞也是这样,即,从初始单个细胞感染没有产生感染性病毒后代。

[0345] 对p-LL19-CHO细胞系的检查示出了截止第1天,初始单细胞感染已经产生了后代感染性病毒,如作为小的感染的焦点观察到的。截止第2天和第3天,这些感染的焦点的大小快速增加,表明病毒扩增随着时间发生。

[0346] 因此,从任何痘苗病毒除去功能痘苗病毒新月支架蛋白编码序列将极大地减毒此病毒。尽管在初始感染时此病毒不能产生感染性后代病毒,其蛋白质,特别是DsRed的表达,在初始的单个感染的细胞中不发生,如从目视红色荧光证明的。独立于痘苗病毒感染,此病毒可以从表达痘苗病毒新月支架蛋白的细胞系被拯救。转基因细胞系对痘苗病毒新月支架蛋白的组成型表达对此细胞系的生长或生理没有不利影响。

[0347] 实施例12

[0348] 表达D13L的C11-LL19-HeLa细胞系的构建

[0349] 为了滴定具有D13L缺失的非-荧光SCV,制造支持溶解噬斑形成、由痘苗病毒允许细胞系组成的表达D13蛋白的救援细胞系。在CHO细胞中,表达CP77的痘苗病毒或CP77的细胞系表达不支持溶解噬斑形成,所述溶解噬斑形成可以用结晶紫染色并且用眼睛计数。为了产生滴定细胞系,转导HeLa细胞以通过表达盒的稳定整合表达D13蛋白,所述表达盒由组成型哺乳动物启动子、包含起始密码子周围的Kozak序列并且以多腺苷酸化信号序列终止的D13-蛋白编码序列组成。然后将此盒克隆到使得能够稳定转移基因到宿主细胞基因组DNA的质粒载体中并且通过抗生素筛选选择成功的整合。然后通过使用抗生素筛选扩增转导的细胞,以选择包含D13表达盒的细胞,并且然后通过用SCV104 (D13L ORF缺失) 在moi 0.001pfu/细胞感染细胞的单层并且选择在3天期间产生最大的噬斑尺寸的细胞系克隆进行噬斑测定。SCV104的D13L ORF已经被2个表达盒代替:1个用于CP77蛋白的表达,并且另一个用于红色荧光蛋白的表达。用SCV104感染表达D13蛋白的HeLa细胞系将产生红色荧光的溶解噬斑。

[0350] pLL19-D13细胞转导载体的构建:pLL19的构建如上所述。在此质粒中,C-末端的加Flag-标签的D13-蛋白编码序列在组成型人延伸因子1 α 启动子的控制下,并且向细胞基因组DNA的稳定基因转移由转座介导。转座还插入新霉素抗生素选择表达盒,这样转导的细胞可以通过向生长培养基中加入G418/遗传霉素被阳性选择。

[0351] 通过用pLL19转导将D13-加Flag标签的表达盒稳定插入到HeLa细胞中:将HeLa细胞接种于T25烧瓶,并且在RPMI-1640/10%FBS/2mM Glutamax/pen-strep生长培养基中培养至大约50%融合度。使用来自Qiagen的Effectene转染试剂(Cat#301425),按照生产商的说明,将1 μ g pLL19转染到T25烧瓶的50%融合的HeLa细胞中。然后,在生长培养基(RPMI1640/10%FBS/2mM Glutamax/Pen-Strep)中将转染的细胞孵育过夜。第二天,使用包含1000 μ g/mL遗传霉素的新鲜生长培养基更换培养基用于选择转导的细胞。当大部分细胞死亡时,通过用TrypLE Select (Gibco-Invitrogen Corp, Cat#12563-029) 将其余细胞从烧瓶表面脱落来回收其余细胞,并且将单细胞分选至包含生长培养基和1000 μ g/mL遗传霉素的96孔板。在37 $^{\circ}$ C/5%CO₂孵育这些板直到在每孔中可以看到细胞的集落。每2至3天更换选

择培养基。回收每个细胞集落并通过在以下培养连续扩增:48孔板至24孔板至6孔板至T75烧瓶,并且以1:5分流比在T75烧瓶中维持直到准备制备冷冻细胞储备。

[0352] 筛选最佳支持SCV104噬斑形成的单克隆细胞系:SCV104是其D13L ORF已经被牛痘病毒025L启动子加编码CP77蛋白的ORF和红色荧光蛋白表达盒(DsRed Express2)取代的痘苗病毒。CP77表达对HeLa细胞中噬斑形成是不重要或不需要的,但此病毒在缺乏D13蛋白的细胞系表达时将不繁殖。但是,在表达D13蛋白的HeLa细胞系中,SCV104应该能够扩增并扩散到其他相邻的细胞,并在这样做时,在细胞单层形成红色荧光的溶解噬斑。在克隆的LL19-HeLa细胞系中的噬斑形成研究中使用此病毒,以选择最佳支持噬斑形成的克隆。

[0353] 细胞系设置:将LL19-HeLa的一些细胞系克隆接种到6孔板的孔中,并且在包含1000ug/mL遗传霉素的生长培养基中、在37°C/5%CO₂培养至100%融合度:

[0354] 病毒感染:通过将SCV104在MM(RPMI-1640/2%FBS/2mM Glutamax/pen-strep)中稀释至10³pfu/ml,以0.001pfu SCV104每细胞感染细胞。将1ml稀释的病毒加入到每孔用于感染。当融合时,6孔板的每孔包含大约1x10⁶细胞,因此1ml 10³pfu/ml产生moi 0.001。0.001的moi将保证从单个感染细胞的噬斑形成。将所有板在室温孵育1小时,这样病毒可以吸附至细胞并且,随后向每孔加入1mL MM,然后将所有板在37°C/5%CO₂孵育,促进病毒同步进入细胞随后病毒扩增,产生随时间的细胞至细胞扩散。每日在荧光显微镜下观察红色荧光噬斑形成。

[0355] 显微镜观察:在3天期间,在带有DsRed滤光器(Cat#U-MRFPHQ,Olympus)的荧光显微镜(Olympus IX51)下观察病毒感染和噬斑形成。使用cellSens数字成像软件(Olympus)捕获图像。

[0356] 结果:感染后第1天,所有感染产生红色荧光的零星的小焦点。截止第3天,所有细胞系产生大小相当的红色荧光溶解噬斑,其中克隆11(C11)产生试验的所有克隆的显著最大的噬斑大小。

[0357] 结论:这些结果证明,C11克隆(C11-LL19-HeLa)表达足够的D13蛋白以支持来自SCV104感染的所有试验的克隆的最大的噬斑形成。此细胞系克隆(C11)用于使用通常用于滴定痘苗病毒的滴定的溶解噬斑计数方法定量D13L缺失的痘苗病毒是良好的。

[0358] 实施例13

[0359] 表达CP77和D13L的p-LL07-LL29-CHO细胞系的构建

[0360] 为了在CHO细胞中救援具有D13L ORF缺失的VACV-COP病毒,将必须构建表达D13和CP77蛋白的CHO细胞系。这通过构建D13哺乳动物表达盒和CP77哺乳动物表达盒进行,所述表达盒由驱动表达的哺乳动物启动子、CHO优选的密码子优化的编码D13或CP77蛋白的DNA序列、随后是多腺苷酸化信号序列组成。然后将这些盒克隆到能够稳定转移基因到宿主细胞基因组DNA的质粒载体中并且通过抗生素筛选选择成功的整合。然后通过使用抗生素筛选扩增转导的细胞,以选择包含D13和CP77表达盒二者的细胞。为了验证CP77的表达,使用表达增强绿色荧光蛋白的痘苗病毒(SCV505)感染该转基因细胞系,其中在4天期间监测荧光噬斑发展。在4天期间内不断扩大的绿色荧光噬斑大小确认了由该细胞系的CP77表达。为了验证D13蛋白的表达,使用表达CP77(SCV104)和DsRed荧光蛋白的D13L缺失的痘苗病毒感染该转基因细胞系,其中在4天期间监测噬斑发展。在4天期间不断扩大的红色荧光噬斑大小确认了由该细胞系的D13蛋白表达。

[0361] pLL07-CP77基因转移质粒的构建

[0362] CP77蛋白编码序列的构建:由GeneArt GmbH (Germany) 通过从由牛痘病毒 Brighton Red毒株UniProtKB/Swiss-Prot:P12932.1的025L ORF编码的CP77的氨基酸序列重建(回译或逆向翻译)DNA序列合成地制造CP77蛋白编码序列,并且为了在中国仓鼠卵巢(CHO)细胞中表达进行密码子优化。

[0363] CP77细胞转导载体的构建:使用PCR引物对将来自pPH51(含有密码子优化的CP77蛋白编码序列的克隆质粒)的密码子优化的CP77蛋白编码序列PCR扩增,其中5'引物被设计为在起始密码子周围加入Kozak序列,并且3'引物被设计为在终止密码子之前加入Flag-标签序列。将扩增的PCR产物通过Bsa I克隆位点亚克隆到从DNA2.0Inc (USA) 购买的转座子piggybac载体pJ507-2 (Hyg+) 中。克隆到BsaI位点中有效地用CP77蛋白编码序列代替cometGFP编码序列以产生pLL07。现在,CP77成为在组成型人延伸因子1 α 启动子(EF1a)的控制下并且与潮霉素抗性基因共表达,一旦二者已经被稳定地整合到转染的细胞的基因组中。通过由piggyBac载体的左转座子边界和右转座子边界结合的DNA序列的转座子整合介导向宿主基因组的稳定整合。

[0364] 用于从pPH51质粒DNA PCR扩增CP77-CHO基因的的PCR引物对

[0365] 正向引物序列:

[0366] AACACGTCTCGGGGGgcccaccATGTTCTGACTACCTGGAAAATGAGGAAGTG (SEQ ID NO:4)

[0367] 反向引物序列:

[0368] CAGGAAGACGCTTTTtcaCTTGTCATCGTCATCCTTGTAATCCTGCTGCTCGAAGATCTTGACT (SEQ ID NO:5)。终止密码子(小写的)之后的Flag标签序列被加下划线。

[0369] pLL29-D13基因转移质粒的构建

[0370] 从pLL19(如上所述的)PCR扩增CHO密码子优化的D13蛋白编码序列以在通过Bsa I克隆位点克隆到从DNA2.0Inc购买的pJ503-02(具有CometGFP和Neo+的pHULK piggyBac哺乳动物表达载体)(Cat#pJ503-2)之前交换C-末端Flag-标签序列为HA-标签序列。

[0371] D13L-HA细胞转导载体的构建:从pLL9PCR扩增不具有C-末端Flag-标签序列的密码子优化的D13L蛋白编码序列,并且通过Bsa I克隆位点亚克隆到从DNA2.0Inc购买的转座子piggyBac载体pJ503-2(具有CometGFP和Neo+的pHULK piggyBac哺乳动物表达载体)(Cat#pJ503-2)以产生pLL29。通过In-Fusion克隆(Clontech:无连接酶克隆)在BsaI位点之间的克隆除去了cometGFP编码序列并通过体外同源重组用新加HA-标签的D13-蛋白编码序列代替它。现在,D13LchoHA-Tagged蛋白编码序列在组成型人延伸因子1 α 启动子(EF1a)的控制下并且将与新霉素抗性基因共表达,一旦二者已经被稳定地整合到转染的细胞的基因组中。通过由piggyBac载体的左转座子边界和右转座子边界结合的DNA序列的转座整合介导向宿主基因组的稳定整合。以下示出了pLL29的质粒图。

[0372] 使用以下引物对进行D13Lcho-HA加标签的序列的PCR扩增:

[0373] 正向引物序列:

[0374] Inf-LL19-D13LC-Fw:5'-AACACGTCTCGGGGGGCCGCCACCATGAACAACACCATCATCAA-3'

[0375] 反向引物序列:

[0376] Inf-LL27-D13L-Rv:5'-CAGGAAGACGCTTTTtaAGCATAATCTGGAACATCATATGGATAGTTGTTATCGCCCATGATCT-3'

[0377] 加下划线的文字中的序列代表HA编码序列。“tta”代表终止密码子。

[0378] 通过用pLL07和pLL29同时转导,CP77-flag加标签的和D13-HA加标签的表达盒在CHO细胞中的稳定双重插入

[0379] 将CHO细胞接种于T25烧瓶,并且在RPMI-1640/10%FBS/2mM Glutamax/pen-strep生长培养基中培养至大约50%融合度。使用来自Qiagen的Effectene转染试剂(Cat# 301425),按照生产商的说明,将1ug pLL07和1ug pLL29转染至T25烧瓶的50%融合的CHO细胞。然后,在生长培养基(RPMI 1640/10%FBS/2mM Glutamax/Pen-Strep)中将转染的细胞孵育过夜。第二天,使用包含500ug/mL遗传霉素和250ug/mL潮霉素B的新鲜生长培养基更换培养基用于选择转导的细胞。每2至3天更换选择培养基,直到大多数细胞死亡并脱离,留下衍生自单个细胞的细胞集落。当转导的细胞生长至超过90至100%融合度时,使用TrypLE Select(Gibco-Invitrogen Corp,Cat#12563-029)将它们回收并接种到T25烧瓶用于进一步细胞扩增。此新的多克隆细胞系被称为p-LL07-LL29-CHO。

[0380] 通过救援痘苗病毒和缺失D13L ORF的痘苗病毒验证D13和CP77表达

[0381] SCV505是表达增强绿色荧光蛋白(EGFP)并且仅在CP77蛋白的存在下在CHO细胞中繁殖的痘苗病毒。此病毒用在p-LL07-LL29-CHO的噬斑传染性研究中以验证此细胞系的CP77表达。SCV104是其D13L ORF已经被牛痘病毒025L启动子加编码CP77蛋白的ORF和红色荧光蛋白表达盒(DsRed Express2)取代的痘苗病毒。此病毒在缺乏D13蛋白的细胞系表达下将不繁殖。此病毒用在p-LL07-LL29-CHO的噬斑传染性研究中以验证此细胞系的D13表达。

[0382] p-LL07-LL29-CHO细胞系噬斑-传染性研究

[0383] 在此研究中使用不同细胞系以帮助使来自SCV505(D13⁺CP77⁻EGFP⁺)和SCV104(D13⁻CP77⁺DsRed⁺)感染的p-LL07-LL29-CHO能够表达D13和CP77。以下细胞系中的预期结果如下:

[0384] Vero:此细胞系通常允许痘苗病毒感染。噬斑形成表明扩增的病毒随着时间的细胞到细胞的扩散,如对SCV505通过绿色荧光蛋白检测的,因为此病毒在此细胞系通常是感染性的。但是,对于SCV104,应见到如通过缺乏红色荧光或仅单细胞红色荧光检测的无噬斑形成,表明由于缺乏细胞系或病毒的D13蛋白表达,不存在扩增的病毒随着时间的细胞到细胞的扩散。

[0385] C11-LL19-HeLa:这是通过pLL19转导表达D13蛋白的允许痘苗病毒的细胞系。如通过绿色荧光蛋白检测的,无论D13的细胞系表达,预期来自SCV505感染的噬斑形成,表明扩增的病毒随着时间的细胞到细胞的扩散。预期来自SCV104感染的噬斑形成,因为由于D13蛋白的细胞系表达,此病毒现在可以在此细胞系中扩增。

[0386] CHO:此细胞系不允许痘苗病毒感染。如通过缺乏红色或绿色荧光或仅单细胞红色或绿色荧光检测的,预期看不到来自SCV505和SCV104感染的噬斑形成。这表明,尽管此病毒可以表达CP77,由于缺乏救援SCV104需要的由细胞系表达的D13蛋白和缺乏救援SCV505需要的由细胞系表达的CP77蛋白,没有扩增的病毒随着时间的细胞到细胞的扩散。

[0387] p-LL07-LL29-CHO:这是表达D13和CP77蛋白二者的CHO细胞系。预期来自SCV505感染的噬斑形成,表明如通过绿色荧光检测的扩增的病毒随着时间的细胞到细胞的扩散,表明CP77蛋白的细胞系表达。预期来自SCV104的噬斑形成,表明如通过红色荧光检测的扩增

