



(12) 发明专利申请

(10) 申请公布号 CN 102239249 A

(43) 申请公布日 2011. 11. 09

(21) 申请号 200980148013. 0

(22) 申请日 2009. 10. 23

(30) 优先权数据

61/108, 362 2008. 10. 24 US

(85) PCT申请进入国家阶段日

2011. 05. 24

(86) PCT申请的申请数据

PCT/US2009/061935 2009. 10. 23

(87) PCT申请的公布数据

W02010/048567 EN 2010. 04. 29

(71) 申请人 威斯康星校友研究基金会

地址 美国威斯康星州

(72) 发明人 J·托马森 俞君英

(74) 专利代理机构 上海专利商标事务所有限公

司 31100

代理人 韦东

(51) Int. Cl.

C12N 5/074 (2006. 01)

权利要求书 2 页 说明书 8 页

序列表 12 页 附图 5 页

(54) 发明名称

通过非病毒重编程获得的多潜能干细胞

(57) 摘要

本发明公开了使用不编码感染性病毒的附加型载体来重编程灵长类体细胞以具有多潜能性的方法。本发明还公开了在所述方法中产生的多潜能细胞。

图1A

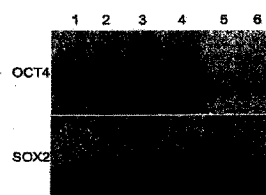
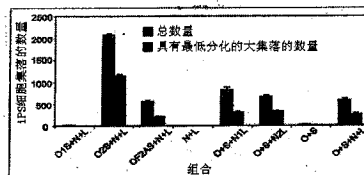


图1B



1. 一种重编程灵长类体细胞的方法,所述方法包括以下步骤:
在足以重编程所述细胞的条件下将所述灵长类体细胞暴露于多种潜能决定因子;和
培养所述暴露细胞来获得具有比所述灵长类体细胞更高的潜能水平的重编程细胞,所述重编程细胞基本不含与将所述潜能决定因子暴露于所述体细胞有关的任何成分。
2. 根据权利要求1所述的方法,其中所述灵长类体细胞是从胎儿或出生后个体中获得的。
3. 根据权利要求1所述的方法,其中所述灵长类是人类。
4. 根据权利要求1所述的方法,其中所述重编程细胞在遗传上与胎儿或出生后个体相同。
5. 根据权利要求1所述的方法,其中所述暴露步骤包括引入至少一种将一种或多种潜能决定因子编码到所述灵长类体细胞中的载体的步骤。
6. 根据权利要求5所述的方法,其中所述载体是非病毒附加型载体。
7. 根据权利要求5所述的方法,其中所述一种或多种潜能决定因子包括人类OCT-4、人类SOX2、人类LIN28、人类NANOG、人类c-Myc和人类KLF4,其中当所述载体包含c-Myc时,所述方法进一步包括将所述灵长类体细胞暴露于SV40T抗原的步骤。
8. 根据权利要求5所述的方法,其中所述至少一种载体是单一载体。
9. 根据权利要求5所述的方法,其中所述单一载体包含至少一个内部核糖体进入位点。
10. 根据权利要求8所述的方法,其中所述单一载体依次包含第一启动子、OCT4、IRES2、SOX2、第二启动子、SV40T抗原、IRES2和KLF4。
11. 根据权利要求8所述的方法,其中所述载体依次包含第一启动子、OCT4、IRES2、SOX2、第二启动子、c-Myc、IRES2、KLF4、第三启动子、NANOG、IRES2和LN28。
12. 根据权利要求5所述的方法,其中第一载体依次包含第一启动子、OCT4、IRES2、SOX2、第二启动子、SV40T抗原、IRES2和KLF4,其中第二载体依次包含第三启动子、OCT4、IRES2、SOX2、第四启动子、NANOG、IRES2和KLF4,并且其中第三载体依次包含第五启动子、c-Myc、IRES2和LN28,其中所述启动子不必相同。
13. 根据权利要求12所述的方法,其中所述第一、第二、第三和第四启动子中的每一个均是延伸因子1 α (EF1 α) 基因启动子。
14. 根据权利要求5所述的方法,其中第一载体依次包含第一启动子、OCT4、IRES2、SOX2、第二启动子、KLF4、IRES2、c-Myc、第三启动子、NANOG、IRES2和LN28。其中第二载体依次包含第四启动子、OCT4、IRES2、SOX2、第五启动子、SV40T抗原、IRES2和KLF4,其中所述启动子不必相同。
15. 根据权利要求14所述的方法,其中所述第一、第三、第四和第五启动子中的每一个均是EF1 α 基因启动子,并且其中所述第二启动子是巨细胞病毒即刻早期基因(CMV) 启动子。
16. 根据权利要求5所述的方法,其中第一载体依次包含第一启动子、OCT4、IRES2、SOX2、第二启动子、NANOG、IRES2和LN28,其中第二载体依次包含第三启动子、OCT4、IRES2、SOX2、第四启动子、SV40T抗原、IRES2和KLF4,并且其中第三载体依次包含第五启动子、OCT4、IRES2、SOX2、第六启动子、c-Myc、IRES2和KLF4,其中所述启动子不必相同。
17. 一种富集的灵长类多潜能细胞群体,其根据包括以下步骤的方法来产生:

在足以表达所述潜能决定因子的条件下将灵长类体细胞暴露于多种潜能决定因子,从而重编程所述体细胞来产生所述灵长类多潜能细胞;

所述细胞基本不含与将所述潜能决定因子暴露于所述体细胞有关的任何成分。

18. 根据权利要求 17 所述的群体,其中所述灵长类体细胞是从胎儿或出生后个体中获得的。

19. 根据权利要求 17 所述的群体,其中所述灵长类是人类。

20. 根据权利要求 17 所述的群体,其中所述多潜能细胞在遗传上与出生后个体基本相同。

21. 根据权利要求 17 所述的群体,其中所述暴露步骤包括以下步骤:引入至少一种将一种或多种潜能决定因子编码到所述灵长类体细胞中的载体。

22. 根据权利要求 21 所述的群体,其中所述载体是非病毒附加型载体。

23. 根据权利要求 21 所述的群体,其中所述一种或多种潜能决定因子包括人类 OCT-4、人类 SOX2、人类 LIN28、人类 NANOG、人类 c-Myc 和人类 KLF4,其中当所述载体包含 c-Myc 时,所述方法进一步包括将所述灵长类体细胞暴露于 SV40T 抗原的步骤。

24. 灵长类诱导的多潜能细胞,其基本不含与多潜能性的诱导有关的任何成分。

通过非病毒重编程获得的多潜能干细胞

[0001] 相关专利申请的交叉引用

[0002] 本专利申请主张 2008 年 10 月 24 日提交的美国临时专利申请 No. 61/108,362 的权益,该美国临时专利申请以引用的方式并入本文,如同对其全文进行叙述。

[0003] 关于联邦政府资助的研究或开发的声明

[0004] 本发明是在由美国国立卫生研究院 (NIH) 授予的编号为 GM081629 和 RR000167 的政府支持下进行的。美国政府具有本发明的某些权利。

背景技术

[0005] 胚胎干 (ES) 细胞因其多潜能性而具有巨大的科学和医疗前景,所述多潜能性,即,以不确定方式复制并且分化成所有三个胚层的细胞的能力。(Thomson 等人,《科学》(Science)282:1145-1147(1998),其以引用方式并入本文,如同对其全文进行叙述)。因为存在被受体免疫系统排斥的可能性,人类 ES 细胞在治疗和再生医学上的应用变得复杂。因此,非常需要对特定受体而言遗传上基本相同的人类多潜能细胞。同样,对于设计患者特异性治疗策略时 ES 细胞的使用,遗传同一性可能是十分重要的。

[0006] 从出生后灵长类个体中产生多潜能细胞的最初尝试采用的是体细胞核移植(参见,例如,Byrne, JA 等人,《自然》(Nature)450:497-502(2007))和细胞融合(参见,例如,Yu, J 等人,《干细胞》(Stem Cells)24:168-176(2006))。然而,体细胞核移植的临床使用由于效率低而不实用,同时细胞融合导致近四倍体细胞。2007 年,两组科学家将来自出生后灵长类个体的体细胞重编程为多潜能干细胞(Yu 等人,《科学》(Science)318:1917-1920(2007),以及Takahashi 等人,《细胞》(Cell)131:861-872(2007)),其各自以引用的方式并入本文,如同对其全文进行叙述。通过用于表达潜能决定转基因的病毒载体系统,两个组均将四个转录因子的 cDNA 导入人类体细胞并在其中表达。Takahashi 等人的转录因子是 OCT4、SOX2、c-Myc 和 KLF4,而 Yu 等人则采用 OCT4、SOX2、NANOG 和 LIN28。这几组转录因子的表达诱导人类体细胞获取 ES 细胞特异性特征,包括形态、增殖以及基因和表面标记物表达。以此方式重编程的体细胞被称为诱导多潜能(iPS)细胞。iPS 细胞的存在消除了对胚胎的需要,并且减少了有关免疫排斥的问题。

[0007] 此后不久,Lowry 等人通过转基因 OCT4、SOX2、c-Myc 和 KLF4 的异位表达产生出患者特异性 iPS 细胞系(lowry 等人,《美国科学院院刊》(PNAS)105:2883-2888(2008))。最近,从若干不同的人类和鼠类体细胞类型产生了 iPS 细胞,诸如上皮细胞、成纤维细胞、肝脏细胞、胃细胞、神经细胞和胰细胞。此外,已成功将 iPS 细胞分化成各种谱系的细胞(例如,Dimos 等人,《科学》(Science)321:1218-1221(2008))。

[0008] 当前用于产生 iPS 细胞的方法采用的是逆转录病毒载体,诸如来源于慢病毒的那些逆转录病毒载体。这些载体在几乎任何染色体位点稳定地整合到靶细胞 DNA 中,并且永久改变该靶细胞 DNA。重编程载体与基因组之间的这种非靶向相互作用伴随异常细胞基因表达以及病毒基因再活化导致的肿瘤生长的风险。(Okita 等人,《自然》(Nature)448:313-317(2007))。

[0009] 此外,转基因的继续存在和表达可能会干扰受体细胞的生理活动。此外,用于重编程体细胞的转录因子(诸如 c-Myc)的异位表达可能诱发编程性细胞死亡(细胞凋亡)(Askew 等人,《致癌基因》(Oncogene)6:1915-1922(1991);Evan 等人,《细胞》(Cell)69:119-128(1992))。此外,诸如 OCT4 之类因子的继续表达可能会干扰后续的 iPS 细胞分化。

[0010] 希望将体细胞重编程到较高潜能的状态,而不使细胞基因构成的改变超出重编程相关改变。最近,Stadtfield 等人使用瞬时表达 OCT4、SOX2、KLF4 和 c-Myc 的非整合腺病毒产生了鼠类 iPS 细胞(Stadtfield 等人,《科学快递》(Sciencexpress),2008 年 9 月 25 日)。目前为止,尚未报道未使用逆转录病毒载体产生的灵长类 iPS 细胞。

发明内容

[0011] 本发明宽泛的概述为涉及重编程分化的灵长类体细胞来产生灵长类多潜能细胞。

[0012] 在第一方面,本发明概述为一种用于产生灵长类多潜能细胞的方法,该方法包括以下步骤:将足以把体细胞重编程为多潜能状态的一组转基因导入灵长类体细胞内并回收多潜能细胞,所述转基因承载于不编码感染性病毒的至少一个附加型载体上。本文所提及的“非病毒”载体或构建体表示该载体或构建体不能编码感染性病毒。

[0013] 在第二方面,本发明涉及灵长类(包括人灵长类)的可补充的重编程多潜能细胞富集群体,其中与现有 iPS 细胞相反,所述多潜能细胞基本上缺乏至少一种载体(包括其具有病毒源或衍生物的任何元件)。如本文所用,这意味着重编程细胞包含每细胞少于一个拷贝的附加型载体,并且优选地在细胞中没有残余附加型载体。因为细胞分裂期间的不对称分割稀释了载体,所以可以容易地获得丢失了载体的重编程细胞。如本文它处所述,在罕见情况下,重编程载体可能整合到细胞的基因组中,但通过针对载体缺失进行筛选,可避免具有整合载体的细胞。此外,与现有 ES 细胞相反,本发明的灵长类多潜能细胞与来自胎儿或出生后个体的体细胞在遗传上基本相同。胎儿细胞可以从例如羊水获得。并不容易从形态上(即,圆形、大核仁和少细胞质)或通过生长特性(即,倍增时间;ES 细胞的倍增时间为约 17 到 18 个小时)来区分富集群细胞与现有灵长类 ES 和 iPS 细胞。如同 iPS 细胞和 ES 细胞,重编程细胞还表达多潜能细胞特异性标记物(例如,OCT-4、SSEA-3、SSEA-4、TRA-1-60、TRA-1-81,但没有 SSEA-1)。不同于 ES 细胞,重编程细胞并非直接来源于胚胎。如本文所用,“不直接来源于胚胎”意味着产生多潜能细胞的起始细胞类型是非多潜能细胞,诸如多潜能细胞或终末分化细胞,如从胎儿或出生后个体获得的体细胞。如同 iPS 细胞,该方法中产生的多潜能细胞可以在衍生期间瞬时表达一个或多个拷贝的所选潜能决定因子。

