



(12) 发明专利

(10) 授权公告号 CN 101522899 B

(45) 授权公告日 2012. 07. 25

(21) 申请号 200780037879. 5

(22) 申请日 2007. 10. 10

(30) 优先权数据

60/850, 269 2006. 10. 10 US

(85) PCT申请进入国家阶段日

2009. 04. 10

(86) PCT申请的申请数据

PCT/KR2007/004940 2007. 10. 10

(87) PCT申请的公布数据

W02008/044869 EN 2008. 04. 17

(73) 专利权人 百疗医株式会社

地址 韩国首尔

(72) 发明人 金善荣 金水晶 李准台

(74) 专利代理机构 永新专利商标代理有限公司

72002

代理人 林晓红

(51) Int. Cl.

C12N 15/64 (2006. 01)

(56) 对比文件

W0 2006078880 A2, 2006. 07. 27, 全文.

Sara J. Cooper 等. Comprehensive analysis.....human genome. 《Genome Research》. 2006, 第 16 卷 (第 1 期), 1-10.

审查员 樊颖

权利要求书 1 页 说明书 35 页

序列表 32 页 附图 8 页

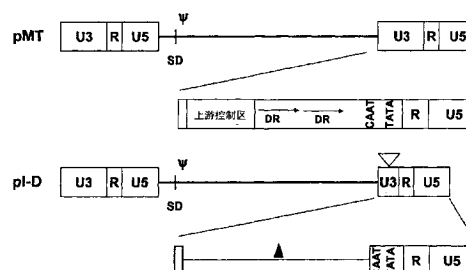
(54) 发明名称

安全性改善的表达载体

(57) 摘要

本发明涉及内部启动子在哺乳动物表达载体上的应用, 其中这些表达载体包括质粒载体和增强子缺失型反转录病毒载体。所述反转录病毒载体具有改善的安全性和最佳的转基因表达水平和载体滴度。

U3-失活的反转录病毒载体



1. 一种反转录病毒载体,其包含 SEQ ID NO:8 的核苷酸序列所表示的 RPL10 启动子或其片段作为内部启动子,其中该载体包含含 5' LTR 和 3' LTR 的核苷酸序列,其中所述 3' LTR 的 U3 区的增强子元件被缺失,其中所述包含片段的载体保留了所述 RPL10 启动子产生基因表达水平的 40-100% 的能力,并且其中所述片段包含相对于 RPL10 启动子转录起始位点从 -350 到 +143 的序列。

2. 如权利要求 1 所述的反转录病毒载体,其中所述片段保留了所述 RPL10 启动子产生病毒滴度的 50-120% 的能力。

3. 如权利要求 1 所述的反转录病毒载体,其中所述载体是质粒载体。

4. 如权利要求 1 所述的反转录病毒载体,其中所述 5' LTR 的 U3 区的增强子元件被缺失。

5. 如权利要求 1 所述的反转录病毒载体,其中所述载体是基于致癌反转录病毒的反转录病毒载体。

6. 如权利要求 1 所述的反转录病毒载体,其中所述载体是基于慢病毒的反转录病毒载体。

7. 如权利要求 1 所述的反转录病毒载体,其中所述启动子包含一个或多个剪接位点。

8. 如权利要求 1 所述的反转录病毒载体,其中所述内部启动子是所述启动子的片段,其中包含所述片段的所述载体保留与所述野生型启动子相同的产生高病毒滴度或高基因表达水平的能力。

9. 如权利要求 1 所述的反转录病毒载体,其中所述载体包含与所述内部启动子可操作连接的目的多核苷酸,其中含有所述内部启动子的所述载体可产生高病毒滴度或所述目的多核苷酸的高水平基因表达。

10. 如权利要求 9 所述的反转录病毒载体,其中所述目的多核苷酸编码多肽。

11. 如权利要求 10 所述的反转录病毒载体,其中所述多肽是治疗用蛋白质或报道蛋白质。

12. 如权利要求 11 所述的反转录病毒载体,其中所述多肽是 eGFP 或 gp91。

13. 如权利要求 9 所述的反转录病毒载体,其中所述目的多核苷酸编码 RNA、反义 RNA、小分子干扰 RNA 或核酶。

14. 如权利要求 9 所述的反转录病毒载体,其中所述目的多核苷酸还包含多聚腺苷酸化序列、内部核糖体进入位点、绝缘子序列、剪接序列或它们的组合。

15. 包含权利要求 1-14 中任一项所述的反转录病毒载体和适宜的载剂的组合物。

16. 包含权利要求 1-14 中任一项所述的反转录病毒载体的细胞。

17. 如权利要求 16 所述的细胞,其中所述细胞是哺乳动物细胞。

18. 如权利要求 17 所述的细胞,其中所述细胞是人细胞。

19. 如权利要求 16 所述的细胞,其中所述细胞是生产细胞系。

20. 包含权利要求 1-14 中任一项所述的反转录病毒载体的试剂盒。

安全性改善的表达载体

发明领域

[0001] 本发明提供了包含可用于表达目的蛋白的内部启动子的表达载体。在一个实施方式中,本发明提供了包含增强子缺失型 U3 区的反转录病毒载体。

[0002] 发明背景

[0003] 基因转移涉及将遗传物质转移到细胞内,通常用于转录和表达。该方法适合于蛋白表达以及治疗用途。已知各种转移方法,例如 DNA 转染和病毒转导。由于转移效率高、转基因表达水平高、以及需要时可通过自然亲和力或假型化而靶向特定受体和 / 或细胞类型的潜力,病毒介导的基因转移有吸引力。

[0004] 具体而言,由于反转录病毒载体可整合到细胞基因组内,它们可用于更长时间的表达。基于鼠白血病病毒 (MLV) 的载体是最常见的反转录病毒载体,其具有适合多数应用场合的许多骨架质粒和包装细胞系 (例如参见 Miller and Buttimore, Mol. Cell. Biol. 6 : 2895 (1986))。与所有的“简单”反转录病毒例如只编码结构病毒蛋白和酶病毒蛋白且不使用病毒辅助蛋白的反转录病毒一样, MLV 载体只可整合到处于分裂的细胞中。其他可能适合于用作载体的简单反转录病毒包括哺乳动物 C 型病毒的其他成员 (例如鼠干细胞病毒、Harvey 鼠肉瘤病毒和脾脏坏死病毒)、B 型病毒 (例如小鼠乳腺癌病毒) 和 D 型病毒 (例如 Mason Pfizer 猴病毒)。适合用作本发明反转录病毒载体的其他反转录病毒包括禽反转录病毒 (例如劳斯肉瘤病毒)、泡沫病毒 (spumaviruses) (例如泡沫病毒) 和 HTLV-BLV 病毒 (例如 HTLV-1)。

[0005] 慢病毒是反转录病毒的一个亚类,它们表达病毒辅助蛋白,并可感染和整合入未分裂细胞以及分裂细胞。源自慢病毒的载体是向靶细胞传递外源基因的理想工具,因为它们可稳定地整合到分裂细胞和未分裂细胞的基因组并介导长期基因表达 ((Gilbert 等人, Somat. Cell Mol. Genet. 26 : 83 (2001) ; Mitrophanous 等人, Gene Ther. 6 : 1808 (1999) ; Naldini 等人, Science 272 : 263 (1996) ; Sauter 等人, Somat. Cell Mol. Genet. 26 : 99 (2001))。

[0006] 业已从许多脊椎动物物种分离出慢病毒,其中包括灵长类例如人和猿猴免疫缺陷病毒 (HIV-1, HIV-2, SIV) 以及非灵长类例如猫免疫缺陷病毒 (FIV)、牛免疫缺陷病毒 (BIV)、马感染病毒 (EIAV)、山羊关节炎脑炎病毒 (CAEV) 和绵羊髓鞘脱落病毒。HIV 和 SIV 是这些病毒中目前了解最多的病毒。在非灵长类慢病毒载体中,源自 FIV (Curran 等人, Curr. Top. Microbiol. Immunol. 261 : 75 (2002)) 和 EIAV (美国专利 2001/0044149) 的载体了解最多。

[0007] 反转录病毒基因治疗领域有两个主要安全方面受到相当大的关注,不论载体是基于鼠白血病病毒 (MLV) 还是慢病毒。具体而言它们是指复制型反转录病毒的出现和插入诱变的发生率。通过开发载体和包装基因组之间不含重叠病毒序列的最小尺寸反转录病毒载体,前一问题已得到明显的改进。但是后一可能性近来引起了严重关注,主要是因为 X-SCID 人试验 (Hacein-Bey-Abina 等人, Science 302 : 415 (2003)) 中发现三例白血病患者。前两例白血病患者回顾分析揭示白血病可能源于整合到染色体上的反转录病毒以及

随后被长末端重复 (LTR) 活化的 LMO2 基因,该基因紧邻整合位点。尽管有人认为由于该疾病及其基因的特殊性,这种载体介导的肿瘤发生可能只限于 X-SCID 基因治疗患者,但现在已清楚反转录病毒载体的安全性需要进一步改善,以成为现实世界可行的治疗方式。

[0008] 已有几种降低载体介导的肿瘤发生几率的方法。一种方法是除去 LTR 的 U3 区 (Yu 等人, Proc. Natl. Acad. Sci. USA 83 :3194(1986) ;Hawley 等人., Proc. Natl. Acad. Sci. USA 84 :2406(1987) ;Yee 等人, Proc. Natl. Acad. Sci. USA 84 :5197(1987))。反转录病毒的 LTR 由 U3、R 和 U5 区组成,而 U3 区含有调控基因表达的增强子和启动子序列 (Sun 等人, J. Virol. 69 :4941(1995) ;Wahlers 等人, Mol. Ther. 6 :313(2002))。因此,除去 U3 区可降低载体引发的插入激活。在该情况下,需向载体提供附加的启动子以驱动靶基因的表达,因为缺失 U3 的载体的 LTR 不再含有启动子序列。

[0009] 正如前述, U3 失活的反转录病毒载体需要用于靶基因表达的内部启动子。反转录病毒载体最常使用的内部启动子之一是人巨细胞病毒 (HCMV) 的即时早期 (IE) 启动子 (Jaalouk 等人, Virol. J. 3 :27(2006) ;来自 BDBiosciences 的 pQCXIN) 或相关内部启动子例如 CA (HCMV IE 增强子 / 鸡 β -肌动蛋白启动子) (Ramezani 等人, Mol. Ther. 14 :245(2006))。但是, HCMV IE 启动子已知在人原代细胞中会快速失活,而且对某些基因不起作用 (Herweijer 等人., J. Gene Med. 3 :280(2001))。因此,业已表明常用的启动子会在某些细胞类型中失活,从而降低异源基因的表达,并可潜在活化 LTR 驱动的转录,所有这些均会降低反转录病毒载体的安全性和有效性。

[0010] 最后,由于常用启动子例如 CMV 和 SV40 的启动子抑制效应,该效应会降低用于包装的基因组 RNA 的转录, U3 失活的反转录病毒载体的滴度极低 (Jaalouk 等人, Virol. J. 3 :27(2006))。事实上,与具有完整 U3 区的类似 MLV 载体相比,基于 MLV 的缺失 U3 的载体的滴度低了甚至 4 个数量级 (Olson 等人, J. Virol. 68 :7060(1994))。因此,如果能够找到既可驱动异源基因高水平转录又可产生高病毒滴度的启动子,将是令人惊讶的。因此,需要开发新的启动子,以用作反转录病毒载体的内部启动子。

[0011] 发明概述

[0012] 本发明提供了包含异源内部启动子的表达载体。在一个实施方式中,所述载体包含一包含 5' LTR 和 3' LTR 的核苷酸序列。在另一实施方式中, 3' LTR 或 3' LTR 和 5' LTR 两者的 U3 区的增强子元件被缺失。在一个实施方式中,所述载体是质粒载体。在另一实施方式中,所述载体是反转录病毒载体,其包含一个或多个增强子缺失型 U3 区,其还包含可操作连接到异源基因的内部启动子,使得所述反转录病毒载体可产生高病毒滴度和所述异源基因的高水平转录。正如在反转录病毒载体的技术领域内众所周知,这类载体还包含反转录、包装等所需的顺式作用元件。在另一实施方式中,所述载体编码包含一个或多个增强子缺失型 U3 区的反转录病毒载体。

[0013] 在本发明的一个实施方式中,所述内部启动子是真核、原核或病毒启动子。在另一实施方式中,所述内部启动子是哺乳动物细胞基因启动子。在另一实施方式中,所述内部启动子选自 RPL10 启动子 (SEQ ID NO :8)、LENG8 启动子 (SEQ ID NO :9)、SNX3 启动子 (SEQ ID NO :10)、UQCQR 启动子 (SEQ ID NO :17) 或 ITGB4BP 启动子 (SEQ ID NO :16)。在另一实施方式中,所述内部启动子是所述全长启动子的片段或变体,并可驱动异源基因的高水平转录,而包含该启动子的载体可产生高病毒滴度。在一个实施方式中,包含所述启动子片段

或变体的载体基本保留与包含野生型启动子的载体相同的能力,即产生高病毒滴度和高水平的转录。在另一实施方式中,所述内部启动子基本由 TATA 盒组成。

[0014] 在一些实施方式中,所述内部启动子还包含用于高水平基因表达的剪接位点。在另一实施方式中,所述载体是还包括附加序列的反转录病毒载体,例如本领域熟知的聚腺苷酸化位点、绝缘子序列、剪接位点、内部核糖体进入位点 (IRES) 和其他的转录和翻译效应物序列。

[0015] 在另一实施方式中,所述载体是包含编码反转录病毒载体的 DNA 的质粒,该反转录病毒载体包含一个带有增强子缺失型 U3 区的 3' LTR。在另一实施方式中,所述质粒编码在 5' LTR 和 3' LTR 均带有增强子缺失型 U3 区的载体。在另一实施方式中,提供了封装载体 RNA 的感染性反转录病毒颗粒,其中该载体 RNA 的两个 LTR 均包含增强子缺失型 U3 区。在另一实施方式中,所述载体或为 RNA 形式或为 DNA 形式,其一个或两个 U3 区的增强子被缺失。

[0016] 在另一实施方式中,所述异源基因编码目的转录物。在另一实施方式中,所述目的转录物是具有生物活性的转录物,例如但不限于小分子干扰 RNA、核酶、反义 RNA 或诱饵 RNA。在另一实施方式中,所述异源基因编码多肽。所述多肽可为任何所需的蛋白质,例如治疗用蛋白质或标记蛋白质。在一个实施方式中,所述异源基因编码 eGFP 或 gp91。

[0017] 本发明还提供了包含所述载体和适宜载剂的组合物。所述组合物可适合于体内施用。

[0018] 提供了包含本发明所述载体的细胞,其中包括以所述载体转化的靶细胞,或包含所需载体和附加序列的生产细胞,其中附加序列编码生成感染性颗粒所需的因子,例如反转录病毒的 env 和 gag-pol 以及其他所需因子。所述靶细胞和生产细胞可为任何适宜的真核细胞类型,例如哺乳动物细胞。在另一实施方式中,所述细胞可源自人、灵长类或鼠类。所述细胞可为原代细胞或细胞系。

[0019] 本发明还提供了一种产生感染性反转录病毒颗粒的方法,其包括培养含有上述反转录病毒载体或编码所述反转录病毒载体的质粒的生产细胞,收集上清,过滤培养基以得到不含细胞的病毒上清。用于构建生产细胞系的包装细胞系可为本领域目前已知的任何细胞系,或可为经转移编码必需病毒蛋白质的基因至细胞系而得到的细胞系,使得包含包装信号的反转录病毒载体一旦在细胞内转录,所述反转录病毒载体即被包装到感染性颗粒内。

[0020] 本发明还提供了一种转导靶细胞的方法,其包括以上述制备的包含依据本发明的感染性反转录病毒颗粒的病毒上清接触所述细胞。正如上述,所述靶细胞可为但不限于为哺乳动物细胞、人细胞、灵长类细胞或鼠类细胞。所述靶细胞可为原代细胞或细胞系。

[0021] 本发明还提供了一种治疗对象的方法,其包括施用包含本发明所述载体和适宜载剂的组合物,其中所述异源基因编码可用于治疗的多肽或转录物。在另一实施方式中,所述方法用于治疗遗传疾病、增殖疾病或感染性疾病。

[0022] 本发明还提供了包含本发明所述多核苷酸和载体的试剂盒。

[0023] 附图简述

[0024] 图 1 是增强子缺失型 U3 反转录病毒载体 pMT 和 pI-D 的示意图。

[0025] 图 2 是增强子缺失型 U3 反转录病毒载体 pI-ND 的示意图。

[0026] 图 3 是增强子缺失型 U3 反转录病毒载体 pI-LND-n 的示意图。

[0027] 图 4 显示了包含所示启动子的增强子缺失型 U3 反转录病毒载体的异源基因 (gp91) 表达水平。对照为 MT-gp91-n 载体。

[0028] 图 5 显示了用于 GAPDH 的启动子序列 (SEQ ID NO :7)。

[0029] 图 6 显示了用于 RPL10 的启动子序列 (SEQ ID NO :8)。

[0030] 图 7 显示了用于 LENG8 的启动子序列 (SEQ ID NO :9)。

[0031] 图 8 显示了用于 SNX3 的启动子序列 (SEQ ID NO :10)。

[0032] 图 9 显示了用于 ITGB4BP 的启动子序列 (SEQ ID NO :16)。

[0033] 图 10 显示了用于 UQCRQ 的启动子序列 (SEQ ID NO :17)。

[0034] 发明详述

[0035] 本发明涉及包含异源内部启动子的表达载体。在一个实施方式中,所述载体是在一个或两个 LTR 包含增强子缺失型 U3 区的反转录病毒载体,该载体具有可操作连接到异源基因的异源启动子,该反转录病毒载体可高水平转录异源基因和产生高病毒滴度。在另一实施方式中,所述载体编码在一个或两个 LTR 包含增强子缺失型 U3 区的反转录病毒载体。

[0036] 载体

[0037] 正如本文使用,术语“反转录病毒”系指在其复制周期内利用反转录酶的 RNA 病毒。反转录病毒的基因组 RNA 被反转录酶转化为双链 DNA。这种双链 DNA 形式的病毒可被整合到受感染细胞的染色体上;一旦整合,该病毒被成为“原病毒”。所述原病毒作为 RNA 聚合酶 II 的模板,指导 RNA 分子的表达,这些 RNA 分子编码结构蛋白质和产生新病毒颗粒所需的酶。

[0038] 正如本文使用,术语“载体”系指任何遗传学元件,例如质粒、噬菌体、转座子、粘粒、染色体、反转录病毒、病毒体等,它们与适宜的调控元件结合时可进行复制,还可在细胞间转移基因序列。因此,该术语包括克隆和表达载体以及 RNA 或 DNA 形式的病毒载体。可使用本领域知晓的许多载体来操纵核酸和将应答元件和启动子整合到基因内等。可能的载体包括例如质粒或修饰病毒,其中包括噬菌体如 λ 衍生物或质粒如 pBR322 或 pUC 质粒衍生物,或 Bluescript 载体。病毒载体特别是反转录病毒载体已被用于细胞以及活实验动物的许多类型的基因传递场合。可使用的病毒载体包括但不限于反转录病毒、腺相关病毒、痘病毒、杆状病毒、痘苗病毒、单纯性疱疹病毒、EB 病毒、腺病毒、双生病毒和花椰菜花叶病毒载体。非病毒载体包括质粒、脂质体、荷电脂质(细胞转染剂)、DNA-蛋白质复合物和生物聚集体。真核生物载体的实例包括但不限于来自 Stratagene 的 pW-LNE0、pSV2CAT、pOG44、pXT1 和 pSG;来自 Amersham Pharmacia Biotech 的 pSVK3、pBPV、pMSG 和 pSVL;和来自 Clontech 的 pCMVDsRed2-express、pIRES2-DsRed2、pDsRed2-Mito 和 pCMV-EGFP。许多其他载体是众所周知的,可从市面购得。可将适宜的 DNA 片段连接到具有互补粘性末端的选定的载体上,从而将对应于应答元件和启动子的 DNA 片段插入到适宜的载体上。或者,可使用酶修饰 DNA 分子的末端,或将核苷酸序列(连接体)连接到 DNA 末端上产生任意位点。这类载体可经改造含有选择性标记基因,以选择已将标记基因整合到细胞基因组上的细胞。这类标记基因允许鉴定和/或选择掺入和表达所述标记基因编码的蛋白质的宿主细胞。

[0039] 正如前述,本发明所述的反转录病毒载体或由本发明载体编码的所述反转录病毒载体可基于简单的反转录病毒,例如 MLV、慢病毒如 HIV、或任何其他反转录病毒。这些载体