的病毒随着时间的细胞到细胞的扩散,表明D13蛋白的细胞系表达。

[0388] 细胞系设置:将Vero、CHO、C11-LL19-HeLa和p-LL07-LL29-CHO细胞系接种在2组的多个6孔板上(一组用于SCV505感染并且另一组用于SCV104感染),并且在相应生长培养基(如下)中在37℃/5%CO₂培养至100%融合度:

[0389] ●Vero、CHO:RPMI-1640/10%FBS/2mM Glutamax/pen-strep

[0390] ●C11-LL19-HeLa:RPMI-1640/10%FBS/2mM Glutamax/pen-strep,加上1000ug/mL遗传霉素

[0391] ●p-LL07-LL29-CHO:RPMI-1640/10%FBS/2mM Glutamax/pen-strep,加上500ug/mL遗传霉素和250ug/mL潮霉素B

[0392] 病毒感染:通过将SCV104和SCV505在MM(RPMI-1640/2%FBS/2mM Glutamax/pen-strep)中稀释至10³pfu/ml以0.001pfu每细胞感染细胞。每板1种病毒:向每孔加入1ml稀释的病毒用于感染。当融合时,6孔板的每个孔包含大约1x10⁶细胞,因此1ml的10³pfu/ml产生moi 0.001。在0.001的moi将保证来自单个感染细胞的噬斑形成。将所有板在室温孵育1小时,这样病毒可以吸附至细胞并且,随后向每孔加入1mL MM,然后将所有板在37℃/5%CO₂孵育,促进病毒同步进入细胞随后病毒扩增,产生随时间的细胞至细胞扩散。每天在荧光显微镜下观察荧光噬斑形成。

[0393] 显微镜观察:在4天的时间段,在荧光显微镜(Olympus IX51)下用DsRed滤光器(Cat#U-MRFPHQ,Olympus)观察SCV104病毒并且用GFP滤光器(Cat#U-MGFPHQ,Olympus)观察SCV505的病毒感染和噬斑形成。使用cellSens数字成像软件(Olympus)捕获图像。

[0394] 结果:

[0395] 在moi 0.001用SCV104和SCV505感染CHO细胞:在感染后第1天仅能看到零星的单细胞荧光,表明两种病毒已经进入细胞,但还没有扩增以将感染扩散至相邻细胞。但是,这从感染后第2天到第4天保持相同的进展,即,只看到单个细胞感染并且随着时间病毒没有向相邻细胞扩散。

[0396] 在moi 0.001用SCV104和SCV505感染Vero细胞:用SCV505的感染如预期地,在感染后第1天形成微小噬斑,截止第4天,所有微小噬斑的大小增加至这样的点,其中所有噬斑合并成细胞单层的一个融合感染。这表明,病毒从第1天开始扩增至截止感染后第4天,将感染扩散至完全。但是,这与SCV104感染完全不同。在感染后第1天仅能看到零星的单细胞荧光,这在之后3天保持相同。这表明,由于病毒缺乏D13L ORF并且不能在病毒基因组复制后开始病毒装配以及此细胞系不表达D13蛋白来帮助救援病毒扩增,SCV104在此细胞系中不能扩增和繁殖。

[0397] 用SCV104和SCV505在0.001感染C11-LL19-HeLa细胞(表达D13蛋白的HeLa细胞系):用SCV505的感染如预期地,在感染后第1天形成小噬斑,截止第4天,小噬斑的大小增加至这样的点,其中所有噬斑合并成细胞单层的一个融合感染。这表明,病毒从第1天开始扩增至截止感染后第4天,将感染扩散至完全。用SCV104的感染也产生相同结果,证明由此细胞系产生的D13蛋白补充SCV104中D13L ORF的缺乏和由此细胞系产生的量足以支持与具有完整的D13L ORF的SCV505相当的病毒扩增和感染的扩散。

[0398] 用SCV104和SCV505在moi 0.001感染p-LL07-LL29-CHO细胞(表达D13和CP77蛋白的CHO细胞系):对于SCV104和SCV505二者,在感染后1天观察到感染的微小焦点。在接下来

的3天,这些感染的焦点扩展成越来越大的噬斑,截止感染后第4天,该噬斑最终合并成融合感染的感染。这证明了CP77被表达,因为它支持SCV505的扩增和繁殖,以及D13蛋白也被表达,因为它支持SCV104的扩增和繁殖。

[0399] 结论:这些结果证明了,表达CP77和D13蛋白的CHO细胞系可被用于D13L缺失的痘苗病毒的生产和制造的细胞基底。在正常允许细胞的感染期间,D13L缺失的病毒将不能引发病毒装配,并且从而不完成其感染性生命周期,因此在安全性方面,使其成为用于人和动物疫苗接种的优秀病毒疫苗递送载体。但是,此高度减毒的痘苗病毒载体可以在生物技术友好的CHO细胞系中制造,如果此细胞系表达CHO宿主范围蛋白(CP77)和缺少的装配蛋白(D13-蛋白)。

[0400] 实施例14

[0401] 通过蛋白质印迹法对p-LL07-LL29-CHO细胞系的D13和CP77蛋白表达分析

[0402] 背景信息

[0403] 由p-LL07-LL29-CHO表达的D13和CP77蛋白被加标签并且可以使用特异性识别D13和CP77蛋白的C-末端的氨基酸标签序列的抗体检测。在缺乏特异性识别D13和CP77蛋白的抗体时,可以使用这些抗-标签抗体进行蛋白质印迹法来证实p-LL07-LL29-CHO细胞系中D13和CP77蛋白的表达。D13蛋白的C末端包含“YPYDVPDYA”的HA-标签氨基酸序列,而CP77蛋白的C末端包含“DYKDDDDK”的flag-标签氨基酸序列。

[0404] 蛋白质印迹分析方法学

[0405] 将以下细胞系接种到T75烧瓶并在包含适当的选择抗生素的生长培养基(RPMI 1640/10%FBS/2mM Glutamax/Pen-Strep)中培养至100%融合度:CHO无选择抗生素、p-LL07-LL29-CHO以500ug/mL遗传霉素和250ug/mL潮霉素B、p-LL07-CHO以250ug/mL潮霉素B、以及C11-LL19-HeLa以1000ug/mL遗传霉素。用TrypLE Select在37℃处理10min,将融合的细胞单层脱离并消化成单细胞悬液。来自每个烧瓶的细胞被回收并通过在300g离心5分钟沉淀并且用PBS洗涤2次。最后一次洗涤后,将细胞沉淀重悬于500uL PBS。向蛋白质提取物加入4^x SDS-PAGE加样缓冲液以得到1^x的终浓度并且在98℃加热2-3分钟。在200V,将变性的蛋白质样品电泳通过Biorad **Mini-PROTEAN®** TGX Stain-Free™ Precast Gel (梯度凝胶)持续30至45分钟。电泳后,在100V,将分散的蛋白质通过电印迹转移到硝酸纤维素膜持续1小时。

[0406] 为了检测蛋白质,在室温将硝酸纤维素膜孵育在于PBS的5%脱脂奶粉中持续1小时,以阻断非特异性抗体结合位点。为了检测加HA标签的蛋白质,在4℃,将膜在抗-HA HRP缀合的(Abcam Cat#AB1265)在阻断缓冲液中的1:1000稀释液中孵育过夜。为了检测加flag-标签的蛋白质,在4℃,将另一个电印迹的硝酸纤维素膜在抗-flag HRP缀合的(Abcam Cat#AB49763)在阻断缓冲液中的1:1000稀释液中孵育过夜。然后,在PBS中洗涤膜3次,每次洗涤5分钟,并且然后在ECL底物(Thermo Scientific Pierce ECL蛋白质印迹底物,Cat No 32106)中孵育数分钟,之后使用BioRad XRS凝胶文件系统采集图像。

[0407] 结果

[0408] 通过HA-标签检测检测D13蛋白:抗-HA-标签抗体能够检测来自从p-LL07-LL29-CHO制备的总细胞蛋白质提取物的具有预期大小的蛋白质,但未能检测来自从CHO细胞制备的总细胞蛋白质提取物的蛋白质。这清楚地表明,p-LL07-LL29-CHO表达D13蛋白。

[0409] 通过Flag-标签检测CP77蛋白:抗-flag-标签抗体能够检测来自从p-LL07-LL29-CHO和p-LL07-CHO (仅表达CP77的CHO细胞系) 制备的总细胞蛋白质提取物的具有预期大小的蛋白质,但未能检测来自从CHO细胞制备的总细胞蛋白质提取物的蛋白质。这清楚地表明,p-LL07-LL29-CHO和p-LL07-CHO表达CP77蛋白。

[0410] 通过Flag-标签检测检测D13蛋白:C11-LL19-HeLa表达具有C-末端flag-标签的D13-蛋白,并且在由此细胞系制成的全细胞蛋白质提取物的蛋白质印迹时,抗-flag-标签抗体能够检测具有预期大小的蛋白质。这清楚地表明,C11-LL19-HeLa表达D13蛋白。

[0411] 对于本领域技术人员来说,很多修改将是明显的而不偏离本发明的范围。

[0412] 表A

[0413] 序列标识符的总结

[0414]

序列 ID NO:	描述
1	被密码子优化以在哺乳动物细胞(CHO)中表达的 CP77 的核苷酸序列
2	由 SEQ ID NO: 1 编码的 CP77 的氨基酸序列
3	pLL07 的核苷酸序列
4	来自 pPH51 质粒的 CP77-CHO 基因的正向引物
5	来自 pPH51 质粒的 CP77-CHO 基因的反向引物
6	痘苗病毒哥本哈根 K1L 基因(从 Genbank M35027.1 提取)的核苷酸序列
7	由 SEQ ID NO: 6 编码的痘苗病毒哥本哈根 K1L (从 Genbank M35027.1 提取)的氨基酸序列
8	VACV-COP-D13L-ORF
9	为了在仓鼠细胞表达进行密码子优化的具有 C-末端 flag-标签(DYKDDDK)的编码 SEQ ID NO:1 的核苷酸序列
10	Inf-LL19-D13LC 的正向引物
11	Inf-LL19-D13LC 的反向引物
12	Flag-标签蛋白
13	D13L 蛋白肽
14	Inf-PCR-D13L-F1 的正向引物
15	Inf-PCR-D13L-F1 的反向引物
16	Inf-PCR-D13L-F2 的正向引物
17	Inf-PCR-D13L-F2 的反向引物
18	CP77+DsRed 表达盒 2903 bp
19	CP77+DsRed 表达盒
20	ID_D12L_LL04 的正向引物
21	ID_A2L_LL04 的反向引物
22	具有 CP77/DsRed 选择的用于 VACV-COP D13L 缺失的 pLL09 同源重组载体
23	用于稳定整合到细胞核基因组 DNA 的 pLL19 加标签的-D13L-CHO 密码子优化的转座子介导的转导载体

[0415] 表1

[0416] 滴定结果

[0417]		24h	48h	72h
	CHO	53pfu/mL	0	0
	Vero	8.5x 10 ² pfu/mL	14.25x 10 ⁶ pfu/mL	27.7x 10 ⁶ pfu/mL

[0418] 病毒产量(输出)

[0419] 用于感染的病毒的量(输入)为4x10⁴pfu/mL。为了比较,以10⁴值表示产量。值代表每时间点的平均产量,即,来自3mL收获(6mL除以2孔)的产量。

[0420]		24h	48h	72h
	CHO	0.0159x10 ⁴ pfu	0	0
	Vero	0.255x 10 ⁴ pfu	4275x 10 ⁴ pfu	8310x 10 ⁴ pfu

[0421] 生产率(输出/输入比)

[0422] 下表示出了在每个收获时间点来自每个细胞系的产生的病毒高于输入水平的产量,即输出/输入比。

[0423]		24h	48h	72h
	CHO	0.004	0	0
	Vero	0.06	1069	2078

[0424] 表2

[0425] 以pfu/mL计的滴度结果

	病毒/细胞系	稀释度	计数	滴度
[0426]	VACV-COP/CHO	10 ⁻¹	0	0pfu/mL
			0	
			0	
			0	
		平均值	0	
	VACV-COP/Vero	10 ⁻⁷	2	2.8x10 ⁷ pfu/mL +/- 34%
			2	
			4	
			3	
		平均值	7.5	

	病毒/细胞系	稀释度	计数	滴度
[0427]	VACV-PH22/Vero	10 ⁻⁶	8	0.75x10 ⁷ pfu/MI +/- 43%
			5	
			12	
			5	
		平均值	7.5	
	VACV-PH22/CHO	10 ⁻⁷	6	7.3x10 ⁷ pfu/mL +/- 20%
			9	
			8	
			6	
		平均值	7.3	

[0428] 来自感染/孔的平均病毒输出(产量)

[0429] 每烧瓶的病毒提取物体积为1mL。用于滴定的铺板体积为1mL。因此,病毒产量等于

以pfu/ml计的滴度乘以1mL (铺板体积)。

[0430]	病毒	细胞系	产量/烧瓶
	VACV-COP	CHO	0pfu
		Vero	2.8×10^7 pfu
	VACV-PH22	Vero	0.75×10^7 pfu
		CHO	7.3×10^7 pfu

[0431] 产量/pfu接种物, 即从1pfu接种物产生的总pfu

[0432] 接种量/烧瓶: 1×10^5 pfu

	病毒	细胞系	产量/pfu 接种物
[0433]	VACV-COP	CHO	0pfu/输入 pfu
		Vero	280 pfu/输入 pfu (对于用于接种的每 pfu, 产生 280pfu)
	VACV-PH22	Vero	75 pfu/输入 pfu (对于用于接种的每 pfu, 产生 75pfu)
		CHO	730 pfu/输入 pfu (对于用于接种的每 pfu, 产生 730pfu)

[0434] 表3

[0435] 这是各1mL病毒提取物的滴度。

[0436]

细胞基底	指示细胞系					
	143B			Vero		
	滴度 pfu/mL	SE	%SE	滴度 pfu/mL	SE	%SE
CHO	5.38E+03	5.74E+02	10.70%	7.25E+02	2.17E+02	30.00%
p-LL07-CHO	1.88E+08	2.45E+07	13.10%	3.95E+07	1.13E+07	28.50%
143B	2.73E+08	3.95E+07	14.50%	3.73E+07	5.45E+06	14.60%

[0437] 表4

[0438] 表4提供了各1mL病毒提取物中病毒的总量。

[0439]	细胞基底	指示细胞系			
		143B		Vero	
		产量 pfu	SE	产量 pfu	SE
	CHO	5.38E+03	5.74E+02	7.25E+02	2.17E+02
	p-LL07-CHO	1.88E+08	2.45E+07	3.95E+07	1.13E+07
	143B	2.73E+08	3.95E+07	3.73E+07	5.45E+06

[0440] 表B

[0441]	毒株	ORF	基因组	蛋白质
	哥本哈根	COP-D13L	M35027	AAA48114
	Lister克隆107	List-114	DQ121394	ABD52596
	LC16m0	m0-149L	AY678277	AAW23819
	LC16m8	M8-149L	AY678275	AAW23537
	WR	WR-118	NC_006998	YP_233000
	Dryvax-3737	VACV-114	DQ377945	ABD57648

Acambis-2000	VACAC2_129	AY313847	AAR17961
Acambis克隆3	VACCL3-129	AY313848	AAQ93215
CVA	CVA-124	AM501482	CAM58288
Tiantan克隆10	TT10-148	JX489137	AGJ91839
牛痘病毒GRI-90毒株	CPXV-GRI-E13L	X94355	CAD90667
牛痘病毒Brighton Red	CPXV-131	NC_003663	NP_619914

[0442] 参考文献

[0443] Altschul et al. (1997) Nucleic Acids Research 25:3389-3402

[0444] Ausubel et al. (1999) Current Protocols in Molecular Biology (Supplement47), John Wiley&Sons, New York

[0445] Boshart et al. (1985) Cell 41:521

[0446] Brooks et al. (1995) J. Virol. 69 (12) :7688-7698

[0447] Dijkema et al. (1985) EMBO J. 4:761

[0448] Drillien R, et al. (1978) J Virol. 28 (3) :843-50

[0449] Gorman et al., (1982) Proc. Natl. Acad. Sci. USA 79:6777

[0450] Ham RG. (1965) Proc. Natl. Acad. Sci. USA 53:288-293

[0451] Hsiao JC, et al. (2006) J. Virol. 80 (15) :7714-28

[0452] Kibler et al. (2011) PLOS ONE 6 (11)

[0453] Meisinger-Henschel et al. (2007) J. Gen. Virol. 88 (12) :3249-3259

[0454] Murphy et al. (1995) Virus Taxonomy Springer Verlag:79-87

[0455] Puck TT, et al. (1958) J. Exp. Med. 108:945-956

[0456] Sambrook et al. (1989) Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Press, Plainville, N.Y.