[0014] 除非另外定义,否则本文所使用的所有科技术语具有与本发明本领域技术人员通常理解的含义相同的含义。虽然用于实施或测试本发明的适合材料和方法在下文进行了描述,但也可使用类似于或等同于本文所述材料和方法的为本领域熟知的其它材料和方法。

[0015] 通过以下结合附图给出的说明,本发明的其它目的、优点和特征将变得明显。

附图说明

[0016] 图 1A-B 示出连接在重编程方法期间导入的载体上的转基因的不同核苷酸序列对重编程效率的作用。

[0017] 图 2A-C 示出在人类新生儿包皮成纤维细胞中 c-Myc、KLF-4 和 SV40 大 T 抗原基因

表达对重编程效率的作用。

[0018] 图 3A-C 示出根据该方法用于将转基因带入体细胞的适合构建体、附加型载体介导的转基因的瞬时表达以及核转染之后载体数目对细胞存活的作用。

[0019] 图 4A-D 示出具有附加型载体介导的转基因表达的人类新生儿包皮成纤维细胞的重编程。

[0020] 图 5A-B 示出携带用于本发明重编程方法的表达盒的相关构建体。

具体实施方式

[0021] 本发明广泛涉及用于通过在非病毒载体上引入潜能决定因子而将分化灵长类体细胞重编程为基本不含在其生产时使用的载体的重编程灵长类细胞的新颖方法,所述非病毒载体在重编程期间存在但在重编程细胞内基本消失。如本文所用,“重编程”是指一种遗传过程,其将分化体细胞转化为去分化细胞,所述去分化细胞比其衍生的细胞具有更高的潜能。

[0022] 有利的是,该方法中产生的更高潜能的细胞是整倍体多潜能细胞。如本文所用,“多潜能细胞”是指表达多潜能细胞特异性标记、具有未分化细胞的细胞形态特征(即,紧密集落、高核质比和明显核仁)并且可以分化为所有三个胚层(例如,内胚层、中胚层和外胚层)的细胞群体。当引入免疫受损动物如 SCID 小鼠中时,所述多潜能细胞形成通常包含所有三个胚层的细胞或组织特征的畸胎瘤。本领域普通技术人员可以使用本领域常用技术来评定这些特征。参见,如,Thomson 等人,同上。多潜能细胞既能够在细胞培养时增殖,也能够分化为显示多潜能特性的各种谱系限制性细胞群体。多潜能细胞比多潜能体细胞具有更高潜能,通过比较,多潜能细胞分化更多,但并不是终末分化。灵长类体细胞重编程方法的多潜能产物在本文中被称为“重编程灵长类多潜能细胞”或诱导多潜能(iPS)细胞。此类细胞适用于当前考虑使用人 ES 细胞或现有 iPS 细胞的研究和治疗应用中。

[0023] 分化体细胞,包括来自胎儿、新生儿、青少年或成年灵长类(包括人类个体)的细胞,是适合本方法的起始细胞。适合的体细胞包括但不限于骨髓细胞、上皮细胞、内皮细胞、成纤维细胞、造血细胞、角化细胞、肝细胞、肠细胞、间充质细胞、脊髓前体细胞和脾细胞。另一种适合的体细胞为附着于基底的 CD29⁺CD44⁺CD166⁺CD105⁺CD73⁺ 和 CD31⁻ 间充质细胞。或者,体细胞可以是能自我增殖和分化为其它类型细胞的细胞,包括血液干细胞、肌肉/骨骼干细胞、脑干细胞和肝脏干细胞。适合的体细胞是接受性的,或可使用科学文献中通常已知的方法使其具有接受性以摄取潜能决定因子,包括编码这些因子的遗传材料。摄取增强方法可根据细胞类型和表达系统而有所不同。本领域普通技术人员了解用于制备具有适合转导率的接受性体细胞的示例性条件。起始体细胞的倍增时间可以是约 24 小时。

[0024] 可以使用科学文献中通常已知的方法来构建和工程改造本文所述载体,以增加其用于治疗的安全性、包括选择和富集标记物以及(如果希望的话)优化其上所含核苷酸序列的表达。所述载体应包含允许载体在起始体细胞中自我复制的结构成分。例如,已知的 Epstein Barr oriP/核抗原-1(EBNA-1)组合(参见,例如,Lindner, S. E. 和 B. Sugden, The plasmid replicon of Epstein-Barr virus: mechanistic insights into efficient, licensed, extrachromosomal replication in human cells, Plasmid 58:1 (2007), 其以引用的方式并入本文,如同对其全文进行叙述。)足以支持载体自我复制,而且也可以使用已

知在哺乳动物（尤其是灵长类）细胞中起作用的其它组合。本领域普通技术人员了解适用于本发明的构建表达载体的标准技术，该技术可见于以下出版物：诸如 Sambrook J 等人，《分子克隆：实验室手册》（第三版）（"Molecular cloning: a laboratory manual," 3rd ed. Cold Spring harbor Press, Cold Spring Harbor, N. Y. 2001），该出版物以引用的方式并入本文，如同对其全文进行叙述。

[0025] 在所述方法中，通过一种或多种重编程载体，将编码一组潜能决定因子的遗传材料导入体细胞。适合的潜能决定因子可以包括但不限于 OCT-4、SOX2、LIN28、NANOG、c-Myc、KLF4 和它们的组合。各潜能决定因子可被引入体细胞中作为编码可操作地连接到异源启动子的潜能决定因子的多核苷酸转基因，所述异源启动子能驱动多核苷酸在体细胞内表达。虽然 SV40T 抗原并不是潜能决定因子本身，但其被有利地引入体细胞，因为其向细胞提供足以在重编程期间促进细胞存活的条件，同时潜能决定因子被表达。表达所述因子的其它充分条件包括在实例中所述的细胞培养条件。

[0026] 适合的重编程载体是附加型载体，诸如质粒，所述附加型载体不会编码足以产生感染性或可复制病毒的所有或部分病毒基因组，但所述载体可以包含获自一种或多种病毒的结构元件。可将一种或多种重编程载体引入单一体细胞。可以在单一重编程载体上提供一种或多种转基因。一种强力的、构成性转录启动子可以向多种转基因提供转录控制，所述转基因可以作为表达盒来提供。载体上的独立表达盒可以受独立强力的、构成性启动子的转录控制，所述启动子可能是多个拷贝的相同启动子，也可能是不同的启动子。各种异源启动子是本领域已知的，并且可根据诸如潜能决定因子的希望表达水平之类的因素来使用。如下文所例示，可能有利的是，在靶标体细胞中使用具有不同浓度的不同启动子来控制独立表达盒的转录。对转录启动子选择的另一考虑为启动子在靶标体细胞内保持沉默的比率。熟练的技术人员将理解，有利的是，在基因产物基本完成了其在重编程方法中的作用之后，减少一种或多种转基因或转基因表达盒的表达。示例性启动子是人类 EF1 α 延伸因子启动子、CMV 巨细胞病毒即刻早期启动子和 CAG 鸡白蛋白启动子、以及其它种类的相应同源启动子。在人类体细胞中，EF1 α 和 CMV 是强启动子，但 CMV 启动子比 EF1 α 启动子能更有效地沉默，以使得在前者控制下的转基因表达比在后者控制下的转基因表达能更快终止。

[0027] 潜能决定因子能够以可被改变以调整重编程效率的相对比例来在体细胞内表达。例如，与在独立载体上提供两种因子时相比，当以 1 : 1 比例在单一载体上的单一转录物内编码 OCT-4 和 SOX2 时，体细胞重编程效率高出四倍，使得进入单一细胞的因子的摄取比例不受控制。优选地，当在单一转录物上编码多种转基因的情况下，在转录启动子远端的转基因的上游提供内部核糖体进入位点。虽然因子的相对比例可根据所导入因子而改变，但本发明所属领域的普通技术人员可以决定因子的最佳比例。

[0028] 熟练的技术人员将理解，通过单一载体而非通过多种载体引入所有因子对效率更有利，但当总的载体大小增加时，则变得越来越难以引入载体。熟练的技术人员还将了解，因子在载体上的位置能影响其瞬时表达，以及所得的重编程效率。这样，申请人在多种载体组合上使用了各种因子组合。若干此类组合在此表明支持重编程。

[0029] 在引入重编程载体之后并当体细胞正在被重编程时，载体能存留在靶细胞内，而引入的转基因则被转录和翻译。在已重编程为多潜能状态的细胞内，可以有利地下调或终止转基因表达。重编程载体可以残留于染色体外。载体可以极低效率整合到细胞的基因组

中。重编程载体与受体细胞的基因组协同地复制,因而,可适度地稳定约两周,这比不能复制其 DNA 的附加型载体稳定时间更长。然而,因为载体在细胞分裂时未等分,在缺乏选择性压力时,细胞丢失附加型载体,所以技术人员可以容易地用该方法回收无载体的多潜能细胞。例如,oriP/EBNA-1- 基附加型质粒通常要用两到三周以稳定维持于体细胞内。在最初两到三周内,细胞快速丢失附加型质粒。一旦细胞稳定,则细胞以每代 5% 的速度继续丢失附加型载体。

[0030] 该方法中产生的多潜能细胞可以在任何支持多潜能细胞生长的培养基中培养,包括但不限于:确定成分培养基,诸如 TeSR™(StemCell Technologies, Inc.; Vancouver, Canada)、mTeSR™(StemCell Technologies, Inc.) 和 StemLine® 无血清培养基 (Sigma; St. Louis, Mo.); 或条件培养基,诸如小鼠胚胎成纤维细胞 (MEF) 条件培养基。如本文所用,“确定成分培养基”是指仅含生化明确组分的生化明确的制剂,可以包括已知化学组成的组分或来源于已知来源的组分。如本文所用,“条件培养基”是指还补充有来自培养基中培养细胞的可溶性因子的生长培养基。或者,细胞可维持于培养基中的 MEF 上。

[0031] 在考虑以下非限制性实例之后,将更完全地理解本发明。

[0032] 实例 1

[0033] 表达盒的设计和构建

[0034] 使用传统方法,通过直接聚合酶链式反应 (PCR) 扩增来自转基因中的一些或所有的可读框 (ORF) (使用编码区的最初和最后 20-22 个碱基作为引物),以及 PCR 扩增来自表 1 中所列内部核糖体进入位点的 ORF,产生了适合的表达盒结构。SV40T 抗原和人类端粒酶逆转录酶、质粒 pBABE-puro SV40LT 和 pBABE-hygro-hTERT 的来源是分别以质粒 13970 和 1773 商购自 Addgene, Inc (Cambridge, MA)。IRES1 和 IRES2 的来源、质粒 pIRESpuro3 和 pIRES2EGFP 的来源是商购自 Clontech Laboratories, Inc. (Mountain View, CA)。以化学方式合成了口蹄疫病毒片段 2。使用下文在表 1 中所述密码描述了框内表达盒。例如,“E-02S”是指一种具有位于 OCT4 和 SOX2 编码区上游且其间有 IRES2 的 EF1 α 启动子的表达盒。同样地,“C-M2K”是指一种具有位于 c-Myc 和 Klf4 编码区上游且其间有 IRES2 的 CMV 启动子的表达盒。在若干构建体内,没有一种用于后续的重编程,而使用了变体 02S 表达盒 (“02S(2)”),该 02S(2) 不同于 02S,因为其包含 TK 启动子 -Hyg-TK 聚腺苷酸盒 (比较图 5A 与图 5B)。通过经验确定各种因子的表达水平,选择具有该所指示结构的盒以便随后用于重编程方法。表示为 EF2 的启动子 (SEQ ID NO:12) 是已知 EF1 α 启动子 (SEQ ID NO:11) 的轻微变体,EF2 在活性上不同于 EF1 α ,并且未在后续附加型载体重编程试验 (如下) 中使用。F2A 是促进共翻译表达自单一转录物的不同编码区的肽接头。测试了 F2A,但并未将其用于后续的使用附加型载体的重编程试验中。测试了 IRES1,但并未将其用于后续的使用附加型载体的重编程试验中。