保留了生成感染性颗粒所需的顺式元件。这类元件包括紧邻反转录病毒基因组 5' LTR 的包装信号,病毒 RNA 进入病毒壳体或颗粒的衣壳化过程需要该包装信号。几种反转录病毒载体使用病毒基因组包壳所需的最短包装信号(也被称为 psi[Ψ] 序列)。附加的顺式元件为本领域熟知,例如引物结合位点、聚嘌呤区和其他序列,它们被纳入在所述反转录病毒载体上。但是,从载体上除去了编码病毒蛋白质的序列,以使没有全长的病毒蛋白质得到表达。通过包装构建体反式提供了生成感染性颗粒所需的任何病毒蛋白质。

[0040] 所述载体还可包含序列如多腺苷酸化序列、绝缘子序列、剪接位点、IRES 和其他转录和翻译效应物序列。

[0041] 本文使用的术语“聚腺苷酸化位点”、“poly A 位点”或“poly A 序列”系指一段指导新生 RNA 转录物终止和多聚腺苷酸化的 DNA 序列。重组转录物的有效多聚腺苷酸化作用是理想的,因为缺乏 poly A 尾的转录物不稳定,很快便会被降解。本载体使用的 poly A 信号可为“异源的”或“内源的”。内源 poly A 信号是天然存在于基因组中某个基因编码区域 3' 末端的 poly A 信号。异源 poly A 信号是分离自一个基因并置于另一基因 3' 端的 poly A 信号。通常使用的异源 poly A 信号为 SV40 poly A 信号。所述 SV40 poly A 信号被包含在一段 237bp 的 BamHI/BclI 限制酶切片段上,可同时指导终止和多聚腺苷酸化。本发明的适宜多聚腺苷酸化序列还包括但不限于牛生长激素 (bGH) 多聚腺苷酸化信号、β-球蛋白 polyA 位点和单纯性疱疹病毒的胸苷激酶的 polyA 位点 (Sambrook 等人., Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory Press, New York, pp. 16.7-16.8 (1989))。

[0042] 本发明所述载体还可含有附加的核酸序列,例如内含子序列、剪接序列、定位序列或信号序列,以足以使细胞有效加工所述载体核酸表达的蛋白质。内含子序列的实例包括 β-球蛋白内含子和人 EF-1α 内含子 (美国专利 7,049,143)。剪接信号介导从原始 RNA 转录物上除去内含子,其由一个剪接供体位点和一个剪接受体位点组成 (Sambrook 等人, Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory Press, New York, pp. 16.7-16.8 (1989))。

[0043] 正如本文使用,术语“内部核糖体进入位点”或“IRES”系指一段位于多个多顺反子基因之间的序列,其可通过双顺反子 mRNA 翻译的内部起始允许生成源自第二基因的表达产物。内部核糖体进入位点的实例包括但不限于源自口蹄疫病毒 (FDV)、脑心肌炎病毒、脊髓灰质炎病毒和 RDV 的内部核糖体进入位点 (Scheper 等人, Biochem. 76:801 (1994); Meyer 等人, J. Virol. 69:2819 (1995); Jang 等人, J. Virol. 62:2636 (1998); Haller 等人, J. Virol. 66:5075 (1995))。正如本领域所知,可组装掺有 IRES 的载体。例如,含有多顺反子序列的载体可含有下列可操作连接的元件:内部启动子、异源基因、内部核糖体进入位点和第二异源基因。

[0044] 这类附加序列被插入到所述载体上,这样若需转录时,它们与启动子序列可操作地连接在一起,或若需翻译和加工时,它们可另外再与启动和加工序列可操作地连接在一起。或者,可将插入序列置于载体上的任何位置。

[0045] 用于构建本发明所述载体的标准技术为本领域普通技术人员所熟知,可见于参考文献如 Sambrook 等人, Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory Press, New York, (1989)。可使用许多策略连接 DNA 片段,其选择取决

于 DNA 片段末端的性质,本领域技术人员可容易地作出这类选择。

[0046] 还预见包含本发明所述表达载体和适宜载剂的组合物。这类载剂为本领域熟知,其系指施用治疗药物时使用的稀释剂、佐剂、赋形剂、或媒剂。这类适宜的载剂可为无菌液体,例如水和油,其中包括石油、动物、植物和合成来源的油,例如花生油、大豆油、矿物油、芝麻油等。水是静脉内施用组合物时的优选载剂。盐水溶液、葡萄糖水溶液和甘油水溶液也可用作液体载剂,尤其是对于可注射溶液。适宜的药物赋形剂包括淀粉、葡萄糖、乳糖、蔗糖、明胶、麦芽、稻米、面粉、白垩、硅胶、硬脂酸钠、单硬脂酸甘油酯、滑石、氯化钠、脱脂奶粉、甘油、丙二醇、水、乙醇等。如需要,所述组合物还可含有少量的润湿剂或乳化剂或 pH 缓冲剂。这些组合物可采取溶液剂、悬浮剂、乳化剂、片剂、丸剂、胶囊剂、粉剂、缓释剂等形式。所述组合物可配制为栓剂,其中加入传统的粘合剂和载剂如甘油三酸酯。口服制剂可包括标准的载剂例如药用级的甘露醇、乳糖、淀粉、硬脂酸镁、糖精钠、纤维素、碳酸镁等。E. W. Martin 在 "Remington's Pharmaceutical Sciences" 中描述了适宜载剂的实例。这类组合物应含有有效治疗量的本发明所述反转录病毒载体以及适宜量的载剂,以提供适宜施用于患者的形式。所述制剂应符合给药方式的要求。

[0047] U3 区

[0048] 正如上文讨论,所述反转录病毒基因组在其基因组的 5' 末端和 3' 末端包含长末端重复 (LTR),其含有反转录病毒生命周期期间复制所需的重要序列。U3 区的启动子和增强子元件对于反转录病毒基因组在复制期间生成全长转录物是重要的。在反转录期间,3' LTR 的一部分作为 3' LTR 以及 5' LTR 的模板,因此 3' LTR 的序列被复制到 5' LTR 中。因此,U3 区的缺失或突变被复制到 5' LTR,使得两个 U3 区基本失活。3' LTR 的这种复制允许载体序列在包装期间含有未改变的 5' U3 区,以允许从待生成和包装的 5' LTR 生成全长转录物。但是,在复制期间,由于包含增强子缺失型 U3 区的 3' U3 区被复制到 5' LTR 中,5' U3 被丢失。这样生成了包装用的全长转录物,但由于一轮复制后两个 LTR 的转录均被沉默,所述载体发生自身失活。或者,两个 LTR 的 U3 区均可以是增强子缺失型的,可使用异源启动子生成全长基因组转录物。

[0049] 本文使用的“增强子缺失型”系指这样的 U3 区,其中通过在增强子内部和 / 或周围进行缺失、添加和 / 置换使得 U3 区内的增强子整个或部分被改变,以使所述增强子基本失活。当增强子的改变足以基本消除 U3 驱动的转录物时,则认为增强子基本失活。在一个实施方式中,与内部启动子驱动基因表达相比,不到 1% 的全部反转录病毒转录物受 LTR 的 U3 区驱动。在一个实施方式中,不到 0.1% 的全部转录物受 LTR 的 U3 区驱动。在另一实施方式中,没有发现可检测的 U3 驱动的转录物。

[0050] 在本发明的一个实施方式中,通过缺失 U3 区的部分或全部增强子和启动子,增强子缺失型 U3 区基本失活。在另一实施方式中,通过在 U3 区置换或插入核苷酸,所述 U3 区基本失活。在又一实施方式中,整个 U3 区被缺失。适合用于本发明的已知反转录病毒载体以及野生型反转录病毒载体的 U3 区为本领域熟知,并可容易地辨识出来 (Coffin, JM, Fundamental Virology, pp. 798-800, Fields 等人 . eds. , 3rd Ed. , Lippincott-Raven Publ. (1996))。

[0051] 内部启动子

[0052] 本发明提供了启动子,其在增强子缺失型 U3 反转录病毒载体环境下可产生高病

毒滴度和高水平的基因表达。在一个实施方式中,所述内部启动子对载体而言是异源的,即不存在于自然界发现的载体上。测试了各种启动子,其中包括 HCMV IE 启动子 (SEQ ID NO:1)、MLV U3 区 (SEQ ID NO:2)、CMV 增强子 / 泛素启动子 (SEQ ID NO:3)、巨细胞病毒增强子 / 鸡 β -肌动蛋白 (CAG) 启动子 (SEQ ID NO:4)、人延伸因子 1 α (EF-1 α) 启动子 (SEQ ID NO:5)、人 β -肌动蛋白 (ACTB) 启动子 (SEQ ID NO:6)、人甘油醛-3-磷酸脱氢酶 (GAPDH) 启动子 (SEQ ID NO:7)、人核糖体蛋白 L10 (RPL10) 启动子 (SEQ ID NO:8)、人白细胞受体簇成员 8 (LENG8) 启动子 (SEQ ID NO:9)、人分选连接蛋白 3 (SNX3) 启动子 (SEQ ID NO:10)、人 CCR4-NOT 转录复合物,亚基 3 (CNOT3) 启动子 (SEQ ID NO:11)、human copine I (CPNE1) 启动子 (SEQ ID NO:12)、人假设蛋白 (HYPO) 启动子 (SEQ ID NO:13)、人先天性角化不良 1, 角化不良蛋白 (DKC1) 启动子 (SEQ ID NO:14)、人囊泡蛋白分选 72 (VPS72) 启动子 (SEQ ID NO:15)、整合蛋白 β -4 结合蛋白 (ITGB4BP) 启动子 (SEQ ID NO:16) 和泛醇-细胞色素 c 还原酶,复合物 III 亚基 VII (UQCRCQ) 启动子 (SEQ ID NO:17)。发现 UQCRCQ、SNX3、ITGB4BP、GAPDH、RPL10 和 LENG8 启动子可产生高病毒滴度和高水平的基因表达。这些数据表明本发明的细胞启动子可驱动增强子缺失型 U3 反转录病毒载体的基因表达。在本发明的一个实施方式中,所述内部启动子可为上面列出的启动子的片段或变体,其可产生高病毒滴度和高水平基因表达。正如本文使用,“启动子的片段”系指一段包含与天然全长启动子相同但短于天然全长启动子的多核苷酸。在一个实施方式中,所述片段具有全长启动子至少约 20% (例如约 30、40、50、60、70、80、85、90 或 95%) 的转录活性。可使用的片段包括但不限于 RPL10 启动子的片段,其包含约 -50 至约 +143 的核苷酸序列 (SEQ ID NO:8 的约 951 至 1143 位核苷酸)、约 -100 至约 +143 的核苷酸序列 (SEQ ID NO:8 的约 901 至 1143 位核苷酸)、约 -200 至约 +143 的核苷酸序列 (SEQ ID NO:8 的约 801 至 1143 位核苷酸)、约 -350 至约 +143 的核苷酸序列 (SEQ ID NO:8 的约 651 至 1143 位核苷酸)、约 -500 至约 +143 的核苷酸序列 (SEQ ID NO:8 的约 501 至 1143 位核苷酸)、约 -1000 至约 +143 的核苷酸序列 (SEQ ID NO:8 的约 1 至 1143 位核苷酸) 或约 -350 至约 +1 的核苷酸序列 (SEQ ID NO:8 的约 651 至 1001 位核苷酸) 或由上述核苷酸序列组成;以及 LENG8 启动子的片段,其包含约 -50 至约 +305 的核苷酸序列 (SEQ ID NO:9 的约 970 至 1325 位核苷酸)、约 -100 至约 +305 的核苷酸序列 (SEQ ID NO:9 的约 920 至 1325 位核苷酸)、约 -200 至约 +305 的核苷酸序列 (SEQ ID NO:9 的约 820 至 1325 位核苷酸)、约 -385 至约 +305 的核苷酸序列 (SEQ ID NO:9 的约 635 至 1325 位核苷酸)、约 -1020 至约 +305 的核苷酸序列 (SEQ ID NO:9 的约 1 至 1325 位核苷酸) 或约 -385 至 +1 的核苷酸序列 (SEQ ID NO:9 的约 635 至 1020 位核苷酸) (其中每个启动子的转录起始位点被认为是 +1)。正如本文使用,“启动子的变体”系指一段多核苷酸,其包含与天然启动子序列至少约 70% (例如至少约 80、85、90、91、92、93、94、95、96、97、98 或 99%) 相同的序列,并具有天然启动子至少 20% (例如约 30、40、50、60、70、80、85、90 或 95%) 的转录活性。启动子的变体也可是该启动子的片段。

[0053] 本文使用的“高病毒滴度”意指增强子缺失型 U3 反转录病毒载体与在两个 LTR 上均具有功能性的未改变的 U3 区的相同载体相比产生至少 10% 的含有所述载体的感染性反转录病毒颗粒。在其他实施方式中,与 U3 功能性载体相比,增强子缺失型 U3 载体的滴度至少为 20%、30%、40%、50%、60%、70%、80%、90%、100%、110%、120%、130%、140%、150%、160%、170%、180%、190%、200% 或更多。可以几种方式进行这类滴度的比较,例如

通过报道基因如 GFP 或荧光素酶在转导细胞内的表达,采用载体表达的选择性标记物选择转导细胞,或通过载体表达治疗基因。用于测量病毒滴度的许多这类方法为本领域所熟知。

[0054] “高水平异源基因转录”或“高水平基因表达”包括异源基因的转录或表达至少占反转录病毒载体产生的总转录物的 70% (例如至少 80、90、95、96、97、98 或 99%),以及异源基因的转录或表达至少是相同载体上相同基因的 10%的转录或表达,其中该相同载体在两个 LTR 均具有功能性的未改变的 U3 区。在其他实施方式中,与 U3 功能性载体上的相同基因相比,增强子缺失型 U3 载体上的异源基因的转录或表达至少为 20%、30%、40%、50%、60%、70%、80%、90%、100%、110%、120%、130%、140%、150%、160%、170%、180%、190%、200%或更多。可通过例如测量报道基因的表达或通过分析异源转录物或蛋白质的量进行这类比较,其中该分析是通过例如标准分子生物学技术如 RNA 印迹分析、RT-PCR、蛋白质印迹分析、免疫组化和酶联免疫吸附测定 (ELISA) 进行的。

[0055] 本发明的启动子可包含源自真核生物、原核生物或病毒的启动子,其足以指导位于远侧的序列 (与启动子序列 3' 末端连接的序列) 在细胞内的转录。所述内部启动子应驱动与其可操作连接的异源基因的高水平转录,同时还允许通过在生产细胞系内生成和包装基因组 RNA 实现高病毒滴度。所述启动子还可包括增强子元件。启动子和增强子由短排列的 DNA 序列组成,它们与参与转录的细胞蛋白质发生特异相互作用 (Maniatis 等人, Science 236:1237 (1987))。已从酵母、昆虫和哺乳动物细胞等各种真核生物来源和病毒 (原核生物也发现有类似调控元件) 分离得到启动子和增强子元件。特定启动子和增强子的选择取决于用于表达目的蛋白的细胞类型。

[0056] 启动子、增强子和其他调控元件可为组织特异性的或细胞特异性的。应用于调控元件的术语“组织特异性”系指可指导异源基因在特定类型的组织 (如肝脏) 内选择性表达的调控元件,其中异源基因在该特定类型组织中的表达大于相同目的核苷酸序列在不同类型组织 (如肺) 中的表达。正如本文使用,“组织特异性” (例如肝脏特异性) 是不需要绝对表达特异性的相对术语。换言之,术语“组织特异性”不需要一种组织有极高水平的表达而另一组织没有表达。一种组织的表达大于另一种组织即可。相反,“严格”或“绝对”的组织特异性表达意指在单一组织类型 (如肝脏) 内的表达,而其他组织没有可检测的表达。同样,应用于调控元件的术语“细胞类型特异性”系指可指导目的核苷酸序列在特异类型的细胞中选择性表达的调控元件,其中目的核苷酸序列在该特异类型细胞中的表达大于相同目的核苷酸序列在相同组织内不同类型细胞 (例如过度增殖细胞如癌细胞) 中的表达。应用于调控元件的术语“细胞类型特异性”还意指可促进目的核苷酸序列在单一组织的某个区域选择性表达的调控元件。

[0057] 尽管本发明设想了可产生高水平转录以及允许高病毒滴度的任何内部启动子,在一个实施方式中,所述内部启动子是细胞启动子。在另一实施方式中,所述内部启动子选自 RPL10 启动子 (SEQ ID NO:8)、LENG8 启动子 (SEQ ID NO:9)、SNX3 启动子 (SEQ ID NO:10)、UQCRCQ 启动子 (SEQ ID NO:17) 或 ITGB4BP 启动子 (SEQ ID NO:16)。在另一实施方式中,所述内部启动子是全长启动子的变体,可驱动异源基因的高水平转录,而包含该启动子的载体可产生高病毒滴度。在另一实施方式中,所述内部启动子基本由 TATA 盒组成。在一些实施方式中,所述内部启动子还包含一个或多个用于高水平基因表达的剪接位点。

[0058] 为了降低内部启动子引起的插入激活的概率,可使用绝缘子序列阻断内部启动子

对邻近基因的活化效应 (Ramezani 等人, Mol. Ther. 14 :245 (2006))。另一种方法是通过插入附加的多聚腺苷酸化信号来修饰载体,以抑制内部启动子的通读 (Ramezani 等人, Mol. Ther. 14 :245 (2006))。

[0059] 异源基因

[0060] 术语“可操作连接”用于描述基因序列与启动子或其他调控或加工序列的连接,该连接使得该基因序列的转录受可操作连接的启动子序列的指导。

[0061] 术语“基因”系指包含产生多肽或转录物所必需的调控和编码序列的 DNA 序列。该术语还包括结构基因的编码区域,并包括距离编码区域 5' 末端和 3' 末端各端约 1kb 或以上的邻近序列,使得该基因对应于全长 mRNA 的长度。位于所述编码区域 5' 末端并出现在 mRNA 上的所述序列被称为 5' 末端未翻译序列。位于所述编码区域 3' 末端或下游并出现在 mRNA 上的所述序列被称为 3' 末端未翻译序列。术语“基因”同时包括基因的 cDNA 和基因组形式。基因的基因组形式或克隆含有被名为“内含子”或“干扰区域”或“干扰序列”的非编码序列中断的编码区域。内含子是被转录到核 RNA (hnRNA) 上的基因片段;内含子可含有调控元件如增强子。内含子从核转录物或初始转录物上被除去或“剪切掉”,因此信使 RNA (mRNA) 转录物不含内含子。mRNA 在翻译期间起着确定新生多肽的氨基酸序列或次序的功能。

[0062] 本文可互换使用的术语“多核苷酸”或“核酸分子”系指任意长度例如两个或多个的核苷酸聚合物,并包括 DNA 和 RNA。核苷酸可为脱氧核糖核苷酸、核糖核苷酸、核苷酸类似物(包括修饰的磷酸盐部分、碱基或糖)、或可被适宜酶如 DNA 聚合酶或 RNA 聚合酶掺入到聚合物中的任何底物。

[0063] 正如会被本领域技术人员所理解,插入的目的多核苷酸的核苷酸序列可为任意的核苷酸序列。例如,多核苷酸序列可为报道基因序列或选择性标记基因序列。正如本文使用,报道基因序列是表达时可产生其存在或活性可被监测的蛋白质的任意基因序列。报道基因的实例包括但不限于荧光素酶(例如参见 deWet 等人, Mol. Cell. Biol. 7 :725 (1987) 和美国专利 6,074,859、5,976,796、5,674,713、和 5,618,682;所有这些通过引用并入本文)、绿色荧光蛋白(例如 GenBank 收录号 U43284;可从 CLONTECH Laboratories, Palo Alto, Calif 购买得到的许多 GFP 变体)、氯霉素乙酰转移酶、 β -半乳糖苷酶、碱性磷酸酶和辣根过氧化物酶。或者,所述报道基因序列可为表达产生影响细胞生理的基因产物的任意基因序列。本发明的多核苷酸序列可包括一个或多个已具有一个或多个启动子、起始序列或加工序列的基因序列。