[0457] Shisler et al. (2004) J. Virol. 78 (7) :3553-3560

[0458] Spehner D, et al. (1988) J Virol., 62 (4) :1297-1304

[0459] Werden SJ, et al. (2008) Chapter 3: Poxvirus Host Range Genes. In: Advances in Virus Research, 71

序列表

<110> 赛门蒂斯有限公司
 <120> 病毒载体制造
 <130> 35226179/WJP
 <160> 11
 <170> PatentIn version 3.5

<210> 1
 <211> 2031
 <212> DNA
 <213> 人工
 <220>
 <223> 为在哺乳动物细胞(CHO)中表达而密码子优化的 CP77 的核苷酸序列

<400> 1
 atgttcgact acctggaaaa tgaggaagtg gccctggacg agctgaagca gatgctgcgg 60
 gaccgggacc ccaacgacac ccggaaccag ttcaagaaca acgccctgca cgcctacctg 120
 tttaacgagc actgcaacaa cgtggaagtg gtcaagctgc tgcctggactc cggcaccaac 180
 cccctgcaca agaactggcg gcagctgacc cccctggcg agtacacca ctcccggcac 240
 ggcaaagtga acaaggatat cgccatggtg ctgctggaag ctaccggcta ctccaacatc 300
 aacgacttca acatcttcac ctacatgaag tccaagaacg tggacatcga cctgatcaag 360
 gtgctggttg aacacggctt cgacttctcc gtgaagtgcg agaagcaca ctccgtgatc 420
 gagaactacg tgatgaccga cgaccccgct cctgagatca tcgacctgtt catcgagaac 480
 ggctgctccg tgatctacga ggacgaggac gacgagtacg gctacgccta cgaggaatac 540
 cactcccaga acgacgacta ccagccccgg aactgcggca ccgtgctgca cctgtacatc 600
 atctcccacc tgtactccga gagcgactcc cggtcttgcg tgaaccccga ggtggtcaag 660
 [0001] tgcctgatea accacggcat caaccctcc agcatcgata agaactactg caccgcctg 720
 cagtactaca tcaagtcctc ccacatcgac atcgatatcg tgaagctgct gatgaagggc 780
 atcgacaaca ccgcctacag ctacatcgac gacctgacct gctgcaccgg gggcatcatg 840
 gccgactacc tgaacagcga ctaccggtac aacaaggacg tggacctgga tctggtcaaa 900
 ctgttcctgg aaaacggcaa gcctcacggc atcatgtgct ccatcgctgc cctgtggcgg 960
 aacgacaaag agacaatctc cctgatcctg aaaacctga actccgacgt gctgcagcat 1020
 atctgatcga agtacatcac ctctctccgat atcgacatct ctctggtcga gtacatgctg 1080
 gaatacggcg ctgtggtcaa caaaggagcc atccacggct acttcaagaa catcaacatc 1140
 gactcctaca ccatgaagta cctgctgaaa aaagagggcg gcgacgccgt caaccacctg 1200
 gacgacggcg agatccccat cggccacctg tgcaagagca actacggcgg gtacaacttt 1260
 tacaccgaca cctaccggca gggtctccgg gacatgtcct acgcctgccc catcctgtcc 1320
 accatcaaca tctgcctgcc ctacctgaag gacatcaata tgatcgataa gcggggcgag 1380
 aactgctgc acaaggccgt gcgctacaac aagcagtccc tgggtgtccct gctgctggaa 1440
 agcggctccg acgtgaacat ccggtccaac aacggctaca cctgtatcgc tatcgccatc 1500
 aacgagtcce ggaacatcga gctgctgaat atgctgctgt gccacaagcc taccctggac 1560
 tgcgtgatcg actccctgcg cgagatcagc aacatcgtgg acaacgccta cgccatcaag 1620
 cagtgcata gatacgccat gattatcgac gactgcatct cctccaagat ccccgagtcc 1680
 atctccaagc actacaacga ctacattgac atctgcaacc aggaactgaa cgagatgaag 1740
 aaaatcatcg tgggcggcaa caccatgttc agcctgatct tcaccgatca cggcgccaag 1800
 atcatccacc gctacgcaa caaccccgag ctgcgggcct actacgagtc caagcagaac 1860
 aagatctacg tcgaggtcta cgacatcatc agcaacgcca tcgtgaagca caacaagatt 1920
 cacaagaaca tcgagtcctg ggacgacaac acctacatct ccaacctgcc ctataccatc 1980
 aagtacaaga tcttcgagca gcaggattac aaggatgacg atgacaagtg a 2031

<210> 2
 <211> 676

<212> PRT
 <213> 人工
 <220>
 <223> 由 SEQ ID NO: 1 编码的 CP77 的氨基酸序列
 <400> 2
 Met Phe Asp Tyr Leu Glu Asn Glu Glu Val Ala Leu Asp Glu Leu Lys
 1 5 10 15
 Gln Met Leu Arg Asp Arg Asp Pro Asn Asp Thr Arg Asn Gln Phe Lys
 20 25 30
 Asn Asn Ala Leu His Ala Tyr Leu Phe Asn Glu His Cys Asn Asn Val
 35 40 45
 Glu Val Val Lys Leu Leu Leu Asp Ser Gly Thr Asn Pro Leu His Lys
 50 55 60
 Asn Trp Arg Gln Leu Thr Pro Leu Gly Glu Tyr Thr Asn Ser Arg His
 65 70 75 80
 Gly Lys Val Asn Lys Asp Ile Ala Met Val Leu Leu Glu Ala Thr Gly
 85 90 95
 Tyr Ser Asn Ile Asn Asp Phe Asn Ile Phe Thr Tyr Met Lys Ser Lys
 100 105 110
 Asn Val Asp Ile Asp Leu Ile Lys Val Leu Val Glu His Gly Phe Asp
 115 120 125
 Phe Ser Val Lys Cys Glu Lys His His Ser Val Ile Glu Asn Tyr Val
 130 135 140
 Met Thr Asp Asp Pro Val Pro Glu Ile Ile Asp Leu Phe Ile Glu Asn
 145 150 155 160
 [0002] Gly Cys Ser Val Ile Tyr Glu Asp Glu Asp Asp Glu Tyr Gly Tyr Ala
 165 170 175
 Tyr Glu Glu Tyr His Ser Gln Asn Asp Asp Tyr Gln Pro Arg Asn Cys
 180 185 190
 Gly Thr Val Leu His Leu Tyr Ile Ile Ser His Leu Tyr Ser Glu Ser
 195 200 205
 Asp Ser Arg Ser Cys Val Asn Pro Glu Val Val Lys Cys Leu Ile Asn
 210 215 220
 His Gly Ile Asn Pro Ser Ser Ile Asp Lys Asn Tyr Cys Thr Ala Leu
 225 230 235 240
 Gln Tyr Tyr Ile Lys Ser Ser His Ile Asp Ile Asp Ile Val Lys Leu
 245 250 255
 Leu Met Lys Gly Ile Asp Asn Thr Ala Tyr Ser Tyr Ile Asp Asp Leu
 260 265 270
 Thr Cys Cys Thr Arg Gly Ile Met Ala Asp Tyr Leu Asn Ser Asp Tyr
 275 280 285
 Arg Tyr Asn Lys Asp Val Asp Leu Asp Leu Val Lys Leu Phe Leu Glu
 290 295 300
 Asn Gly Lys Pro His Gly Ile Met Cys Ser Ile Val Pro Leu Trp Arg
 305 310 315 320
 Asn Asp Lys Glu Thr Ile Ser Leu Ile Leu Lys Thr Met Asn Ser Asp
 325 330 335
 Val Leu Gln His Ile Leu Ile Glu Tyr Ile Thr Phe Ser Asp Ile Asp
 340 345 350
 Ile Ser Leu Val Glu Tyr Met Leu Glu Tyr Gly Ala Val Val Asn Lys
 355 360 365

Glu Ala Ile His Gly Tyr Phe Lys Asn Ile Asn Ile Asp Ser Tyr Thr
 370 375 380
 Met Lys Tyr Leu Leu Lys Lys Glu Gly Gly Asp Ala Val Asn His Leu
 385 390 395 400
 Asp Asp Gly Glu Ile Pro Ile Gly His Leu Cys Lys Ser Asn Tyr Gly
 405 410 415
 Arg Tyr Asn Phe Tyr Thr Asp Thr Tyr Arg Gln Gly Phe Arg Asp Met
 420 425 430
 Ser Tyr Ala Cys Pro Ile Leu Ser Thr Ile Asn Ile Cys Leu Pro Tyr
 435 440 445
 Leu Lys Asp Ile Asn Met Ile Asp Lys Arg Gly Glu Thr Leu Leu His
 450 455 460
 Lys Ala Val Arg Tyr Asn Lys Gln Ser Leu Val Ser Leu Leu Leu Glu
 465 470 475 480
 Ser Gly Ser Asp Val Asn Ile Arg Ser Asn Asn Gly Tyr Thr Cys Ile
 485 490 495
 Ala Ile Ala Ile Asn Glu Ser Arg Asn Ile Glu Leu Leu Asn Met Leu
 500 505 510
 Leu Cys His Lys Pro Thr Leu Asp Cys Val Ile Asp Ser Leu Arg Glu
 515 520 525
 Ile Ser Asn Ile Val Asp Asn Ala Tyr Ala Ile Lys Gln Cys Ile Arg
 530 535 540
 Tyr Ala Met Ile Ile Asp Asp Cys Ile Ser Ser Lys Ile Pro Glu Ser
 545 550 555 560
 Ile Ser Lys His Tyr Asn Asp Tyr Ile Asp Ile Cys Asn Gln Glu Leu
 565 570 575
 Asn Glu Met Lys Lys Ile Ile Val Gly Gly Asn Thr Met Phe Ser Leu
 580 585 590
 Ile Phe Thr Asp His Gly Ala Lys Ile Ile His Arg Tyr Ala Asn Asn
 595 600 605
 Pro Glu Leu Arg Ala Tyr Tyr Glu Ser Lys Gln Asn Lys Ile Tyr Val
 610 615 620
 Glu Val Tyr Asp Ile Ile Ser Asn Ala Ile Val Lys His Asn Lys Ile
 625 630 635 640
 His Lys Asn Ile Glu Ser Val Asp Asp Asn Thr Tyr Ile Ser Asn Leu
 645 650 655
 Pro Tyr Thr Ile Lys Tyr Lys Ile Phe Glu Gln Gln Asp Tyr Lys Asp
 660 665 670
 Asp Asp Asp Lys
 675

[0003]

<210> 3

<211> 11742

<212> DNA

<213> 人工

<220>

<223> pLL07 的核苷酸序列

<400> 3

tcagaattgg ttaattggtt gtaacactga cccctatttg tttatttttc taaatacatt	60
caaatatgta tccgctcatg agacaataac cctgataaat gcttcaataa tattgaaaaa	120
ggaagaatat gagccatatt caacgggaaa cgctcaggcc gcgattaaat tccaacatgg	180

	atgctgattt	atatgggtat	aaatgggctc	gcgataatgt	cgggcaatca	ggtgcgacaa	240
	tctatcgctt	gtatgggaag	cccgatgcgc	cagagttgtt	tctgaaacat	ggcaaaggta	300
	gcgttgccaa	tgatgttaca	gatgagatgg	tcagactaaa	ctggctgacg	gaatttatgc	360
	cacttccgac	catcaagcat	tttatccgta	ctcctgatga	tgcattggtta	ctcaccactg	420
	cgatcccccg	aaaaacagcg	ttccaggat	tagaagaata	tcctgattca	ggtgaaaata	480
	ttgttgatgc	gctggcagtg	ttcctgcgcc	ggttgcactc	gattcctgtt	tgtaatgttc	540
	cttttaacag	cgatcgcgta	tttcgcctcg	ctcaggcgca	atcacgaatg	aataacgggt	600
	tggttgatgc	gagtgatatt	gatgacgagc	gtaatggctg	gcctgttgaa	caagtctgga	660
	aagaaatgca	taaacttttg	ccattctcac	cggattcagt	cgctactcat	ggtgatttct	720
	cacttgataa	ccttattttt	gacgagggga	aattaatagg	ttgtattgat	gttggacgag	780
	tcggaatcgc	agaccgatac	caggatcttg	ccatcctatg	gaactgcctc	ggtgagtttt	840
	ctccttcatt	acagaaacgg	ctttttcaaa	aatatgggat	tgataatcct	gatatgaata	900
	aattgcagtt	tcatttgatg	ctcgatgagt	ttttctaact	catgaccaa	atcccttaac	960
	gtgagttacg	cgcgcgtcgt	tccactgagc	gtcagacccc	gtagaaaaga	tcaaaggatc	1020
	ttcttgagat	cctttttttc	tgcgcgtaat	ctgctgcttg	caaacaaaaa	aaccaccgct	1080
	accagcggtg	gtttgtttgc	cggatcaaga	gtaccaact	ctttttccga	aggttaactgg	1140
	cttcagcaga	gcgcagatac	caaatactgt	tcttctagt	tagccgtagt	tagcccacca	1200
	cttcaagaac	tctgtagcac	cgcctacata	cctcgtctg	ctaactcctg	taccagtggc	1260
	tgctgccagt	ggcgataagt	cgtgtcttac	cgggttggac	tcaagacgat	agttaccgga	1320
	taaggcgcag	cggtcgggct	gaacgggggg	ttcgtgcaca	cagcccagct	tggagcgaac	1380
	gacctacacc	gaactgagat	acctacagcg	tgagctatga	gaaagcgcca	cgttccccga	1440
	agggagaaaag	gcggacaggt	atccggtaag	cggcagggtc	ggaacaggag	agcgcacgag	1500
	ggagcttcca	gggggaaacg	cctggtatct	ttatagtcct	gtcgggtttc	gccacctctg	1560
	acttgagcgt	cgatttttgt	gatgctcgtc	aggggggagg	agcctatgga	aaaacgccag	1620
	caacgcggcc	tttttacggt	tcctggcctt	ttgctggcct	tttctcaca	tgttctttcc	1680
[0004]	tgcttatcc	cctgattctg	tggataaccg	tattaccgcc	tttgagttag	ctgataccgc	1740
	tctcaatatt	ggccattagc	catattatc	attggttata	tagcataaat	caatattggc	1800
	tattggccat	tgcatacgtt	gtatctatat	cataatatgt	acatttatat	tggctcatgt	1860
	ccaatatgac	cgccatgttg	gcattgatga	ttgactagtt	attaatagta	atcaattacg	1920
	gggtcatttag	ttcatagccc	atatatggag	ttccgcgtta	cataacttac	ggtaaatggc	1980
	ccgcctggct	gaccgcccac	cgacccccgc	ccattgacgt	caataatgac	gtatgttccc	2040
	atagtaacgc	caatagggac	tttccattga	cgtcaatggg	tggagtattt	acggtaaac	2100
	gcccacttgg	cagtacatca	agtgtatcat	atgccaaagc	cgccccctat	tgacgtcaat	2160
	gacggtaaat	ggcccgcctg	gcattatgcc	cagtacatga	ccttacggga	cttctctact	2220
	tggcagtaca	tctacgtatt	agtcacgcgt	attaccatgc	tgatgcggtt	ttggcagtac	2280
	accaatgggc	gtggatagcg	gtttgactca	cggggatttc	caagtctcca	ccccattgac	2340
	gtcaatggga	gtttgttttg	gcacaaaaat	caacgggact	ttccaaaatg	tcgtaataac	2400
	cccgcgccgt	tgacgcaaat	gggcggtagg	cgtgtacggt	gggaggtcta	tataagcaga	2460
	gctcgttttag	tgaaccgtca	gatcactaga	agctttattg	cggtagttta	tcacagttaa	2520
	attgctaacg	cagtcaggcc	aacatgggct	ctagcctgga	cgacgagcac	atcctgagcg	2580
	ccctgctgca	gagcgacgac	gaactgggtg	gcgaggacag	cgacagcgag	gtcagcgacc	2640
	acgtgtccga	ggacgacgtg	cagtcgcaca	ccgaggaagc	cttcacgcac	gaggtgcacg	2700
	aagtgcagcc	taccagcagc	ggctccgaga	tcctggacga	gcagaacgtg	atcgagcagc	2760
	ctggcagctc	cctggccagc	aacagaatcc	tgacctgcc	ccagagaacc	atcagaggca	2820
	agaacaagca	ctgctgggtc	acctccaaga	gcaccaggcg	gagcagagtg	tccgccctga	2880
	acatcgtgcg	gagccagagg	ggccccacca	gaatgtgcag	aaacatctac	gaccccctgc	2940
	tgtgcttcaa	gctgttcttc	accgacgaga	tcacagcga	gatcgtgaag	tggaccaacg	3000
	ccgagatcag	cctgaagagg	cgggagagca	tgaccagcgc	caccttcaga	gacaccaacg	3060
	aggacgagat	ctacgccttc	ttcggcatcc	tgggtgatgac	cgccgtgaga	aaggacaacc	3120
	acatgagcac	cgacgacctg	ttcgacagat	ccctgagcat	ggtgtacgtg	tccgtgatga	3180
	gcagagacag	attcgacttc	ctgatcagat	gcctgagaat	ggacgacaag	agcatcagac	3240