[0035] 通过分别将各种转基因表达盒克隆到 pSin4 (基于改性慢病毒的载体) 中来评估各种启动子、IRES 测序和转基因排列对上游和下游 ORF 的表达的相对作用,从而测试其在转染之后重编程人类体细胞的能力,如前文所述 (Yu 等人,上文)。使用 SuperFect (Qiagen, Valencia, CA),使 293FT 细胞转染了表达由 IRES1 或 IRES2 连接的 OCT4 和 SOX2 的慢病毒质粒载体,如下文所述。转染后两天收集细胞。图 1A 示出 293FT 细胞中 OCT-4 和 SOX2 的 Western 印迹分析。泳道 1,

[0036] pSin4-EF2-Oct4-IRES1-Sox2 ;泳道 2,

[0037] pSIN4-EF2-OCT4-IRES2-SOX2 ;泳道 3,

[0038] pSIN4-EF2-OCT4-F2A-SOX2 ;泳道 4,

[0039] pSIN4-EF2-OCT4-IRES1-PUR0 ;泳道 5,

[0040] pSIN4-EF2-SOX2-IRES1-PUR0 ;泳道 6, 无质粒 (对照)。将小鼠抗人 OCT4 单克隆抗体 (1 : 500, Santa Cruz Biotechnology, Inc., Santa Cruz, CA, sc-5279) 和山羊抗人 SOX2 多克隆抗体 (1 : 500, R&D Systems, Minneapolis, MN AF2018) 分别用于检测 OCT4 和 SOX2 的相对表达。

[0041] 图 1B 示出使用连接的潜能决定因子使来源于 OCT4 敲入人 ES 细胞的 0.2×10^6 间充质细胞重编程 (美国专利申请 No. 2006/0128018 和 Zwaka 和 Thomson, 《自然生物技术》(Nature Biotechnology) 21 : 319-321 (2003), 各自以引用的方式并入本文, 如同对其全文进行叙述)。这种细胞系在新霉素筛选下得以维持 (遗传霉素 : $100 \mu\text{g/ml}$, Invitrogen Corp.)。转导 16 天后计数人类 iPS 细胞集落。这些基因组合是 pSIN4-EF2-OCT4-IRES1-SOX2 (01S) ; pSIN4-EF2-OCT4-IRES2-SOX2 (02S) ; pSIN4-EF2-OCT4-F2A-SOX2 (0F2AS) ;

[0042] pSIN4-EF2-NANOG-IRES1-LIN28 (N1L) ;

[0043] pSIN4-EF2-NANOG-IRES2-LIN28 (N2L) ;

[0044] pSIN4-EF2-OCT4-IRES1-PUR0 (0) ; pSIN4-EF2-SOX2-IRES1-PUR0 (S) ; pSIN4-EF2-NANOG-IRES1-PUR0 (N) ;

[0045] pSIN4-EF2-LIN28-IRES1-PUR0 (L)。用于各慢病毒质粒载体的缩写载体名称之后的括号中示出。

[0046] 实例 2

[0047] 使用慢病毒构建体重编程人类新生儿包皮成纤维细胞

[0048] 通过将慢病毒载体引入到人类初生儿包皮成纤维细胞中进行初步重编程实验。图 2A 示出当还引入了 OCT4、SOX2、LIN28 和 c-MYC 时 NANOG 对重编程效率具有重要积极作用, 以及结合 OCT4、SOX2 和 LIN28, NANOG 可以支持重编程, 即便缺少 c-MYC 或 KLF4。所用的慢病毒构建体是 pSIN4-EF2-OCT4-IRES2-SOX2 (02S) ; pSIN4-EF2-NANOG-IRES2-LIN28 (N2L) ; pSIN4-EF2-LIN28-IRES1-PUR0 (L) ; pSIN4-CMV-c-Myc-IRES1-PUR0 (M) ; pSIN4-EF2-KLF4-IRES1-PUR0 (K)。转导后 21 天, 计数碱性磷酸酶阳性的人类 iPS 细胞集落。iPS 细胞集落的数量来源于输入的 2.5×10^4 个人类新生儿包皮成纤维细胞 (第 9 代)。浅灰色条表示形成的具有典型人 ES 细胞形态的重编程集落的总数 ; 深灰色条表示具有最低分化的大集落的数量。

[0049] 图 2B 显示了使用连接的潜能决定因子的重编程。所用的慢病毒构建体是 pSIN4-EF2-c-Myc-IRES2-KLF4 (EF2-M2K) ; pSIN4-CMV-c-Myc-IRES2-KLF4 (CMV-M2K) ; pSIN4-EF2-KLF4-IRES2-c-Myc (EF2-K2M) ; pSIN4-CMV-KLF4-IRES2-c-Myc (CMV-K2M) ; pSIN4-CMV-c-Myc-IRES2-LIN28 (M2L) ; pSIN4-EF2-NANOG-IRES2-KLF4 (N2K)。转导后 14 天, 计数碱性磷酸酶阳性的人类 iPS 细胞集落。iPS 细胞集落的数量来源于输入的大约 7.0×10^4 个人类新生儿包皮成纤维细胞 (第 12 代)。星号表示大部分碱性磷酸酶阳性的集落出现形态松散。

[0050] 图 2C 示出 SV40 大 T 抗原基因表达对重编程效率的作用。SV40 大 T 抗原鼠类成纤

维细胞被 c-Myc 诱导 (Hermeking 等人, PNAS 91:10412-10416 (1994)) 并且增强重编程效率 (Hanna 等人, 《细胞》(Cell) 133:250-264 (2008); Mali 等人, 《干细胞》(Stem Cells) doi: 10.1634/stemcells.2008-0346 (2008))。基因组合的缩写与图 2B 中相同, 但增加了 SV40 大 T 抗原 (T)。c-Myc 还促进细胞增殖。转导后 12 天, 计数碱性磷酸酶阳性的人类 iPS 细胞集落。iPS 细胞集落的数量来源于输入的大约 3.5×10^4 个人类新生儿包皮成纤维细胞 (第 17 代)。图 2C 证实如果以基于慢病毒的重编程期间达到的水平存在, 则 T 抗原会抑制 iPS 细胞衍生的最后阶段。相反, 参见下文, 其中当 T 抗原以在使用附加型载体的重编程期间达到的瞬时表达时间和 / 或水平存在时, 则其不具有这种作用。另外, T 抗原防止 c-Myc 诱导的编程性细胞死亡, 但不会不利地影响 c-Myc 诱导的细胞增殖。

[0051] 实例 3

[0052] 使用非病毒附加型构建体重编程人类新生儿包皮成纤维细胞

[0053] 将人类新生儿包皮成纤维细胞 (目录号 CRL-2097™, ATCC) 维持于包皮成纤维细胞培养基内 (DMEM (目录号 11965, Invitrogen), 该培养基补充有 10% 热灭活的胎牛血清 (FBS, HyClone Laboratories, Logan, UT)、2mM Glutamax、0.1mM 非必需氨基酸和 0.1mM β -巯基乙醇)。

[0054] 使用附加型构建体 pCEP4-EGFP (如图 3A 所示), 将作为如实例 1 中和如下表 3 中详述那样构建的转基因表达盒来提供的潜能决定因子的各种组合引入体细胞, 从而产生具有不同效率的重编程。通过在 pCEP4BamHI 位点与 NheI 位点之间插入 EGFP 编码区, 由市售的哺乳动物附加型表达载体 pCEP4 (Invitrogen Corp., Carlsbad, CA) 产生 pCEP4-EGFP。通过将指定表达盒插入到 pCEP4-EGFP 中或插入到缺少 P_{CMV} 的相关骨架 (指定的 pEP4) 中来产生表 2 的附加型载体。参见图 3A 和表 2 中表达盒插入的克隆位点的脚注。

[0055] 将载体引入到成纤维细胞中是通过单一核转染活动, 使用人类真皮成纤维细胞细胞核转染试剂盒 (Human Dermal Fibroblasts Nucleofector Kit) (普通人类真皮成纤维细胞 (Normal Human Dermal Fibroblasts), Amaxa, Inc. 目录号 VPD-1001), 并且根据制造商说明书来进行的。在核转染之后, 将转染的成纤维细胞 (每次约 0.8 到 1.0×10^6 个细胞) 即刻接种到已接种了辐照小鼠胚胎成纤维细胞 (MEF) 的 10cm 培养皿上。包皮成纤维细胞培养基每隔一天更换。四天之后, 将包皮成纤维细胞培养基更换为人 ES 细胞培养基 (DMEM/F12 培养基, 补充有 20% 敲除血清替代物、0.1mM 非必需氨基酸 (均来自 Invitrogen Corp.)、1mM Glutamax、0.1mM β -巯基乙醇和 100ng/ml 斑马鱼碱性成纤维细胞生长因子 (zbFGF), 如前文所述 (Amit 等人, 《发育生物学》(Developmental Biology) 227:271-278 (2006); Ludwig 等人, 《自然方法》(Nature Methods) 3:637-646 (2006), 其中每个均以引用的方式并入本文, 如同对其全文进行叙述)。当接种的 MEF 不能再维持重编程培养物时, 即接种后约 8 到 10 天, 使用经辐照的 MEF 调节的人 ES 细胞培养基替换。在恰当时 (转染后 2-3 周), 将培养物染色为碱性磷酸酶以指示人类 iPS 集落发育。

[0056] 为决定引入转基因构建体的适合参数, 在评估了 EGFP 从 pEGFP-N2 (对照) 和 pCEP4-EGFP 附加型载体到 293FT 细胞中的引入之后, 通过测量随时间的 EGFP 水平来最初地评估瞬时表达 (图 3B)。

[0057] 还在初级实验中评估了引入的转基因构建体量对人类新生儿包皮成纤维细胞存活的作用。图 3C 示出所用 pCEP4-EGFP 附加型载体量对核转染效率和人类新生儿包皮成纤

维细胞存活的作用,这是通过核转染后当天的细胞汇合来评估的。将大约 1×10^6 核转染包皮成纤维细胞接种到 6 孔板的每个孔中。灰色线表示未转染的对照成纤维细胞;黑色线表示转染的成纤维细胞。

[0058] 图 4A 示出表 3 的包含各种表达盒的示意性转基因表达构建体,当所述表达盒以某些组合引入到人类新生儿包皮成纤维细胞中时,导致成纤维细胞重编程为多潜能细胞。引入的附加型重编程载体的三种组合已产出了重编程多潜能细胞:(1)pEP4-E-02S-E-T2K, pEP4-E-02S-E-N2K 和 pCEP4-C-M2L;(2)pEP4-E-02S-C-K2M-E-N2L 和 pEP4-E-02S-E-T2K;以及(3)pEP4-E-02S-E-N2L、pEP4-E-02S-E-T2K 和 pEP4-E-02S-E-M2K。表 3 示出用于每种成功组合的各载体量。每种成功重编程组合中的一种载体在 EF1 α 启动子的控制下编码 T 抗原。

[0059] 图 4B 示出具有在附加型载体转染后 18 天观察到的形态变化的典型集落的明视野显微镜图像。图 4C 示出附加型载体后转染 18 天碱性磷酸酶阳性的集落的明视野显微镜图像。

[0060] 在转染后 25 到 30 天,由于存在形态类似于人类 iPS 细胞集落的许多非 iPS 细胞集落,所以将重编程培养物传代一次到新鲜 10cm MEF 培养皿(1:3 比例)。随后挑选集落进一步分析。图 4D 示出在转染重编程培养物后 28 天的第一次传代之后人类 iPS 细胞集落的明视野显微镜图像。标尺表示 0.1mm。将重编程细胞维持在具有条件培养基的基质胶(Matrigel)(BD Biosciences, Bedford, MA)上的无饲养细胞培养物中进行后续分析,如前文所述(Xu 等人,《自然生物技术》(Nat. Biotechnol.) 19:971(2001),其以引用的方式并入本文,如同对其全文进行叙述)。