[0064] 报道基因序列可为选择性标记物,它是可表达蛋白质的任意基因序列,其中该蛋白质的存在允许选择性增殖含有该蛋白质的细胞。选择性标记物可为“显性的”,显性选择性标记物编码可在任何真核细胞系检测到的酶活性。显性选择性标记物的实例包括但不限于赋予哺乳动物细胞抗药物 G418 抗性的细菌氨基葡萄糖 3' 磷酸转移酶基因(也被称为 neo 基因)、赋予抗潮霉素抗性的细菌潮霉素 G 磷酸转移酶 (hyg) 基因和赋予在霉酚酸存在下生长的能力的细菌黄嘌呤鸟嘌呤磷酸核糖转移酶(也被称为 gpt 基因)。其他选择性标记物不是显性的,因为它们必须与缺乏相关酶活性的细胞系联用。非显性选择性标记物的实例包括与 tk⁻ 细胞系联用的胸苷激酶 (tk) 基因、与 CAD 缺陷型细胞联用的 CAD 基因和与 hpRT⁻ 细胞系联用的哺乳动物次黄嘌呤鸟嘌呤磷酸核糖转移酶 (hpRT) 基因。Sambrook, J 等

人在Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory Press, New York pp. 16.9-16.15 (1989) 中提供了用于哺乳动物细胞系的选择性标记物的综述。

[0065] 报道基因序列足以允许识别或选择正常细胞内的载体。在本发明的一个实施方式中,所述报道基因可编码哺乳动物细胞通常缺乏的酶或其他蛋白质,因此它们的存在可最终确定载体在该细胞中的存在。

[0066] 本发明所述的反转录病毒载体用于将异源基因掺入病毒颗粒,从而提供了一种扩增含有异源核酸的受感染宿主细胞数目的方法。异源基因的掺入促进了该异源基因在病毒颗粒内的复制以及病毒颗粒内异源转录物或蛋白质的随后生成。如果一个基因没有天然存在于用于递送该基因至细胞的野生型载体上,即被称为异源的。正如本文使用,术语“异源基因”意指核酸分子。

[0067] 所述异源基因还可包括所需产物的编码序列,例如适宜的生物活性蛋白质或多肽、免疫原性或抗原性蛋白质或多肽、或有治疗活性的蛋白质或多肽。所述多肽可补充宿主细胞内源蛋白质的缺陷表达或不表达。这类基因序列可源自许多来源,其中包括DNA、cDNA、合成DNA、RNA或它们的组合。这类基因序列可包括基因组DNA,其可包括或不包括天然存在的内含子。而且,可获得与启动子序列或多聚腺苷酸化序列结合的这类基因组DNA。本发明所述的基因序列优选为cDNA。基因组DNA或cDNA可以任意数目的方式获得。基因组DNA可通过本领域熟知的方法从适宜细胞提取和纯化。或者,mRNA可从细胞分离并以反转录或其他方法制备为cDNA。或者,所述多核苷酸序列可包含与RNA序列互补的序列,例如反义RNA序列,该反义序列可施用给个体以抑制互补多核苷酸在该个体细胞内的表达。

[0068] 所述异源基因的表达可提供免疫原性或抗原性蛋白质或多肽,以实现抗体应答。可从动物体液如血液、血清或腹水收集如此产生的抗体。

[0069] 所述异源基因还可为可转录的任何目的核酸。外源基因一般编码多肽。优选情况下所述多肽具有一些治疗益处。所述多肽可补充宿主细胞内源蛋白质的缺陷表达或不表达。所述多肽可赋予宿主细胞新的属性,例如嵌合信号转导受体,参见美国专利5,359,046。普通技术人员根据本文教导和本领域知晓的技术可确定异源基因的适宜性。例如,该技术人员会知道异源基因是否具有用于包壳的适宜大小以及该异源基因产物是否得到合适地表达。

[0070] 可用于本发明的特定异源基因不是本发明的关键之处。但是,在一个实施方式中,所述异源基因编码细胞因子、趋化因子、激素、抗体、遗传工程化的免疫球蛋白样分子、单链抗体、融合蛋白、酶、免疫共刺激分子、免疫调节分子、反义RNA、核酶、RNA外部指导序列、靶蛋白的反式显性负性突变体、毒素、条件性毒素、抗原、肿瘤抑制蛋白和生长因子、膜蛋白、血管活性蛋白和肽、抗病毒蛋白或它们的变体。

[0071] 在又一实施方式中,所述异源基因可编码多肽,其包括但不限于免疫球蛋白、促红细胞生成素、 α -干扰素、 α -1蛋白酶抑制剂、血管生成素、抗凝血酶III、 β -酸脱羧酶、人生长激素、牛生长激素、猪生长激素、人血清白蛋白、 β -干扰素、小牛肠碱性磷酸酶、囊性纤维化跨膜传导调节蛋白、因子VIII、因子IX、因子X、胰岛素、乳铁蛋白、组织纤溶酶原激活物、髓鞘碱性蛋白、胰岛素、胰岛素原、促乳素、乙型肝炎抗原,免疫球蛋白片段(例如FABs)、单克隆抗体CTLA4Ig、Tag 72单克隆抗体、Tag 72单链抗原结合蛋白、蛋白C、细胞

因子及其受体例如肿瘤坏死因子 α 和 β 、其受体及其衍生物；肾素；生长激素释放因子；甲状旁腺激素；促甲状旁腺激素；脂蛋白； α -1-抗胰蛋白酶；促卵泡激素；降钙素；黄体化激素；胰高血糖素；血管性血友病因子；心房利钠因子；肺表面活性物质；尿激酶；铃蟾肽；凝血酶；造血生长因子；脑啡肽酶；人巨噬细胞炎症蛋白 (MIP-1- α)，血清白蛋白如苗勒管抑制物质；松弛素 A 链；松弛素 B 链；松弛素原；小鼠促性腺激素相关肽； β -内酰胺酶；DNA 酶；抑制素；激活素；血管内皮生长因子 (VEGF)；激素或生长因子受体；整合蛋白；蛋白 A 或 D；类风湿因子；神经营养因子如骨源性神经营养因子 (BDNF)，神经营养因子 -3、-4、-5 或 -6 (NT-3、NT-4 和 NT-5 或 NT-6)，或神经生长因子，如 NGF- β ；血小板衍生生长因子 (PDGF)；成纤维细胞生长因子如 aFGF 和 bFGF；表皮生长因子 (EGF)；转化生长因子 (TGF) 如 TGF- α 和 TGF- β ，其中包括 TGF- β 1、TGF- β 2、TGF- β 3、TGF- β 4 或 TGF- β 5；胰岛素样生长因子 I 和 II (IGF-I 和 IGF-II)；des(1-3)-IGF-I (脑 IGF-I)、胰岛素样生长因子结合蛋白；CD 蛋白如 CD-3、CD-4、CD-8 和 CD-19；骨诱导因子；免疫毒素；骨形态生成蛋白 (BMP)；干扰素如 α -、 β - 和 γ -干扰素；集落刺激因子 (CSF)，例如 M-CSF、GM-CSF 和 G-CSF；白细胞介素 (IL)，例如 IL-1 到 IL-12；超氧化物歧化酶；T 细胞受体；表面膜蛋白；衰变加速因子；病毒抗原，例如 AIDS 包膜蛋白的一部分；运输蛋白；归巢受体；地址素；调节蛋白；抗体；嵌合蛋白例如免疫粘合素，以及任何上述所列多肽的片段或融合物。在又一实施方式中，所述异源基因编码氧化酶，特别是 NADPH 氧化酶，例如 NADPH 氧化酶的 gp91 亚基。这些蛋白的核酸和蛋白序列可从公共数据库例如 GenBank 获得。

[0072] 当特定蛋白具有多于一个的亚基时（例如免疫球蛋白），编码该序列的基因可在载体上排列成多顺反子序列，以一个或多个 IRES 元件隔开。或者，编码蛋白不同亚基的基因可在不同的载体上引入宿主细胞。根据本发明，编码目的蛋白的所述基因优选包含一个或多个内含子。内含子可为正常情况下与该基因连接的内含子，或可为合成或外源内含子。在一些实施方式中，所述基因可包含少于其正常数目的内含子。例如，可除去该基因上一些天然存在的内含子，同时保留其他内含子，或一个或多个天然存在的内含子可被一个或多个外源内含子取代。

[0073] 所谓“野生型”或天然的，意指核苷酸或氨基酸序列与天然存在的序列相同。

[0074] 所谓“变体”意指包括基本类似的序列。因此，对于核苷酸序列或氨基酸序列，变体包括功能等同的序列，例如保留至少 20%（例如 30、40、50、60、70、80 或 90%）野生型序列的一种或多种活性。变体核苷酸序列还包括通过例如定向诱变生成的但仍保留天然序列功能的合成核苷酸序列。一般而言，本发明的核苷酸序列变体或氨基酸序列变体具有与其各自天然核苷酸序列至少 70%，一般为 80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98% 或 99% 的序列相同性。

[0075] 技术人员会理解公开的核酸构建体的许多保守变动会得到功能相同的构建体。特定核酸序列的保守变动系指编码与基本相同的序列相同或基本相同的氨基酸序列的这些核酸，或所述核酸不编码氨基酸序列。由于遗传密码的简并性，许多功能相同的核酸编码任意给定的多肽。例如，由于遗传密码的简并性，“沉默置换”（例如不会引起所编码多肽变化的核酸序列的置换）是编码氨基酸的每个核酸序列的隐含特点。同样，“保守氨基酸置换”，即以性质高度类似的不同氨基酸取代包装构建体或可包装构建体的氨基酸序列的一个或几个氨基酸，也被容易地鉴定为高度类似于公开的构建体。例如，密码子 CGU、CGC、CGA、CGG、

AGA 和 AGG 都编码精氨酸。因此,在密码子所指定的每处精氨酸位置,该密码子可被变动为任何描述的对应该密码子,而不改变编码的多肽。这类核酸变动是“沉默变动”,是“保守修饰变动”的一类。本文每个编码多肽的核酸序列还描述了每种可能的沉默变动。技术人员会认识到可经标准技术修饰核酸的每个密码子(AUG 除外,它通常是甲硫氨酸的唯一密码子)以生成功能相同的分子。因此,编码多肽的核酸的每个“沉默变动”暗含在任何描述的序列中。而且,技术人员会认识到通过变更、添加或缺失编码序列的单个氨基酸或少量的氨基酸(典型情况下小于 5%,更典型情况下小于 1%)而作出的个别置换、缺失或添加是“保守修饰变动”,其中这些变更导致氨基酸被化学性质类似的氨基酸置换。提供功能类似的氨基酸的保守置换表为本技术领域熟知。下列 6 组每组含有互为保守置换的氨基酸:

[0076] 1) 丙氨酸(A)、丝氨酸(S)、苏氨酸(T);

[0077] 2) 天冬氨酸(D)、谷氨酸(E);

[0078] 3) 天冬酰胺(N)、谷氨酰胺(Q);

[0079] 4) 精氨酸(R)、赖氨酸(K);

[0080] 5) 异亮氨酸(I)、亮氨酸(L)、甲硫氨酸(M)、缬氨酸(V);以及

[0081] 6) 苯丙氨酸(F)、酪氨酸(Y)、色氨酸(W)。

[0082] 另外参见 Creighton(1984) *Proteins* W.H. Freeman and Company。最后,添加不改变核酸分子活性的序列如非功能性序列是该基本核酸的保守修饰。每个公开序列的这类保守性置换变动是本发明的一个特点。

[0083] 就用于本发明载体体系的各种全长或成熟多肽的氨基酸序列而言,变体包括源自天然多肽的这些多肽,其中通过在天然多肽的 N 末端和 / 或 C 末端缺失(即所谓的截短)或添加一个或多个氨基酸、在天然多肽上的一个或多个位点处缺失或添加一个或多个氨基酸、或在天然多肽的一个或多个位点处置换一个或多个氨基酸从天然多肽衍生得到这些多肽。这类变体可源自例如遗传多态性或人工操作。用于这类操作的方法一般为本领域知晓。

[0084] 技术人员会识别出在给定核酸构建体上产生变更的许多方式。这类熟知的方法包括定向诱变、采用简并寡核苷酸进行的 PCR 扩增、将含有所述核酸的细胞暴露于诱变剂或辐射、所需寡核苷酸的化学合成(例如联用连接和 / 或克隆以生成大核酸)和其他熟知的技术。

[0085] 天然核苷酸序列或天然多肽的变体与天然序列或天然多肽基本相同。变体的不同之处可少至 1 到 10 个氨基酸残基,例如 6-10、少至 5、少至 4、3、2 或甚至 1 个氨基酸残基。核苷酸序列的变体的不同之处可少至 1 到 30 个核苷酸,例如 6 到 20、少至 5、少至 4、3、2 或甚至 1 个核苷酸残基。

[0086] 所谓“序列相同性”意指当变体的核苷酸序列或氨基酸序列的一段指定的连续片段与参照序列的核苷酸序列或氨基酸序列进行比对和比较时,发现变体序列与参照序列有相同的核苷酸或氨基酸残基。用于序列比对的方法和用于测定序列间相同性的方法为本领域熟知。就两个核苷酸序列的最佳比对而言,与参照核苷酸序列相比,变体核苷酸序列的连续片段可具有附加的核苷酸或缺失的核苷酸。同样,就两个氨基酸序列的最佳比对的目的而言,与参照氨基酸序列相比,变体氨基酸序列的连续片段可具有附加的氨基酸残基或缺失的氨基酸残基。用于与参照核苷酸序列或与参照氨基酸序列比较的连续片段应包括至少 20 个连续核苷酸或氨基酸残基,并可为 30、40、50、100 或更多的核苷酸或氨基酸残基。可通

过指定空位罚分来校正变体核苷酸序列或氨基酸序列因含有空位而增加的序列相同性。

[0087] 细胞

[0088] 还包括了包含本发明所述表达载体的细胞。在一个实施方式中,还包括了包含增强子缺失型 U3 反转录病毒载体的细胞。以本发明所述表达载体转导的靶细胞可为不论位于体外还是体内均可被转导的任何真核细胞类型,其中包括哺乳动物细胞、禽细胞、两栖动物细胞、植物细胞、鱼细胞和昆虫细胞。在一个实施方式中,所述靶细胞是哺乳动物细胞,特别是人细胞。所述靶细胞可为原代细胞或细胞系。

[0089] 在一个实施方式中,本发明包括生产细胞系,其还包含产生含本发明载体的感染性反转录病毒颗粒必需的反式病毒蛋白。这类生产细胞系产生反转录病毒的壳体,并转录反转录病毒载体,随后通过包装信号将该反转录病毒载体募集到病毒壳体。构建生产细胞系所需的包装细胞系为本领域所知晓,一般包含反转录病毒的 gag-pol 和 env 基因,这些基因提供了感染性反转录病毒颗粒所需的酶(例如反转录酶)和结构蛋白(例如 Gag 和 Env)。已知许多这类包装细胞,例如 PG13、ΨCRIP、PA317、GP+envAm12、FLYA13、FLYRD18、Phoenix-Ampho、Phoenix-Eco、Phoenix-GALV、PE501、GP+E86、PT67、BING、BOSC23、ProPak-A 和其他包装细胞以及慢病毒包装细胞系(Logan 等人, J Virol. 78 :8421-8436(2004))。而且,包装细胞系可被瞬时转染以供短期使用,或具有整合到其基因组上的病毒基因以供长期使用。

[0090] 许多包装细胞系采用反转录病毒的天然包膜蛋白,所述反转录病毒载体即基于反转录病毒。还有可能通过使用来自另一密切相关病毒的包膜蛋白基因改变本发明所述病毒载体可感染的细胞宿主范围。换言之,通过利用某些病毒包膜蛋白可参与其他病毒包壳的能力,有可能扩大本发明所述反转录病毒载体的宿主范围。源自反转录病毒的基因的实例包括但不限于:水泡性口炎病毒的 G 蛋白(VSV-G)、长臂猿白血病病毒(GaLV)、猫内源病毒 RD114、劳斯肉瘤病毒(RSV)、双嗜性莫洛尼鼠白血病病毒(MoMuLV)、亲嗜性莫洛尼鼠白血病病毒(MoMuLV)、10A1 鼠白血病病毒、莫洛尼貂致细胞灶病毒(MCFV)、Mus dunni 内生性病毒(MDEV)、小鼠乳腺癌病毒(MMTV)和人免疫缺陷病毒(HIV)。所有这些病毒包膜蛋白有效地形成具有其他病毒基因组和基质组分的假型病毒体。正如本文使用,术语“假型”系指含有一种病毒的核酸和另一种病毒包膜蛋白的病毒颗粒。一般而言,VSV-G 或 GaLV 假型载体具有很宽的宿主范围,可通过超速离心沉淀成高浓度的滴度,同时仍保留高水平的感染性。

[0091] 本发明的方法

[0092] 本发明还提供了产生包含本发明所述反转录病毒载体的感染性反转录病毒颗粒的方法,其包括培养上述的生产细胞系、收集细胞培养的上清以及过滤上清以获得不含细胞的病毒上清。本领域技术人员应可容易地优化获得良好病毒滴度的条件,例如使用适宜的培养基,确定最佳的收集时间段和培养物的细胞密度。

[0093] 本文还提供了一种转导靶细胞的方法,其包括以上述制备的包含依据本发明的感染性反转录病毒颗粒的病毒上清接触所述细胞。正如上述,所述靶细胞可为但不限于为哺乳动物细胞、人细胞、灵长类细胞或鼠类细胞。所述靶细胞可为原代细胞或细胞系。所述方法还可包括加入物质以增加对病毒上清的转导,例如聚凝胺、Retronectin 和 / 或硫酸精蛋白。此外,所述方法还可包括施加病毒上清后低速离心细胞。这些以及其他转导优化技术

是本领域熟知且常规的技术。

[0094] 本发明还提供了一种通过施用以本发明所述反转录病毒载体转导的细胞来治疗对象的方法,其中所述异源基因编码可用于治疗的多肽或转录物。在一个实施方式中,体外或离体转导细胞。

[0095] 本发明还提供了一种治疗对象的方法,其包括施用包含本发明所述表达载体和适宜载剂的组合物,其中所述异源基因编码可用于治疗的多肽或转录物。在一个实施方式中,所述目的核酸编码治疗剂。术语“可用于治疗的”取其普通意义,包括处理剂、预防剂和置换剂。如果治疗剂改善或预防疾病或疾患的至少一种症状,则该治疗剂可视为有治疗效果。可以本发明载体和/或方法治疗的遗传疾病包括期望长期表达治疗性核酸的这些疾病。在另一实施方式中,所述方法用于治疗遗传疾病、增殖疾病或感染性疾病。另一实施方式包括治疗一种或多种疾病、障碍或疾患的方法,这些疾病、障碍或疾患包括但不限于:神经疾病、免疫系统疾病、肌肉疾病、生殖疾病、胃肠道疾病、肺疾病、心血管疾病、肾疾病、增殖疾病和/或癌性疾病或疾患。另外的实施方式包括治疗需取代、上调或下调目的基因的神经退行性疾病或病症、阿尔茨海默氏病、精神分裂症、癫痫、新生物、癌和 AIDS 或其他疾病。

[0096] 施用本发明所述表达载体的方法包括但不限于皮内、肌内、腹膜内、静脉内、皮下、鼻内、硬膜外和口服途径。所述载体或组合物可以任何方便的途径施用,例如输注或弹丸注射、经上皮或粘膜层(例如口腔粘膜、直肠和小肠粘膜等)吸收,并可与其他生物活性剂一同施用。施用可为全身或局部的。此外,通过任何适宜的途径向中枢神经系统引入本发明所述载体或组合物是理想的,这些途径包括脑室内注射或鞘内注射;可以例如连接到囊(reservoir)如 Ommaya 囊的脑室内导管促进脑室内注射。也可使用肺部给药,例如通过使用吸入器或雾化器和含气雾剂的制剂。

[0097] 试剂盒

[0098] 本发明的另一目的是提供一种包含用于上述方法的载体的试剂盒或药物输送系统。施用靶向反转录病毒颗粒所需的全部必需材料和试剂可组装在一个试剂盒内(例如包装细胞构建体或细胞系)。该试剂盒的组分可以多种配方提供。本发明所述的一种或多种反转录病毒载体可与一种或多种药剂(例如化学治疗剂)配制成单一的药用组合物或不同的药用组合物。

[0099] 这些试剂盒或药物输送系统的组分还可为干燥或冻干形式。当试剂或组分以干燥形式提供时,一般通过加入也可提供在另一容器装置中的适宜溶剂进行重建。本发明所述试剂盒还可包括剂量和/或给药信息方面的说明。本发明所述的试剂盒或药物输送系统一般还包括一种商业销售用的封闭容纳小瓶的装置,例如容纳所需小瓶的注射容器或吹塑塑料容器。不管使用多少容器或使用何种类型容器,所述试剂盒还可包含或包装有一种辅助注射/施用或置放最终复合物到对象体内的仪器。这类仪器可为敷料器、吸入器、注射器、吸量管、镊子、量匙、滴眼瓶或准予在医学上使用的任何输送媒介。