	ccacctgctg	ggagaacgac	gtgttcaccc	ccgtgcggaa	gatctgggac	ctgttcaccc	3300
	accagtgcac	ccagaactac	acccctggcg	cccacctgac	catcgatgag	cagctgctgg	3360
	gcttcagagg	cagatgcccc	ttcagagtgt	acatccccaa	caagcccagc	aagtacggca	3420
	tcaagatcct	gatgatgtgc	gacagcggca	ccaagtacat	gatcaacggc	atgccctacc	3480
	tgggcagagg	cacccagaca	aacggcgtgc	ccctgggcga	gtactacgtg	aaagaactga	3540
	gcaagcctgt	gcatggcagc	tgcaggaaca	tcacatgcga	caactggttc	accagcatcc	3600
	ccctggccaa	gaacctcctg	caggaaccct	acaagctgac	catcgtgggc	accgtgcgga	3660
	gcaacaagcg	ggagatccca	gaggtgctga	agaacagcag	atccagacct	gtgggaacaa	3720
	gcatgttctg	cttcgacggc	cccctgaccc	tgggtgtccta	caagcccag	cccgccaa	3780
	tgggtgtacct	cctgtccagc	tgcgacgagg	acgccagcat	caacgagagc	accggcaagc	3840
	cccagatggt	gatgtactac	aaccagacca	agggcggcgt	ggacacctg	gaccagatgt	3900
	gcagcgtgat	gacatgcagc	agaaagacca	acagatggcc	tatggccctg	ctgtacggca	3960
	tgatcaatat	cgctgcac	aacagcttca	tcactacag	ccacaacgtg	tccagcaagg	4020
	gcgagaaggt	gcagagccgg	aagaaattca	tgcggaacct	gtacatgagc	ctgacctcca	4080
	gcttcatgag	aaagagactg	gaagccccc	ccctgaagag	atacctccgg	gacaacatca	4140
	gcaacatcct	gcccgaaggaa	gtgccaggaa	caagcgacga	cagcaccgag	gaacccgtga	4200
	tgaagaagag	gacctactgc	acctactgtc	ccagcaagat	cagaagaaag	gccaacgcca	4260
	gctgcaagaa	atgcaaaaaa	gtgatctgcc	gggagcaca	catcgacatg	tgccagagct	4320
	gtttctgatt	cagacatgat	aagatacatt	gatgagtttg	gacaaaccac	aactagaatg	4380
	cagtgaaaaa	aatgctttat	ttgtgaaatt	tgtgatgcta	ttgctttatt	tgttaaccatt	4440
	ataagctgca	ataaacaagt	taacaacaac	aattgcattc	attttatgtt	tcaggttcag	4500
	ggggagatgt	gggaggtttt	ttaaagcaag	taaaacctct	acaaatgtgg	taagcaggtt	4560
	taaccttaga	aagatagtct	gcgtaaaatt	gacgcatgca	ttcttgaaat	attgctctct	4620
	ctttctaaat	agcgcgaatc	cgctcgctgtg	catttaggac	atctcagtcg	ccgcttgagg	4680
	ctcccgtgag	gcgtgcttgt	caatgcggta	agtgtcactg	attttgaact	ataacgaccg	4740
[0005]	cgtgagtcaa	aatgacgcat	gattatcttt	tacgtgactt	ttaagattta	actcatacga	4800
	taattatatt	gtttattcat	gttctactta	cgtgataact	tattatata	atattttctt	4860
	gttatagata	tctttcttta	tgttttaaat	gcaatgacct	cccacattcc	ctttttagta	4920
	aaatatccag	aaataattta	aatacatcat	tgcaatgaaa	ataaatgttt	tttattaggc	4980
	agaatccaga	tgctcaaggc	ccttcataat	atccccagct	ttagtagttg	gacttaggga	5040
	acaaaggaac	ctttaataga	aattggacag	caagaaagcg	agctattcct	ttgccctcgg	5100
	acgagtgtctg	gggcgtcggg	ttccactatc	ggcgagtact	tctacacagc	catcggtcca	5160
	gacggcccg	cttctgcggg	cgatttgtgt	acgcccagca	gtcccggctc	cggatcggac	5220
	gattgcgtcg	catcgacct	gcgcccagc	tgcacatcg	aaattgccgt	caaccaagct	5280
	ctgatagagt	tggatcaagc	caatgcggag	catatacgcc	cggagccg	gcgacccgc	5340
	aagctccgga	tgcctccgct	cgaagtagcg	cgtctgctgc	tccatacaag	ccaaccagg	5400
	cctccagaag	aagatgttgg	cgacctcgta	ttgggaatcc	ccgaacatcg	cctcgctcca	5460
	gtcaatgacc	gctgttatgc	ggccattgtc	cgtcaggaca	ttgttgagc	cgaatccgc	5520
	gtgcacgagg	tgcgggactt	cggggcagtc	ctcgcccaa	agcatcagct	catcgagagc	5580
	ctgcgcgacg	gacgcactga	cgggtgtcgtc	catcacagtt	tgccagtgat	acacatgggg	5640
	atcagcaatc	gcgcatatga	aatcacgcca	tgtagtgtat	tgaccgatc	cttgcggtcc	5700
	gaatgggccc	aacccgctcg	tctggctaag	atcgccgca	gcgacgcat	ccatggcctc	5760
	cgcgaccggc	tgcagaacag	cgggcagttc	ggtttcaggc	agatcttgca	acgtgacacc	5820
	ctgtgcacgg	cgggagatgc	aataggtcag	gctctcgtcg	aattcccaa	tgtcaagcac	5880
	ttccggaatc	gggagcgcgg	ccgatgcaaa	gtgccgataa	acataacgat	ctttgtagaa	5940
	accatcggcg	cagctattta	ccgcaggac	atatccacgc	cctctacat	cgaagctgaa	6000
	agcacgagat	tcttcgccct	ccgagagctg	catcaggtcg	gacacgctgt	cgaacttttc	6060
	gatcagaaac	ttctcgacag	acgtcgcggg	gagttcaggc	tttttcatgg	tggcggacga	6120
	aaggcccgga	gatgaggaag	aggagaacag	cgcggcagac	gtgcgctttt	gaagcgtgca	6180
	gaatgccggg	cctccggagg	accttcgggc	gcccccccg	cccctgagcc	cgcctctgag	6240
	cccccccg	gacccacccc	ttccagcct	ctgagcccg	aaagcgaagg	agcaaagctg	6300

	ctattggccg	ctgccccaaa	ggcctacccg	cttccattgc	tcagcgggtgc	tgtccatctg	6360
	cacgagacta	gtgagtcgtg	ctacttccat	ttgtcacgtc	ctgcacgacg	cgagctgcgg	6420
	ggcggggggg	aacttcctga	ctaggggagg	agtagaaggt	ggcgcgaagg	ggccacaaaa	6480
	gaacggagcc	ggttggcgcc	taccggtgga	tgtggaatgt	gtgcgaggcc	agaggccact	6540
	tgtgtagcgc	caagtgccca	gcggggctgc	taaagcgcac	gtcccagact	gccttgggaa	6600
	aagcgctec	cctaccgggt	agagaaactt	gatctgtcgc	cgcaattcaa	gcttcgtgag	6660
	gctccgggtgc	ccgtcagtga	cctgctatac	tctggagacg	gcacatcgcc	cacagtcccc	6720
	gagaagttag	gaggggtcgg	caattgaacc	ggtgcctaga	gaaggtggcg	cggggtaaacc	6780
	tgggaaagtg	atgtcgtgta	ctggctccgc	ctttttcccg	aggggtggggg	agaaccgtat	6840
	ataagtgcag	tagtcgccgt	gaacgttctt	tttcgcaacg	ggtttgccgc	cagaacacag	6900
	gtaagtgcg	tgtgtggttc	ccgcgggcct	ggcctcttta	cgggttatgg	cccttgcgtg	6960
	ccctgaatta	cttccacctg	gtctccagtac	gtgattcttg	atcccagact	ggagccaggg	7020
	gcgggccttg	cgctttagga	gccccctcgc	ctcgtgcttg	agttgaggcc	tggcctgggc	7080
	gctggggccg	ccgcgtgcga	atctgggtgc	accttcgcgc	ctgtctcgct	gctttcgata	7140
	agtctctagc	catttaaaat	ttttgatgac	gtgctgcgac	gctttttttc	tggcaagata	7200
	gtcttgtaaa	tgcgggccag	gatctgcaca	ctggtatttc	ggtttttggg	cccgcggccg	7260
	gcgacggggc	ccgtgcgtcc	cagcgcacat	gttcggcgag	gcggggccctg	cgagcgcggc	7320
	caccgagaat	cggacggggg	tagtctcaag	ctggccggcc	tgtcttggtg	cctggcctcg	7380
	cgccgccgtg	tatcgccccg	ccctgggcgg	caaggctggc	ccggtcggca	ccagttgcgt	7440
	gagcggaaa	atggccgctt	cccggccctg	ctccaggggg	ctcaaaatgg	aggacgcggc	7500
	gctcgggaga	gcgggcgggt	gagtcaccca	cacaaaggaa	aagggccttt	ccgtcctcag	7560
	ccgtcgcttc	atgtgactcc	acggagtacc	gggcgcctgc	caggcacctc	gattagttct	7620
	ggagcttttg	gagtacgtcg	tcttttaggtt	ggggggagggg	gttttatgcg	atggagtttc	7680
	cccacactga	gtgggtggag	actgaagtta	ggccagcttg	gcacttgatg	taattctcct	7740
	tggaaatttg	cttttttag	tttggtatct	ggttcattct	caagcctcag	acagtgggtc	7800
[0006]	aaagtttttt	tcttccattt	caggtgtcgt	gaacacgtct	cggggggccg	ccaccatgtt	7860
	cgactacctg	gaaaaatgag	aagtggccct	ggacgagctg	aagcagatgc	tgcgggaccg	7920
	ggacccccac	gacacccgga	accagttaaa	gaacaacgcc	ctgcacgcct	acctgtttaa	7980
	cgagcactgc	aacaacgtgg	aagtggtaaa	gctgctgctg	gactccggca	ccaacccccct	8040
	gcacaagaac	tggcggcagc	tgacccccct	gggcgagtac	accaactccc	ggcacggcaa	8100
	agtgaacaag	gatatcgcca	tgggtgctgt	ggaagctacc	ggctactcca	acatcaacga	8160
	cttcaacatc	ttcacctaca	tgaagtccaa	gaacgtggac	atcgacctga	tcaaggtgct	8220
	ggtggaacac	ggcttcgact	tctccgtgaa	gtgcgagaag	caccactccg	tgatcgagaa	8280
	ctacgtgatg	accgacgacc	ccgtgcctga	gatcatcgac	ctgttcacgc	agaacggctg	8340
	ctccgtgatc	tacgaggacg	aggacgacga	gtacggctac	gcctacgagg	aataccactc	8400
	ccagaacgac	gactaccage	cccggaaactg	cggcacccgtg	ctgcacctgt	acatcatctc	8460
	ccacctgtac	tccgagagcg	actcccggtc	ttgcgtgaac	cccgaggtgg	tcaagtgcct	8520
	gatcaaccac	ggcatcaacc	cctccagcat	cgataagaac	tactgcaccg	ccctgcagta	8580
	ctacatcaag	tctctccaca	tcgacatcga	tatcgtgaag	ctgctgatga	agggcacatga	8640
	caacaccgcc	tacagctaca	tcgacgacct	gacctgtgc	acccggggca	tcatggccga	8700
	ctacctgaac	agcgactacc	ggtacaacaa	ggacgtggac	ctggatctgg	tcaaactgtt	8760
	cctggaaaac	ggcaagcctc	acggcatcat	gtgctccatc	gtccccctgt	ggcggaacga	8820
	caaagagaca	atctccctga	tctgaaaac	catgaactcc	gacgtgctgc	agcatatcct	8880
	gatcgagtac	atcaccttct	ccgatatcga	catctctctg	gtcgagtaca	tgttggaata	8940
	cggcgctgtg	gtcaacaaa	aggccatcca	cggctacttc	aagaacatca	acatcgactc	9000
	ctacaccatg	aagtacctgc	tgaaaaaaga	gggcggcgac	gccgtcaacc	acctggacga	9060
	cggcgagatc	cccatcgccc	acctgtgcaa	gagcaactac	ggccggtaca	acttttacac	9120
	cgacacctac	cggcagggct	tccgggacat	gtcctacgcc	tgccccatcc	tgtccacct	9180
	caacatctgc	ctgccttacc	tgaaggacat	caatatgac	gataagcggg	gcgagacact	9240
	gctgcacaag	gccgtgcgct	acaacaagca	gtccctgggtg	tccctgctgc	tggaaagcgg	9300
	ctccgacgtg	aacatccggt	ccaacaacgg	ctacacctgt	atcgctatcg	ccatcaacga	9360