[0061] 有利的是,用比使用四种基因组合所用重编程时间显著更少的重编程时间,实现了新生儿包皮成纤维细胞的大于 1% 的重编程效率。

[0062] 应理解,本公开中所描述的本发明的某些改进是本领域技术人员的常规优化实践,并且可以在不脱离本发明精神或所附权利要求书范围的情况下加以实施。以上说明书中提及的所有出版物和专利以引用方式并入本文。在不脱离本发明的范围和精神的情况下,本发明的所述方法和系统的各种修改和变型将对本领域技术人员而言是明显的。然而,应理解,上文所述的本发明的实例和实施例是说明性的,而并非旨在限制本发明。本发明涵盖与下述权利要求书的范围相符的实例和实施例的所有修改形式。

[0001]

序列表

<110> 威斯康星校友研究基金会(Wisconsin Alumni Research Foundation)

<120> 非病毒重编程

<130> 960296.00915

<140> PCT/US 09/061935

<141> 2009-10-23

<160> 12

<170> PatentIn version 3.5

<210> 1

<211> 1411

<212> DNA

<213> 智人 (Homo sapiens)

<400> 1

ccttcgcaag ccctcatttc accaggcccc cggttgggg cgccttcctt ccccatggcg	60
ggacacctgg cttcgatttt cgccttctcg cccctccag gtggtggagg tgatgggcca	120
ggggggccgg agccgggctg ggttgatcct cggacctggc taagcttcca aggccctcct	180
ggagggccag gaatcgggccc gggggttggg ccaggctctg aggtgtgggg gattccccc	240
tgcccccgcg cgtatgagtt ctgtgggggg atggcgact gtgggcccc ggttggagtg	300
gggctagtgc cccaaggcgg cttggagacc tctcagcctg agggcgaagc aggagtcggg	360
gtggagagca actccgatgg ggctccccg gagccctgca ccgtcacccc tggcgccgtg	420
aagctggaga aggagaagct ggagcaaac ccggaggagt cccaggacat caaagctctg	480
cagaaagaac tcgagcaatt tgccaagctc ctgaagcaga agaggatcac cctgggatat	540
acacaggccg atgtggggct caccctgggg gttctatttg ggaaggtatt cagccaaacg	600
accatctgcc gctttgaggc tctgcagctt agcttcaaga acatgtgtaa gctgcggccc	660
ttgctgcaga agtgggtgga ggaagctgac aacaatgaaa atcttcagga gatatgaaa	720
gcagaaaccc tcgtgcaggc ccgaaagaga aagcgaacca gtatcgagaa ccgagtgaga	780
ggcaacctgg agaatttggt cctgcagtgc ccgaaacca cactgcagca gatcagccac	840
atcgcccagc agcttgggct cgagaaggat gtggtccgag tgtggttctg taaccggcgc	900
cagaagggca agcgatcaag cagcgactat gcacaacgag aggattttga ggctgctggg	960
tctcctttct cagggggacc agtgtccttt cctctggccc cagggcccca ttttgggtacc	1020
ccaggctatg ggagccctca cttcactgca ctgtactcct cggtccttt ccctgagggg	1080
gaagcctttc cccctgtctc cgtcaccact ctgggctctc ccatgcatc aaactgaggt	1140
gcctgccctt ctaggaatgg gggacagggg gaggggagga gctagggaaa gaaaacctgg	1200
agtttgtgcc aggttttttg ggattaagtt cttcattcac taaggaagga attgggaaca	1260
caaagggtgg gggcagggga gtttggggca actggttggg gggaaggtga agttcaatga	1320
tgctcttgat tttaatccca catcatgtat cacttttttc ttaaataaag aagcctggga	1380
cacagtagat agacacactt aaaaaaaaaa a	1411

[0002]

<210> 2
 <211> 2518
 <212> DNA
 <213> 智人 (Homo sapiens)

<400> 2

```

ctattaactt gttcaaaaaa gtatcaggag ttgtcaaggc agagaagaga gtgtttgcaa      60
aagggggaaa gtagtttgct gcctcttttaa gactaggact gagagaaaga agaggagaga      120
gaaagaaagg gagagaagtt tgagccccag gcttaagcct ttccaaaaaa taataataac      180
aatcatcggc ggcggcagga tcggccagag gaggaggga gcgctttttt tgatcctgat      240
tccagtttgc ctctctcttt ttttcccca aattattctt cgcctgattt tcctcgcgga      300
gccttcgct cccgacaccc ccgcccgcct cccctcctcc tctcccccg cccgcgggcc      360
ccccaaagtc ccggccgggc cgagggtcgg cggccgcccg cgggcccggc ccgcgcacag      420
cgcccgcatg tacaacatga tggagacgga gctgaagccg ccgggcccgc agcaaaactt      480
ggggggcggc ggcggaact ccaccgcggc ggcgccggc ggcaaccaga aaaacagccc      540
ggaccgcgtc aagcgcccca tgaatgcctt catggtgttg tcccgcgggc agcggcgcaa      600
gatggcccag gagaacccca agatgcacaa ctcgagatc agcaagcgcc tgggcgcgga      660
gtggaactt ttgtcggaga cggagaagcg gccgttcac gacgaggcta agcggctgcg      720
agcgtgcac atgaaggagc acccgatta taaataccgg cccgcggcga aaaccaagac      780
gctcatgaag aaggataagt acacgtgcc cggcgggctg ctggcccccg gcggcaatag      840
catggcgagc ggggtcgggg tgggcgcgg cctgggcgcg ggcgtgaacc agcgcattga      900
cagttacgcg cacatgaacg gctggagcaa cggcagctac agcatgatgc aggaccagct      960
gggctacccg cagcaccgg gcctcaatgc gcacggcgca gcgcagatgc agcccatgca      1020
ccgctacgac gtgagcgccc tgcagtacaa ctccatgacc agctcgaga cctacatgaa      1080
cggctcgccc acctacagca tgctctactc gcagcagggc acccctggca tggtcttgg      1140
ctccatgggt tcggtggtca agtccgaggc cagctccagc cccctgttg ttacctctt      1200
ctccactcc agggcgccct gccaggcgg ggacctcgg gacatgatca gcatgtatct      1260
ccccggcgcc gaggtgccgg aaccgcgcg cccagcaga cttcacatgt cccagcacta      1320
ccagagcggc ccggtgcccg gcacggccat taacggcaca ctgccctct cacacatgtg      1380
agggccggac agcgaactgg aggggggaga aattttcaaa gaaaaacgag ggaaatggga      1440
ggggtgcaaa agaggagagt aagaaacagc atggagaaaa cccggtacgc tcaaaaagaa      1500
aaaggaaaaa aaaaaatccc atcaccaca gcaaatgaca gctgcaaaag agaacaccaa      1560
tccatccac actcacgcaa aaaccgcgat gccacaaga aaacttttat gagagagatc      1620
ctggacttct ttttggggga ctattttgt acagagaaaa cctggggagg gtggggaggg      1680
cgggggaatg gacctgtat agatctggag gaaagaaagc tacgaaaaac tttttaaag      1740
ttctagtgg acggtaggag ctttgcagga agtttgcaa agtctttacc aataatattt      1800
agagctagtc tccaagcgac gaaaaaatg ttttaatat tgcaagcaac tttgtacag      1860
tatttatcga gataaacatg gcaatcaaaa tgtccattgt ttataagctg agaatttgc      1920
aatattttt aaggagaggc ttcttgctga attttgattc tgcagctgaa atttaggaca      1980
gttgcaaacg tgaaaagaag aaaattattc aaatttgac attttaattg tttaaaaatt      2040
gtacaaaagg aaaaaattag aataagtact ggcgaccat ctctgtggtc ttgtttaaaa      2100
agggcaaaag ttttagactg tactaaattt tataacttac tgttaaaagc aaaaatggcc      2160
atgcaggttg acaccgttgg taatttataa tagcttttgt tcgatcccaa ctttccattt      2220
tgttcagata aaaaaacca tgaaattact gtgttgaaa tatttctta tggttttaa      2280
tatttctgta aatttattgt gatattttaa ggtttcccc ctttatttt ccgtagttgt      2340
attttaaag attcgctct gtattatttg aatcagctc cggagaatcc atgtatatat      2400
ttgaactaat atcatcctta taacaggtac attttcaact taagttttta ctccattatg      2460

```

[0003]

cacagtttga gataaataaa tttttgaaat atggacactg aaaaaaaaaa aaaaaaaa 2518

<210> 3

<211> 2098

<212> DNA

<213> 智人 (Homo sapiens)

<400> 3

attataaatc tagagactcc aggattttaa cgttctgctg gactgagctg gttgcctcat	60
gttattatgc aggcaactca ctttatccca atttcttgat acttttcctt ctggaggctc	120
tattttctta acatcttcca gaaaagtctt aaagctgcct taaccttttt tccagtccac	180
ctcttaaatt ttttctcctt cttcctctat actaacatga gtgtggatcc agcttgtccc	240
caaagcttgc cttgctttga agcatccgac tgtaaagaat cttcacctat gcctgtgatt	300
tgtgggcctg aagaaaacta tccatccttg caaatgtctt ctgctgagat gcctcacacg	360
gagactgtct ctctcttccc ttcctccatg gatctgctta ttcaggacag ccctgattct	420
tccaccagtc ccaaaggcaa acaacccact tctgcagaga agagtgtcgc aaaaaaggaa	480
gacaaggctc cggtaagaa acagaagacc agaactgtgt tctcttcac ccagctgtgt	540
gtactcaatg atagatttca gagacagaaa tacctcagcc tccagcagat gcaagaactc	600
tccaacatcc tgaacctcag ctacaaacag gtgaagacct ggttccagaa ccagagaatg	660
aaatctaaga ggtggcagaa aaacaactgg ccgaagaata gcaatggtgt gacgcagaag	720
gcctcagcac ctacctccc cagcctttac tcttctacc accagggatg cctgggtgaac	780
ccgactggga acctccaat gtggagcaac cagacctgga acaattcaac ctggagcaac	840
cagaccaga acatccagtc ctggagcaac cactcctgga acactcagac ctggtgcacc	900
caatcctgga acaatcaggc ctggaacagt cccttctata actgtggaga ggaatctctg	960
cagtctgca tgcagttcca gccaaattct cctgccagt acttgaggc tgccttgaa	1020
gctgctgggg aaggccttaa tgtaatacag cagaccacta ggtatttttag tactccacaa	1080
accatggatt tattcctaaa ctactccatg aacatgcaac ctgaagacgt gtgaagatga	1140
gtgaaactga tattactcaa ttccagtctg gacactggct gaatcctcc tctcccctcc	1200
tcccatccct cataggattt ttcttgttg gaaaccacgt gttctggtt ccatgatgcc	1260
catccagtc atctcatgga ggggtggagta tggttggagc ctaatcagcg aggtttcttt	1320
ttttttttt ttctatttg atcttcttg agaaaatact ttttttttt tttttttga	1380
aacggagtct tgctctgtcg cccaggttg agtgcagtgg cgcggtcttg gctcactgca	1440
agctccgtct cccgggttca cgccattctc ctgcctcagc ctcccagca gctgggacta	1500
caggcggccg ccacctgcc cggctaatat tttgtatttt tagtagagac ggggtttcac	1560
tgtgttagcc aggatggtct cgatctcctg acctgtgat ccaccgcct cggcctccct	1620
aacagctggg atttacaggc gtgagccacc gcgcctgcc tagaaaagac attttaataa	1680
ccttggtgc cgtctctggc tatagataag tagatctaact actagtttg atattcttag	1740
ggtttagaat ctaacctcaa gaataagaaa tacaagtaca aattggtgat gaagatgtat	1800
tcgtattgtt tgggattggg aggccttctg ttttttttaa aaactattga ggtaaagggt	1860
taagctgtaa catacttaat tgatttctta ccgttttttg ctctgttttg ctatatcccc	1920
taatttggtg gttgtgctaa tctttgtaga aagaggtctc gtatttgctg catcgtaatg	1980
acatgagtac tgctttagtt ggtttaagtt caaatgaatg aaacaactat ttttcttta	2040
gttgatttta ccctgatttc accgagtgtt tcaatgagta aatatacagc ttaaacad	2098

<210> 4

[0004]

<211> 4014
 <212> DNA
 <213> 智人 (Homo sapiens)