[0100] 必须注意的是,正如本说明书和所附权利要求所使用,除非上下文清楚表明其他含义,单数形式的“一”(a, an)、“所述(the)”等包括复数引用。因此,例如提及“一多核苷酸”包括多个多核苷酸,而提及“一个细胞”则包括多个细胞。

[0101] 正如本文使用,术语“体外”系指人工环境和发生在人工环境内的过程或反应。体外环境可由试管和细胞培养组成,但不限于此。术语“体内”系指天然环境(例如动物或细

胞)和发生在天然环境内的过程或反应。

[0102] 下列实施例用于例示而非限制本发明所述方法。对医学治疗和制药业通常碰到的各种条件和参数进行的其他适宜修改和改动对本领域技术人员是显而易见的,它们位于本发明的精神和范围内。

[0103] 实施例 1

[0104] 本实施例提供了各种启动子的构建过程。

[0105] 1) 人巨细胞病毒 (HCMV) 即时早期 (IE) 启动子 (SEQ ID NO :1)

[0106] HCMV IE 启动子得自 pCN 质粒 (Lee 等人., Biochem. Biophys. Res. Commun. 272 : 230 (2000))。

[0107] 2) MLV (SEQ ID NO :2) 的 U3

[0108] MLV LTR 的 U3 区含有强增强子和启动子序列。以 MLV 载体 MT 作为模板经 PCR 扩增 MLV 3' LTR 的 U3 区 (Hong 等人, J. Gene Med. 6 :724 (2004) ;美国专利 6,451,595)。下列引物对被用于 PCR。

[0109] ME5 :ACGCGTGCAAGGCATGGAAAAA (SEQ ID NO :18)

[0110] MluI

[0111] MP3 :ACGCGTAGATCTGAATTCTACCCGGGCGACGCAGT (SEQ ID NO :19)

[0112] MluI BglIII EcoRI

[0113] 采用 Expend 高保真 PCR 体系 (Cat# 92351824, Roche) 对含有 200ng 质粒 DNA 模板和每种引物各 1 μ l (10pmol/ μ l) 的 100 μ l PCR 反应液进行 35 个循环的 PCR 扩增反应。每次循环条件 :95℃变性 30 秒,55℃退火 30 秒,72℃聚合 30 秒。

[0114] 扩增出的 455bp 片段被克隆到 pGEM T easy 载体上 (Cat#A1360, Promega, WI, USA),生成 pGEM T-MTU3。

[0115] 3) CMV/ 泛素启动子 (SEQ ID NO :3)

[0116] A. CMV 增强子

[0117] 以 pCK (PCT/KR99/00855) 作为模板经 PCR 扩增 CMV 增强子。下列引物对被用于 PCR。

[0118] CMV5 :ACGCGTTGACATTGATTATTG (SEQ ID NO :20)

[0119] MluI

[0120] KMD1 :TCTAGAGCCAAAACAACTCCCAT (SEQ ID NO :21)

[0121] XbaI

[0122] 采用 Expend 高保真 PCR 体系对含有 200ng 质粒 DNA 模板和每种引物各 2 μ l (5pmol/ μ l) 的 50 μ l PCR 反应液进行 30 个循环的 PCR 扩增反应。每次循环条件 :94℃变性 1 分钟,55℃退火 1 分钟,72℃聚合 1 分钟。

[0123] 扩增片段被克隆到 pGEM T easy 载体上,生成 pGEM T-Enh。核苷酸序列经测序确认。

[0124] B. 人聚泛素 C 启动子 (Gill 等人, Gene Ther. 8 :1539 (2001))

[0125] 以分离自 HT1080 细胞的基因组 DNA 作为模板扩增人聚泛素 C 启动子 (-333 ~ +877)。下列引物对被用于 PCR。

[0126] KMD4 :GCTAGCGGCTCCGCGCCGGGTTT (SEQ ID NO :22)

[0127] NheI

[0128] KMD5 :ACGCGTAGATCTGAATTCGTCTAACAAAAAGCCAA (SEQ ID NO :23)

[0129] MluI BglII EcoRI

[0130] 采用 Expend 高保真 PCR 体系对含有 200ng DNA 模板和每种引物各 2 μ l (5pmol/ μ l) 的 50 μ l PCR 反应液进行 30 个循环的 PCR 扩增反应。每次循环条件 :94℃变性 1 分钟,55℃退火 1 分钟,72℃聚合 1 分半钟。

[0131] 扩增出的 1230bp 片段被克隆到 pGEM T easy 载体上,生成 pGEMT-UbC。核苷酸序列经测序确认。

[0132] C. CMV 增强子 /UbC 启动子

[0133] 为了构建由 CMV 增强子和 UbC 启动子组成的杂合启动子,从 pGEM TEasy-Enh 切割下 SalI-XbaI 片段,将其插入到 pGEM T Easy-UbC 的 SalI-XbaI 位点,以生成 pGEM T Easy-Enh+UbC。

[0134] 4) CAG (巨细胞病毒增强子,鸡 β -肌动蛋白启动子) 启动子 (SEQ IDNO :4)

[0135] 为了获得 CAG 启动子 (巨细胞病毒增强子,鸡 β -肌动蛋白启动子) 启动子 (SEQ ID NO :4),将来自 pAxCawt (Takara Bio, Otsu, Japan) 以 Klenow 片段处理的 SalI-SwaI 片段克隆到 pGEM T easy (Promega, WI, USA),以生成 pGEM T easy-CAG。核苷酸序列经测序确认。

[0136] 5) 人 延 伸 因 子 1 α (EF1- α) 启 动 子 (SEQ ID NO :5) (Kim 等 人, Gene 91 : 217 (1990))

[0137] 以分离自 HT1080 细胞 (人纤维肉瘤细胞系, ATCC CCL-121) 的基因组 DNA 作为模板扩增人延伸因子 1 α (EF1- α) 启动子 (-341 ~ +1007)。用于 PCR 的引物对的核苷酸序列如下 :

[0138] EEF1A1F :ACGCGTGTAAGCCAGCAATGGTAGAGGGAAGATTCTGCACG

[0139] MluI (SEQ ID NO :24)

[0140] EEF1A1R :GGATCCTTTTGGCTTTTAGGGGTAGTTTTACGACACC

[0141] BamHI (SEQ ID NO :25)

[0142] 采用 Expend 高保真 PCR 体系 (Cat#92351824, Roche) 对含有 500ng 基因组 DNA 模板、各 1 μ l 的每种引物 (10pmol/ μ l) 和 550 μ l PCR 反应液进行 30 个循环的 PCR 扩增反应。每次循环条件 :95℃变性 1 分钟,55℃退火 1 分钟,72℃聚合 1 分半钟。

[0143] 扩增片段起初被克隆到 pGEM T easy 上以生成 pGEM T easy-EF。核苷酸序列经测序确认。

[0144] 6) 人 β -肌动蛋白启动子 (SEQ ID NO :6) (Nakajima-Iijima 等人, Proc. Natl. Acad. Sci. USA 82 :6133 (1985) ;Miyamoto, Nucleic Acids Res. 15 :9095 (1987))

[0145] 以分离自 K562 细胞 (人髓系细胞系, ATCC CCL-243) 的基因组 DNA 作为模板扩增人 β -肌动蛋白启动子 (-387 ~ +944)。用于 PCR 的引物对的核苷酸序列如下 :

[0146] BApF :ACGCGTGAGATGTCCACACCTAGGATGTCC (SEQ ID NO :26)

[0147] MluI

[0148] BApR :GGATCCGGTGAGCTGCGAGAATAGCCG (SEQ ID NO :27)

[0149] BamHI

[0150] 采用 Expend 高保真 PCR 体系对含有 500ng 基因组 DNA 模板、每种引物各 1 μ l (10pmol/ μ l) 和 5 μ l DMSO 的 50 μ l PCR 反应液进行 30 个循环的 PCR 扩增反应。每次循环条件:95℃变性 1 分钟,55℃退火 1 分钟,72℃聚合 1 分半钟。

[0151] 扩增片段被克隆到 pGEM T easy 上,生成 pGEM T easy-BA。核苷酸序列经测序确认。

[0152] 7) 人甘油醛-3-磷酸脱氢酶 (GAPDH) 启动子 (SEQ ID NO:7) (Ercolani 等人, J. Biol. Chem. 263:15335(1988))

[0153] 以分离自 HT1080 细胞的基因组 DNA 作为模板扩增人 GAPDH(甘油醛-3-磷酸脱氢酶)启动子(-350 ~ +315)。用于 PCR 的引物对的核苷酸序列如下:

[0154] GAPDHF: ACGCGTTTCATCCAAGCGTGAAGGG (SEQ ID NO:28)

[0155] MluI

[0156] GAPDHR: GTTTAAACGGTGTCTGAGCGATGTGGCT (SEQ ID NO:29)

[0157] PmeI

[0158] 采用 Expend 高保真 PCR 体系对含有 500ng 基因组 DNA 模板、每种引物各 1 μ l (10pmol/ μ l) 和 5 μ l DMSO 的 50 μ l PCR 反应液进行 30 个循环的 PCR 扩增反应。每次循环条件:95℃变性 1 分钟,55℃退火 1 分钟,72℃聚合 1 分半钟。

[0159] 扩增片段被克隆到 pGEM T easy 上,生成 pGEM T easy-GAPDH。核苷酸序列经测序确认。

[0160] 8) 人核糖体蛋白 L10(RPL10) 启动子 (SEQ ID NO:8) (NCBI 收录号:NM_006013, NT_011726;Bignon 等人, Biochem. Biophys. Res. Commun. 184:1165(1992))

[0161] 以分离自 HT1080 细胞的基因组 DNA 作为模板扩增人 RPL10(核糖体蛋白 L10)启动子(-350 ~ +143) (SEQ ID NO:8 的 651 到 1143 位核苷酸)。用于 PCR 的引物对的核苷酸序列如下:

[0162] RPLF: ACGCGTAGGCCCACCTAGGGTACTTTCCTTT (SEQ ID NO:30)

[0163] MluI

[0164] RPLR: GGATCCGGCGACACCAGGATCTTCAGTGGCT (SEQ ID NO:31)

[0165] BamHI

[0166] 采用 Expend 高保真 PCR 体系对含有 500ng 基因组 DNA 模板、每种引物各 1 μ l (10pmol/ μ l) 和 5 μ l DMSO 的 50 μ l PCR 反应液进行 30 个循环的 PCR 扩增反应。每次循环条件:95℃变性 1 分钟,55℃退火 1 分钟,72℃聚合 1 分半钟。

[0167] 扩增片段被克隆到 pGEM T easy 上以生成 pGEM T easy-RPL。核苷酸序列经测序确认。

[0168] 9) 人白细胞受体簇成员 8 (LENG8) 启动子 (SEQ ID NO:9) (NCBI 收录号:AL834532, NT_011109;Cooper 等人, Genome Res. 16:1(2006))

[0169] 以分离自 HT1080 细胞的基因组 DNA 作为模板扩增人 LENG8(白细胞受体簇 (LRC) 成员 8) 启动子(-385 ~ +305, +1908 ~ +2121) (SEQ ID NO:9 的 635 到 1538 位核苷酸)。用于 PCR 的引物对的核苷酸序列如下:

[0170] LENG8F1: ACGCGTAGAATTGTTTGAACCCAGGAGGCGG (SEQ ID NO:32)

[0171] MluI

[0172] LENG8R1 :GTTTAAACAAAAGTAGAAGACGACGGCGCACGCG

[0173] PmeI (SEQ ID NO :33)

[0174] LENG8F2 :GTTTAAACCCCACACCCAGAACTCTTCAGATCCT

[0175] PmeI (SEQ ID NO :34)

[0176] LENG8R2 :GAATTCCTGGACCTTGGGGTATAAGGGGTGG (SEQ ID NO :35)

[0177] EcoRI

[0178] 采用 Expend 高保真 PCR 体系对含有 500ng 基因组 DNA 模板、每种引物各 1 μ l (10pmol/ μ l) 和 5 μ l DMSO 的 50 μ l PCR 反应液进行 30 个循环的 PCR 扩增反应。每次循环条件 :95℃变性 1 分钟,55℃退火 1 分钟,72℃聚合 1 分半钟。

[0179] 扩增片段起初被克隆到 pGEM T easy 上,分别生成 pGEM Teasy-LENG1 和 pGEM Teasy-LENG2。确认其核苷酸序列后,pGEM Teasy-LENG1 的 MluI-PmeI 片段被克隆到 pGEM Teasy-LENG2 的 MluI-PmeI 位点上,以生成 pGEM T easy-LENG8。

[0180] 10) 人分选连接蛋白 3 (SNX3) 启动子 (SEQ ID NO :10) (NCBI 收录号 :NM 152828, NT 025741 ;Haft 等人, Mol. Cell. Biol. 18 :7278-87 (1998))

[0181] 以分离自 HT1080 细胞的基因组 DNA 作为模板扩增人 SNX3 (分选连接蛋白 3) 启动子 (-353 ~ +338)。用于 PCR 的引物对的核苷酸序列如下 :

[0182] SNX3F :GAATTCAATCCAGACGCGTGTCTGGTGCAA (SEQ ID NO :36)

[0183] EcoRI

[0184] SNX3R :GGATCCTTCGCTGTAGCTGCTG (SEQ ID NO :37)

[0185] BamHI

[0186] 采用 Expend 高保真 PCR 体系对含有 500ng 基因组 DNA 模板、每种引物各 1 μ l (10pmol/ μ l) 和 5 μ l DMSO 的 50 μ l PCR 反应液进行 30 个循环的 PCR 扩增反应。每次循环条件 :95℃变性 1 分钟,55℃退火 1 分钟,72℃聚合 1 分半钟。

[0187] 扩增片段被克隆到 pGEM T easy 上以生成 pGEM T easy-SNX。核苷酸序列经测序确认。

[0188] 11) 人 CNOT3 (SEQ ID NO :11) (CCR4-NOT 转录复合物,亚基 3) 启动子 (NCBI 收录号 :NM 014516 ;Albert 等人, Nucleic Acids Res. 28 :809 (2000))

[0189] 以分离自 HT1080 细胞的基因组 DNA 作为模板扩增人 CNOT3 (CCR4-NOT 转录复合物,亚基 3) 启动子 (-350 ~ +654, +5076 ~ +5266)。用于 PCR 的引物对的核苷酸序列如下 :

[0190] CNOT3F1 :ACGCGTGTAGCTCCTCCCCAGACCAATTGTTTAAAG

[0191] MluI (SEQ ID NO :38)

[0192] CNOT3R1 :GGATCCTCCATCCTTCCAGCCAGGAGCCAATACCGAC

[0193] BamHI (SEQ ID NO :39)

[0194] CNOTF2 :AGATCTTGGGGCTGGTCTCTTGTCAGATAGC (SEQ ID NO :40)

[0195] BglIII

[0196] CNOTR2 :GGATCCCTTCCCTGCCCTACAGACGCACTCT (SEQ ID NO :41)

[0197] BamHI

[0198] 采用 Expend 高保真 PCR 体系对含有 500ng 基因组 DNA 模板、每种引物各 1 μ l (10pmol/ μ l) 和 5 μ l DMSO 的 50 μ l PCR 反应液进行 30 个循环的 PCR 扩增反应。每

次循环条件:95℃变性 1 分钟,55℃退火 1 分钟,72℃聚合 1 分半钟。

[0199] 扩增片段起初被克隆到 pGEM T easy 上,分别生成 pGEM Teasy-CNOT1 和 pGEM Teasy-CNOT2。确认其核苷酸序列后,pGEM Teasy-CNOT 1 的 MluI-BamHI 片段被克隆到 pGEM Teasy-CNOT2 的 MluI-BglIII 位点上,以生成 pGEM T easy-CNOT3。

[0200] 12) 人 CPNE1(copine I) 启动子 (SEQ ID NO:12) (NCBI 收录号:NM 152926;Creutz 等人, J. Biol. Chem. 273:1393(1998))

[0201] 以分离自 HT1080 细胞的基因组 DNA 作为模板扩增人 CPNE1(copine I) 启动子 (-300 ~ +489, +5612 ~ +5999)。用于 PCR 的引物对的核苷酸序列如下:

[0202] CPNE1F1: ACGCGTGTCCATTTAATCCTCAAAAACTTA (SEQ ID NO:42)

[0203] MluI

[0204] CPNE1R1: GGATCCTTTTTACTGCAGTCCCCGTTATTAGCTC

[0205] BamHI (SEQ ID NO:43)

[0206] CPNE1F2: AGATCTAGCTGTGAAGCTGAGCTTTATGACT (SEQ ID NO:44)

[0207] BglIII

[0208] CPNE1R2: GGATCCCTGATAAACAAGAGATGAATTTCC (SEQ ID NO:45)

[0209] BamHI

[0210] 采用 Expend 高保真 PCR 体系对含有 500ng 基因组 DNA 模板、每种引物各 1 μl (10pmol/μl) 和 5 μl DMSO 的 50 μl PCR 反应液进行 30 个循环的 PCR 扩增反应。每次循环条件:95℃变性 1 分钟,55℃退火 1 分钟,72℃聚合 1 分半钟。

[0211] 扩增片段起初被克隆到 pGEM T easy 上,分别生成 pGEM Teasy-CPNEF 和 pGEM Teasy-CPNER。确认其核苷酸序列后,pGEM Teasy-CPNEF 的 MluI-BamHI 片段被克隆到 pGEM Teasy-CPNER 的 MluI-BglIII 位点上,以生成 pGEM T easy-CPNE1。

[0212] 13) 人 HYPO(假设蛋白) 启动子 (SEQ ID NO:13) (NCBI 收录号:AF351613)

[0213] 以分离自 HT1080 细胞的基因组 DNA 作为模板扩增人 HYPO(假设蛋白) 启动子 (-350 ~ +66)。用于 PCR 的引物对的核苷酸序列如下:

[0214] HYPOF: ACGCGTTCTTTTACACGTTTGGTTTATGGT (SEQ ID NO:46)

[0215] MluI

[0216] HYPOR: GGATCCGGCTGCAACAGGCCAGGAAACCTTC (SEQ ID NO:47)

[0217] BamHI

[0218] 采用 Expend 高保真 PCR 体系对含有 500ng 基因组 DNA 模板、每种引物各 1 μl (10pmol/μl) 和 5 μl DMSO 的 50 μl PCR 反应液进行 30 个循环的 PCR 扩增反应。每次循环条件:95℃变性 1 分钟,55℃退火 1 分钟,72℃聚合 1 分半钟。

[0219] 扩增片段起初被克隆到 pGEM T easy 上以生成 pGEM T easy-HYPO。核苷酸序列经测序确认。

[0220] 14) 人 DKC1(先天性角化不良 1,角化不良蛋白) 启动子 (SEQ ID NO:14) (NCBI 收录号:BC 009928;Strausberg 等人, Proc. Natl. Acad. Sci. USA99:16899(2002))

[0221] 以分离自 HT1080 细胞的基因组 DNA 作为模板扩增人 DKC1(先天性角化不良 1,角化不良蛋白) 启动子 (-473 ~ +91)。用于 PCR 的引物对的核苷酸序列如下:

[0222] DKC1F: ACGCGTGCACTACTCCTATTGGC (SEQ ID NO:48)

[0223] MluI

[0224] DKC1R :GAATTCGTTACCCTGCACCGCGTGC (SEQ ID NO :49)

[0225] EcoRI

[0226] 采用 Expend 高保真 PCR 体系对含有 500ng 基因组 DNA 模板、每种引物各 1 μ l (10pmol/ μ l) 和 5 μ l DMSO 的 50 μ l PCR 反应液进行 30 个循环的 PCR 扩增反应。每次循环条件 :95℃变性 1 分钟,55℃退火 1 分钟,72℃聚合 1 分半钟。

[0227] 扩增片段起初被克隆到 pGEM T easy 上以生成 pGEM T easy-DKC1。核苷酸序列经测序确认。

[0228] 15) 人 VPS72(囊泡蛋白分选 72) 启动子 (SEQ ID NO :15) (NCBI 收录号 :NM_005997 ;Horikawa 等人, Biochem. Biophys. Res. Commun. 208 :999(1995))

[0229] 以分离自 HT1080 细胞的基因组 DNA 作为模板扩增人 VPS72(囊泡蛋白分选 72) 启动子 (-466 ~ +43)。用于 PCR 的引物对的核苷酸序列如下 :

[0230] VPS72F :ACGCGTACAAAAATTAGTTGGGCAT (SEQ ID NO :50)

[0231] MluI

[0232] VPS72R :GAATTCACCGCCTACCGAGACTGCG (SEQ ID NO :51)