	gtcccgaac atcgagctgc tgaatatgct gctgtgccac aagcctaccc tggactgcgt	9420
	gacgactcc ctgcgcgaga tcagcaacat cgtggacaac gcctacgcca tcaagcagtg	9480
	catcagatac gccatgatta tcgacgactg catctctcc aagatccccg agtccatctc	9540
	caagcactac aacgactaca ttgacatctg caaccaggaa ctgaacgaga tgaagaaaat	9600
	catcgtgggc ggcaacacca ttttcagcct gatcttcacc gatcacggcg ccaagatcat	9660
	ccaccgctac gccacaaccc ccgagctgcg ggccactac gagtccaagc agaacaagat	9720
	ctacgtcgag gtctacgaca tcatcagcaa cgccatcgtg aagcacaaca agattcacia	9780
	gaacatcgag tccgtggacg acaacacctt catctccaac ctgccctata ccatcaagta	9840
	caagatcttc gagcagcagg attacaagga tgacgatgac aagtgaaaaa gcgtcttcct	9900
	gttctcatca catcatatca aggttatata ccatcaatat tgccacagat gttacttagc	9960
	cttttaatat ttctctaatt tagtgtatat gcaatgatag ttctctgatt tctgagattg	10020
	agtttctcat gtgtaatgat tatttagagt ttctctttca tctgttcaaa tttttgtcta	10080
	gttttatttt ttactgattt gtaagacttc tttttataat ctgcatatta caattctctt	10140
	tactggggtg ttgcaaatat tttctgtcat tctatggcct gacttttctt aatggttttt	10200
	taattttaaa aataagtctt aatattcatg caatctaatt aacaatcttt tctttgtggt	10260
	taggactttg agtcataaga aatttttctc tacactgaag tcatgatggc atgcttctat	10320
	attattttct aaaagattta aagttttgcc ttctccattt agacttataa ttcactggaa	10380
	ttttttgtg tgtatgggat gacatatggg ttccctttta tttttacat ataaatatat	10440
	ttccctgttt ttctaaaaaa gaaaaagatc atcattttcc cattgtaaaa tgccatattt	10500
	ttttcatagg tcacttacat atatcaatgg gtctgtttct gagctctact ctattttatc	10560
	agcctcactg tctatcccca cacatctcat gctttgctct aaatcttgat atttagtgga	10620
	acattctttc ccattttgtt ctacaagaat atttttgta ttgtctttgg gctttctata	10680
	tacattttga aatgaggttg acaagttaat aatcaacctc tggattacaa aatttgtgaa	10740
	agattgactg gtattcttaa ctatgttgct ccttttacgc tatgtggata cgctgcttta	10800
	atgcctttgt atcatgctat tgcttcccgat atggctttca tttctctctc cttgtataaa	10860
[0007]	tcctggttgc tgtctcttta tgaggagtig tgccccgtg tcaggcaacg tggcgtggtg	10920
	tgcactgtgt ttgtgacgc aacccccact ggttggggca ttgccaccac ctgtcagctc	10980
	ctttccggga ctttcgcttt cccctccct attgccacgg cggaactcat cgccgcctgc	11040
	cttgcgccgt gctggacagg ggctcggtcg ttgggcactg acaattccgt ggtgttgtcg	11100
	gggaaatcat cgtcctttcc ttggctgctc gcctgtgttg ccacctggat tctgcgcggg	11160
	acgtccttct gctacgtccc ttcggccctc aatccagcgg accttcttc ccgcggcctg	11220
	ctgcggctc tgcggcctct tccgctctt cgcttccgcc ctcagacgag tcggatctcc	11280
	ctttgggccc cctccccgca tcgcttgeta ttgtcttgcc aatcctcccc cttgctgtcc	11340
	tgccccaccc cccccccag aatagaatga cacctactca gacaatgcga tgcaatttcc	11400
	tcattttatt aggaaaggac agtgggagtg gcaccttcca gggtaagga aggcacgggg	11460
	gaggggcaaa caacagatgg ctggcaacta gaaggcacat ttgttacttt atagaagaaa	11520
	ttttgagttt ttgttttttt ttaataaata aataaacata aataaattgt ttgttgaatt	11580
	tattattagt atgtaagtgt aaatataata aaacttaata tctattcaaa ttaataaata	11640
	aacctcgata tacagaccga taaaacacat gcgtcaattt tacgcatgat tatctttaac	11700
	gtacgtcaca atatgattat ctttctaggg ttaagaagac tg	11742
	<210> 4	
	<211> 54	
	<212> DNA	
	<213> 人工	
	<220>	
	<223> 来自 pH51 质粒的 CP77-CHO 基因的正向引物	
	<400> 4	
	aacacgtctc ggggggccgc caccatgttc gactacctgg aaaatgagga agtg	54
	<210> 5	

	<211> 65	
	<212> DNA	
	<213> 人工	
	<220>	
	<223> 来自 pPH51 质粒的 CP77-CHO 基因的反向引物	
	<400> 5	
	caggaagacg ctttttcaact tgtcatcgctc atccttgtaa tcttctgtct cgaagatctt	60
	gtact	65
	<210> 6	
	<211> 855	
	<212> DNA	
	<213> 人工	
	<220>	
	<223> 哥本哈根痘苗病毒 K1L 基因的核苷酸序列(从 Genbank M35027.1 提取)	
	<400> 6	
	atggatctgt cacgaattaa tacttggaag tetaagcagc tgaaaagctt tctctctagt	60
	aaagatacat ttaaggcgga tgtccatgga catagtgcct tgtattatgc aatagctgat	120
	aataacgtgc gtctagtatg tacgttggtg aacgctggag cattgaaaaa tcttctagag	180
	aatgaatttc cattacatca ggcagccaca ttagaagata ccaaaatagt aaagattttg	240
	ctattcagtg gaatggatga ttcacaattt gatgacaaag gaaacaccgc attgtattat	300
	gcggttgata gtggtaacat gcaaacggtg aaactgtttg ttaagaaaaa ttggagactg	360
	atgttctatg ggaaaactgg atggaaaact tcattttatc atgccgtcat gcttaatgat	420
	gtaagtattg tatcatactt tctttcagaa ataccatcta cttttgatct ggctattctc	480
	cttagttgta ttcacaccac tataaaaaat ggacacgtgg atatgatgat tctcttgctc	540
[0008]	gactatatga cgtcgacaaa caccaataat tcccttctct tcattccgga cattaattg	600
	gctatagata ataaagacat tgagatgtta caggtctctg tcaaatacga cattaatata	660
	tactctgtta atctggaaaa tgtactattg gatgatgccg aaataactaa gatgattata	720
	gaaaagcatg ttgaatacaa gtctgactcc tatacaaaag atctcgatat cgtcaagaat	780
	aataaattgg atgaataat tagcaaaaac aaggaaactca gactcatgta cgtcaattgt	840
	gtaaagaaaa actaa	855
	<210> 7	
	<211> 284	
	<212> PRT	
	<213> 人工	
	<220>	
	<223> 由 SEQ ID NO: 6 编码的哥本哈根痘苗病毒 K1L 的氨基酸序列(从 Genbank M35027.1 提取)	
	<400> 7	
	Met Asp Leu Ser Arg Ile Asn Thr Trp Lys Ser Lys Gln Leu Lys Ser	
	1 5 10 15	
	Phe Leu Ser Ser Lys Asp Thr Phe Lys Ala Asp Val His Gly His Ser	
	20 25 30	
	Ala Leu Tyr Tyr Ala Ile Ala Asp Asn Asn Val Arg Leu Val Cys Thr	
	35 40 45	
	Leu Leu Asn Ala Gly Ala Leu Lys Asn Leu Leu Glu Asn Glu Phe Pro	
	50 55 60	
	Leu His Gln Ala Ala Thr Leu Glu Asp Thr Lys Ile Val Lys Ile Leu	
	65 70 75 80	
	Leu Phe Ser Gly Met Asp Asp Ser Gln Phe Asp Asp Lys Gly Asn Thr	
	85 90 95	

```

Ala Leu Tyr Tyr Ala Val Asp Ser Gly Asn Met Gln Thr Val Lys Leu
      100                      105                      110
Phe Val Lys Lys Asn Trp Arg Leu Met Phe Tyr Gly Lys Thr Gly Trp
      115                      120                      125
Lys Thr Ser Phe Tyr His Ala Val Met Leu Asn Asp Val Ser Ile Val
      130                      135                      140
Ser Tyr Phe Leu Ser Glu Ile Pro Ser Thr Phe Asp Leu Ala Ile Leu
      145                      150                      155                      160
Leu Ser Cys Ile His Thr Thr Ile Lys Asn Gly His Val Asp Met Met
      165                      170                      175
Ile Leu Leu Leu Asp Tyr Met Thr Ser Thr Asn Thr Asn Asn Ser Leu
      180                      185                      190
Leu Phe Ile Pro Asp Ile Lys Leu Ala Ile Asp Asn Lys Asp Ile Glu
      195                      200                      205
Met Leu Gln Ala Leu Phe Lys Tyr Asp Ile Asn Ile Tyr Ser Val Asn
      210                      215                      220
Leu Glu Asn Val Leu Leu Asp Asp Ala Glu Ile Thr Lys Met Ile Ile
      225                      230                      235                      240
Glu Lys His Val Glu Tyr Lys Ser Asp Ser Tyr Thr Lys Asp Leu Asp
      245                      250                      255
Ile Val Lys Asn Asn Lys Leu Asp Glu Ile Ile Ser Lys Asn Lys Glu
      260                      265                      270
Leu Arg Leu Met Tyr Val Asn Cys Val Lys Lys Asn
      275                      280

```

[0009]

```

<210> 8
<211> 551
<212> PRT
<213> 人工
<220>
<223> VACV-COP-D13L ORF
<400> 8
Met Asn Asn Thr Ile Ile Asn Ser Leu Ile Gly Gly Asp Asp Ser Ile
1      5      10      15
Lys Arg Ser Asn Val Phe Ala Val Asp Ser Gln Ile Pro Thr Leu Tyr
      20      25      30
Met Pro Gln Tyr Ile Ser Leu Ser Gly Val Met Thr Asn Asp Gly Pro
      35      40      45
Asp Asn Gln Ala Ile Ala Ser Phe Glu Ile Arg Asp Gln Tyr Ile Thr
      50      55      60
Ala Leu Asn His Leu Val Leu Ser Leu Glu Leu Pro Glu Val Lys Gly
65      70      75      80
Met Gly Arg Phe Gly Tyr Val Pro Tyr Val Gly Tyr Lys Cys Ile Asn
      85      90      95
His Val Ser Ile Ser Ser Cys Asn Gly Val Ile Trp Glu Ile Glu Gly
      100     105     110
Glu Glu Leu Tyr Asn Asn Cys Ile Asn Asn Thr Ile Ala Leu Lys His
      115     120     125
Ser Gly Tyr Ser Ser Glu Leu Asn Asp Ile Ser Ile Gly Leu Thr Pro
      130     135     140
Asn Asp Thr Ile Lys Glu Pro Ser Thr Val Tyr Val Tyr Ile Lys Thr

```

	145				150					155					160	
	Pro	Phe	Asp	Val	Glu	Asp	Thr	Phe	Ser	Ser	Leu	Lys	Leu	Ser	Asp	Ser
					165					170					175	
	Lys	Ile	Thr	Val	Thr	Val	Thr	Phe	Asn	Pro	Val	Ser	Asp	Ile	Val	Ile
					180					185					190	
	Arg	Asp	Ser	Ser	Phe	Asp	Phe	Glu	Thr	Phe	Asn	Lys	Glu	Phe	Val	Tyr
					195					200					205	
	Val	Pro	Glu	Leu	Ser	Phe	Ile	Gly	Tyr	Met	Val	Lys	Asn	Val	Gln	Ile
					210					215					220	
	Lys	Pro	Ser	Phe	Ile	Glu	Lys	Pro	Arg	Arg	Val	Ile	Gly	Gln	Ile	Asn
	225					230					235					240
	Gln	Pro	Thr	Ala	Thr	Val	Thr	Glu	Val	His	Ala	Ala	Thr	Ser	Leu	Ser
					245						250					255
	Val	Tyr	Thr	Lys	Pro	Tyr	Tyr	Gly	Asn	Thr	Asp	Asn	Lys	Phe	Ile	Ser
					260					265					270	
	Tyr	Pro	Gly	Tyr	Ser	Gln	Asp	Glu	Lys	Asp	Tyr	Ile	Asp	Ala	Tyr	Val
					275					280					285	
	Ser	Arg	Leu	Leu	Asp	Asp	Leu	Val	Ile	Val	Ser	Asp	Gly	Pro	Pro	Thr
					290					295					300	
	Gly	Tyr	Pro	Glu	Ser	Ala	Glu	Ile	Val	Glu	Val	Pro	Glu	Asp	Gly	Ile
	305					310					315					320
	Val	Ser	Ile	Gln	Asp	Ala	Asp	Val	Tyr	Val	Lys	Ile	Asp	Asn	Val	Pro
					325						330					335
	Asp	Asn	Met	Ser	Val	Tyr	Leu	His	Thr	Asn	Leu	Leu	Met	Phe	Gly	Thr
					340					345					350	
[0010]	Arg	Lys	Asn	Ser	Phe	Ile	Tyr	Asn	Ile	Ser	Lys	Lys	Phe	Ser	Ala	Ile
					355					360					365	
	Thr	Gly	Thr	Tyr	Ser	Asp	Ala	Thr	Lys	Arg	Thr	Ile	Phe	Ala	His	Ile
					370					375					380	
	Ser	His	Ser	Ile	Asn	Ile	Ile	Asp	Thr	Ser	Ile	Pro	Val	Ser	Leu	Trp
	385					390					395					400
	Thr	Ser	Gln	Arg	Asn	Val	Tyr	Asn	Gly	Asp	Asn	Arg	Ser	Ala	Glu	Ser
					405					410					415	
	Lys	Ala	Lys	Asp	Leu	Phe	Ile	Asn	Asp	Pro	Phe	Ile	Lys	Gly	Ile	Asp
					420					425					430	
	Phe	Lys	Asn	Lys	Thr	Asp	Ile	Ile	Ser	Arg	Leu	Glu	Val	Arg	Phe	Gly
					435					440					445	
	Asn	Asp	Val	Leu	Tyr	Ser	Glu	Asn	Gly	Pro	Ile	Ser	Arg	Ile	Tyr	Asn
					450					455					460	
	Glu	Leu	Leu	Thr	Lys	Ser	Asn	Asn	Gly	Thr	Arg	Thr	Leu	Thr	Phe	Asn
	465					470					475					480
	Phe	Thr	Pro	Lys	Ile	Phe	Phe	Arg	Pro	Thr	Thr	Ile	Thr	Ala	Asn	Val
					485					490					495	
	Ser	Arg	Gly	Lys	Asp	Lys	Leu	Ser	Val	Arg	Val	Val	Tyr	Ser	Thr	Met
					500					505					510	
	Asp	Val	Asn	His	Pro	Ile	Tyr	Tyr	Val	Gln	Lys	Gln	Leu	Val	Val	Val
					515					520					525	
	Cys	Asn	Asp	Leu	Tyr	Lys	Val	Ser	Tyr	Asp	Gln	Gly	Val	Ser	Ile	Thr
					530					535					540	
	Lys	Ile	Met	Gly	Asp	Asn	Asn									
	545					550										

<210>	9	
<211>	1680	
<212>	DNA	
<213>	人工	
<220>		
<223>	编码 SEQ ID NO:1 的 CHO 密码子优化的核苷酸序列	
<400>	9	
	atgaacaaca ccatcatcaa ctccctgata ggcggcgacg actccatcaa gcggtccaac	60
	gtgttcgccg tggactccca gatccccacc ctgtacatgc ccagtacat ctccctgtcc	120
	ggcgtgatga ccaacgacgg ccctgacaac caggctatcg cctccttcga gatccgggat	180
	cagtacatca ccgccctgaa ccacctggtg ctgtccctgg aactgcccga agtgaagggc	240
	atgggcagat tcggctacgt gccctacgtg ggctacaagt gcatcaacca cgtgtccatc	300
	tccagctgca acggcgtgat ctgggaaatc gagggcgagg aactgtacaa caactgcatt	360
	aacaacacaa tcgccctgaa gcactccggc tactcctccg agctgaacga catctccatc	420
	ggcctgaccc ccaacgacac catcaaagaa cctccaccg tgtacgtgta catcaagacc	480
	cccttcgacg tgaagatac cttctccagc ctgaagctgt ccgactccaa gatcacctg	540
	accgtgacct tcaacctgt gtccgacatc gtgacccggg actccagctt cgacttcgag	600
	acattcaaca aagaatttgt gtacgtgcc gagctgtcct tcatcggtca catggtcaag	660
	aacgtgcaga tcaagcccag cttcatcgag aagcctcgga gagtgatcgg ccagatcaac	720
	cagcctaccg ccaccgtgac agagggtcac gccgccacat ccctgagcgt gtacaccaag	780
	ccctactacg gcaacaccga caacaagttc atctcctacc ccggtacag ccaggacgag	840
	aaggactaca tcgacgccta cgtgtcccgg ctgctggacg acctcgtgat cgtgtctgat	900
	ggcccccta ccggctaccc tgagtctgcc gagatcgtgg aagtgcccga ggacggcatc	960
	gtcagcatcc aggacgccga tgtgtatgtg aagatcgaca acgtgccaga caacatgtcc	1020
[0011]	gtgtacctgc acaccaacct gctgatgttc ggcaccgga agaattcctt catctacaac	1080
	atctccaaga agttctccgc catcaccggc acctactccg acgccacca gcggaccatc	1140
	ttcggccaca tctccacag catcaacatc atcgacacct ccatccccgt gtccctgtgg	1200
	acctctcaga gaaacgtgta caacggcgac aacagatccg ccgagtccaa ggccaaggac	1260
	ctgttcatca acgacccctt catcaagggc atcgacttca agaacaagac cgacatcatc	1320
	tcccggtcgg aagtgcgctt cggcaacgac gtgctgtact ccgagaacgg ccctatcagc	1380
	cggatctaca acgagctgct gaccaagtcc aacaacggca ccagaacctt gacctttaac	1440
	tttaccacca agatcttctt ccggcccacc accatcacgg ctaacgtgtc cagaggcaag	1500
	gacaagctga gcgtgcgggt ggtgtactcc accatggacg tgaaccaccc catctactac	1560
	gtgcagaaac agctggtggt cgtgtgcaac gatctgtaca aggtgtccta cgaccagggc	1620
	gtgtccatta ccaagatcat gggcgataac aacgactaca aggacgacga cgacaagtga	1680
<210>	10	
<211>	44	
<212>	DNA	
<213>	合成引物	
<400>	10	
	aacacgtctc ggggggccgc caccatgaac aacaccatca tcaa	44
<210>	11	
<211>	40	
<212>	DNA	
<213>	合成引物	
<400>	11	
	caggaagacg ctttttcaact tgtcgtcgtc gtcctttag	40