<400> 4

gtgcggggga	agatgtagca	gcttcttctc	cgaaccaacc	ctttgccttc	ggacttctcc	60
ggggccagca	gccgcccgc	caggggccc	gggcccacgg	ctcagccgac	gaccatgggc	120
tccgtgtcca	accagcagtt	tgcaggtggc	tgcgccaagg	cggcagaaga	ggcgcccag	180
gaggcgccgg	aggacgcggc	ccgggcggcg	gacgagcctc	agctgctgca	cggtgcgggc	240
atctgtgaag	ggttcaacgt	gcgcatgggg	ttcggtcttc	tgtccatgac	cgcccgcgcc	300
ggggtcgcgc	tcgaccccc	agtggatgtc	tttgtgcacc	agagtaagct	gcacatggaa	360
gggttccgga	gcttgaagga	gggtgaggca	gtggagttca	cctttaagaa	gtcagccaag	420
ggtctggaat	ccatccgtgt	caccggacct	ggtggagtat	tctgtattgg	gagtgaagag	480
cggccaaaag	gaaagagcat	gcagaagcgc	agatcaaaa	gagacagggt	ctacaactgt	540
ggaggtctag	atcatcatgc	caaggaatgc	aagctgccac	cccagcccaa	gaagtgccac	600
ttctgccaga	gcacagccca	tatggtagcc	tcatgtccgc	tgaaggccca	gcagggccct	660
agtgcacagg	gaaagccaac	ctactttcga	gaggaagaag	aagaaatcca	cagccctacc	720
ctgctcccgg	aggcacagaa	ttgagccaca	atgggtgggg	gctattcttt	tgctatcagg	780
aagtttttag	gagcaggcag	agtggagaaa	gtgggaatag	ggtgcattgg	ggctagttag	840
cactgccatg	tatctcaggc	ttgggttcac	accatcaccc	tttcttcctt	ctaggtgggg	900
ggaaagggtg	agtcaaagga	actccaacca	tgctctgtcc	aaatgcaagt	gagggttctg	960
ggggcaacca	ggagggggga	atcaccttac	aacctgcata	ctttgagtct	ccatccccag	1020
aattttccagc	ttttgaaagt	ggcctggata	gggaagttgt	tttcttttta	aagaaggata	1080
tataataatt	cccattgccag	agtgaatga	ttaagtataa	gaccagattc	atggagccaa	1140
gccactacat	tctgtggaag	gagatctctc	aggagtaagc	attgtttttt	ttcacatct	1200
tgtatctctc	tacccacttt	tgggtagggg	tgctggcagc	tgtcccaagc	aatgggtaat	1260
gatgatggca	aaaagggtgt	ttgggggaac	agctgcagac	ctgctgctct	atgctcaccc	1320
ccgccccatt	ctgggccaat	gtgattttat	ttatttgctc	ccttggatac	tgacaccttg	1380
gtccactttt	ctccaggatg	ccaactgcac	tagctgtgtg	cgaatgacgt	atcttgtgca	1440
ttttaacttt	ttttccttaa	tataaatatt	ctggttttgt	atttttgtat	attttaatct	1500
aaggccctca	tttctgcac	tgtgttctca	ggtacatgag	caatctcagg	gatagccagc	1560
agcagctcca	ggtctgcgca	gcaggaatta	ctttttgttg	tttttgccac	cgtggagagc	1620
aactattttg	agtgcacagc	ctattgaact	acctattttt	tgccaataag	agctggcttt	1680
tctgccatag	tgctctcttg	aaacccccct	tgccctgaaa	atgttttatg	ggagactagg	1740
ttttaactgg	gtggcccat	gacttgattg	ccttctactg	gaagattggg	aattagtcta	1800
aacaggaaat	ggtggtacac	agaggctagg	agaggctggg	cccggtgaaa	aggccagaga	1860
gcaagccaag	attaggtgag	ggttgtctaa	tcctatggca	caggacgtgc	tttacatctc	1920
cagatctgtt	cttcaccaga	ttaggttagg	cctaccatgt	gccacagggt	gtgtgtgtgt	1980
ttgtaaaact	agagttgcta	aggataagtt	taaagaccaa	taccctgtga	cttaatcctg	2040
tgtgtgcgag	ggatggatat	atgaagtaag	gtgagatcct	taacctttca	aaattttcgg	2100
gttccaggga	gacacacaag	cgagggtttt	gtggtgcctg	gagcctgtgt	cctgccctgc	2160
tacagtagtg	attaatagtg	tcatggtagc	taaaggagaa	aaagggggtt	tcgtttacac	2220
gctgtgagat	caccgcaaac	ctaccttact	gtgttgaaac	gggacaaatg	caatagaacg	2280
cattgggtgg	tgtgtgtctg	atcctgggtt	cttgtctccc	ctaaatgctg	ccccccaagt	2340
tactgtattt	gtctgggctt	tgtaggactt	cactacgttg	attgctaggt	ggcctagttt	2400
gtgtaaatat	aatgtattgg	tctttctccg	tgttctttgg	gggttttgtt	tacaaacttc	2460
tttttgtatt	gagagaaaaa	tagccaaagc	atctttgaca	gaaggttctg	caccaggcaa	2520

[0005]

aaagatctga	aacattagtt	tggggggccc	tcttcttaaa	gtggggatct	tgaaccatcc	2580
tttcttttgt	attcccttc	ccctattacc	tattagacca	gatcttctgt	cctaaaaact	2640
tgtcttctac	cttgcctct	tttctgttca	ccccaaaag	aaaacttaca	caccacacaca	2700
catacacatt	tcatgcttgg	agtgtctcca	caactcttaa	atgatgtatg	caaaaatact	2760
gaagctagga	aaaccttcca	tcccttggtc	ccaacctcct	aagtcaagac	cattaccatt	2820
tctttcttct	ttttttttt	ttttttaaaa	tggagtctca	ctgtgtcacc	caggctggag	2880
tgcagtggca	tgatcggtc	actgcagcct	ctgcctcttg	ggttcaagtg	attctcctgc	2940
ctcagcctcc	tgagtagctg	ggatttcagg	caccgcgcac	actcagctaa	tttttgtatt	3000
tttagtagag	acggggtttc	accatgttgt	ccaggctggt	ctggaactcc	tgacctcagg	3060
tgatctgccc	accttggctt	cccaaagtgc	tgggattaca	ggcatgagcc	accatgctgg	3120
gccaacatt	tcttgggtga	ttcatgcaa	acacttaaga	caactgtgta	gccaggcgc	3180
ggtggctcac	acctgtaate	ccagcacttt	ggaaggctga	ggcgggcgga	tcacaaggctc	3240
acgagttcaa	aactatcctg	gccaacacag	tgaaccccg	tctctactaa	aatacaaaaa	3300
aattagccgg	gtgtgggtgt	gcatgccttt	agtcctagct	attcaggagg	ctgaggcagg	3360
ggaatcgctt	gaacccgaga	ggcagagggt	gcagttagct	gagatcgac	cactgcactc	3420
cagcctggtt	acagagcaag	actctgtctc	aaacaaaaca	aaacaaaaca	aaaacacact	3480
actgtatttt	ggatggatca	aacctcctta	attttaattt	ctaataccta	agtaaagaga	3540
tgcaattggg	ggccttccat	gtagaaagtg	gggtcaggag	gccaaagaa	ggaatatgaa	3600
tgtatatcca	agtcactcag	gaacttttat	gcagggtgcta	gaaactttat	gtcaaagtgg	3660
ccacaagatt	gtttaatagg	agacgaacga	atgtaactcc	atgtttactg	ctaaaaacca	3720
aagctttgtg	taaaatcttg	aatttatggg	gcgggagggt	aggaaagcct	gtacctgtct	3780
gtttttttcc	tgatcctttt	ccctcattcc	tgaactgcag	gagactgagc	ccctttgggc	3840
tttggtgacc	ccatcactgg	ggtgtgttta	tttgatggtt	gattttgctg	tactgggtac	3900
ttcctttccc	atttttcta	cattttttta	cacaagctga	ctcttcctt	cccttctcct	3960
ttccttgga	aaatacaatg	aataaataaa	gacttatttg	tacgcaaact	gtca	4014

<210> 5

<211> 2377

<212> DNA

<213> 智人 (Homo sapiens)

<400> 5

acccccgagc	tgtgctgctc	gcggccgcca	ccgccgggcc	ccggccgtcc	ctggctcccc	60
tcctgcctcg	agaagggcag	ggcttctcag	aggcttggcg	ggaaaaagaa	cggaggggagg	120
gategcgctg	agtataaaag	ccggttttcg	gggttttate	taactcgctg	tagtaattcc	180
agcgagaggc	agaggggagc	agcggggcgc	cggctagggt	ggaagagccg	ggcagacaga	240
gctgcgctgc	gggcgtcctg	ggaaggggaga	tccggagcga	atagggggct	tcgcctctgg	300
cccagccctc	ccgtgatcc	cccagccagc	ggtcgcgaac	ccttgcgcga	tccacgaaac	360
tttgcctata	gcagcgggcg	ggcactttgc	actggaactt	acaacacccg	agcaaggacg	420
cgactctccc	gacgcgggga	ggctattctg	cccatttggg	gacacttccc	cgcgcctgcc	480
aggaccgctt	tctctgaaag	gctctccttg	cagctgctta	gacgctggat	tttttcggg	540
tagtggaana	ccagcagcct	cccgcgacga	tgccctctca	cgtagcttc	accaacagga	600
actatgacct	cgactacgac	tcgggtgcagc	cgtattttcta	ctgcgacgag	gaggagaact	660
tctaccagca	gcagcagcag	agcgagctgc	agccccgggc	gcccagcgag	gatatctgga	720
agaaattcga	gctgctgccc	accccgcctc	tgctccctag	ccgccgctcc	gggctctgct	780
cgccctccta	cgttgcggtc	acacccttct	cccttcgggg	agacaacgac	ggcgggtggcg	840

[0006]

ggagcttctc	cacggccgac	cagctggaga	tggtgaccga	gctgctggga	ggagacatgg	900
tgaaccagag	tttcatctgc	gacccggacg	acgagacctt	catcaaaaac	atcatcatcc	960
aggactgtat	gtggagcggc	ttctcggccg	ccgccaagct	cgtctcagag	aagctggcct	1020
cctaccaggc	tgcgcgcaaa	gacagcggca	gcccgaaccc	cgcccgcggc	cacagcgtct	1080
gctccacctc	cagcttgtag	ctgcaggatc	tgagcgccgc	cgctcagag	tgcacgacc	1140
cctcggtggt	cttcccctac	cctctcaacg	acagcagctc	gcccaggtcc	tgcgcctcgc	1200
aagactccag	cgcttctctt	ccgtcctcgg	attctctgct	ctcctcgacg	gagtcctccc	1260
cgcagggcag	ccccgagccc	ctggtgctcc	atgaggagac	accgcccacc	accagcagcg	1320
actctgagga	ggaacaagaa	gatgaggaag	aaatcgatgt	tgtttctgtg	gaaaagaggc	1380
aggctcctgg	caaaagggtca	gagtcctggat	caccttctgc	tggaggccac	agcaaacctc	1440
ctcacagccc	actggtcctc	aagaggtgcc	acgtctccac	acatcagcac	aactacgcag	1500
cgctccctc	cactcggaag	gactatcctg	ctgccaagag	ggtcaagttg	gacagtgtca	1560
gagtcctgag	acagatcagc	aacaaccgaa	aatgcaccag	cccaggtcc	tggacaccg	1620
aggagaatgt	caagaggcga	acacacaacg	tcttgagcgc	ccagaggagg	aacgagctaa	1680
aacggagctt	ttttgccctg	cgtgaccaga	tcccggagtt	ggaaaacaat	gaaaaggccc	1740
ccaaggtagt	tatccttaaa	aaagccacag	catacatcct	gtccgtccaa	gcagaggagc	1800
aaaagctcat	ttctgaagag	gacttgttgc	ggaaacgacg	agaacagttg	aaacacaaac	1860
ttgaacagct	acggaactct	tgtgcgtaag	gaaaagtaag	gaaaacgatt	ccttctaaca	1920
gaaatgtcct	gagcaatcac	ctatgaactt	gtttcaaatg	catgatcaaa	tgcaacctca	1980
caaccttggc	tgagtcttga	gactgaaaga	tttagccata	atgtaaactg	cctcaaattg	2040
gactttgggc	ataaaagaac	ttttttatgc	ttaccatctt	tttttttctt	ttaacagatt	2100
tgtatttaag	aattgttttt	aaaaaathtt	aagatttaca	caatgtttct	ctgtaaatat	2160
tgccattaaa	tgtaataaac	tttaataaaa	cgtttatagc	agttacacag	aattttcaatc	2220
ctagtatata	gtacctagta	ttataggtac	tataaacctt	aatttttttt	atttaagtac	2280
attttgcttt	ttaaagttga	tttttttcta	ttgttttttag	aaaaaataaa	ataactggca	2340
aatatatcat	tgagccaaaa	aaaaaaaaaa	aaaaaaa			2377