[0233] EcoRI

[0234] 采用 Expend 高保真 PCR 体系对含有 500ng 基因组 DNA 模板、每种引物各 1 μ l (10pmol/ μ l) 和 5 μ l DMSO 的 50 μ l PCR 反应液进行 30 个循环的 PCR 扩增反应。每次循环条件 :95℃变性 1 分钟,55℃退火 1 分钟,72℃聚合 1 分半钟。

[0235] 扩增片段起初被克隆到 pGEM T easy 上以生成 pGEM T easy-VPS72。核苷酸序列经测序确认。

[0236] 16) 人 ITGB4BP(InTeGrin Beta 4 结合蛋白) 启动子 (SEQ ID NO :16) (NCBI 收录号 :BC011845, NT_028392 ;Strausberg 等人., Proc. Natl. Acad. SciUSA 99 :16899(2002))

[0237] 以分离自 HT1080 细胞的基因组 DNA 作为模板扩增人 ITGB4BP(InTeGrin Beta 4 结合蛋白) 启动子 (-350 ~ +304)。用于 PCR 的引物对的核苷酸序列如下 :

[0238] ITGB4BPF :ACGCGTTCTGTCCCTCAAGGCACAGCT (SEQ ID NO :52)

[0239] MluI

[0240] ITGB4BPR :GTTTAAACGAGGCCTAGGGGCGGCGGAGGCGGGAGTTCAA

[0241] PmeI (SEQ ID NO :53)

[0242] 采用 Expend 高保真 PCR 体系对含有 500ng 基因组 DNA 模板、每种引物各 1 μ l (10pmol/ μ l) 和 5 μ l DMSO 的 50 μ l PCR 反应液进行 30 个循环的 PCR 扩增反应。每次循环条件 :95℃变性 1 分钟,55℃退火 1 分钟,72℃聚合 1 分半钟。

[0243] 扩增片段起初被克隆到 pGEM T easy 上以生成 pGEM Teasy-ITGB4BP。核苷酸序列经测序确认。

[0244] 17) 人 UQCRQ(泛醇-细胞色素 c 还原酶,复合物 III,亚基 VII) promoter (SEQ ID NO :17) (NCBI 收录号 :BC090048, NT_034772 ;Strausberg 等人, Proc. Natl. Acad. Sci. USA 99 :16899(2002))

[0245] 以分离自 HT1080 细胞的基因组 DNA 作为模板扩增人 UQCRQ(泛醇-细胞色素 c 还原酶,复合物 III,亚基 VII) 启动子 (-350 ~ +217)。用于 PCR 的引物对的核苷酸序列如

下：

[0246] UQCRQF :ACGCGTGTCACCTTTTGTTCCTCCC (SEQ ID NO :54)

[0247] MluI

[0248] UQCRQR :GTTTAAACTGTGGCGGCGGCCCTGCAGG (SEQ ID NO :55)

[0249] PmeI

[0250] 采用 Expend 高保真 PCR 体系对含有 500ng 基因组 DNA 模板、每种引物各 1 μ l (10pmol/ μ l) 和 5 μ l DMSO 的 50 μ l PCR 反应液进行 30 个循环的 PCR 扩增反应。每次循环条件 :95℃变性 1 分钟,55℃退火 1 分钟,72℃聚合 1 分半钟。

[0251] 扩增片段起初被克隆到 pGEM T easy 上以生成 pGEM T easy-UQCRQ。核苷酸序列经测序确认。

[0252] 实施例 2

[0253] 在本实施例中,比较了实施例 1 制备的所述启动子对表达载体上增强型绿色荧光蛋白 (eGFP) 基因表达的效应。

[0254] 1. eGFP 表达载体的构建

[0255] 首先,通过 PCR 从 pAxCawt (Takara Bio, Japan) 获得兔 β -球蛋白的 polyA 序列。用于 PCR 的引物的核苷酸序列如下：

[0256] RGpAF :GGATCCTTTTCCCTCTGCCAAA (SEQ ID NO :56)

[0257] BamHI

[0258] RGpAR :ACTAGTATAAAGAGAAGAGGGACAGC (SEQ ID NO :57)

[0259] SpeI

[0260] 采用 Expend 高保真 PCR 体系对含有 100ng pAxCawt DNA 模板 (1 μ l)、每种引物各 1 μ l (10pmol/ μ l) 和 5 μ l dNTP (10mM) 的 50 μ l PCR 反应液进行 30 个循环的 PCR 扩增反应。每次循环条件 :95℃变性 1 分钟,50℃退火 1 分钟,72℃聚合 1 分半钟。

[0261] 扩增片段起初被克隆到 pGEM T easy 上以生成 pGEM T easy-RGpA。核苷酸序列经测序确认。

[0262] 随后,来自 pC-LND-GFP-n、pG-LND-GFP-n、pR-LND-GFP-n、pL-LND-GFP-n、pIT-LND-GFP-n 和 pU-LND-GFP-n (参见实施例 7) 的 MluI-BglIII 片段被克隆到 pGEM T easy-RGpA 的 MluI-BamHI 位点,生成 pC-GFP-RGpA、pG-GFP-RGpA、pR-GFP-RGpA、pL-GFP-RGpA、pIT-GFP-RGpA 和 pU-GFP-RGpA。为了构建 pS-GFP-RGpA,将 pS-LND-GFP-n 的 EcoRI-XhoI 片段变为平端,随后将其克隆到 pGEM Teasy-RGpA 的平端 BamHI 位点。

[0263] 2. eGFP 表达的分析

[0264] 根据生产商的说明,使用 FuGene6 将含有 eGFP 基因的表达载体转移到 293T 细胞。通过流式细胞术分析测量 GFP 的表达水平。流式细胞术按如下方法进行:转染 48 小时后,收集 293T 细胞,以含 0.1%叠氮化钠 (FACS 缓冲液) 的磷酸缓冲盐水洗涤一次。随后细胞被重新悬浮于 PBS,借助 CellQuest (Becton Dickinson) 数据采集和分析软件以 FACSsort (Becton Dickinson, Los Angeles, CA, USA) 分析细胞。表 1 给出了该结果。

[0265] 表 1. GFP 表达的比较

[0266]

启动子	相对平均荧光强度
CMV	100
GAPDH	102.5
RPL10	81.9
LENG8	66.6
SNX3	36.4
ITGB4BP	76.4
UQCRQ	47.1

[0267] 这些数据显示 HCMV 和 GAPDH 启动子产生了可比较的 GFP 表达水平。其他启动子例如 RPL10、LENG8 和 ITGB4BP 也产生了显著的 GFP 表达,表明它们有可能用作真核基因表达体系的启动子。

[0268] 实施例 3

[0269] 从实施例 1 制备的各种启动子中选择 RPL10 和 LENG8 启动子进行进一步的鉴定,以分析基因表达所需的启动子序列。

[0270] 1. 一系列 RPL10 启动子的构建

[0271] 1) 一系列 RPL10 启动子的构建

[0272] 通过 PCR 产生各种长度的 RPL10 启动子。以 HT1080 的基因组 DNA 作为模板扩增 RPL 启动子。用于 PCR 的引物对的核苷酸序列如下:

[0273] RPLF50 :ACGCGTACGCGCGCAGACAGACCGCCTATATA AGCCAT

[0274] MluI (SEQ ID NO :58)

[0275] RPLF100 :ACGCGTTGACGTCTGACAGAGCGTCCACCCGTCTTCG

[0276] MluI (SEQ ID NO :59)

[0277] RPLF200 :ACGCGTCTGGCCGCCCGCGGCCCTGGTACCCGGTCACC

[0278] MluI (SEQ ID NO :60)

[0279] RPLF500 :ACGCGTGTCTCCCCCTCCGGCCTCCCGGGTTGACAAAGG

[0280] MluI (SEQ ID NO :61)

[0281] RPLF1000 :ACGCGTGTGCGCTCGAGCAGGATTCCTCCCGTCCTTCC

[0282] MluI (SEQ ID NO :62)

[0283] RPLR :GGATCCGGCGACACCAGGATCTTCAGTGGCT (SEQ ID NO :31)

[0284] BamHI

[0285] RPLRTSS :GGATCCGGCGCTCCTCCGCCTGCGCATGGCTTATATA

[0286] BamHI (SEQ ID NO :63)

[0287] 采用 Expend 高保真 PCR 体系对含有 1 基因组 DNA 模板、各 1 μ l 的每种引物 (10pmol/ μ l) 和 5 μ l dNTP (10mM) 的 50 μ l PCR 反应液进行 30 个循环的 PCR 扩增反应。每

次循环条件:95℃变性 1 分钟,60℃退火 1 分钟,72℃聚合 1 分半钟。

[0288] 实施例 1(8) 描述了 RPL 启动子(-350 ~ +143)(SEQ ID NO:8 的 651 到 1143 位核苷酸)。

[0289] RPL50 启动子(-50 ~ +143)(SEQ ID NO:8 的 951 到 1143 位核苷酸)以引物对 RPL F50 和 RPLR 扩增,并被克隆到 pGEM T easy,生成 pGem Teasy-pRPL50。

[0290] RPL100 启动子(-100 ~ +143)(SEQ ID NO:8 的 901 到 1143 位核苷酸)以引物对 RPL F100 和 RPLR 扩增,并被克隆到 pGEM T easy,生成 pGem Teasy-pRPL100。

[0291] RPL200 启动子(-200 ~ +143)(SEQ ID NO:8 的 801 到 1143 位核苷酸)以引物对 RPL F200 和 RPLR 扩增,并被克隆到 pGEM T easy,生成 pGem Teasy-pRPL200。

[0292] RPL500 启动子(-500 ~ +143)(SEQ ID NO:8 的 501 到 1143 位核苷酸)以引物对 RPL F500 和 RPLR 扩增,并被克隆到 pGEM T easy,生成 pGem Teasy-pRPL500。

[0293] RPL1000 启动子(-1000 ~ +143)(SEQ ID NO:8 的 1 到 1143 位核苷酸)以引物对 RPL F1000 和 RPLR 扩增,并被克隆到 pGEM T easy,生成 pGem Teasy-pRPL 1000。

[0294] RPL TSS 启动子(-350 ~ -1)(SEQ ID NO:8 的 651 到 1000 位核苷酸)以引物对 RPL F 和 RPL R TSS 扩增,并被克隆到 pGEM T easy,生成 pGem Teasy-pRPL TSS。

[0295] 核苷酸序列经测序确认。

[0296] 2) 含有 RPL 启动子的一系列 eGFP 表达载体的构建

[0297] 通过将来自 pGem T easy-GFP(参见实施例 4)的 BamHI 片段插入到 pGem T easyRGpA(参见实施例 2)的 BamHI 位点,构建 pGem Teasy-GFP-RGpA。随后, pGem T easy-pRPL 启动子的 MluI-BamHI 片段被克隆到 pGem T easy-GFP-RGpA 的 MluI-BamHI 位点上,生成 pRPL-、pRPL50-、pRPL100-、pRPL200-、pRPL500-、pRPL 1000- 和 pRPLTSS-GFP-RGpA。

[0298] 2. 一系列 LENG8 启动子的构建

[0299] 1) 一系列 LENG8 启动子的构建

[0300] 正如实施例 1(9) 描述,通过连接 2 个片段(LENG1 和 LENG2)产生 LENG8 启动子。为了产生各种长度的 LENG8 启动子,通过 PCR 获得一系列的 LENG1 片段。随后,LENG1 片段被连接到实施例 1(9) 描述的 LENG2 片段上,以生成最终的 LENG 启动子。HT1080 的基因组 DNA 被用作模板。用于 PCR 的引物对的核苷酸序列如下:

[0301] LENGF50 : ACGCGTGTGACGTCAGGACGCCGCGTCAGG(SEQ ID NO:64)

[0302] MluI

[0303] LENGF100 :ACGCGTTGGCGTTCATTGGCTGTGCAGGGCC(SEQ ID NO:65)

[0304] MluI

[0305] LENGF200 :ACGCGTTTGTCCCCCTCGGGGCCACCGTCCCC(SEQ ID NO:66)

[0306] MluI

[0307] LENGF1000 :ACGCGTTTGTATCAGAGTCCTGGACGGAAAC(SEQ ID NO:67)

[0308] MluI

[0309] LENG8R1 :GTTTAAACAAAGTAGAAGACGACGGCGCACGCG(SEQ ID NO:33)

[0310] PmeI

[0311] LENGRTSS :GTTTAAACCTCTGGTCTTCTTTGGCTTCGACGT(SEQ ID NO:68)

[0312]PmeI

[0313]采用 Expend 高保真 PCR 体系对含有 1 基因组 DNA 模板、各 1 μ l 的每种引物 (10pmol/ μ l) 和 5 μ l dNTP(10mM) 的 50 μ l PCR 反应液进行 30 个循环的 PCR 扩增反应。每次循环条件 :94℃变性 1 分钟,55℃退火 1 分钟,72℃聚合 2 分钟。

[0314]L50 启动子 (-50 ~ +305) (SEQ ID NO :9 的 970 到 1325 位核苷酸) 以引物对 LENG F50 和 LENG8R1 扩增,并被克隆到 pGEM T easy,生成 pGem Teasy-pL50。

[0315]L100 启动子 (-100 ~ +305) (SEQ ID NO :9 的 920 到 1325 位核苷酸) 以引物对 LENG F100 和 LENG8R1 扩增,并被克隆到 pGEM T easy,生成 pGemT easy-pL100。

[0316]L200 启动子 (-200 ~ +305) (SEQ ID NO :9 的 820 到 1325 位核苷酸) 以引物对 LENG F200 和 LENG8R1 扩增,并被克隆到 pGEM T easy,生成 pGemT easy-pL200。

[0317]L1000 启动子 (-1020 ~ +305) (SEQ ID NO :9 的 1 到 1325 位核苷酸) 以引物对 LENG F1000 和 LENG8R1 扩增,并被克隆到 pGEM T easy,生成 pGEMT easy-pL1000。

[0318]随后, pGEM T easy-pL50、-pL100、-pL200 和 -pL1000 的 MluI-PmeI 片段被克隆到实施例 1(9) 描述的 pGEM T easy-LENG2 的 MluI-PmeI 位点,以生成 pGEM T easy-pLENG50P、-pLENG 100P、-pLENG200P 和 -pLENG1000P。

[0319]LENG TSS 启动子 (-385 ~ -1) (SEQ ID NO :9 的 635 到 1019 位核苷酸) 以引物对 LENG F 和 LENG R TSS 扩增,并被克隆到 pGEM T easy,生成 pGem T easy-pLENG TS S。

[0320]实施例 1(9) 描述了 LENG8 启动子 (-385 ~ +305,+1908 ~ +2121) (SEQ IDNO :9 的 635 到 1538 位核苷酸)。

[0321]2) 含有 LENG 启动子的一系列 eGFP 表达载体的构建

[0322]为了构建含有 eGFP 基因的表达载体,将 pGEM T easy-pLENG 启动子的 MluI-EcoRI 片段 (EcoRI 位点变为平端) 克隆到 pGem T easy-GFP-RGpA 的 MluI-BamHI 位点 (BamHI 位点变为平端),生成 pLENG-、pLENG50-、pLENG 100-、pLENG200- 和 pLENG 1000-GFP-RGpA。将 pGem T easy-pLENGTS S 的 MluI-PmeI 片段克隆到 pGem T easy-GFP-RGpA 的 MluI-BamHI 位点 (BamHI 位点变为平端),生成 pLENG TSS-GFP-RGpA。

[0323]3. eGFP 表达的分析

[0324]根据生产商的说明,使用 FuGene6 (Roche,Germany) 以 eGFP 表达载体转染 293T 细胞,培养 48 小时。通过流式细胞术分析测量 GFP 的表达水平。流式细胞术按如下方法进行 :转染 48 小时后,收集 293T 细胞,以含 0.1%叠氮化钠的 PBS 洗涤一次。随后细胞被重新悬浮在 PBS 中,借助 CellQuest 数据采集和分析软件以 FACSort 分析细胞。表 2 给出了该结果。

[0325]表 2. eGFP 表达的比较

[0326]A. RPL10 启动子

[0327]

	相对平均荧光强度
RPL-GFP	100
RPL50-GFP	20

RPL100-GFP	40
RPL200-GFP	70
RPL500-GFP	150
RPL1000-GFP	210
RPL TSS-GFP	20

[0328] B. LENG8 启动子

[0329]

	相对平均荧光强度
LENG-GFP	100
LENG50-GFP	14
LENG100-GFP	42
LENG200-GFP	87
LENG1000-GFP	170
LENG TSS-GFP	39

[0330] 最短的启动子 RPL50 或 LENG50 可驱动显著高于背景的 GFP 表达,表明转录起始位点上游 50bp 的存在足以具有基本活性。但是,对于 RPL10 和 LENG8 启动子,使用更长的启动子进行基因表达时,eGFP 表达更高。最长的 RPL10 启动子 RPL1000 可驱动比 RPL500 或其他启动子更高水平的 eGFP 表达。在各种 LENG 启动子中,最长的 LENG8 启动子 LENG1000 也表现出最高的启动子活性。而且,当包括位于转录起始位点(TSS)和翻译起始位点之间的元件时,启动子的活性会更佳。

[0331] 实施例 4

[0332] 在本实施例中,所述内部启动子被克隆到反转录病毒载体上,并比较了它们对病毒滴度和 eGFP 基因表达水平的效应。

[0333] 1. 表达 eGFP 的反转录病毒载体的构建

[0334] 1) 反转录病毒载体的构建

[0335] 1-1) I-D

[0336] 构建了缺失 U3 的反转录病毒质粒。首先,以 pMT(Hong 等人, J. GeneMed. 6 : 724(2004) ;US 6, 451, 595) 作为模板扩增 MLV 的正常 3' LTR。用于 PCR 的引物对的核苷酸序列如下 :

[0337] SCV3LB :GGATCCCTCGAGCGATAAAAATAAAAGATTTTATTTAGTCTCC

[0338] BamHI XhoI (SEQ ID NO :69)

[0339] SCV3LRI :GAATTCGTCGACTGAAAGACCCCGCTGACGG (SEQ ID NO :70)

[0340] EcoRI SalI

[0341] 扩增片段起初被克隆到 pGEM T easy 上,生成 pGEM T easy-3' LTR。

[0342] 以 pMT 为模板扩增缺失形式的 3' LTR。用于 PCR 的引物的核苷酸序列如下:

[0343] 3' LTR-1 :GCTAGCCCCTGTGCCTTATTTGAA (SEQ ID NO :71)

[0344] NheI

[0345] SCV3LRI :GAATTCGTCGACTGAAAGACCCCGCTGACGG (SEQ ID NO :70)

[0346] EcoRI SalI

[0347] 扩增片段起初被克隆到 pGEM T easy,生成 pGEM T easy-3' dLTR-1。NheI-SalI 片段被克隆到 pGEM T easy-3' LTR 的 NheI-SalI 位点,生成 pPreSIN2。来自 pPreSIN2 的 BamHI-SalI 片段被克隆到 pMT 的 BamHI-SalI 位点,生成 I-D。

[0348] 1-2) I-ND

[0349] 为了促进克隆,将新的多克隆位点 (MCS) 引入到反转录病毒质粒 I-D 上。在没有模板的情况下通过聚合酶反应生成新 MCS 的 52bp 长的片段。使用了下列引物对:

[0350] NEWMCSF :ACGCGTTTAAACCGCGGAATTCGGATCCACATCGTG

[0351] MluI SacII BamHI (SEQ ID NO :72)

[0352] NEWMCSR :CTCGAGATCTAGGCCTCACGATGTGGATCCGAATTC

[0353] XhoI StuI DraIII EcoRI (SEQ ID NO :73)

[0354] 含有 MluI、PmeI、SacII、EcoRI、BamHI、DraIII、StuI、BglIII 和 XhoI 限制酶切位点的扩增片段起初被克隆到 pGEM T easy。确认核苷酸序列后,将 MluI-XhoI 片段克隆到 I-D 的 MluI-XhoI 位点,生成 I-ND。

[0355] 2) 含 eGFP 基因的反转录病毒载体的构建

[0356] 2-1) eGFP 基因

[0357] 为了构建表达 eGFP 基因的反转录病毒载体,采用下列引物对从 pIRES2-EGFP (CLONTECH Laboratory, Palo Alto, CA, USA, Cat. #6029-1) 扩增 eGFP 基因:

[0358] eGFP5 :ACGCGTGGATCCATGGTGAGCAAGGGCGAG3'

[0359] MluI BamHI (SEQ ID NO :74)

[0360] eGFP3 :CTCGAGAGATCTTTACTTGTACAGCTCGTC3' (SEQ ID NO :75)