	<210> 12	
	<211> 8	
	<212> PRT	
	<213> 人工	
	<220>	
	<223> Flag-标签蛋白	
	<400> 12	
	Asp Tyr Lys Asp Asp Asp Asp Lys	
	1 5	
	<210> 13	
	<211> 16	
	<212> PRT	
	<213> 人工	
	<220>	
	<223> D13L 肽	
	<400> 13	
	Cys Tyr Asp Gln Gly Val Ser Ile Thr Lys Ile Met Gly Asp Asn Asn	
	1 5 10 15	
	<210> 14	
	<211> 35	
	<212> DNA	
	<213> 合成引物	
	<400> 14	
[0012]	cggtaccg ggcacgaa aaataatagt aacca	35
	<210> 15	
	<211> 38	
	<212> DNA	
	<213> 合成引物	
	<400> 15	
	aatttagtgt gcgcgtggaa aaagcttaca ataaactc	38
	<210> 16	
	<211> 36	
	<212> DNA	
	<213> 合成引物	
	<400> 16	
	atatttaa at gcgcgcaata atggaacaag aaccct	36
	<210> 17	
	<211> 36	
	<212> DNA	
	<213> 合成引物	
	<400> 17	
	cgactctaga ggcgcgcgt gaggtcggca actacg	36
	<210> 18	
	<211> 668	
	<212> PRT	

<213> 人工

<220>

<223> 包含 CPXV-125L Brighton Red 的序列的 CP77 蛋白

<400> 18

```

Met Phe Asp Tyr Leu Glu Asn Glu Glu Val Ala Leu Asp Glu Leu Lys
1           5           10           15
Gln Met Leu Arg Asp Arg Asp Pro Asn Asp Thr Arg Asn Gln Phe Lys
           20           25           30
Asn Asn Ala Leu His Ala Tyr Leu Phe Asn Glu His Cys Asn Asn Val
           35           40           45
Glu Val Val Lys Leu Leu Leu Asp Ser Gly Thr Asn Pro Leu His Lys
           50           55           60
Asn Trp Arg Gln Leu Thr Pro Leu Gly Glu Tyr Thr Asn Ser Arg His
65           70           75           80
Gly Lys Val Asn Lys Asp Ile Ala Met Val Leu Leu Glu Ala Thr Gly
           85           90           95
Tyr Ser Asn Ile Asn Asp Phe Asn Ile Phe Thr Tyr Met Lys Ser Lys
           100          105          110
Asn Val Asp Ile Asp Leu Ile Lys Val Leu Val Glu His Gly Phe Asp
           115          120          125
Phe Ser Val Lys Cys Glu Lys His His Ser Val Ile Glu Asn Tyr Val
           130          135          140
Met Thr Asp Asp Pro Val Pro Glu Ile Ile Asp Leu Phe Ile Glu Asn
145          150          155          160
Gly Cys Ser Val Ile Tyr Glu Asp Glu Asp Asp Glu Tyr Gly Tyr Ala
           165          170          175
Tyr Glu Glu Tyr His Ser Gln Asn Asp Asp Tyr Gln Pro Arg Asn Cys
           180          185          190
Gly Thr Val Leu His Leu Tyr Ile Ile Ser His Leu Tyr Ser Glu Ser
           195          200          205
Asp Ser Arg Ser Cys Val Asn Pro Glu Val Val Lys Cys Leu Ile Asn
           210          215          220
His Gly Ile Asn Pro Ser Ser Ile Asp Lys Asn Tyr Cys Thr Ala Leu
225          230          235          240
Gln Tyr Tyr Ile Lys Ser Ser His Ile Asp Ile Asp Ile Val Lys Leu
           245          250          255
Leu Met Lys Gly Ile Asp Asn Thr Ala Tyr Ser Tyr Ile Asp Asp Leu
           260          265          270
Thr Cys Cys Thr Arg Gly Ile Met Ala Asp Tyr Leu Asn Ser Asp Tyr
           275          280          285
Arg Tyr Asn Lys Asp Val Asp Leu Asp Leu Val Lys Leu Phe Leu Glu
           290          295          300
Asn Gly Lys Pro His Gly Ile Met Cys Ser Ile Val Pro Leu Trp Arg
305          310          315          320
Asn Asp Lys Glu Thr Ile Ser Leu Ile Leu Lys Thr Met Asn Ser Asp
           325          330          335
Val Leu Gln His Ile Leu Ile Glu Tyr Ile Thr Phe Ser Asp Ile Asp
           340          345          350
Ile Ser Leu Val Glu Tyr Met Leu Glu Tyr Gly Ala Val Val Asn Lys
           355          360          365
Glu Ala Ile His Gly Tyr Phe Lys Asn Ile Asn Ile Asp Ser Tyr Thr

```

[0013]

	370	375	380
	Met Lys Tyr Leu Leu Lys Lys Glu Gly Gly Asp Ala Val Asn His Leu		
	385	390	395
	Asp Asp Gly Glu Ile Pro Ile Gly His Leu Cys Lys Ser Asn Tyr Gly		400
	405	410	415
	Arg Tyr Asn Phe Tyr Thr Asp Thr Tyr Arg Gln Gly Phe Arg Asp Met		
	420	425	430
	Ser Tyr Ala Cys Pro Ile Leu Ser Thr Ile Asn Ile Cys Leu Pro Tyr		
	435	440	445
	Leu Lys Asp Ile Asn Met Ile Asp Lys Arg Gly Glu Thr Leu Leu His		
	450	455	460
	Lys Ala Val Arg Tyr Asn Lys Gln Ser Leu Val Ser Leu Leu Leu Glu		
	465	470	475
	Ser Gly Ser Asp Val Asn Ile Arg Ser Asn Asn Gly Tyr Thr Cys Ile		
	485	490	495
	Ala Ile Ala Ile Asn Glu Ser Arg Asn Ile Glu Leu Leu Asn Met Leu		
	500	505	510
	Leu Cys His Lys Pro Thr Leu Asp Cys Val Ile Asp Ser Leu Arg Glu		
	515	520	525
	Ile Ser Asn Ile Val Asp Asn Ala Tyr Ala Ile Lys Gln Cys Ile Arg		
	530	535	540
	Tyr Ala Met Ile Ile Asp Asp Cys Ile Ser Ser Lys Ile Pro Glu Ser		
	545	550	555
	Ile Ser Lys His Tyr Asn Asp Tyr Ile Asp Ile Cys Asn Gln Glu Leu		
	565	570	575
[0014]	Asn Glu Met Lys Lys Ile Ile Val Gly Gly Asn Thr Met Phe Ser Leu		
	580	585	590
	Ile Phe Thr Asp His Gly Ala Lys Ile Ile His Arg Tyr Ala Asn Asn		
	595	600	605
	Pro Glu Leu Arg Ala Tyr Tyr Glu Ser Lys Gln Asn Lys Ile Tyr Val		
	610	615	620
	Glu Val Tyr Asp Ile Ile Ser Asn Ala Ile Val Lys His Asn Lys Ile		
	625	630	635
	His Lys Asn Ile Glu Ser Val Asp Asp Asn Thr Tyr Ile Ser Asn Leu		
	645	650	655
	Pro Tyr Thr Ile Lys Tyr Lys Ile Phe Glu Gln Gln		
	660	665	

<210> 19

<211> 2903

<212> DNA

<213> 人工

<220>

<223> CP77+DsRed 表达盒序列

<400> 19

gcgcgcacac taaattat	60
ttttttaata attgtacaag tttttgatct ggtataaata	
cattcaaaaa tgataattta atgacattag ttgtgcgggt gtatagagtt cacagtagct	120
cattcacttc tattcagtca aaatgtttga ttatctggaa aatgaggagg tggctctcga	180
tgaacttaaa cagatgttga gagatagaga tcctaagat accaggaacc aattcaagaa	240
taatgctcta catgcatacc ttttcaatga gcaactgtaaat aatgttgagg ttgtcaaact	300
actactagac agtgggtacta atccattaca caaaaattgg agacagctta ctccattagg	360

	ggaatacaca aatagtagac atggtaaagt taataaggat atagcgaagg ttctactaga	420
	agctactgga tatagcaaca taaatgactt taatatattc acctatatga aatccaaaaa	480
	tgtagatatt gacttgataa aggtattggg agaacatgga ttgtatttta gtgttaaag	540
	cgaaaaacat cattcagtta tagaaaatta tgtaatgaca gatgatcctg ttccctgaaat	600
	tattgatttta ttcatagaaa atggatgcag tggtatttat gaggacgagg atgatgagta	660
	cggatacgcg tatgaagaat atcactcaca aaatgacgat tatcaaccac gaaattgcgg	720
	tacagtatta catctgtata tcactctctca tctgtattca gaggcggatt cgagatcatg	780
	tgtgaacccg gaagttgtta aatgtctgat taatcatgga atcaaccat cttctataga	840
	taaaaactat tgtacagctc ttcaatatta tattaagtca tctcatatag atatagacat	900
	cgttaaattg ttaatgaaag gaatagataa cacggcttat tcatatatag acgatctaac	960
	atgttgact cgaggaatta tggctgatta tctaaatagt gattatagat acaataaaga	1020
	tgtagatttta gatttggtca aattgttttt ggagaatgga aaaccgcacg gaataatgtg	1080
	tagtattgta ccactatgga gaaatgataa ggaaaccatc tctttgatat tgaaaacaat	1140
	gaactcggat gtcctccaac atatacta atagtagata acattcagcg atatcgatat	1200
	ctctctagtg gaatacatgt tggaaatagg agctgtggta aataaagagg ctattcacgg	1260
	atactttaa aatattaata ttgattctta cagcatgaaa tatctactaa aaaaggaagg	1320
	gggagatgcc gtcaatcatc tcgatgatgg agagatcccg attggacacc tatgtaaatc	1380
	caactatgga cgttataatt tctacactga tacatacaga cagggttttc gtgatatgtc	1440
	ttatgcttgc ccaattctta gtactataaa catttgccca ccttatctta aagacattaa	1500
	catgattgac aaacgaggag aaacacttct tcacaaggct gttagatata ataacaatc	1560
	tctagtgtct ttactgctag aatccgggtc agatgtcaac attagatcaa ataacggata	1620
	tacatgtata gccattgcaa tcaacgaatc tagaaacatt gaactgctga acatgctatt	1680
	atgtcataaa cctacattag attgtgtgat tgattcattg agagaaatat ctaacatagt	1740
	agataatgcc tatgctataa aacaatgtat tagatatgcc atgattatag atgactgtat	1800
[0015]	atcgtctaag attccagagt ccataagtaa acactataat gattatatag atatttgcaa	1860
	tcaagaattg aacgagatga aaaaaataat agtgggaggc aacactatgt tctcattaat	1920
	atttactgat catggagcta aaattattca tcggtatgcc aataatccag aattacgtgc	1980
	gtattatgag tcaaaacaaa ataaaatata cgtggaagta tatgatatta ttccaatgc	2040
	gatagtgaag cataataaaa ttcataaaaa catagaatca gttgatgaca atacctacat	2100
	ttctaatttg ccttatacca tcaaatacaa aatattcgag caacaataag tattttttat	2160
	caaaaaattg aaattttatt tttttttttt ggaatataaa taatatggat agcactgaga	2220
	acgtcatcaa gcccttcatg cgcttcaagg tgcacatgga gggctccgtg aacggccacg	2280
	agttcgagat cgagggcgag ggcgagggca agccctacga gggcaccag accgccaagc	2340
	tgcaggtagc caagggcggc cccctgccct tcgcttggga catcctgtcc cccagattcc	2400
	agtacggctc caaggtgtac gtgaagcacc ccgcccacat ccccgactac aagaagctgt	2460
	ccttccccga gggcttcaag tgggagcgcg tgatgaactt cgaggacggc ggcgtggtga	2520
	ccgtgaccca ggactcctca ctgcaggacg gcaccttcat ctaccacgtg aagttcatcg	2580
	gcgtgaactt cccctccgac ggccccgtaa tgcagaagaa gactctgggc tgggagccct	2640
	ccaccgagcg cctgtacccc cgcgacggcg tgctgaaggg cgagatccac aaggcgctga	2700
	agctgaaggg cgggcgccac tacctggtgg agttcaagtc aatctacatg gccagaagc	2760
	ccgtgaagct gcccggctac tactacgtgg actccaagct ggacatcacc tcccacaacg	2820
	aggactacac cgtggtggag cagtacgaac gcgcccaggc ccgccaccac ctgttccagt	2880
	agtttttata tttaaatgcg cgc	2903
	<210> 20	
	<211> 25	
	<212> DNA	
	<213> 合成引物	
	<400> 20	
	tacaaaatca aataatggtc gaaac	25