<210> 6

<211> 2949

<212> DNA

<213> 智人 (Homo sapiens)

<400> 6

agtttcccga	ccagagagaa	cgaacgtgtc	tgcgggcgcg	cggggagcag	aggcgggtggc	60
ggcgggcggc	ggcaccggga	gccgccgagt	gaccttcccc	cgccctcttg	gccccccacc	120
ctcccaccog	cccgtggccc	gcgcccattg	ccgcgcgcgc	tccacacaac	tcaccggagt	180
ccgcgccttg	cgccgcccgc	cagttcgcag	ctccgcgcca	cggcagccag	tctcacctgg	240
cggcaccgcc	cgcccaccgc	ccgggccaca	gcccctgcgc	ccacggcagc	actcgaggcg	300
accgcgacag	tggtggggga	cgctgctgag	tggaagagag	cgagcccggg	ccaccggacc	360
tacttactcg	ccttgctgat	tgtctathtt	tgcgtttaca	actttttctaa	gaacttttgt	420
atacaaagga	actttttaaa	aaagacgctt	ccaagttata	tttaatecaa	agaagaagga	480
tctcggccaa	tttggggttt	tgggttttgg	cttcgtttct	tctcttcgtt	gactttgggg	540
ttcaggtgcc	ccagctgctt	cgggctgcgc	aggaccttct	gggccccac	attaatgagg	600
cagccacctg	gcgagtctga	catggctgtc	agcgacgcgc	tgctcccatc	tttctccacg	660
ttcgctctg	gcccggcggg	aaggggagaag	acactgcgtc	aagcaggtgc	cccgaataac	720
cgctggcggg	aggagctctc	ccacatgaag	cgacttcccc	cagtgtcttc	cgccgcggcc	780

[0007]

tatgacctgg	cggcggcgac	cgtggccaca	gacctggaga	gcggcggagc	cggtgcggct	840
tgcggcggtg	gcaacctggc	gcccctacct	cggagagaga	ccgaggagtt	caacgatctc	900
ctggacctgg	actttattct	ctccaattcg	ctgacctatc	ctccggagtc	agtggccgcc	960
accgtgtcct	cgtcagcgtc	agcctcctct	tcgtcgtcgc	cgtcgagcag	cggccctgcc	1020
agcgcgcctt	ccacctgcag	cttcacctat	ccgatccggg	ccgggaacga	cccgggcgtg	1080
gcgcggggcg	gcacggggcg	aggcctcctc	tatggcaggg	agtccgctcc	ccctccgacg	1140
gctcccttca	acctggcgga	catcaacgac	gtgagccctt	cgggcggctt	cgtggccgag	1200
ctcctgcggc	cagaattgga	cccgggtgtac	attccgccgc	agcagccgca	gccgccaggt	1260
ggcgggctga	tgggcaagtt	cgtgctgaag	gcgtcgctga	gcgcccctgg	cagcgagtac	1320
ggcagcccgt	cggtcacatc	cgtcagcaaa	ggcagccctg	acggcagcca	cccggtggtg	1380
gtggcgccct	acaacggcgg	gccgcgcgcg	acgtgcccc	agatcaagca	ggaggcggtc	1440
tcttcgtgca	cccacttggg	cgtcggaccc	cctctcagca	atggccaccg	gccggctgca	1500
cacgacttcc	ccctggggcg	gcagctcccc	agcaggacta	ccccgaccct	gggtcttgag	1560
gaagtgtctg	gcagcaggga	ctgtcacctt	gccctgccgc	ttcctcccgg	cttccatccc	1620
cacccggggc	ccaattaccc	atccttcctg	cccgatcaga	tgcagccgca	agtcccgcgg	1680
ctccattacc	aagagctcat	gccacccggt	tcttgcctgc	cagaggagcc	caagccaaag	1740
aggggaagac	gatcgtggcc	ccggaaaagg	accgccaccc	acacttgtga	ttacgcgggc	1800
tgcggcaaaa	cctacacaaa	gagttcccat	ctcaaggcac	acctgcgaac	ccacacaggt	1860
gagaaacctt	accactgtga	ctgggacggc	tgtggatgga	aattcgcccc	ctcagatgaa	1920
ctgaccaggc	actaccgtaa	acacacgggg	caccgcccgt	tccagtgcga	aaaatgcgac	1980
cgagcatttt	ccaggtcgga	ccacctcgcc	ttacacatga	agaggcattt	ttaaatccca	2040
gacagtggat	atgaccaca	ctgccagaag	agaattcagt	attttttact	tttcacactg	2100
tcttcccgat	gaggggaagga	gcccagccag	aaagcactac	aatcatggtc	aagttcccaa	2160
ctgagtcatc	ttgtgagtgg	ataatcagga	aaaatgagga	atccaaaaga	caaaaatcaa	2220
agaacagatg	gggtctgtga	ctggatcttc	tatcattcca	attctaaatc	cgacttgaat	2280
attcctggac	ttacaaaatg	ccaagggggt	gactggaagt	tgtggatatc	agggtataaa	2340
ttatatccgt	gagttggggg	agggaagacc	agaattccct	tgaattgtgt	attgatgcaa	2400
tataagcata	aaagatcacc	ttgtattctc	tttaccttct	aaaagccatt	attatgatgt	2460
tagaagaaga	ggaagaaatt	caggtacaga	aaacatgttt	aaatagccta	aatgatgggtg	2520
cttgggtgagt	cttggttcta	aaggtaccaa	acaaggaagc	caaagttttc	aaactgctgc	2580
atactttgac	aaggaaaatc	tatatattgtc	ttccgatcaa	catttatgac	ctaagtcagg	2640
taatatacct	ggtttacttc	tttagcattt	ttatgcagac	agtctgttat	gcactgtggt	2700
ttcagatgtg	caataatttg	tacaatgggt	tattcccaag	tatgccttaa	gcagaacaaa	2760
tgtgtttttc	tatatagtgc	cttgccttaa	taaatatgta	atataaattt	aagcaaactg	2820
ctattttgta	tatttgtaaa	ctacaaagta	aaatgaacat	tttgtggagt	ttgtattttg	2880
catactcaag	gtgagaatta	agttttaaat	aaacctataa	tattttatct	gaaaaaaaaa	2940
aaaaaaaaa						2949

<210> 7

<211> 2949

<212> DNA

<213> 猿猴病毒 40 (Simian virus 40)

<400> 7

agtttcccg	ccagagagaa	cgaacgtgtc	tgcgggcgcg	cggggagcag	aggcgggtggc	60
gggcggcggc	ggcaccggga	gccgcgcagt	gaccctcccc	cggccctctg	gccccccacc	120

[0008]

ctcccaccg	cccggtggccc	gcgcccattg	ccgcgcgcgc	tccacacaac	tcaccggagt	180
ccgcgccttg	cgccgcccac	cagttcgcag	ctccgcgcc	cggcagccag	tctcacctgg	240
cggcaccgcc	cgcccaccgc	ccgggccaca	gcccctgcgc	ccacggcagc	actcgaggcg	300
accgcgacag	tgggtggggga	cgctgctgag	tggaagagag	cgcagcccgg	ccaccggacc	360
tacttactcg	ccttgctgat	tgtctatttt	tgcgtttaca	actttttctaa	gaacttttgt	420
atacaaagga	actttttaaa	aaagacgctt	ccaagttata	tttaatccaa	agaagaagga	480
tctcgcccaa	tttggggttt	tgggttttgg	cttcgtttct	tctcttcgtt	gactttgggg	540
ttcaggtgcc	ccagctgctt	cggtctgccg	aggaccttct	gggccccac	attaatgagg	600
cagccacctg	gcgagtctga	catggctgtc	agcgacgcgc	tgctcccatc	tttctccacg	660
ttcgctctg	gcccggcggg	aaggagagaag	acactgcgtc	aagcaggtgc	cccgaataac	720
cgctggcggg	aggagctctc	ccacatgaag	cgacttcccc	cagtgtcttc	cggcgccccc	780
tatgacctgg	cggcggcgac	cgtggccaca	gacctggaga	gcggcggagc	cggtgcggct	840
tgcgcggtga	gcaacctggc	gcccctacct	cggagagaga	ccgaggagtt	caacgatctc	900
ctggacctgg	actttattct	ctccaattcg	ctgacctatc	ctccggagtc	agtggccgcc	960
accgtgtcct	cgtcagcgtc	agcctcctct	tcgtcgtcgc	cgtcgagcag	cgccctgcc	1020
agcgcgccct	ccacctgcag	cttcacctat	ccgatccggg	ccgggaacga	cccgggcgtg	1080
gcgcggggcg	gcacgggcgg	aggcctctct	tatggcaggg	agtccgctcc	ccctccgacg	1140
gctcccttca	acctggcggg	catcaacgac	gtgagccctt	cgggcggctt	cgtggccgag	1200
ctctgcggc	cagaattgga	cccgtgttac	attccgcgc	agcagccgca	gccgccaggt	1260
ggcgggctga	tgggcaagtt	cgtgctgaag	gcgtcgtga	gcgcccctgg	cagcgagtac	1320
ggcagcccgt	cggtcatcag	cgtcagcaaa	ggcagccctg	acggcagcca	cccggtggtg	1380
gtggcgccct	acaacggcgg	gcccgcgcgc	acgtgcccc	agatcaagca	ggaggcggtc	1440
tcttcgtgca	cccacttggg	cgctggacct	cctctcagca	atggccaccg	gccggctgca	1500
cacgacttcc	ccctggggcg	gcagctcccc	agcaggacta	ccccgacctt	gggtcttgag	1560
gaagtgtga	gcagcaggga	ctgtcacctt	gccctgccgc	ttcctcccgg	cttccatccc	1620
caccgggggc	ccaattaccc	atccttctct	cccgatcaga	tgacagccga	agtcccggcg	1680
ctccattacc	aagagctcat	gccacccggg	tcttgcatgc	cagaggagcc	caagccaaag	1740
aggggaagac	gatcgtggcc	ccggaaaagg	accgccacc	acacttgtga	ttacgcgggc	1800
tgcggcaaaa	cctacacaaa	gagttcccat	ctcaaggcac	acctgcgaac	ccacacaggt	1860
gagaaacctt	accactgtga	ctgggacggc	tgtggatgga	aattcgcccc	ctcagatgaa	1920
ctgaccaggc	actaccgtaa	acacacgggg	caccgcccgt	tccagtgcc	aaaatgcgac	1980
cgagcatttt	ccaggtcgga	ccacctgcc	ttacacatga	agaggcattt	ttaaatccca	2040
gacagtggat	atgaccaca	ctgccagaag	agaattcagt	attttttact	tttcacactg	2100
tcttcccgat	gagggaagga	gcccagccag	aaagcactac	aatcatggtc	aagttcccaa	2160
ctgagtcac	ttgtgagtgg	ataatcagga	aaaatgagga	atccaaaaga	caaaaatcaa	2220
agaacagatg	gggtctgtga	ctggatcttc	tatcattcca	attctaaatc	cgacttgaat	2280
attcctggac	ttacaaaatg	ccaagggggg	gacttggaagt	tgtggatata	agggtataaa	2340
ttatatccgt	gagttggggg	agggaagacc	agaattccct	tgaatttgtgt	attgatgcaa	2400
tataagcata	aaagatcacc	ttgtattctc	tttaccttct	aaaagccatt	attatgatgt	2460
tagaagaaga	ggaagaaatt	caggtacaga	aaacatgttt	aaatagccta	aatgatgggtg	2520
cttgggtgagt	cttggttcta	aaggtaccaa	acaaggaagc	caaagttttc	aaactgctgc	2580
atactttgac	aaggaaaatc	tatatattgtc	ttccgatcaa	catttatgac	ctaagtcagg	2640
taatatacct	ggtttacttc	tttagcattt	ttatgcagac	agtctgttat	gcactgtggt	2700
ttcagatgtg	caataatttg	tacaatgggt	tattcccaag	tatgccttaa	gcagaacaaa	2760
tgtgtttttc	tatatagtgc	cttgccttaa	taaatatgta	atataaattt	aagcaaacgt	2820
ctattttgta	tattttgtaa	ctacaaagta	aaatgaacat	tttgtggagt	ttgtattttg	2880
catactcaag	gtgagaatta	agtttttaaat	aaacctataa	tatttttatct	gaaaaaaaaa	2940