[0361] XhoI BglIII

[0362] 扩增的 eGFP 被克隆到 pGEM T easy 上以生成 pGEM T easy-eGFP。BamHI-BglIII 片段被克隆到反转录病毒载体 pI-D 的 BamHI 位点,生成 pI-D-GFP;BamHI-BglIII 片段被插入到反转录病毒载体 I-ND 的 BamHI 位点,生成 I-ND-GFP。

[0363] 2-2) 含 eGFP 基因的反转录病毒载体的构建

[0364] 通过将 pGEM T easy-eGFP 的 BamHI-BglIII 片段插入到含有野生型 LTR 的反转录病毒载体 pMT 的 BamHI 位点,构建 pMT-GFP (Kim 等人, Biochem. Biophys. Res. Commun. 343 : 1017 (2006))。

[0365] 对于带有缺陷型 LTR 的反转录病毒载体,实施例 1 的各种内部启动子分别被克隆到含有 GFP 序列的反转录病毒载体上。

[0366] 来自 pCN 质粒的 MluI-BamHI 片段即 HCMV IE 启动子被克隆到 pI-D-GFP 的

MluI-BamHI 位点,生成 pC-D-GFP。来自 pGEM T Easy-MTU3 的 MluI 片段被插入到 pI-D-GFP 的 MluI 位点,生成 pM-D-GFP。来自 pGEM T Easy-Enh+UbC 的 MluI 片段被插入到 pI-D-GFP 的 MluI 位点,生成 pCU-D-GFP。

[0367] 来自 pCN 质粒的 MluI-BamHI 片段即 HCMV IE 启动子被克隆到 pI-ND-GFP 的 MluI-BamHI 位点,生成 C-ND-GFP。来自 pAxCawt (TakaraBio, Otsu, Japan) 的以 Klenow 片段处理的 SalI-SwaI 片段被克隆到 pI-ND-GFP 的 PmeI 位点,生成 pCA-ND-GFP。来自 pGEM T easy-EF 的 MluI-BamHI 片段被克隆到 pI-ND-GFP 的 MluI-BamHI 位点,生成 pE-ND-GFP。来自 pGEM T easy-BA 的 MluI-BamHI 片段被克隆到 pI-ND-GFP 的 MluI-BamHI 位点,生成 pB-ND-GFP。来自 pGEM Teasy-GAPDH 的 MluI-PmeI 片段被克隆到 pI-ND-GFP 的 MluI-PmeI 位点,生成 pG-ND-GFP。来自 pGEM T easy-RPL 的 MluI-BamHI 片段被克隆到 pI-ND-GFP 的 MluI-BamHI 位点,生成 pR-ND-GFP。来自 pGEM Teasy-LENG8 的 MluI-EcoRI 片段被克隆到 pI-ND-GFP 的 MluI-EcoRI 位点,生成 pLe-ND-GFP。来自 pGEM T easy-SNX 的 EcoRI-BamHI 片段被克隆到 pI-ND-GFP 的 EcoRI-BamHI 位点,生成 pS-ND-GFP。来自 pGEM Teasy-CNOT3 的 MluI-BamHI 片段被克隆到 pI-ND-GFP 的 MluI-BamHI 位点,生成 pCo-ND-GFP。来自 pGEM T easy-CPNE1 的 MluI-BamHI 片段被克隆到 pI-ND-GFP 的 MluI-BamHI 位点,生成 pCP-ND-GFP。来自 pGEM Teasy-HYPO 的 MluI-BamHI 片段被克隆到 pI-ND-GFP 的 MluI-BamHI 位点,生成 pHY-ND-GFP。来自 pGEM T easy-DKC 1 的 MluI-EcoRI 片段被克隆到 pI-ND-GFP 的 MluI-EcoRI 位点,生成 pD-ND-GFP。来自 pGEM T easy-VPS 72 的 MluI-EcoRI 片段被克隆到 pI-ND-GFP 的 MluI-EcoRI 位点,生成 pV-ND-GFP。来自 pGEM T easy-ITGB4BP 的 MluI-PmeI 片段被克隆到 pI-ND-GFP 的 MluI-PmeI 位点,生成 pIT-ND-GFP。来自 pGEM Teasy-UQCRQ 的 MluI-PmeI 片段被克隆到 pI-ND-GFP 的 MluI-PmeI 位点,生成 pU-ND-GFP。

[0368] 2. eGFP 表达的分析

[0369] 以含有 eGFP 基因的相应反转录病毒载体以及双嗜性包装构建体 pVM-GP 和 pVM-AE (Yu 等人, Gene Ther. 10 :706 (2003)) 转染 293T 细胞,随后培养 48 小时。以 0.45 μ m 滤纸过滤培养上清,制备不含细胞的病毒,将其用于分别转导 2×10^5 的 HT1080 细胞和 2×10^5 的 K562 细胞。温育细胞 48 小时,随后收集细胞进行检测。通过 FACS 分析 (参见表 3 和表 4) 测量 GFP 阳性细胞的比例和 GFP 表达的水平 (平均荧光强度)。

[0370] FACS 分析按如下方法进行:转导 48 小时后,收集 HT1080 细胞和 K562 细胞,以含 0.1% 叠氮化钠 (FACS 缓冲液) 的磷酸缓冲盐水洗涤一次。随后细胞被重新悬浮于 PBS,借助 CellQuest (Becton Dickinson) 数据采集和分析软件以 FACS Sort (Becton Dickinson, Los Angeles, CA, USA) 分析细胞。

[0371] 首先,比较带有缺陷型 LTR 的反转录病毒载体的 GFP 表达和含有野生型 LTR 的反转录病毒载体 MT-GFP 的 GFP 表达。正如表 3 所示, C-D-GFP 载体的表现与 MT-GFP 同样好。在 HT1080 细胞中,来自 C-D-GFP 载体的 GFP 阳性细胞的比例高于来自 MT-GFP 载体的 GFP 阳性细胞比例,而在 K562 细胞中也超过来自 MT-GFP 载体的 GFP 阳性细胞比例的 80%。在 HT1080 细胞中, C-D-GFP 载体驱动的 GFP 表达水平高于 MT-GFP 载体驱动的 GFP 表达水平,而在 K562 细胞中约为 MT-GFP 载体驱动的 GFP 表达水平的 70%。从这些实验我们证实了 HCMV 启动子对 GFP 表达有良好效果。但是,我们没有发现与 HCMV 启动子效果一样好的其他启动子。CMV/ 泛素的杂合启动子也产生了高病毒滴度和高水平的 GFP 表达,不过比 HCMV

启动子低。

[0372] 表 3. GFP 表达的比较

[0373]

	HT1080		K562	
	pf GFP ⁺ 细胞 相对百分数	相对平均荧光 强度	pf GFP ⁺ 细胞相 对百分数	相对平均荧光 强度
MT-GFP	100	100	100	100
C-D-GFP	110.2	243.3	81.4	67.2
M-D-GFP	29.6	54.4	21.7	19.9
CU-D-GFP	68.2	59.4	46.6	56.1

[0374] 作为下一个步骤,除 HCMV 启动子外我们还测试了含有其他各种内部启动子的更多反转录病毒载体。表 2 给出了该结果。

[0375] 表 4. GFP 表达的比较

[0376]

	HT1080		K562	
	pf GFP ⁺ 细胞 相对百分数	相对平均荧光 强度	pf GFP ⁺ 细胞相 对百分数	相对平均荧光 强度
C-ND-GFP	100	100	100	100
CA-ND-GFP	7.9	37.1	3.9	94.5
E-ND-GFP	28.4	21.0	19.2	157.1
B-ND-GFP	15.3	1.3	1.5	10.4
G-ND-GFP	84.1	9.8	68.9	51.6
R-ND-GFP	109.7	9.1	107.4	42.6
L-ND-GFP	112.2	6.1	124.4	33.3
S-ND-GFP	83.8	2.1	57.1	12.5
CN-ND-GFP	52.5	1.9	31.4	11.7
CP-ND-GFP	48.2	1.4	40.1	12.2
HY-ND-GFP	19.7	6.5	14.9	44.3
D-ND-GFP	42.2	1.7	44.7	17.0
V-ND-GFP	17.5	1.2	18.0	12.4
IT-ND-GFP	96.0	3.0	79.5	18.1
U-ND-GFP	89.7	2.6	82.9	17.7

[0377] 正如表 4 所示, HCMV 启动子在 HT108 细胞和 K562 细胞均产生最高水平的基因表达 (参见平均荧光强度) 和大量的 GFP 阳性细胞。但是, LENG8 和 RPL10 启动子在 HT1080 细胞和 K562 细胞均产生了最高的病毒滴度 (被转导细胞的比例)。这两种启动子驱动的 GFP

表达水平在 HT1080 细胞中相对较低,但在 K562 细胞中,与 HCMV 启动子相比高于 30%。因此, LENG8 和 RPL10 启动子可用于反转录病毒载体体系在某些细胞类型中的基因表达。此外, GAPDH、UQCRCQ、ITGB4BP 和 SNX3 启动子产生了相对较高的病毒滴度 (CMV 启动子的 80% 以上)。

[0378] CA-ND-GFP 和 E-ND-GFP 载体在 K562 细胞内产生了最高水平的 GFP 表达,但是源自这些载体的病毒滴度极低,因此难以应用这些载体。

[0379] 实施例 5

[0380] 在本实施例中,比较了内部启动子对病毒滴度和 gp91 基因表达水平的效应。

[0381] 1. 表达 gp91 的反转录病毒载体的构建

[0382] 1) gp91-phox 基因 (NCBI 收录号 :NM_000397)

[0383] 为了构建表达人 gp91-phox 的反转录病毒载体,通过 RT-PCR 从人外周血淋巴细胞的总 RNA 克隆 gp91cDNA。用于该步骤的引物核苷酸序列如下:

[0384] GP91F :GGATCCATGGGGAACTGGGCTGTGAAT (SEQ ID NO :76)

[0385] BamHI

[0386] GP91R :GGATCCCTCGAGTTAGAAGTTTTCCTTGTGAAAA

[0387] BamHI XhoI (SEQ ID NO :77)

[0388] 扩增片段起初被克隆到 pGEM T easy 上以生成 pGEM T easy-gp91,并确认了该扩增片段的核苷酸序列。

[0389] 2) pPromoter-ND

[0390] 来自 pCN 质粒的 MluI-BamHI 片段即 HCMV IE 启动子被克隆到 pI-ND 的 MluI-BamHI 位点,生成 pC-ND。来自 pGEM T easy-GAPDH 的 MluI-PmeI 片段被克隆到 pI-ND 的 MluI-PmeI 位点,生成 pG-ND。来自 pGEM Teasy-RPL 的 MluI-BamHI 片段被克隆到 pI-ND 的 MluI-BamHI 位点,生成 pR-ND。来自 pGEM T easy-LENG8 的 MluI-EcoRI 片段被克隆到 pI-ND 的 MluI-EcoRI 位点,生成 pL-ND。来自 pGEM T easy-SNX 的 EcoRI-BamHI 片段被克隆到 pI-ND 的 EcoRI-BamHI 位点,生成 pS-ND。来自 pGEM Teasy-ITGB4BP 的 MluI-PmeI 片段被克隆到 pI-ND 的 MluI-PmeI 位点,生成 pIT-ND。

[0391] 3) 表达 gp91 的反转录病毒载体的构建

[0392] 通过将 pGEM T easy-gp91 的 BamHI 片段插入到 pMT,构建 MT-gp91 载体。

[0393] 通过将来自 pGEM T easy-gp91 的 BamHI-XhoI 片段插入到 pC-ND、pG-ND、pR-ND、pL-ND、pS-ND 和 pIT-ND 的 BamHI-XhoI 位点,分别生成 pC-ND-gp91、pG-ND-gp91、pR-ND-gp91、pL-ND-gp91、pS-ND-gp91 和 pIT-ND-gp91,构建 gp91-phox 表达受内部启动子驱动的反转录病毒载体。

[0394] 实施例 7(3) 描述了反转录病毒载体 pR-LND-gp91-phox-n、pR1000-LND-gp91-n 和 pR1000-LND-gp91-pA-n-rev 的构建过程。

[0395] 2. gp91 表达的分析

[0396] 采用含有 gp91 基因的相应反转录病毒载体以及包装构建体 pVM-GP 和 pVM-GeR (Kim 等人., Biochem. Biophys. Res. Commun. 343 :1017 (2006)) 以磷酸钙沉淀方法转染 293T 细胞,随后培养 48 小时。以 0.45 μ m 滤纸过滤培养上清,制备不含细胞的病毒,将其用于转导 K562 细胞。

[0397] 对于K562细胞的转导,转导前一天在6孔板上每孔接种 2.5×10^5 个细胞。在 $8 \mu\text{g/ml}$ 聚凝胺的存在下每孔加入相同体积的病毒上清,培养板在 32°C 下以 2800rpm 离心 90 分钟 (Eppendorfcentrifuge 5810R)。转导后,细胞在 37°C 的 CO_2 培养箱内温育 2 天。

[0398] 通过流式细胞术分析 gp91 的蛋白表达。转导 2 天后,收集 K562 或 PLB-985/gp91^{-/-} 细胞,以 PBS 洗涤。细胞随后被重新悬浮在 $100 \mu\text{l}$ PBS 中,以 $1 \mu\text{l}$ 抗-gp91 的抗体 (7D5 ;MBL, Japan) 在 4°C 下染色 30 分钟。随后细胞以 PBS 洗涤两次,重新悬浮于 $100 \mu\text{l}$ PBS,再以 $1 \mu\text{l}$ 结合 FITC 的山羊抗小鼠抗体 (Southern Biotechnology Associates, Inc, Birmingham,AL,USA) 在 4°C 下染色 30 分钟。随后细胞以 PBS 洗涤 3 次,重新悬浮于 $500 \mu\text{l}$ 的 PBS。借助 CellQuest (BD) 数据采集和分析软件以 FACSsort (BD, San Jose,CA) 进行流式细胞术。

[0399] 首先,比较来自反转录病毒载体 MT-gp91、C-ND-gp91、G-ND-gp91、R-ND-gp91、L-ND-gp91、S-ND-gp91 和 IT-ND-gp91 的 gp91 表达。表 5 给出了该结果。在各种缺失 U3 的反转录病毒载体当中, R-ND-gp91 和 S-ND-gp91 产生了比其他反转录病毒载体更高的病毒滴度 (达到 MT-gp91 在 K562 细胞中的 70% 以上)。在 K562 细胞中, R-ND-gp91 载体驱动的 gp91 表达比 S-ND-gp91 载体驱动的 gp91 表达高。C-ND-gp91 载体没有产生 gp91 阳性细胞,不过 ND-eGFP 可产生高水平的 eGFP 表达 (表 4)。

[0400] 表 5. gp91 表达的比较

	K562	
	gp91 ⁺ 细胞的相对百分数	相对平均荧光强度
MT-gp91	100	100
C-ND-gp91	2.4	1.9
G-ND-gp91	14.0	3.3
R-ND-gp91	75.5	18.9
L-ND-gp91	35.4	5.8
S-ND-gp91	79.4	9.9
IT-ND-gp91	47.4	5.5

[0402] 在实施例 3 中观察到与短形式的 RPL10 启动子相比,更长形式的 RPL10 启动子可驱动更高水平的 GFP 基因表达。我们检测了该观察结果是否适用于 gp91 的表达。我们构建了含有最长 RPL10 启动子即 RPL1000 (pR1000-LND-gp91-n) 并表达 gp91 的反转录病毒载体,证实了其对 gp91 基因表达的效应。我们还构建了反转录病毒载体 pR1000-LND-gp91-n-rev,其中受 RPL1000 启动子驱动的 gp91 基因表达盒以反向插入,并比较了 gp91 的基因表达。表 6 给出了该结果。

[0403] 表 6. gp91 表达的比较

[0404]

	K562	
	gp91 ⁺ 细胞的相对百分数	相对平均荧光强度
R-LND-gp91-n	100	100
R1000-LND-gp91-n	95.5	145.6
R1000-LND-gp91-n-rev	58.6	193.9

[0405] 与 R-LND-gp91-n 载体相比, R1000-LND-gp91-n 载体可产生可比较量的病毒滴度 (gp91⁺ 细胞的百分数) 和更高水平的 gp91 基因表达 (相对平均荧光强度)。R1000-LND-gp91-n-rev 载体可驱动最高水平的 gp91 基因表达, 尽管来自它的病毒滴度只是 R-LND-gp91-n 的一半。

[0406] 实施例 6

[0407] 本发明的反转录病毒载体可用于离体基因传递。

[0408] 从对象体内收集 CD34⁺ 造血干细胞。CD34⁺ 细胞的来源可为骨髓抽吸物或动员的外周血。收集的 CD34⁺ 细胞以无血清培养基 (Cell Gro, Germany) 在 Vuelife 培养袋 (37°C, 5% CO₂) 中培养两天, 该培养基含有 300ng/ml 的人干细胞因子 (SCF)、300ng/ml 的人 FLT-3L、100ng/ml 的人血小板生成素 (TPO) 和 20ng/ml 的人 IL-3 (预刺激)。采用预包装人纤连蛋白 CH296 片段 (Retronectin, TaKaRa Bio) 的 Vuelife 培养袋进行转导。预刺激的细胞被转移到包被 retronectin 的 Vuelife 培养袋, 2 天内分 3 次加入反转录病毒的上清。细胞随后收集, 盐水洗涤 3 次, 重悬浮于输注溶液 (含 1% 人血清白蛋白的盐水), 再输注给对象。

[0409] 实施例 7

[0410] 在本实施例中, 考察了内部启动子对包装细胞系 PG13 内病毒滴度和基因表达水平的效应。

[0411] 1. 表达 eGFP 的反转录病毒载体的构建

[0412] 1) 反转录病毒载体的构建

[0413] 1-1) pI-LND

[0414] 缺失 U3 的反转录病毒载体不能在第一轮反转录病毒转导后被动员, 因为转导后 5' LTR 和 3' LTR 均成为缺陷型。因此, 通过以 pDNA 转染反转录病毒的包装细胞, 以使所述载体稳定整合到反转录病毒的包装细胞的基因组上, 从而建立了稳定的生产细胞系。载体 DNA 的线性化对于所有转染子在其染色体上含有适宜的 DNA 排列是重要的。

[0415] 为了构建便于线性化的反转录病毒载体, 将用于线性化的两个限制酶切位点引入到反转录病毒质粒 pI-LND。通过将片段 L1 插入到 5' LTR 之前约 200bp 处引入限制酶切位点中的一个, 并将 L2 插入到 3' LTR 之后引入另一位点。在没有模板的情况下通过聚合酶反应产生 L1 片段。以 pUC18 (Promega, WI, USA) 作为模板扩增 L2 片段。引物的核苷酸序列如下:

[0416] L1F : 5' GCTCTTCCGCTCACGTGTGATCAATTTAAATTTTCGAA

[0417] SapI PmlI BclI SwaI BstBI (SEQ ID NO : 78)

[0418] L1R : 5' AGCGGAAGAGCTTCGAAATTTAAATTGATCACACGTG

[0419] SapI BstBI SwaI BclI PmlI (SEQ ID NO : 79)

[0420] L2F : 5' AGGCCTGGTCACCGGCCATTATGGCCACGTGATCATTAAATTTG

[0421] StuI BstEII SfiI BclI

[0422] AAGCATTTATCAGGGTTA (SEQ ID NO :80)

[0423] L2R :5' TATTCGCGCGTTTCGGTGATGAATATT (SEQ ID NO :81)

[0424] SspI

[0425] 含有 SapI、PmlI、BclI、SwaI、BstBI 和 SapI 限制酶切位点的扩增 L1 片段起初被克隆到 pGEM T easy (Promega, WI, USA), 生成 pGem T easy-L1。确认核苷酸序列后, 将 pGem T easy-L1 释放的 SapI 片段插入到 pI-ND 的 SapI 位点, 生成 pI-L1ND。含有 StuI、BstEII、SfiI、PmlI、BclI、SwaI 和 SspI 限制酶切位点的扩增 L2 片段起初被克隆到 pGEM T easy (Promega, WI, USA), 生成 pGem T easy-L2。确认核苷酸序列后, 将 pGem T easy-L2 释放的 StuI-SspI 片段克隆到 pI-L1ND 的 SspI 位点, 生成 pI-LND。

[0426] 1-2)pI-LND-n

[0427] 为了构建具有高病毒滴度的生产细胞系, 选择含有反转录病毒载体 DNA 的转染子是很重要的。出于该目的, 通常使用抗药性基因。但是, 抗药性基因位于载体基因组内不是有利的, 因为如果抗药性基因被包括在载体的基因组上, 该基因会在体内表达。因此, 我们制备了在反转录病毒基因组外携带有抗药性基因的载体构建体。