	<210>	21	
	<211>	29	
	<212>	DNA	
	<213>	合成引物	
	<400>	21	
		tgccaagaaa acactccttc taagacaat	29
	<210>	22	
	<211>	6809	
	<212>	DNA	
	<213>	人工	
	<220>		
	<223>	具有 CP77/DsRed 选择的用于 VACV-COP D13L 缺失的 pLL09 同源重组载体	
	<400>	22	
		tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccc gagacgggtca	60
		cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg	120
		ttggcgggtg tcggggctgg cttaactatg cgcatcaga gcagattgta ctgagagtgc	180
		accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc	240
		attcgccatt caggctgcgc aactgttggg aaggcgatc ggtgcgggcc tcttcgctat	300
		tacgccagct ggcgaaagg ggatgtctg caaggcgatt aagtgggta acgccagggt	360
		tttcccagtc acgacgttgt aaaacgacgg ccagtgaatt cgagctcggg acccggggat	420
		cacgaaaaat aatagtaacc aattagtatg ggaaaacttt ttgacacata tgagatctaa	480
		aaaacgtgtt actatggtag aggattatgg acacgagtac gttttttag atgagaggtt	540
		ttctacttgc tcattagaag tataaaaaaa tagttccgta attaaatggc taagcgagta	600
		agccttcag atgtggttat ttcagacct aaagcagctt ttaagccgc taaagaagaa	660
[0016]		gcactcgctt gtatactacc aaagtattat aaatctatgg cagatgtgtc tattaagaca	720
		aatagtgtaa ttgataagtg ttggttttgt aatcaagatt tggtttttag acctattagt	780
		attgagacat acaagggtgg tgaagtaggg tatctctgtt ctaaaatatg tagggattcg	840
		ttgcttcta tggttaaagtc tcacgtagct cttagagaag aaccaaaaat ttctttgtt	900
		cctttagtat tctatgaaga taaggaaaag gttataaata caataaactt actaagagat	960
		aaagacggcg ttacggaag ctgttacttt aaggaaaact cacaaattat agatatttct	1020
		ctacggagtt tattgtaagc tttttccacg cgcacactaa attatttttt attaataatt	1080
		gtacaagttt ttgatctggt ataaatacat tcaaaaatga taatttaattg acattagtgt	1140
		tgcgggtgta tagagttcac agtagctcat tcacttctat tcagtcaaaa tgtttgatta	1200
		tctgaaaaat gaggaggtgg ctctcgatga acttaaacag atgttgagag atagagatcc	1260
		taatgatacc aggaaccaat tcaagaataa tgctctacat gcataccttt tcaatgagca	1320
		ctgtaataat gttgaggttg tcaaaactact actagacagt ggtactaatc cattacacaa	1380
		aaattggaga cagcttactc cattagggga atacacaaat agtagacatg gtaaagttaa	1440
		taaggatata gcgatggttc tactagaagc tactggatat agcaacataa atgactttta	1500
		tatatccacc tatatgaaat ccaaaaatgt agatattgac ttgataaagg tattggtaga	1560
		acatggattt gatttttagt ttaaatgcga aaaacatcat tcagttatag aaaattatgt	1620
		aatgacagat gatctgttc ctgaaattat tgattttatc atagaaaatg gatgcagtgt	1680
		tattttatgag gacgaggatg atgagtacgg atacgcgtat gaagaatata actcacaaaa	1740
		tgacgattat caaccacgaa attgcggtac agtattacat ctgtatatca tctctcatct	1800
		gtattcagag tcggattcga gatcatgtgt gaacccggaa gttgttaaat gtctgattaa	1860
		tcatggaatc aaccatctt ctatagataa aaactattgt acagctcttc aatattatat	1920
		taagtcactc catatagata tagacatcgt taaattgtta atgaaaggaa tagataacac	1980
		ggcttattca tatatagacg atctaactatg ttgtactcga ggaattatgg ctgattatct	2040
		aaatagtgat tatagataca ataaagatgt agatttagat ttgggtcaaat tgtttttgga	2100
		gaatggaaaa ccgcacggaa taatgtgtag tattgtacca ctatggagaa atgataagga	2160
		aaccatctct ttgatattga aaacaatgaa ctcggtatgc ctccaacata tactaattga	2220

	gtatataaca ttcagcgata tcgatatctc tctagtggaa tacatgttgg aatatggagc	2280
	tgtggtaaat aaagaggcta ttcacggata ctttaaaaat attaatatg attcttacac	2340
	gatgaaatat ctactaaaaa aggaaggggg agatgccgtc aatcatctcg atgatggaga	2400
	gatcccgatt ggacacctat gtaaattcaa ctatggacgt tataatttct acactgatac	2460
	atacagacag ggttttcgtg atatgtctta tgettgccea attcttagta ctataaacat	2520
	ttgcctacct tatcttaaag acattaacat gattgacaaa cgaggagaaa cacttcttca	2580
	caaggctgtt agatataata aacaatctct agtgtcttta ctgctagaat ccggttcaga	2640
	tgtcaacatt agatcaaata acggatatac atgtatagcc attgcaatca acgaatctag	2700
	aaacattgaa ctgctgaaca tgctattatg tcataaacct acattagatt gtgtgattga	2760
	ttcattgaga gaaatatcta acatagtaga taatgcctat gctataaaac aatgtattag	2820
	atatgccatg attatagatg actgtatatc gtctaagatt ccagagtcca taagtaaca	2880
	ctataatgat tatatagata ttgcaatca agaattgaac gagatgaaaa aaataatagt	2940
	gggaggcaac actatgttct cattaatatt tactgatcat ggagctaaaa ttattcatcg	3000
	gtatgccaat aatccagaat tacgtgcgta ttatgagtca aaacaaaata aaatatacgt	3060
	ggaagtatat gatattatct ccaatgcgat agtgaagcat aataaaattc ataaaaacat	3120
	agaatcagtt gatgacaata cctacatttc taatttgcct tataccatca aatacaaaat	3180
	attcgagcaa caataagtat tttttatcaa aaaattgaaa ttttattttt tttttttgga	3240
	atataaataa tatggatagc actgagaacg tcatcaagcc ctctcatgcgc ttcaaggtgc	3300
	acatggaggg ctccgtgaac ggccacgagt tcgagatcga gggcgagggc gagggcaagc	3360
	cctacgaggg caccagacc gccaaagctgc aggtgaccaa gggcgggccc ctgcccttcg	3420
	cctgggacat cctgtccccc cagttccagt acggtcccaa ggtgtacgtg aagcaccctg	3480
	ccgacatccc cgactacaag aagctgtcct tccccgaggg cttcaagtgg gagcgcgtga	3540
	tgaacttcca ggacggcggc gtggtgaccg tgacccagga ctccctactg caggacggca	3600
	ccttcatcta ccacgtgaag ttcacggcg tgaaactccc ctccgacggc cccgtaatgc	3660
	agaagaagac tctgggctgg gagccctcca ccgagcgct gtacccccgc gacggcgtgc	3720
[0017]	tgaaggcgga gatccacaag gcgctgaagc tgaaggcgcg cgccactac ctggtggagt	3780
	tcaagtcaat ctacatggcc aagaagcccg tgaagctgcc cgctactac tacgtggact	3840
	ccaagctgga catcacctcc cacaacgagg actacaccgt ggtggagcag tacgaacgcg	3900
	ccgagggccc ccaccacctg ttccagtagt ttttatattt aaatgcgcgc aataatggaa	3960
	caagaaccct aacttttaac tttacacca agatattctt taggccgaca actattacgg	4020
	ccaatgtatc tagggggaaa gataaactat ctgttcgagt agtttatccc accatggatg	4080
	tcaaccatcc aatctattat gtacaaaaac aattggtagt tgtatgtaat gacctgtata	4140
	aggtatctta cgatcaaggg gtaagtatta ccaagattat gggagataat aactaataat	4200
	aatgaaaaca aactatagag ttgtaaatgg atgaaattgt aaaaaatc cgggagggaa	4260
	cgcattgctt tcttccattt tatgaaacat tgccagaact taatctgtct ctaggtaaaa	4320
	gccattacc tagtctggaa tacggagcta attactttct tcagatttct agagttaatg	4380
	atctaaatag aatgccgacc gacatgttaa aactttttac acatgatac atgttaccag	4440
	aaagcgatct agataaagtc tatgaaattt taaagattaa tagcgtaaag tattatggga	4500
	ggagtactaa agcggacgcc gtagttgccg acctcagcgc gatcctctag agtcgacctg	4560
	caggcatgca agcttgccgt aatcatggtc atagctgttt cctgtgtgaa attgttatcc	4620
	gctcacaatt ccacacaaca tacgagccgg aagcataaag tgtaaagcct ggggtgccta	4680
	atgagtgagc taactcacat taattgcgtt gcgctcactg cccgctttcc agtcgggaaa	4740
	cctgtcgtgc cagctgcatt aatgaatcgg ccaacgcgcg gggagaggcg gtttgcgtat	4800
	tgggcgctct tccgttctct cgctcactga ctgctgcgc tcggctgttc ggtgcggcg	4860
	agcggtatca gctcactcaa agcggttaat acggttatcc acagaatcag gggataacgc	4920
	aggaaagaac atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt	4980
	gctggcggtt ttccataggg tccgcccccc tgacgagcat cacaaaaatc gacgtcaag	5040
	tcagagggtg cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc	5100
	cctcgtgcgc tctctgttc cgacctgcc gcttaccgga tacctgtccg ctttctccc	5160
	ttcggaagc gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cgggtgtaggt	5220
	cgttcgctcc aagctgggct gtgtgcacga accccccgtt cagcccgacc gctgcgcctt	5280

	atccggtaac tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc	5340
	agccactggt aacaggatta gcagagcgag gtaigttaggc ggtgctacag agttcttgaa	5400
	gtgggtggcct aactacggct aactagaag gacagtattt ggtatctgcg ctctgctgaa	5460
	gccagttacc ttcggaaaaa gagttagtag ctcttgatcc ggcaaaaaa ccaccgctgg	5520
	tagcgggtggt ttttttggtt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga	5580
	agatcccttg atcttttcta cggggctctga cgctcagtgg aacgaaaact cacgttaagg	5640
	gattttggtc atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg	5700
	aagttttaa tcaatctaaa gtatatatga gtaaaccttg tctgacagtt accaatgctt	5760
	aatcagttag gcacctatct cagcgatctg tctatttctg tcatccatag ttgcctgact	5820
	ccccgtcgtg tagataacta cgatacggga gggtttacca tctggcccca gtgctgcaat	5880
	gataccgcga gaccacgct caccggctcc agatttatca gcaataaacc agccagccgg	5940
	aagggccgag cgcagaagtg gtcttgcaac tttatccgcc tccatccagt ctattaattg	6000
	ttgccgggaa gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg ttgttgccat	6060
	tgctacaggc atcgttggtg cacgctcgtc gtttggtatg gcttcattca gctccggttc	6120
	ccaacgatca aggcgagtta catgatcccc catgttggtc aaaaaagcgg ttagctcctt	6180
	cggctctccg atcgttggtc gaagtaagt ggccgcagt ttatcactca tggttatggc	6240
	agcactgcat aattctctta ctgtcatgcc atccgtaaga tgcttttctg tgactggtga	6300
	gtactcaacc aagtcattct gagaatagt tatgcccga ccgagttgct cttgcccggc	6360
	gtcaatacgg gataataacc cgccacatag cagaacttta aaagtgtca tcatggaaa	6420
	acgtttctcg gggcgaaaac tctcaaggat cttaccgctg ttgagatcca gttcgatgta	6480
	accactcgt gcacccaact gatcttcagc atcttttact ttcaccagcg tttctgggtg	6540
	agcaaaaaca ggaaggcaaa atgccgcaaa aaagggaata agggcgacac ggaaatgttg	6600
	aatactcata ctcttctttt ttcaatatta ttgaagcatt tatcagggtt attgtctcat	6660
	gagcggatac atatttgaat gtatttagaa aaataaaca ataggggttc cgcgacatt	6720
	tccccgaaaa gtgccacctg acgtctaaga aaccattatt atcatgacat taacctataa	6780
[0018]	aaataggcgt atcacgaggc ctttcgtc	6809

<210> 23

<211> 11161

<212> DNA

<213> 人工

<220>

<223> 用于整合到细胞核基因组 DNA 的 pLL19 加标签的-D13L-CHO 密码子优化的转座子介导的转导载体

<400> 23

tcagaattgg ttaattggtt gtaacactga cccctatttg tttatttttc taaatacatt	60
caaatatgta tccgtcatg agacaataac cctgataaat gttcaataa tattgaaaaa	120
ggaagaatat gagccatatt caacgggaaa cgtcgaggcc gcgattaaat tccaacatgg	180
atgctgattt atatgggtat aaatgggctc gcgataatgt cgggcaatca ggtgcgacaa	240
tctatcgctt gtatgggaag cccgatgcgc cagagttggt tctgaaacat ggcaaaggta	300
gcgttgccaa tgatgttaca gatgagatgg tcagactaaa ctggetgacg gaatttatgc	360
cacttccgac catcaagcat tttatccgta ctctgatga tgcattggtt ctcaccactg	420
cgatccccgg aaaaacagcg ttccaggtat tagaagaata tcttgattca ggtgaaaata	480
ttgttgatgc gctggcagtg ttctgcgcc ggttgcactc gattcctggt tgtaattgtc	540
cttttaacag cgatcgcgta tttcgctcgc ctcaggcgca atcacgaatg aataacggtt	600
tggttgatgc gagtgatttt gatgacgagc gtaatggctg gcctgttgaa caagtctgga	660
aagaaatgca taaacttttg ccattctcac cggattcagt cgtcactcat ggtgatttct	720
cacttgataa ccttattttt gacgagggga aattaatagg ttgtattgat gttggacgag	780
tcggaatcgc agaccgatac caggatcttg ccatcctatg gaactgcctc ggtgagtttt	840
ctccttcatt acagaaacgg ctttttcaaa aatatggtat tgataatcct gatatgaata	900
aattgcagtt tcatttgatg ctcgatgagt ttttctaact catgacaaa atcccttaac	960

	gtgagttacg	cgcgcgtcgt	tccactgagc	gtcagacccc	gtagaaaaga	tcaaaggatc	1020
	ttcttgagat	cctttttttc	tgcgcgtaat	ctgctgcttg	caaacaaaaa	aaccaccgct	1080
	accagcgggtg	gtttgtttgc	cggatcaaga	gctaccaact	ctttttccga	aggtaactgg	1140
	cttcagcaga	gcgcagatac	caaatactgt	tcttctagt	tagccgtagt	tagccccacca	1200
	cttcaagaac	tctgtagcac	cgcctacata	cctcgctctg	ctaatectgt	taccagtggc	1260
	tgctgccagt	ggcgataagt	cgtgtcttac	cgggttggac	tcaagacgat	agttaccgga	1320
	taaggcgcag	cggtcgggct	gaacgggggg	ttcgtgcaca	cagcccagct	tggagcgaac	1380
	gacctacacc	gaactgagat	acctacagcg	tgagctatga	gaaagcgcca	cgttccccga	1440
	agggagaaaag	gcggacaggt	atccggttaag	cggcagggtc	ggaacaggag	agcgcacgag	1500
	ggagcttcca	gggggaaacg	cctggtatct	ttatagtcct	gtcgggtttc	gccacctctg	1560
	acttgagcgt	cgatttttgt	gatgctcgtc	aggggggcgg	agcctatgga	aaaacgccag	1620
	caacgcggcc	tttttacggt	tcttggcctt	ttgctggcct	tttctcaca	tgttctttcc	1680
	tgcgttatcc	cctgattctg	tggataaccg	tattaccgcc	tttgagttag	ctgataccgc	1740
	tctcaatatt	ggccattagc	catattatc	attggttata	tagcataaat	caatattggc	1800
	tattggccat	tgcatacgtt	gtatctatat	cataatatgt	acatttatat	tggctcatgt	1860
	ccaatatgac	cgccatgttg	gcattgatta	ttgactagtt	attaatagta	atcaattacg	1920
	gggtcattag	ttcatagccc	atataatggag	ttccgcgtta	cataacttac	ggtaaatggc	1980
	ccgcctggct	gaccgcccac	cgacccccgc	ccattgacgt	caataatgac	gtatgttccc	2040
	atagtaacgc	caatagggac	tttccattga	cgtcaatggg	tggagtattt	acggtaaact	2100
	gcccacttgg	cagtacatca	agtgtatcat	atgccaagtc	cgccccctat	tgacgtcaat	2160
	gacggtaaat	ggcccgcctg	gcattatgcc	cagtacatga	ccttacggga	ctttcctact	2220
	tggcagtaca	tctacgtatt	agtcacgcgt	attaccatgc	tgatgcggtt	ttggcagtac	2280
	accaatgggc	gtggatagcg	gtttgactca	cggggatttc	caagtctcca	ccccattgac	2340
	gtcaatggga	gtttgttttg	gcacaaaat	caacgggact	ttccaaaatg	tcgtaataac	2400
	cccgcctcgt	tgacgcaaat	gggcggtagg	cgtgtacggt	gggaggtcta	tataagcaga	2460
[0019]	gctcgtttag	tgaaccgtca	gatcactaga	agctttattg	cggtagttta	tcacagttaa	2520
	attgctaacg	cagtcaggcc	aacatgggct	ctagcctgga	cgacgagcac	atcctgagcg	2580
	ccttctgca	gagcgacgac	gaactggtgg	gcgaggacag	cgacagcgag	gtcagcgacc	2640
	acgtgtccga	ggacgacgtg	cagtccgaca	ccgaggaagc	cttcatcgac	gaggtgcacg	2700
	aagtgcagcc	taccagcagc	ggctccgaga	tcctggacga	gcagaacgtg	atcgagcagc	2760
	ctggcagctc	cctggccagc	aacagaaatc	tgacctgcc	ccagagaacc	atcagaggca	2820
	agaacaagca	ctgctgggtc	acctccaaga	gcaccaggcg	gagcagagtg	tccgccctga	2880
	acatcgtcg	gagccagagg	ggccccacca	gaatgtgcag	aaacatctac	gacccccctgc	2940
	tgtgtctcaa	gctgttcttc	accgacgaga	tcacagcgga	gatcgtgaag	tggaccaacg	3000
	ccgagatcag	cctgaagagg	cgggagagca	tgaccagcgc	caccttcaga	gacaccaacg	3060
	aggacgagat	ctacgccttc	ttcggcatcc	tgggtgatgac	cgccgtgaga	aaggacaacc	3120
	acatgagcac	cgacgacctg	ttcgacagat	ccctgagcat	ggtgtacgtg	tccgtgatga	3180
	gcagagacag	attcgacttc	ctgatcagat	gcctgagaat	ggacgacaag	agcatcagac	3240
	ccacctgcg	ggagaacgac	gtgttcaccc	ccgtgcggaa	gatctgggac	ctgttcaccc	3300
	accagtgcac	ccagaactac	accctggcgc	cccacctgac	catcgatgag	cagctgctgg	3360
	gcttcagagg	cagatgcccc	ttcagagtgt	acatcccca	caagcccagc	aagtacggca	3420
	tcaagatcct	gatgatgtgc	gacagcggca	ccaagtacat	gatcaacggc	atgccctacc	3480
	tgggcagagg	caccagaca	aacggcgtgc	ccctgggcga	gtactacgtg	aaagaactga	3540
	gcaagcctgt	gcatggcagc	tgcaggaaca	tcacatgcga	caactggttc	accagcatcc	3600
	ccctggccaa	gaacctcctg	caggaaccct	acaagctgac	catcgtgggc	accgtgcgga	3660
	gcaacaagcg	ggagatccca	gaggtgctga	agaacagcag	atccagacct	gtgggaacaa	3720
	gcatgttctg	cttcgacggc	cccctgacct	tgggtgtccta	caagcccaag	cccgccaaaga	3780
	tgggtgtacct	cctgtccagc	tgcgacgagg	acgccagcat	caacgagagc	accggcaagc	3840
	cccagatgggt	gatgtactac	aaccagacca	agggcgggct	ggacaccttg	gaccagatgt	3900
	gcagcgtgat	gacatgcagc	agaaagacca	acagatggcc	tatggccctg	ctgtacggca	3960
	tgatcaatat	cgctgcac	aacagcttca	tcactacag	ccacaacgtg	tccagcaagg	4020