[0009]

aaaaaaaa

2949

<210> 8

<211> 4018

<212> DNA

<213> 智人 (Homo sapiens)

<400> 8

```

caggcagcgc tgcgtcctgc tgcgcacgtg ggaagccctg gccccggcca cccccgcgat      60
gccgcgcgct ccccgctgcc gagccgtgcg ctccctgctg cgcagccact accgcgaggt      120
gctgccgctg gccacgttcg tgcggcgcct ggggccccag ggctggcggc tggcgcagcg      180
cggggaccgc gcggtttcc gcgcgttgtt gggccagtgc ctggtgtgcg tgccctggga      240
cgcacggccg cccccgcgcg cccctcctt cggccaggtg tctgcctga aggagctggt      300
ggccccagtg ctgcagaggc tgtgcgagcg cggcgcgaag aacgtgctgg ccttcggctt      360
cgcgtgctgc gacggggccc gcggggggccc ccccgaggcc ttcaccacca gcgtgcgcag      420
ctacctgcc aacacggtga ccgacgcact gcgggggagc ggggcgtggg ggctgctgct      480
gcgccgcgtg ggcgacgacg tgcgtgttca cctgctggca cgctgcgcgc tctttgtgct      540
ggtggctccc agctgcccct accaggtgtg cgggcccgcg ctgtaccagc tcggcgtgc      600
cactcaggcc cggccccgc caccgctag tggaccccga aggcgtctgg gatgcgaacg      660
ggcctggaac catagcgtca gggaggcccg ggtccccctg ggctgccag ccccgggcgc      720
gaggaggcgc gggggcagtg ccagccgaag tctgccgttg cccaagaggc ccaggcgtgg      780
cgctgccct gagccggagc ggacgccctg tgggcagggg tctggggccc acccgggcag      840
gacgcgtgga ccgagtgacc gtggtttctg tgtggtgtca cctgccagac ccgccgaaga      900
agccacctct ttggagggtg cgctctctgg caccgcccac tcccacccat ccgtgggccc      960
ccagcaccac gcgggcccc catccacatc gcggccacca cgccctggg acacgccttg      1020
tccccgggtg tacgccgaga ccaagcactt ccttactcc tcaggcgaca aggagcagct      1080
gcggccctcc ttcctactca gctctctgag gccagccctg actggcgctc ggaggtcgt      1140
ggagaccatc tttctgggtt ccaggccctg gatgccaggg actccccgca ggttgccccg      1200
cctgccccag cgctactggc aaatgcggcc cctgtttctg gagctgcttg ggaaccacgc      1260
gcagtgcccc tacggggtgc tctcaagac gcaactgccc ctgcgagctg cggtcacccc      1320
agcagccggt gtctgtgccc gggagaagcc ccagggtctt gtggcgccc cagaggagga      1380
ggacacagac cccgtcgcc tggcgcagct gctccgccag cacagcagcc cctggcaggt      1440
gtacggcttc gtgcgggccc gcctgcgccg gctggtgccc ccaggcctct ggggtccag      1500
gcacaacgaa cgccgcttcc tcaggaacac caagaagttc atctccctgg ggaagcatgc      1560
caagctctcg ctgcaggagc tgacgtggaa gatgagcgtg cgggactgcg cttggctgcg      1620
caggagccca ggggttggtg gtgttcggc cgcagagcac cgtctgcgtg aggagatcct      1680
ggccaagtcc ctgcactggc tgatgagtgt gtacgtcgtc gagctgctca ggtctttctt      1740
ttatgtcacg gagaccacgt ttcaaaagaa caggctcttt ttctaccgga agagtgtctg      1800
gagcaagttg caaagcattg gaatcagaca gcacttgaag aggggtgcagc tgcgggagct      1860
gtcgggaagca gaggtcaggc agcatcgcca agccaggccc gccctgctga cgtccagact      1920
ccgcttcac cccaagcctg acgggctgcg gccgattgtg aacatggact acgtcgtggg      1980
agccagaacg ttccgcagag aaaagagggc cgagcgtctc acctcgaggg tgaaggcact      2040
gttcagcgtg ctcaactacg agcgggcgcg gcgccccggc ctctggggcg cctctgtgct      2100
gggcctggac gatatccaca gggcctggcg caccttcgtg ctgcgtgtgc ggcccagga      2160
ccgcccgcct gagctgtact ttgtcaaggt ggatgtgacg ggcggtacg acaccatccc      2220
ccaggacagg ctacaggagg tcacgccag catcatcaa cccagaaca cgtactgcgt      2280

```

[0010]

gcgtcggtat gccgtggtcc agaaggccgc ccatggggcac gtccgcaagg ccttcaagag	2340
ccacgtctct accttgacag acctccagcc gtacatgcga cagtctgtgg ctcacctgca	2400
ggagaccagc ccgctgaggg atgccgtcgt catcgagcag agctcctccc tgaatgaggc	2460
cagcagtggc ctcttcgacg tcttctctacg cttcatgtgc caccacgccg tgcgcatcag	2520
gggcaagtcc tacgtccagt gccaggggat cccgcagggc tccatcctct ccacgctgct	2580
ctgcagcctg tgctacggcg acatggagaa caagctgttt gcggggattc ggcgggacgg	2640
gctgctcctg cgtttggtgg atgatttctt gttggtgaca cctcacctca cccacgcgaa	2700
aaccttcctc aggaccctgg tccgaggtgt ccctgagtat ggctgcgtgg tgaacttgcg	2760
gaagacagtg gtgaacttcc ctgtagaaga cgaggccctg ggtggcacgg cttttgttca	2820
gatgccggcc caccgcctat tcccctggtg cggcctgctg ctggataccc ggaccctgga	2880
ggtgcagagc gactactcca gctatgcccg gacctccatc agagccagtc tcaccttcaa	2940
ccgcggcttc aaggctggga ggaacatgcg tcgcaaactc tttggggtct tgcggctgaa	3000
gtgtcacagc ctgtttctgg atttgagggt gaacagcctc cagacgggtg gcaccaacat	3060
ctacaagatc ctctgtctgc aggcgtacag gtttcacgca tgtgtgctgc agtccccatt	3120
tcatcagcaa gtttggaaga accccacatt tttcctgcgc gtcattctctg acacggcctc	3180
cctctgctac tccatcctga aagccaagaa cgcagggatg tcgctggggg ccaagggcgc	3240
cgccggccct ctgccctccg aggccgtgca gtggctgtgc caccaagcat tctgtctcaa	3300
gtgactcga caccgtgtca cctacgtgcc actcctgggg tactcagga cagcccagac	3360
gcagctgagt cggaaagctcc cggggacgac gctgactgcc ctggaggccg cagccaaccc	3420
ggcactgccc tcagacttca agaccatcct ggactgatgg ccacccgccc acagccaggc	3480
cagagcagca caccagcagc cctgtcacgc cgggctctac gtcccaggga gggaggggcg	3540
gcccacaccc agggccgcac cgctgggagt ctgaggcctg agtgagtgtt tggccgaggc	3600
ctgcatgtcc ggctgaaggc tgagtgtccg gctgaggcct gacgagtggt ccagccaagg	3660
gctgagtgtc cagcacacct gccgtcttca cttccccaca ggctggcgct cggctccacc	3720
ccagggccag cttttcctca ccaggagccc ggcttccact cccacatag gaatagtcca	3780
tccccagatt cgccattgtt caccctcgc cctgccctcc tttgccttc accccacca	3840
tccaggtgga gaccctgaga aggaccctgg gagctctggg aatttgaggt gaccaaaggt	3900
gtgccctgta cacaggcgag gaccctgcac ctggatgggg gtccctgtgg gtcaaattgg	3960
ggggaggtgc tgtgggagta aaatactgaa tatatgagtt tttcagtttt gaaaaaaa	4018

<210> 9

<211> 90

<212> DNA

<213> 口蹄疫病毒

<400> 9

ggttccggaa aaattgtcgc tctgtcaaa caaactctta actttgattt actcaaactg	60
gctgggggatg tagaaagcaa tccaggtcca	90

<210> 10

<211> 589

<212> DNA

<213> 细胞巨化病毒 (Cytomegalovirus)

<400> 10

[0011]

tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata tggagttccg	60
cgttacataa cttacggtaa atggccccgcc tggctgaccg cccaacgacc cccgcccatt	120
gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca	180
atgggtggag tatttacggt aaactgcca cttggcagta catcaagtgt atcatatgcc	240
aagtacgcc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta	300
catgacctta tgggactttc ctacttgga gtacatctac gtattagtca tcgctattac	360
catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg	420
atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg	480
ggactttcca aaatgtcgta acaactccgc ccattgacg caaatgggcg gtaggcgtgt	540
acgggtgggag gtctatataa gcagagctgg tttagtgaac cgtcagatc	589

<210> 11

<211> 1192

<212> DNA

<213> 智人 (Homo sapiens)

<400> 11

ctagcttcgt gaggtccgg tgcccgtcag tgggcagagc gcacatcgcc cacagtcccc	60
gagaagttgg ggggaggggt cggcaattga accggtgcct agagaagggtg gcgcggggta	120
aactgggaaa gtgatgtcgt gtactggctc cgcctttttc ccgagggtgg gggagaaccg	180
tatataagtg cagtagtcgc cgtgaacgtt ctttttcgca acgggtttgc cgccagaaca	240
caggtaaagt ccgtgtgtgg ttcccgcggg cctggcctct ttacgggtta tggcccttgc	300
gtgccttgaa ttacttcac ctggctccag tacgtgattc ttgatccga gctggagcca	360
ggggcgggcc ttgccttta ggagcccctt cgcctcgtgc ttgagttgag gcctggcctg	420
ggcgtcgggg ccgcccgtg cgaatctggt ggcaccttcg cgcctgtctc gctgctttcg	480
ataagtctct agccatttaa aatttttgat gacctgtgc gacgttttt ttctggcaag	540
atagcttgt aaatgcgggc caggatctgc aactggtat ttcggttttt gggcccgcgg	600
ccggcgacgg ggcccgtg cccagcgca catgttcggc gaggcggggc ctgcgagcgc	660
ggccaccgag aatcggacgg gggtagtctc aagctggccg gcctgctctg gtgcctggcc	720
tcgcgccgcc gtgtatcgcc ccgccctggg cggcaaggct ggcccgtcg gcaccagttg	780
cgtgagcgga aagatggccg cttcccggcc ctgctccagg gggctcaaaa tggaggacgc	840
ggcgtcggg agagcggcg ggtgagtcac ccacacaaag gaaaagggcc tttccgtcct	900
cagccgtcgc ttcatgtgac tccacggagt accgggcgcc gtccaggcac ctcgattagt	960
tctggagctt ttggagtac tcgtctttag gttgggggga ggggttttat gcgatggagt	1020
ttccccacac tgagtgggtg gagactgaag ttaggccagc ttggcacttg atgtaattct	1080
cgttggaatt tgcccttttt gagtttgat cttggttcat tctcaagcct cagacagtgg	1140
ttcaaagttt tttcttcca tttcaggtgt cgtgaacacg tggtcgcggc ca	1192

<210> 12

<211> 1192

<212> DNA

<213> 智人 (Homo sapiens)