[0428] 首先, 我们使用新霉素抗性基因, 用于选择转染子。为了表达新霉素抗性基因, 将人 β -肌动蛋白启动子和多聚腺苷酸化序列连接到细菌的 Neo 编码序列。

[0429] 使用来自 K562 的基因组扩增人用于 PCR 的引物的核苷酸序列如下:

[0430] BApF :5' GTCGACATTAATGCCGGTGAGTGAGCGCGCGGGGCCAA

[0431] SalI PshBI (SEQ ID NO :82)

[0432] BApR :5' GGATCCGGTGGCGCGTCGCGCCGCTGGGTTTT (SEQ ID NO :83)

[0433] BamHI

[0434] 扩增片段被克隆到 pGEM T easy 上, 生成 pGEM T easy-pBA。

[0435] 使用 pcDNA 3.1 (Invitrogen, CA, USA) 扩增细菌的 Neo 编码基因。用于 PCR 的引物的核苷酸序列如下:

[0436] NeoF :5' AGATCTATGGGATCGGCCATTGAACAA (SEQ ID NO :84)

[0437] BglII

[0438] pAR :5' CATATGTCATAATCAGCCATACCACATTT (SEQ ID NO :85)

[0439] NdeI

[0440] 扩增片段被克隆到 pGEM T easy 上, 生成 pGEM T easy-Neo。

[0441] 使用 pTet-On (Clontech, TAKARA bio, Japan) 扩增多聚腺苷酸化信号序列。用于 PCR 的引物的核苷酸序列如下:

[0442] SVpAF :CTCGAGATGGGATCGGCCATTGAACAA (SEQ ID NO :86)

[0443] XhoI

[0444] SVpAR :CATATGAGTAATCAGCCATACCACATTT (SEQ ID NO :87)

[0445] NdeI

[0446] 扩增片段被克隆到 pGEM T easy 上, 生成 pGEM T easy-pA。

[0447] 确认核苷酸序列后, pGEM T easy-pA 释放的 XhoI-NdeI 片段被插入到 pGEM T easy-Neo 的 XhoI-NdeI 位点, 生成 pGEM T easy-NeopA。BglIII-NdeI 片段被克隆到 pGEM T

easy-pBA 的 BamHI-NdeI 位点,生成 pGEM Teasy-pBA-Neo-pA。MluI-EcoRI-Klenow 处理的片段被克隆到 pI-LND 的 SspI 位点,生成 pI-LND-n。

[0448] 2) 含 eGFP 基因的反转录病毒载体的构建

[0449] 2-1)pI-LND-GFP-n

[0450] pGem T easy-eGFP(参见实施例 2)释放的 BamHI-BglIII 片段被克隆到 pI-LND-n 的 BamHI 位点,生成 pI-LND-GFP-n。

[0451] 2-2) 含 eGFP 基因的反转录病毒载体的构建

[0452] 实施例 1 的各种内部启动子分别被克隆到含有 GFP 序列的反转录病毒载体。

[0453] 来自 pCN 质粒的 MluI-BamHI 片段即 HCMV IE 启动子被克隆到 pI-LND-GFP-n 的 MluI-BamHI 位点,生成 pC-LND-GFP-n。来自 pGEM Teasy-GAPDH 的 MluI-PmeI 片段被克隆到 pI-LND-GFP-n 的 MluI-PmeI 位点,生成 pG-LND-GFP-n。来自 pGEM T easy-RPL 的 MluI-BamHI 片段被克隆到 pI-LND-GFP-n 的 MluI-BamHI 位点,生成 pR-LND-GFP-n。来自 pGEM Teasy-LENG8 的 MluI-EcoRI 片段被克隆到 pI-LND-GFP-n 的 MluI-EcoRI 位点,生成 pL-LND-GFP-n。来自 pGEM T easy-SNX 的 EcoRI-BamHI 片段被克隆到 pI-LND-GFP-n 的 EcoRI-BamHI 位点,生成 pS-LND-GFP-n。来自 pGEM T easy-ITGB4BP 的 MluI-PmeI 片段被克隆到 pI-LND-GFP-n 的 MluI-PmeI 位点,生成 pIT-LND-GFP-n。来自 pGEM T easy-UQCRQ 的 MluI-PmeI 片段被克隆到 pI-LND-GFP-n 的 MluI-PmeI 位点,生成 pU-LND-GFP-n。

[0454] 3) 含 gp91-phox 基因的反转录病毒载体的构建

[0455] 3-1)pPromoter-LND

[0456] 实施例 1 的各种内部启动子被克隆到 pI-LND。

[0457] 来自 pCN 质粒的 MluI-BamHI 片段即 HCMV IE 启动子被克隆到 pI-LND 的 MluI-BamHI 位点,生成 pC-LND。来自 pGEM T easy-GAPDH 的 MluI-PmeI 片段被克隆到 pI-LND 的 MluI-PmeI 位点,生成 pG-LND。来自 pGEM T easy-RPL 的 MluI-BamHI 片段被克隆到 pI-LND 的 MluI-BamHI 位点,生成 pR-LND。来自 pGEM T easy-LENG8 的 MluI-EcoRI 片段被克隆到 pI-LND 的 MluI-EcoRI 位点,生成 pL-LND。来自 pGEM T easy-SNX 的 EcoRI-BamHI 片段被克隆到 pI-LND 的 EcoRI-BamHI 位点,生成 pS-LND。来自 pGEM T easy-ITGB4BP 的 MluI-PmeI 片段被克隆到 pI-LND 的 MluI-PmeI 位点,生成 pIT-LND。来自 pGEM T easy-UQCRQ 的 MluI-PmeI 片段被克隆到 pI-LND 的 MluI-PmeI 位点,生成 pU-LND。来自 pGEM Teasy-pR1000 的 MluI-BamHI 片段被克隆到 pI-LND 的 MluI-BamHI 位点,生成 pR1000-LND。

[0458] 3-2)pPromoter-LND-n

[0459] 来自 pGEM T easy-pBA-Neo-pA 的 MluI-EcoRI-Klenow 处理的片段被克隆到 pC-LND、pG-LND、pR-LND、pL-LND、pS-LND、pIT-LND、pU-LND 和 pR1000-LND 的 SspI 位点,生成 pC-LND-n、pG-LND-n、pR-LND-n、pL-LND-n、pS-LND-n、pIT-LND-n、pU-LND-n 和 pR1000-LND-n。

[0460] 3-3) 含 gp91-phox 基因的反转录病毒载体的构建

[0461] 来自 pGEM T easy-gp91 的 BamHI 片段被克隆到 pC-LND-n、pG-LND-n、pR-LND-n、pL-LND-n、pS-LND-n、pIT-LND-n、pU-LND-n 和 pR1000-LND-n 的 BamHI 位点,生成 pC-LND-gp91-phox-n、pG-LND-gp91-phox-n、pR-LND-gp91-phox-n、pL-LND-gp91-phox-n、pS-LND-gp91-phox-n、pIT-LND-gp91-phox-n、pU-LND-gp91-phox-n

和 pR 1000-LND-gp91-phox-n。

[0462] 还构建了 gp91 基因表达盒反向插入的反转录病毒载体。pC-LND-gp91-pA-n-rev 按如下步骤构建;i) 将来自 pC-LND-gp91-n 的 MluI-BamHI 片段插入到 pI-LND-n 的 BamHI-StuI 位点,生成 pC-LND-n-rev;ii) 将 pGemT easy-gp91 的 BamHI 片段插入到 pC-LND-n-rev 的 BamHI 位点,生成 pC-LND-gp91-n-rev;iii) 随后将 pGEM T easy-RGpA 的 EcoRI 片段插入到 pC-LND-gp91-n-rev 的 Pme I 位点,生成 pC-LND-gp91-pA-n-rev。pR1000-LND-gp91-pA-n-rev 按如下步骤构建;i) 将来自 pR1000-LND-gp91-n 的 MluI-BamHI 片段插入到 pI-LND-n 的 BamHI-StuI 位点,生成 pR1000-LND-n-rev;ii) 将 pGemT easy-gp91 的 BamHI 片段插入到 pR1000-LND-n-rev 的 BamHI 位点,生成 pR1000-LND-gp91-n-rev;iii) 随后将 pGEM T easy-RGpA 的 EcoRI 片段插入到 pR 1000-LND-gp91-n-rev 的 Pme

[0463] I 位点,生成 pR1000-LND-gp91-pA-n-rev。

[0464] 将 MluI-EcoRI 片段插入到 pMT-gp91 的 SapI 位点,构建作为对照的 pMT-gp91-n。

[0465] 2. 生产细胞系的构建

[0466] 1) 反转录病毒载体的线性化

[0467] 为了线性化含有 eGFP 或 gp91 的反转录病毒载体,10 μ g 的反转录病毒质粒以限制性内切酶 (SwaI) 处理 16 小时。含有反转录病毒载体的 DNA 片段从琼脂凝胶上洗脱,沉淀,重悬浮于 30 μ l 的蒸馏水。测量 DNA 浓度后,其被用于电穿孔。

[0468] 2) 电穿孔

[0469] PG13 细胞系被用于电穿孔。将含有 7.5×10^5 个细胞的 0.5ml 无血清 DMEM 培养基加入到 0.4-cm 的电转杯 (Bio-Rad Laboratories, Hercules, CA),细胞与 10 μ l (1 μ g/ μ l) 的线性化反转录病毒质粒一同温育 5min,随后进行电穿孔。使用 Gene Pulser Xcell™ (Bio-Rad) 以 200V 电压进行 20msec 的电穿孔。电穿孔后,悬浮液立即铺板到含 10% 特级 FBS 的 DMEM 培养基。

[0470] 3) 筛选

[0471] 电穿孔后使用 G-418 (终浓度 1 μ g/ml) 筛选抗新霉素的细胞。筛选 10 天后,获得带有整合质粒的细胞。

[0472] 3. eGFP 表达的分析

[0473] PG13 生产细胞系的 5×10^6 个细胞被铺板到 100mm 培养皿上,培养基为含 10% 特级 FBS 的 DMEM 培养基。48h 后,收集上清,以 0.45 μ m 滤膜过滤。这些上清分别被用于实时 PCR 测量病毒滴度和转导 HT1080 细胞和 K562 细胞。温育细胞 48 小时,随后收集细胞进行检测。通过 FACS 分析测量 GFP 阳性细胞的比例和 GFP 的表达水平 (平均荧光强度)。

[0474] 4. gp91 表达的分析

[0475] PG13 生产细胞系的 5×10^6 个细胞被铺板到 100mm 培养皿上,培养基为含 10% 特级 FBS 的 DMEM 培养基。48h 后,收集含有病毒的上清,以 0.45 μ m 滤膜过滤。这些上清被用于转导 K562 细胞。

[0476] 对于 K562 细胞的转导,转导前一天在 6 孔板上每孔接种 2.5×10^5 个细胞。在 8 μ g/ml 聚凝胺的存在下每孔加入相同量的病毒上清,培养板在 32°C 下以 2800rpm 离心 90 分钟 (Eppendorf centrifuge 5810R)。转导后,细胞在 37°C 的 CO₂ 培养箱内温育 2 天。

[0477] 进行 FACS 定量分析 gp91 的蛋白表达。收集 K562 细胞, PBS 洗涤。细胞随后重新悬浮在 100 μ l PBS 中, 以 1 μ l 抗 -gp91 的抗体 (7D5) 在 4°C 下染色 30 分钟。随后细胞以 PBS 洗涤两次, 重新悬浮于 100 μ l PBS, 再以 1 μ l 结合 PE 的山羊抗小鼠抗体 (Southern Biotechnology Associates, Inc, Birmingham, AL, USA) 在 4°C 下染色 30 分钟。随后细胞以 PBS 洗涤 3 次, 重新悬浮于 500 μ l 的 PBS。借助 CellQuest 数据采集和分析软件以 FACSort 进行流式细胞术。图 4 给出了该结果。该图显示了 gp91 阳性细胞的比例和平均荧光强度 (括号内的数值)。正如图 4 所示, R1000 启动子可驱动高水平的 gp91 基因表达。而且, 当表达盒反向插入时, R1000 启动子可产生更高的病毒滴度和更高水平的基因表达。事实上, 含有反向 R1000 启动子的载体所产生的病毒滴度比具有野生型 MLV LTR 的 MT-gp91-n 更高 (参见 gp91 阳性细胞比例)。

[0478] 以 0.5% DMF (二甲基甲酰胺, C_3H_7NO) 处理细胞, 诱导细胞分化, 随后通过二羟罗丹明 -123 (DHR) 测定测量 NADPH 氧化酶的活性水平。细胞在 DMF 的存在下温育 6 天, 随后收集进行检测。收集的细胞以磷酸缓冲盐水 (PBS) 洗涤 2 次。细胞以 PBS 悬浮, 随后与 1.8 μ l DHR (Molecular Probes, USA, 29mM) 混合。37°C 温育 5 分钟后, 细胞以 10 μ l 的乙酰肉豆蔻佛波醇 (PMA, 1 μ g/ml) 进行刺激。随后通过 FACS 分析测量活性吞噬细胞的比例及其氧化酶活性。

[0479] 至此已充分地描述了本发明, 因此本领域的普通技术人员应理解, 在不影响本发明或其任意实施方式范围的情况下可在宽泛和等价范围内的条件、配方和其他参数下进行本发明。本文引用的所有专利、专利申请和出版物均通过引用的方式整体并入本文。

序列表

<110> 百疗医株式会社

<120> 安全性改善的表达载体

<130> PCA70979VML

<150> 60/850, 269

<151> 2006-10-10

<160> 87

<170> PatentIn version 3.3

<210> 1

<211> 1567

<212> DNA

<213> Artificial Sequence

<220>

<223> Virus HCMV IE promoter

<400> 1

tgacattgat tattgactag ttattaatag taatcaatta cggggtcatt agttcatagc	60
ccatataatgg agttccgcgt tacataactt acggtaaata gcccgcctgg ctgaccgccc	120
aacgaccccc gcccatigac gtcaataatg acgtatgttc ccatagtaac gccaataggg	180
actttccatt gacgtcaatg ggtggagtat ttacggtaaa ctgcccactt ggcagtacat	240
caagtgtatc atatgccaag tacgccccct attgacgtca atgacggtaa atggccccgc	300
tggcattatg ccagtagcat gaccttatgg gactttccta cttggcagta catctacgta	360
ttagtcatcg ctattaccat ggtgatgcgg ttttggcagt acatcaatgg gcgtggatag	420
cggtttgact cacggggatt tccaagtctc caccctattg acgtcaatgg gagtttgttt	480
tggcaccaaa atcaacggga ctttccaaaa tgtcgtaaca actccgcccc attgacgcaa	540
atgggcggta ggcgtgtacg gtgggaggtc tatataagca gagctcgttt agtgaaccgt	600
cagatgcctt ggagacgcca tccacgtgtt ttgacctc atagaagaca cgggaccga	660
tccagctcc gcgccggga acggtgcatt ggaacgcgga ttccccgtgc caagagtac	720
gtaagtaccg cctatagagt ctatagcccc acccccttgg cttcttatgc atgtataact	780
gtttttggct tggggtctat acacccccgc ttcctcatgt tataggtgat ggtatagctt	840
agcctatagg tgtgggttat tgaccattat tgaccactcc cctattgggt acgatacttt	900
ccattactaa tccataacat ggctctttgc cacaactctc tttattggct atatgccaat	960

```

acactgtcct tcagagactg acacggactc tgtatTTTTa caggatgggg tctcatttat 1020
tatttacaaa ttcacatata caacaccacc gtccccagtg cccgcagttt ttattaaaca 1080
taacgtggga tctccacgcg aatctcgggt acgtgttccg gacatgggct cttctccggt 1140
agcgcgaggag cttctacatc cgagccctgc tcccatgcct ccagcgactc atggtcgctc 1200
ggcagctcct tgctcctaac agtggaggcc agacttaggc acagcacgat gcccaccacc 1260
accagtgtgc cgcacaaggc cgtggcggtg gggatatgtg ctgaaaatga gctcggggag 1320
cgggcttgca ccgctgacgc atttggaaga cttaaggcag cggcagaaga agatgcaggc 1380
agctgagttg ttgtgttctg ataagagtca gaggttaactc ccgttgcggt gctgttaacg 1440
gtggagggca gtgtagtctg agcagtactc gttgctgccg cgcgcgccac cagacataat 1500
agctgacaga ctaacagact gtccctttcc atgggtcttt tctgcagtca ccgtccttga 1560
cacgatg 1567

```

<210>2

<211>431

<212>DNA

<213>Artificial Sequence

<220>

<223>Virus MLV U3 region

<400>2

```

gcaaggcatg gaaaaataca taactgagaa tagaaaagtt cagatcaagg tcaggaacag 60
atggaacagc tgaatatggg ccaaacagga tatctgtggt aagcagttcc tgccccggct 120
cagggccaaag aacagatgga acagctgaat atgggccaaa caggatatct gtggtaagca 180
gttcttgccc cggctcaggg ccaagaacag atggccccca gatgcgggtcc agccctcagc 240
agttttctaga gaaccatcag atgtttccag ggtgccccaa ggacctgaaa tgacctgtg 300
ccttatttga actaaccaat cagttegtt ctcgcttctg ttcgcgcgct tctgtcccc 360
gagctcaata aaagagccca caacccctca ctcggcgcgc cagtcttccg atagactgcg 420
tcgcccggt a 431

```

<210>3

<211>1703

<212>DNA

<213>Artificial Sequence

<220>

<223>Synthetic CMV enhancer/ubiquitin promoter

<400>3

ttgacattga ttattgacta gttattaata gtaatcaatt acggggtcat tagttcatag	60
cccatatatg gagttccgcg ttacataact tacggtaaatt ggcccgcctg gctgaccgcc	120
caacgacccc cgcccattga cgtcaataat gacgtatgtt cccatagtaa cgccaatagg	180
gactttccat tgacgtcaat ggggtggacta ttacggtaa actgcccact tggcagtaca	240
tcaagtgtat catatgccaa gtacgcccc tattgacgtc aatgacggta aatggccccgc	300
ctggcattat gcccagtaca tgacctatg ggactttcct acttggcagt acatctacgt	360
attagtcac gctattacca tggatgacg gttttggcag tacatcaatg ggcgtggata	420
gcggtttgac tcacggggat ttccaagtct ccacccatt gacgtcaatg ggagtttgtt	480
ttggctctag cggcctccgc gccgggtttt ggcgccctcc gcgggcgccc cctcctcac	540
ggcgagcgct gccacgtcag acgaaggcg cagcgagcgt cctgacctt ccgcccggac	600
gctcaggaca gcggcccgt gtcataaga ctggcctta gaaccccagt atcagcagaa	660
ggacatttta ggacgggact tgggtgactc tagggcactg gttttctttc cagagagcgg	720
aacaggcgag gaaaagtagt cccttctcgg cgattctcgc gagggatctc cgtggggcgg	780
tgaacccga tgattatata aggacgcgc ggggtgtggca cagctagttc cgtcgagcc	840
gggatttggg tcgcggttct tgtttgtgga tcgctgtgat cgtcacttgg tgagtagcgg	900
gctgctgggc tggccggggc ttctgtggc gccgggcccgc tcggtgggac ggaagcgtgt	960
ggagagaccg ccaagggtg tagtctgggt ccgcgagcaa ggttgccctg aactgggggt	1020
tggggggagc gcagcaaat ggcggctgtt cccgagctt gaatggaaga cgcttgtgag	1080
gcggcctgtg aggtcgttga aacaagggtg ggggcatggt gggcggcaag aacccaaggt	1140
cttgaggcct tcgctaattgc gggaaagctc ttattcgggt gagatgggct ggggcacat	1200
ctggggaccc tgacgtgaag ttgtcactg actggagaac tcggtttgtc gtctgttgcg	1260
ggggcggcag ttatggcggt gccgttgggc agtgcacccg tacctttggg agcgcgcgc	1320
ctcgtcgtgt cgtgacgtca cccgttctgt tggcttataa tgcagggtgg ggccacctgc	1380
cggtaggtgt gcggtaggct ttctccgtc gcaggacgca gggttcgggc ctaggtagg	1440
ctctcctgaa tcgacaggcg ccggacctct ggtgagggga gggataagt aggcgtcagt	1500
ttctttggtc ggttttatgt acctatctt ttaagtagct gaagctccgg ttttgaacta	1560
tgcgctcggg gttggcgagt gtgtttgtg aagttttta ggcacctttt gaaatgtaat	1620
catttgggtc aatatgtaat ttccagtgtt agactagtaa attgtccgct aaattctggc	1680
cgtttttggc tttttgtta gac	1703