[0020]	gcgagaaggt	gcagagccgg	aagaaattca	tgcggaacct	gtacatgagc	ctgacctcca	4080
	gcttcatgag	aaagagactg	gaagccccc	ccctgaagag	atacctccgg	gacaacatca	4140
	gcaacatcct	gcccgaaggaa	gtgccaggaa	caagcgacga	cagcaccgag	gaacccgtga	4200
	tgaagaagag	gacctactgc	acctactgtc	ccagcaagat	cagaagaaag	gccaacgcca	4260
	gctgcaagaa	atgcaaaaaa	gtgatctgcc	gggagcacaa	catcgacatg	tgccagagct	4320
	gtttctgatt	cagacatgat	aagatacatt	gatgagtttg	gacaaaccac	aactagaatg	4380
	cagtgaaaaa	aatgctttat	ttgtgaaatt	tgtgatgcta	ttgctttatt	tgttaaccatt	4440
	ataagctgca	ataaacaagt	taacaacaac	aattgcatc	attttatggt	tcaggttcag	4500
	ggggagatgt	gggaggtttt	ttaaagcaag	taaaacctct	acaaatgtgg	taagcaggtt	4560
	taaccttaga	aagatagtct	gcgtaaaatt	gacgcatgca	ttcttgaaat	attgctctct	4620
	ctttctaaat	agcgcgaatc	cgtcgctgtg	catttaggac	atctcagtcg	ccgcttgag	4680
	ctcccgtag	gcgtgcttgt	caatgcggtg	agtgtcactg	attttgaact	ataacgaccg	4740
	cgtgagtcaa	aatgacgc	gattatcttt	tacgtgactt	ttaagattta	actcatacga	4800
	taattatatt	gttatttcat	gttctactta	cgtgataact	tattatata	atattttctt	4860
	gttatagata	tctttcttta	tgttttaaat	gcactgacct	cccacattcc	ctttttagta	4920
	aaatattcag	aaataattta	aatacatcat	tgcaatgaaa	ataaatgttt	tttattaggc	4980
	agaatccaga	tgctcaaggc	ccttcataat	atccccaggt	ttagtagttg	gacttaggga	5040
	acaaaggaac	ctttaataga	aattggacag	caagaaagcg	agtcagaaga	actcgtcaag	5100
	aaggcgatag	aaggcgatgc	gctgcgaatc	gggagcggcg	ataccgtaaa	gcacgaggaa	5160
	gcggtcagcc	cattcgccgc	caagttcttc	agcaatatca	cgggtagcca	acgctatgtc	5220
	ctgatagcgg	tccgccacac	ccagccggcc	acagtcgatg	aatccagaaa	agcggccatt	5280
	ttccaccatg	atattcgcca	agcaggcatc	gccatgggtc	acgacgagat	cctcgccgtc	5340
	gggcatgcgc	gccttgagcc	tggcgaacag	ttcggtggc	gcgagccctt	gatgttcttc	5400
	gtccagatca	tcctgatcga	caagaccggc	ttccatccga	gtacgtgtc	gtcgtatgcg	5460
	atgtttcgct	tgggtggtcga	atggacaggt	agccggatca	agcgtatgca	gccgccgat	5520
	tgcatcagcc	atgatggata	ctttctcggc	aggagcaagg	tgagatgaca	ggagatcctg	5580
	ccccggcact	tcgcccaata	gcagccagtc	ccttcccgt	tcagtacaa	cgtcgagcac	5640
	agctgcgcaa	ggaacgccc	tcgtggccag	ccacgatagc	cgcgtgcct	cgtcctgcag	5700
	ttcattcagg	gcaccggaca	ggtcggtctt	gacaaaaaga	accgggcgcc	cctgcgtga	5760
	cagccggaac	acggcggcat	cagagcagcc	gattgtctgt	tgtgcccagt	catagccgaa	5820
	tagcctctcc	acccaagcgg	ccggagagcc	tgcgtgcaat	ccatcttgtt	caatcatggt	5880
	ggcggacgaa	aggcccggag	atgaggaaga	ggagaacagc	gcggcagacg	tgcgtttttg	5940
	aagcgtgcag	aatgccgggc	ctccggagga	ccttcgggcg	cccgcgccgc	ccctgagccc	6000
	gcccctgagc	ccgcccccg	acccaccct	tcccagcctc	tgagcccaga	aagcgaagga	6060
	gcaaagctgc	tattggccgc	tgccccaaag	gcctaccgc	ttccattgct	cagcgggtgt	6120
	gtccatctgc	acgagactag	tgagtcgtgc	tacttccatt	tgtcacgtcc	tgacgacgc	6180
	gagctgcggg	gcggggggga	acttccctgac	taggggagga	gtagaagggtg	gcgcgaaggg	6240
	gccaccaaag	aacggagccg	gttggcgcct	accggtggat	gtggaatgtg	tgcgaggcca	6300
	gaggccactt	gtgtagcgcc	aagtgtcccag	cggggctgct	aaagcgcgtg	ctccagactg	6360
	ccttgggaaa	agcgcctccc	ctaccggta	gagaaacttg	atctgtcgcc	gcaattcaag	6420
	cttcgtgagg	ctccgggtgc	cgtcagtgac	ctgctatact	ctggagacgg	cacatcgccc	6480
	acagtccccg	agaagtggg	aggggtcggc	aattgaaccg	gtgcctagag	aaggtggcgc	6540
	ggggtaaaact	gggaaagtga	tgtcgtgtac	tggctccgcc	tttttcccga	gggtggggga	6600
	gaaccgtata	taagtgcagt	agtcgccgtg	aacgttcttt	ttcgcaacgg	gtttgccgcc	6660
	agaacacagg	taagtccgt	gtgtggttcc	cgcgggcctg	gcctctttac	gggttatggc	6720
	ccttgcgtgc	cttgaattac	ttccacctgg	ctccagtacg	tgattcttga	tcccagactg	6780
	gagccagggg	cgggccttgc	gcttttaggag	ccccttcgcc	tcgtgcttga	gttgaggcct	6840
	ggcctgggcg	ctggggccgc	cgcgtgcgaa	tctggtggca	ccttcgcgcc	tgtctcgtg	6900
	ctttcgataa	gtctctagcc	attttaaatt	tttgatgacg	tgtcgcgacg	cttttttct	6960
	ggcaagatag	tcttgtaaat	gcgggccagg	atctgcacac	tggtatttcg	gtttttgggc	7020
	ccgcggccgg	cgacggggcc	cgtgcgtccc	agcgcacatg	ttcggcgagg	cggggcctgc	7080

	gagcgcggcc accgagaatc ggacgggggt agtctcaagc tggccggcct gctctggtgc	7140
	ctggcctcgc gccgccgtgt atcgccccgc cctgggcggc aaggctggcc cggtcggcac	7200
	cagttgcgtg agcggaaaga tggccgcttc ccggccctgc tccagggggc tcaaatgga	7260
	ggacgcggcg ctcgggagag cgggcgggtg agtcaccac acaaaggaaa agggcctttc	7320
	cgtcctcagc cgtcgttca tgtactcca cggagtaccg ggcgcctcc aggcacctcg	7380
	attagtcttg gagcttttgg agtacgtcgt ctttaggttg gggggagggg ttttatgcga	7440
	tggagtttcc ccacactgag tgggtggaga ctgaagttag gccagcttgg cacttgatgt	7500
	aattctcctt ggaatttggc ctttttgagt ttggatcttg gtccattctc aagcctcaga	7560
	cagtgtttca aagttttttt ctccatttc aggtgtcgtg aacacgtctc ggggggcccgc	7620
	cacctatgaa caacaccatc atcaactccc tgatcggcgg cgacgactcc atcaagcggc	7680
	ccaactgttt cgccgtggac tcccagatcc ccacctgta catgcccag tacatctccc	7740
	tgtccggcgt gatgaccaac gacggccctg acaaccaggc tatcgctcc ttcgagatcc	7800
	gggatcagta catcaccgcc ctgaaccacc tgggtcgtgc cctggaactg cccgaagtga	7860
	agggcattgg cagattcggc tacgtgccct acgtgggcta caagtgcac aaccacgtgt	7920
	ccatctccag ctgcaacggc gtgatctggg aaatcgaggg cgaggaaactg tacaacaact	7980
	gcattaacaa cacaatcgcc ctgaagcact ccggctactc ctccgagctg aacgacatct	8040
	ccatcggcct gaccccaac gacaccatca aagaaccctc caccgtgtac gtgtacatca	8100
	agacccctt cgacgtggaa gataccttct ccagcctgaa gctgtccgac tccaagatca	8160
	ccgtgaccgt gacctcaac cctgtgtccg acatcgtgat ccgggactcc agcttcgact	8220
	tcgagacatt caacaaagaa tttgtgtacg tgcccgagct gtccttcac ggctacatgg	8280
	tcaagaacgt gcagatcaag ccagcttca tcgagaagcc tcggagagtg atcggccaga	8340
	tcaaccagcc taccgccacc gtgacagagg tgcacgccgc cacatccctg agcgtgtaca	8400
	ccaagcccta ctacggcaac accgacaaca agttcatctc ctaccccgcc tacagccagg	8460
	acgagaagga ctacatcgac gcctacgtgt ccggctgct ggacgacctc gtgatcgtgt	8520
	ctgatggccc ccctaccggc taccctgagt ctgccgagat cgtggaagtg cccgaggacg	8580
[0021]	gcatcgtcag catccaggac gccgatgtgt atgtgaagat cgacaacgtg ccagacaaca	8640
	tgtccgtgta cctgcacacc aacctgctga tgttcggcac ccggaagaat tcttcatct	8700
	acaacatctc caagaagtte tccgccatca ccggcaccta ctccgacgcc accaagcgga	8760
	ccatcttcgc ccacatctcc cacagcatca acatcatcga caccctccac cccgtgtccc	8820
	tgtggacctc tcagagaaac gtgtacaacg gcgacaacag atccgccgag tccaaggcca	8880
	aggacctgtt catcaacgac cccttcatca agggcatcga ctcaagaac aagaccgaca	8940
	tcatctccc gctggaagtg cgcttcggca acgacgtgt gtactccgag aacggcccta	9000
	tcagccgat ctacaacgag ctgctgacca agtccaacaa cggcaccaga accctgacct	9060
	ttaactttac ccccaagatc ttcttcggc ccaccacat caccgctaac gtgtccagag	9120
	gcaaggacaa gctgagcgtg cgggtggtgt actccacat ggacgtgaac caccctatct	9180
	actacgtgca gaaacagctg gtggtcgtgt gcaacgatct gtacaagggt tcttacgacc	9240
	agggcgtgtc cattaccaag atcatggcg ataacaacga ctacaaggac gacgacgaca	9300
	agtgaaaaag cgtcttctg tttctcatc atcatatcaa ggttatatac catcaatatt	9360
	gccacagatg ttacttagcc ttttaattt tctctaattt agtgtatatg caatgatagt	9420
	tctctgattt ctgagattga gtttctcatg tgtaattgatt atttagagtt tctcttcat	9480
	ctgttcaaat ttttctctag ttttatttt tactgatttg taagacttct ttttataatc	9540
	tgcataattac aattctcttt actgggtgt tgcaaatatt tctgtcatt ctatggcctg	9600
	actttcttta atggtttttt aattttaaaa ataagtctta atattcatgc aatctaatta	9660
	acaatctttt ctttgtggtt aggactttga gtcataagaa atttttctct aactgaagt	9720
	catgatggca tgcttctata ttattttcta aaagatttaa agttttgcct tctccattta	9780
	gacttataat tcaactggaat ttttttgtgt gtatggtatg acatatgggt tcccttttat	9840
	ttttacata taaatatatt tccctgtttt tctaaaaaag aaaaagatca tcatttccc	9900
	attgtaaaat gccatatttt tttcataggt cacttacata tatcaatggg tctgtttctg	9960
	agctctactc tattttatca gctcactgt ctatccccac acatctcatg ctttgcctta	10020
	aatcttgata tttagtggaa cattctttcc cattttgttc tacaagaata ttttgttat	10080
	tgtctttggg ctttctatat acattttgaa atgaggttga caagttaata atcaacctct	10140

	ggattacaaa atttgtgaaa gattgactgg tattcttaac tatgttgctc cttttacgct	10200
	atgtggatac gctgctttta tgcctttgta tcatgctatt gcttcccgtg ttgctttcat	10260
	tttctcctcc ttgtataaat cctggttgct gtctctttat gaggagtgtg ggcccgttgt	10320
	caggcaacgt ggcggtggtg gcactgtgtt tgcigacgca acccccactg gttggggcat	10380
	tgccaccacc tgcagctcc ttccgggac ttctgcttc cccctcccta ttgccacggc	10440
	ggaactcacc gccgcctgcc ttgcccgtg ctggacaggg gctcggtgtg tgggcaactga	10500
	caattccgtg gtgttgctgg ggaatcacc gtcccttcct tggtgctcg cctgtgttgc	10560
	cacctggatt ctgcgcggga cgtccttctg ctacgtccct tcggccctca atccagcgga	10620
	ccttccttcc cgcggcctgc tgcggctct ggcgcctctt ccgcctcttc gccttcgccc	10680
	tcagacgagt cggatctccc ttggggccgc ctccccgcat cgcctgctat tgtcttgcca	10740
	atctccccc ttgctgtcct gccccacccc acccccaga atagaatgac acctactcag	10800
[0022]	acaatgcgat gcaatttcct cattttatta ggaaaggaca gtgggagtgg caccttccag	10860
	ggtcaaggaa ggcacggggg aggggcaaac aacagatggc tggcaactag aaggcacatt	10920
	tgttacttta tagaagaaat tttagtitt ttgttttttt taataaataa ataaacataa	10980
	ataaattgtt tgttgaattt attattagta tgtaagtgt aatataataa aacttaatat	11040
	ctattcaaat taataaataa acctgatat acagaccgat aaaacacatg cgtcaatttt	11100
	acgcatgatt atctttaacg tacgtcacia tatgattatc ttctagggt taagaagact	11160
	g	11161
\\davies.local\meldfs\redirected\dx\Desktop\AAA Sementis PCT\Final Sequence		
Listing.doc-3/11/2014		