<400> 12

ctagcttcgt gaggtccgg tgcccgtcag tgggcagagc gcacatcgcc cacagtcccc	60
--	----

[0012]

gagaagttgg	ggggaggggt	cggcaattga	accggtgcct	agagaaggtg	gcgcggggta	120
aactgggaaa	gtgatgtcgt	gtactggctc	cgcctttttc	ccgaggggtg	gggagaaccg	180
tatataagtg	cagtagtcgc	cgtgaacgtt	ctttttcgca	acgggtttgc	cgccagaaca	240
caggtaaagt	ccgtgtgtgg	ttcccgcggg	cctggcctct	ttacgggtta	tggcccttgc	300
gtgccttgaa	ttacttccac	ctggctccag	tacgtgattc	ttgateccga	gctggagcca	360
ggggcggggc	ttgcgtttta	ggagcccctt	cgcctcgtgc	ttgagttgag	gcctggcctg	420
ggcgtgggg	ccgccgcgtg	cgaatctggt	ggcaccttcg	cgcctgtctc	gctgctttcg	480
ataagtctct	agccatttaa	aatttttgat	gacctgctgc	gacgtttttt	ttctggcaag	540
atagtcttgt	aaatcggggc	caggatctgc	acactggtat	ttcggttttt	gggcccgcg	600
ccggcgacgg	ggcccgtgcg	tcccagcgca	catgttcggc	gaggcggggc	ctgcgagcgc	660
ggccaccgag	aatcggacgg	gggtagtctc	aaagtggccg	gcctgctctg	gtgcctggcc	720
tcgcgccgcc	gtgtatcgcc	ccgccctggg	cggcaaggct	ggcccggtcg	gcaccagttg	780
cgtgagcgga	aagatggccg	cttcccggcc	ctgctccagg	gggctcaaaa	tggaggacgc	840
ggcgctcggg	agagcgggcg	ggtgagtcac	ccacacaaag	gaaaagggcc	tttccgtcct	900
cagccgtcgc	ttcatgtgac	tccacggagt	accgggcgcc	gtccaggcac	ctcgattagt	960
tctggagctt	ttggagtacg	tcgtcttttag	gttgggggga	ggggttttat	gcgatggagt	1020
ttccccacac	tgagtgggtg	gagactgaag	ttaggccagc	ttggcacttg	atgtaattct	1080
cgttgaatt	tgcccttttt	gagtttgat	cttggttcat	tctcaagcct	cagacagtgg	1140
ttcaaagttt	ttttcttcca	tttcagggtg	cgtgaacacg	tggtcgcggc	ca	1192

图1A

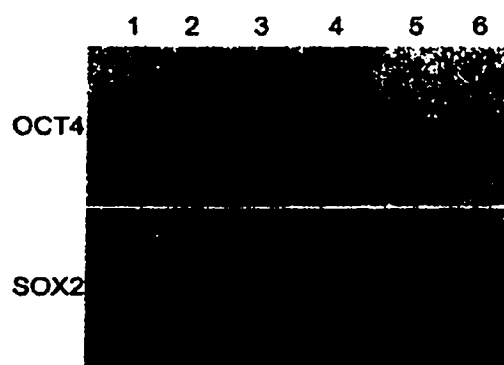


图1B

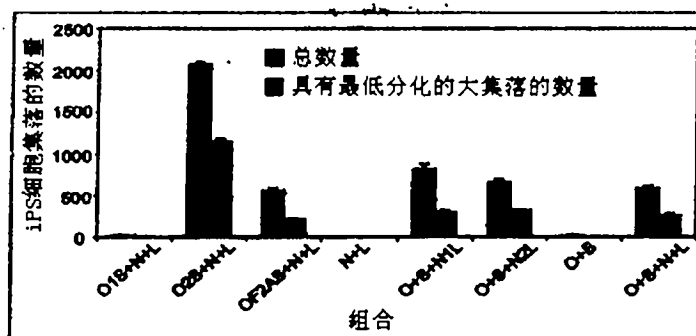


图 1

图2A

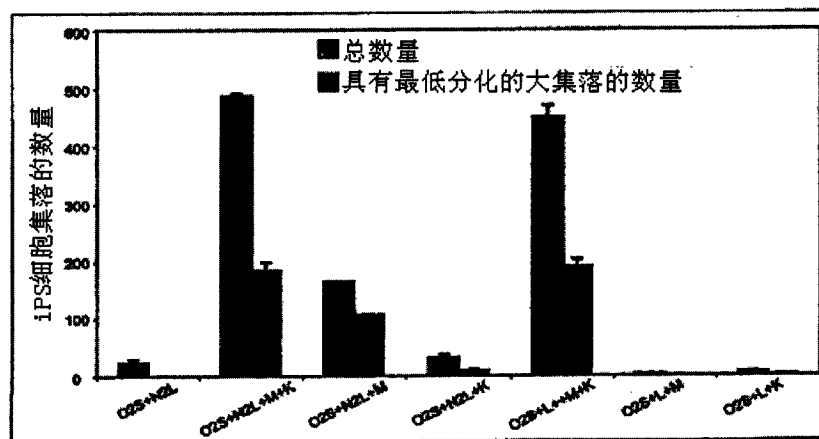


图2B

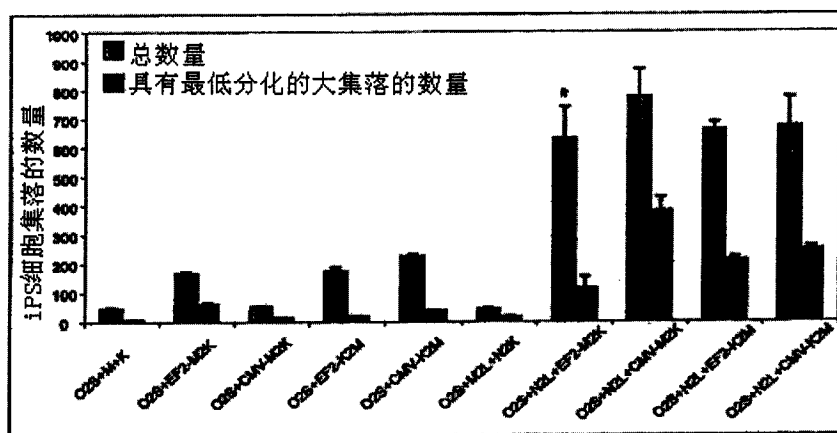


图2C

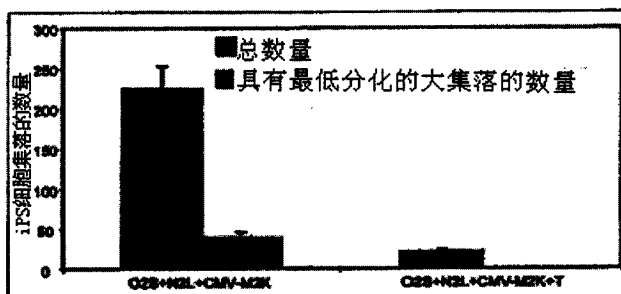


图 2

图3A

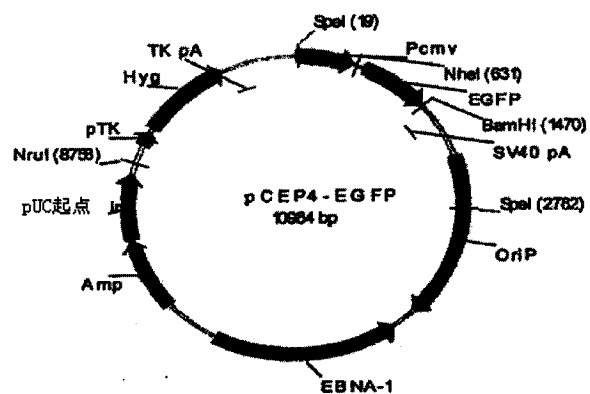


图3B

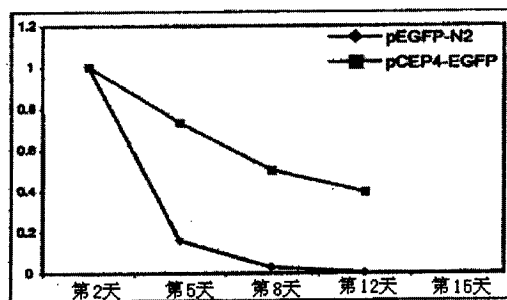


图3C

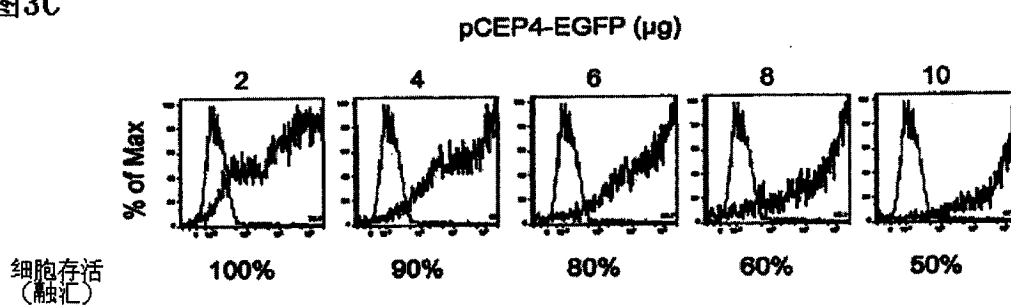


图 3

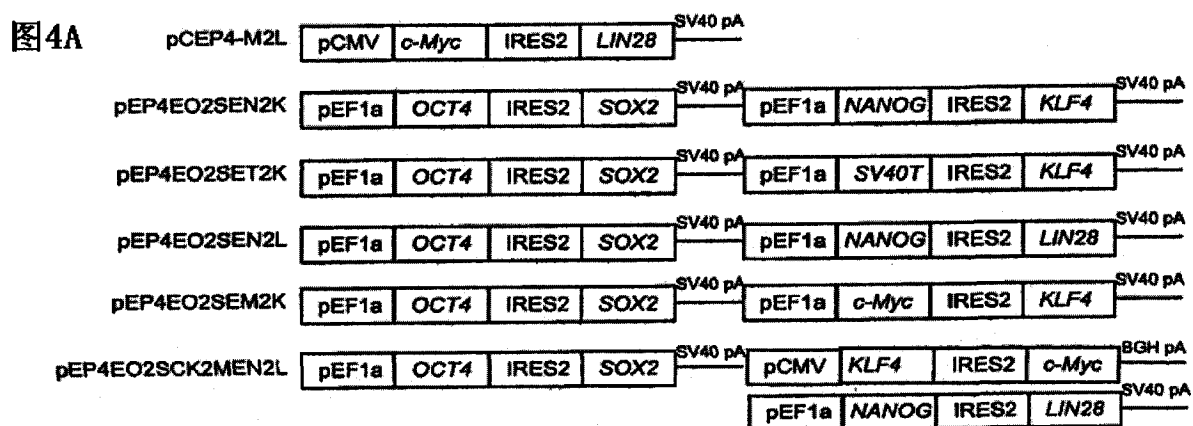


图4B



图4C

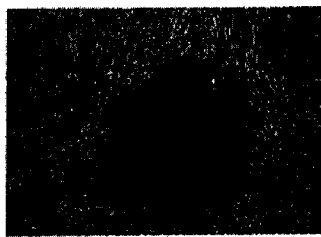


图4D

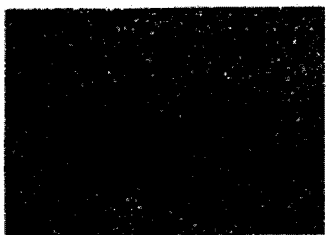


图 4

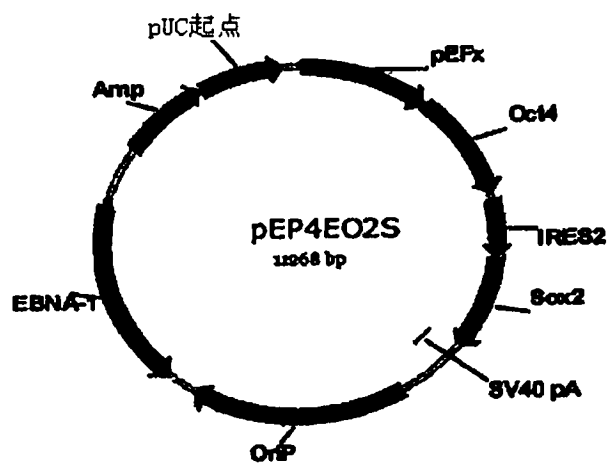


图5A

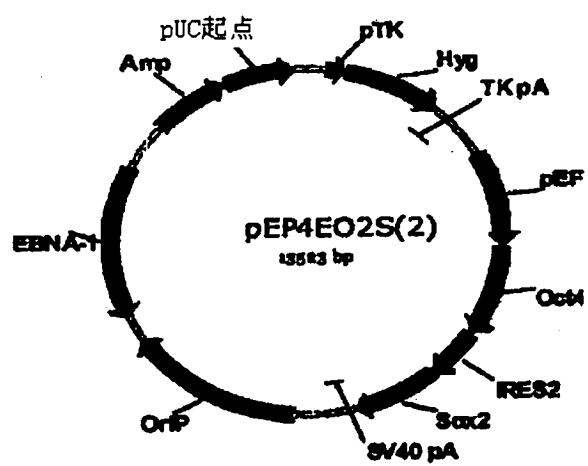


图5B

图 5