<210>4

<211>1735

<212>DNA

<213>Artificial Sequence

<220>

<223>Synthetic CAG promoter

<400>4

attgattatt gactagttat taatagtaat caattacggg gtcattagtt catagcccat	60
---	----

atatggagtt ccgcgttaca taacttacgg taaatggccc gcctggctga ccgccaacg	120
acccccgccc attgacgtca ataatgacgt atgttcccat agtaacgcca atagggactt	180
tccattgacg tcaatgggtg gagtatttac ggtaaactgc ccacttggca gtacatcaag	240
tgtatcatat gccaagtacg cccctattg acgtcaatga cggtaaatgg cccgcctggc	300
attatgcca gtacatgacc ttatgggact ttcctacttg gcagtacatc tacgtattag	360
tcatcgctat taccatggtc gaggtgagcc ccacgttctg cttcactctc cccatctccc	420
ccccctcccc accccaatt ttgtatttat ttatttttta attattttgt gcagcgatgg	480
ggcggggggg gggggggggg cgcgcgccag gcggggcggg gcggggcgag gggcggggcg	540
gggcgaggcg gagaggtgcg gcggcagcca atcagagcgg cgcgctccga aagtttcctt	600
ttatggcgag gcggcgggcg cggcgccct ataaaaagcg aagcgcgcgg cggcggggag	660
tcgctgcgcg ctgccttcgc cccgtgcccc gctccgcgc cgcctcgcgc cgcccgcgcc	720
ggctctgact gaccgcgtta ctccacagg tgagcgggcg ggacggccct tctcctccgg	780
gctgtaatta gcgcttggtt taatgacggc ttgtttcttt tctgtggctg cgtgaaagcc	840
ttgaggggct ccgggagggc cttttgtgcg gggggagcgg ctcggggggg gcgtgcgtgt	900
gtgtgtgcgt ggggagcgcc gcgtgcggct ccgcgtgcc cggcggtgt gagcgctgcg	960
ggcgcggcgc ggggctttgt gcgctccga gtgtgcgcga ggggagcgcg gccgggggcg	1020
gtgccccgcg gtgcgggggg ggctgcgagg ggaacaaagg ctgcgtgcgg ggtgtgtgcg	1080
tgggggggtg agcagggggg gtgggcgcgt cggctgggct gcaaccccc ctgcaccccc	1140
ctccccgagt tgctgagcac ggcccggctt cgggtgcggg gctccgtacg gggcgtggcg	1200
cggggctcgc cgtgcggggc ggggggtggc ggacggtggg ggtgccgggc gggcggggc	1260
cgcctcgggc cggggagggc tcgggggagg gcgcggcg ccccgagac gccggcggt	1320
gtcagagcgc ggcgagcgc agccattgcc ttttatggta atcgtgcgag agggcgagg	1380
gacttccttt gtcccaaate tgtgcggagc cgaaatctgg gaggcgccgc cgcaccccc	1440
ctagcgggcg cggggcgaag cgggtgcggc cggcgaggaa ggaaatgggc ggggagggc	1500
ttcgtgcgtc gccgcgccgc cgtcccttc tccctctcca gcctcggggc tgtccgcggg	1560
gggacggctg ctttcggggg ggacggggca gggcggggtt cggcttctgg cgtgtgaccg	1620
gcggtctag agcctctgct aaccatgttc atgccttctt ctttttcta cagtccttg	1680
gcaacgtgct ggttattgtg ctgtctcacc attttggcaa agaattgatt tatcg	1735

<210>5

<211>1450

<212>DNA

<213>Artificial Sequence

<220>

<223>Homo sapiens EF-lalpha promoter

<400>5

gtaagccagc aatggtagag ggaagattct gcacgtccct tccaggcggc ctccccgtca	60
ccaccccccc caaccgccc cgaccggagc tgagagtaat tcatacaaaa ggactcgccc	120

ctgccttggg gaatcccagg gaccgtcggtt aaactccac taacgtagaa cccagagatc	180
gctgcgttcc cgccccctca cccgcccgt ctcgtcatca ctgagggtgga gaagagcatg	240
cgtgaggctc cgggtgcccgt cagtgggcag agcgcacatc gccacagtc cccgagaagt	300
tggggggagg ggtcggcaat tgaaccggtg cctagagaag gtggcgcggg gtaaactggg	360
aaagtgatgt cgtgtactgg ctccgccttt tccccagggtg tgggggagaa ccgtatataa	420
gtgcagtagt cgccgtgaac gttctttttc gcaacgggtt tgccgccaga acacaggtaa	480
gtgccgtgtg tggttcccgc gggcctggcc tctttacggg ttatggccct tgcgtgcctt	540
gaattacttc cacgcccctg gctgcagtac gtgattcttg atcccagct tcgggttggg	600
agtgggtggg agagttcgag gccttgcgt taaggagccc cttcgcctcg tgcttgagtt	660
gaggcctggc ctgggcgctg gggccgcgc gtgcgaatct ggtggcacct tcgcgcctgt	720
ctcgtgctt tcgataagtc tctagccatt taaaattttt gatgacctgc tgcgacgctt	780
tttttctggc aagatagtct tgtaaatgcg ggccaagatc tgcacactgg tatttcggtt	840
tttggggccg cgggcggcga cggggcccgt gcgtcccagc gcacatgttc ggcgaggcgg	900
ggcctgcgag cgcggccacc gagaatcgga cgggggtagt ctcaagctgg ccggcctgct	960
ctggtgcctg gcctcgcgc gccgtgtatc gccccgcctt gggcggcaag gctggcccgg	1020
tcggcaccag ttgcgtgagc ggaaagatgg ccgcttcccg gccctgctgc agggagctca	1080
aaatggagga cgcggcgctc gggagagcgg gcgggtgagt caccacaca aaggaaaagg	1140
gcctttccgt cctcagcgt cgcttcatgt gactccacgg agtaccgggc gccgtccagg	1200
cacctcgatt agttctcgag cttttggagt acgtcgtctt taggttgggg ggaggggttt	1260
tatgcgatgg agtttccca cactgagtgg gtggagactg aagttaggcc agcttggcac	1320
ttgatgtaat tctccttggg atttgccctt tttaggttg gatcttggtt cattctcaag	1380
cctcagacag tggttcaaag ttttttctt ccatttcagg tgtcgtgaaa actacccta	1440
aaagccaaaa	1450

<210>6

<211>1330

<212>DNA

<213>Artificial Sequence

<220>

<223>Homo sapiens ACTB promoter

<400>6

gagatgtcca cacctaggat gtcccgcggt ggggtgggggg cccgagagac gggcaggccg	60
ggggcaggcc tggccatgcg gggccgaacc gggcactgcc cagcgtgggg cgcgggggcc	120
acggcgcgcg cccccagccc ccgggcccag caccccaagg cggccaacgc caaaactctc	180
cctcctctc ttcctcaatc tcgtctctgc tctttttttt ttctgcaaaa ggaggggaga	240
gggggtaaaa aaatgtgca ctgtgcggcg aagccggtga gtgagcggcg cggggccaat	300
cagcgtgcgc cgttccgaaa gttgcctttt atggctcgag cggccgcggc ggcgcctat	360
aaaaccacgc ggcgcgacgc gccaccaccg ccgagaccgc gtccgccccg cgagcacaga	420

gcctcgccctt tgccgateccg ccgcccgtcc acacccgccg ccaggtaagc ccggccagcc	480
gaccggggca ggccggtcac ggcccggccg caggcgccg cgcccccttc gcccgtgcag	540
agccgccgtc tgggcccag cgggggggcgc atgggggggg aaccggaccg ccgtgggggg	600
cgccggagaa gcccctgggc ctccggagat gggggacacc ccacgccagt tcggaggcgc	660
gaggccgcgc tcgggaggcg cgctccgggg gtgccgtctt cggggcgggg gcaaccggcg	720
gggtctttgt ctgagccggg ctcttgccaa tggggatcgc aggggtggcg cggcggagcc	780
cccgccaggc ccggtggggg ctggggcgcc attgcgcgtg cgcgctggtc ctttgggcgc	840
taactgcgtg cgcgctggga attggcgcta attgcgcgtg cgcgctggga ctcaaggcgc	900
taactgcgcg tgcgttcttg ggcccggggg gccgcggcct gggctggggc gaaggcgggc	960
tcggccgga ggggtggggg cgcccgggct cccgggcgct tgcgcgact tcctgcccga	1020
gccgctggcc gcccagagggt gtggccgctg cgtgcgcgcg cgccgacccg gcgctgtttg	1080
aaccgggcgg aggcggggct ggcgcccggt tgggaggggg ttggggcctg gcttccctgc	1140
gcgcgccgcg gggacgcctc cgaccagtgt ttgcctttta tggtaataac gcggccggcc	1200
cggcttcctt tgtccccaat ctgggcgcgc gccggcgccc cctggcggcc taaggactcg	1260
gcgcgccgga agtggccagg gcgggggcga cctcggtca cagcgcgccc ggctattctc	1320
gcagtcacc	1330

<210>7

<211>665

<212>DNA

<213>Artificial Sequence

<220>

<223>Homo sapiens GAPDH promoter

<400>7

ttcatccaag cgtgtaaggg tccccgtcct tgactcccta gtgtcctgct gcccacagtc	60
cagtcctggg aaccagcacc gatcacctcc catcgggcca atctcagtc cttccccct	120
acgtcggggc ccacacgctc ggtgcgtgcc cagttgaacc aggcggctgc ggaaaaaaaa	180
aagcggggag aaagtagggc ccggctacta gcggttttac gggcgcacgt agtcaggcc	240
tcaagacctt gggctgggac tggctgagcc tggcgggagg cggggtccga gtcaccgcct	300
gccgcgcgc ccccggtttc tataaattga gccgcagcc tcccgcttcg ctctctgctc	360
ctcctgttcg acagtcagcc gcatcttctt ttgcgtcgcc aggtgaagac gggcggagag	420
aaaccggga ggctagggac ggcctgaagg cggcaggggc gggcgaggc cgcatgtgtt	480
cgcgccgctg cgggggtggg ccgggcggcc tccgcatgac aggggcgggc ggaggacgtg	540
atgcggcgcg ggctgggcat ggaggcctgg tgggggaggg gaggggaggc gtgtgtgtcg	600
gccggggcca ctaggcgctc actgttctct ccctccgcgc agccgagcca catcgctcag	660
acacc	665

<210>8

<211>1143

<212>DNA

<213>Artificial Sequence

<220>

<223>Homo sapiens RPL10 promoter

<400>8

```

gtgcgctcga gcaggatttc ctcccgtcct tcctgtcaaa ggacgggaag actttgttac      60
cccaccgcgc ccacactgca gaatgggtgga cagatacctc cagatgccac ttccccagg      120
acgcccgcct gctctgcgca cctctccccg gatgctgccc cgtgggcggg tgggggcggc      180
cctgcttccc cacgaccccc agacgcaccc ggagggactc ttgagcacag tggagtggga      240
agggcgaggt ggggcggtgc ccaggcgaga gcggtcatg ggaggcggcg cccgagacgc      300
agctggtcgg gacggtgcgg gtcagggtgg gcgagcggg gctagagatg ccccggggtt      360
tcccaggcca tgagtctccg tggagatttc tcctcgacct cttccccgcg gcaatgtgcg      420
aaccctgggt ctccaggaaa cggggatacg gggcatggct ccagcaagg cctggtccag      480
cctctccggt aggggaatgg gtctccccct ccggcctccc gggttgacaa aggaacgcgg      540
gccagatcc ccgtatggcg cttcacgcgc ggggcctcta gcctagaagg aggcacggag      600
cgcggtgtcc agaccctgc aagctcaggg acactctcgc ggtcgccggg agggccacct      660
agggtacttt ctttttttcc actctcagaa atatacgtct gtcacagtta acggcaaagc      720
ctagggcaag agttctacgc ccaagatggc cagccggaag cgggcttctc gcgaccatgt      780
ggcgaagccc cattcgtcag ctggccgccc gcgcccttgg taccgggtca cctctctgat      840
ctgcgcatgt gctgggctac gcccgggcgc aagcgccaag agcggtgcg tctatggtca      900
tgacgtctga cagagcgtec acccgcttc gacaggactc tatggttctt acgcgcgcag      960
acagaccgcc tatataagcc atgcgcaggc ggaggagcgc ctctttccct tcggtgtggt      1020
gagtaagcgc agttgtcgtc tcttgcggtg ccgttgctgg ttctcacacc ttttaggtct      1080
gttctcgtct tccgttccga ctctctcttt ttcgttgcag ccactgaaga tcctggtgtc      1140
gcc                                                  1143

```

<210>9

<211>1538

<212>DNA

<213>Artificial Sequence

<220>

<223>Homo sapiens LENG8 promoter

<400>9

```

ttgtatcaga gtctggacg gaaacagatg gcactcaaaa ggtggcgcgc agttcagaga      60

```

aatgcctatg tacggatttg gtccaatgcc tcagcctgac ctcagggacc ttcgggggtc	120
tgtctccgcgc ccacccttac acatctgtga cccacacac ttccacccca gcgccactgc	180
caacagctac acccateccc ctccaaccgc gtcagcttcc agcctcggtc catctgaact	240
cgcctgccc cctccccctgc gcccttcag attcatttgc tagggaagcc cgctcttccg	300
ggtggagctg ttcctcatec cttttcttta tcattctctc cccagggctt ccacatcacc	360
gtgctgtgga caatcccgga actcctgtca cgccagttaa catttaggaa cagtaatggc	420
tccactgac tcagtcaaaa caaggctgcg gccgggcacg gtggctcacg cccgtaatcc	480
cagcactttg ggaagccgag acggagggat cacgaggtca ggagttcgag aacagcctgg	540
ccgacatggt gaaaccccg ttttcaaaa acacaaaaat tagccgggca tagtggcgcg	600
cgcctgtaat cccagctact ccggaggctg aggcagaatt gtttgaacc aggaggcgga	660
ggttgcagtg agcagagatc acgccactgt actctatcgt gggcgacgac agagcaagag	720
caagactccg tctccgagaa caacaacaac agcaacaaga aaacaacaat aaaaaaata	780
aggctgcgtg ggaggcagaa agagctaatg cgccacgct tgtccctcgc gggccaccgt	840
ccccaccag acttccggtc tgccttaaaa tgttcatgcg taagtgcgtg ggcaggaagg	900
cgggctcaag cgcagctcgt ggcgttcatt ggctgtgcag ggccgaggga ggcggtgcaa	960
ggccgcccgc tgacgtcagg acgccgcggt caggacgtcg aagccaaaga agaccagagc	1020
cagccgggtg gcacagcggg gtcgtggccg tgttctgat cgcctgggtg gttgttggcg	1080
tgtccctgca gcgaaggatc ctggttggta aggggagcgg cgggcgagca ggcgggcggg	1140
gatagcatct ctttttggtc ttgcgccccg cgagccccga ggccttctcg gccgtcgcag	1200
cagcagacgc cgcgcgcgag cgtcgacagg gtgtggcggc gcaggggcag ccactgcgcc	1260
tgcgcaccgg gcctggggcc gcgcgttcgg gcactagcgc gcgtgcgccg tcgtcttcta	1320
ctttccacac ccagaactct tcagatcctt gaccccagtg gcttttcagt cagcctcccc	1380
ttttctgccc agcttctctt gagtccatct acttttcttc cccactttgt gacgtgtttt	1440
tagctcccc ttaagtctcc ctaactcatt ctttttctca taggcagtga aaaagcagtc	1500
tggtctccga ggtccacccc ttatacccca aggtccag	1538

<210>10

<211>691

<212>DNA

<213>Artificial Sequence

<220>

<223>Homo sapiens SNX3 promoter

<400>10

aatccagacg cgtgtctggt gcaacgctcg ggtttatggc aaaatcatct caggcatttg	60
cttaaccttc tccagaaagg cattttcagg ggttcacagt gagacgggtc acaggttggc	120
acagagttag taggggcagt ttigtcttga ttgctgggca aatctctaag atctctccgt	180
ttaactttcg cccgcaattc ccaaagccgc taaagccgtt tccggcgctc taccgccgcc	240

caggccgagg	ctggcgcaga	gagacaggaa	gcgccagctc	tgggcgtctg	ggtcctcgcc	300
tcctcggccg	cagcccccg	gcggcgcgct	cgcggtgcat	tgtgggcgct	gtagtccggc	360
cggaacctgt	ttgcgacccc	gagtccecatg	acaccgcttc	tcctcacacc	ccagtccgca	420
gtgcccctcc	ccagcctcgg	ccgggcctcc	cgggagccgg	gcgtggcggt	ccagctagt	480
agccgtttct	cccctgggct	cggaggcgga	agcttgaggg	gcgcggggag	gagcttcgcg	540
tgcggggtga	acgcccgcctc	tacgtgctcg	ttctcttcgc	gaccgctgcg	cgcgagcccc	600
gtgtccccac	ggcgggcagc	agcggcgggc	gcggcggtcg	aacgcggagg	gggcggagg	660
agcccgcggc	ggcggcagca	gctacagcga	a			691

<210>11

<211>1022

<212>DNA

<213>Artificial Sequence

<220>

<223>Homo sapiens CNOT3 promoter

<400>11

gtagctcctc	ccccagacca	attgttttaa	gagagggggg	cggatacatc	caatcagcac	60
gacacaggtc	tcttgattga	cgttcgggtc	ctcgcgctgg	cgtgttgtgc	cctgaggcgg	120
gaggaggagg	aggagcgggg	aggaaaacct	gagccaatcc	tagcagcctg	cgcgggaggc	180
caatcgaacg	ccgcgccttg	gagcgatcac	ccaatccgcg	aaagggggca	gggcgcatcc	240
ctgccaggaa	ccaatagaaa	gcctccaagg	gtcaggagcg	acgttcagca	ggagcaatga	300
ctggcctata	ttcgggactc	gggggcgggt	cggcgccaga	gacgagaaga	gaggagggga	360
ggcctcctcc	gccgcgcga	tcttgaccgc	ggcccgggtc	gcttcgcggg	agccatcggc	420
agacgccgcg	gcctcccttg	agccccgacc	cccgtcgtca	gaacaacccc	ggggccactc	480
ccccaacccc	acttcgcctt	cgcgcgcgta	tcgcgatagc	gcccggggcc	ggggcgcgag	540
aaaaaggcgg	cgggcgcctg	cctccccgcg	ctgtcgcgat	acgtcctca	gcggcgggcg	600
cagctcctgt	ggtgagagcg	tcaggctega	ctgggccgga	ccccttcctt	tcctcccccc	660
ggcgccatcg	gccgcctcc	ccgcgcctc	ccgccctggc	gacaccgccc	tctgtcgcga	720
catggcctcc	cctcgcctgc	ccccctgcgc	cgcctctgca	gcgcggggct	cccggcgggg	780
ggcggtctcc	tcctctcgc	cctcccgctt	ctgcgcctct	ttcacgttcc	ttggggctgg	840
tctcttgtea	gatagcaaat	gtttcttctc	tttaccagtc	ccacctacct	cactatgctg	900
actaggtcca	tgtctctggg	tttttaccag	ccagggaata	cgtgttaatt	cctctccaat	960
ctctcctage	agcgtccgct	tccaagagag	tatgaagaga	gtgcgtctgt	agggcaggga	1020
ag						1022

<210>12

<211>1175

<212>DNA

<213>Artificial Sequence

<220>

<223>Homo sapiens CPNE1 promoter

<400>12

```
gtccatttaa tcctcaaaaa acttagaggt aagcacaatt cgctccttcc gcggaaactg    60
gaagacgtga agcaatttgc acagtcttgc agtgaagggc caagctagac actcataccc    120
agcccacccg ccagggccac tagcccaagg tcggtcctgt gcgccaacag tcgcaagcag    180
ggactgaata tccccctctt agcacaaaac ccagtaactt ccgggagatc gcccagagacc    240
gtcagcgcgc ttgcgcggcc ccttgacgtc aaccacgccg ttcttcccc ttctggaacg    300
cggagccgct gagcgggccg acggccatct tgtgaagcgg cgaaggaggt ggtggctgcg    360
ttgggctccg ggaagccgtt cgggctgggg ctgtcggccg cggggcggag gcactcgcgc    420
ggggggtaat tcggggtctg ggttctgggt cgcgcagct ttccccggta agactcccgc    480
agcccctgaa cgggtggggc tgtgcggggc tcgtggtgcc ccttggtggc ccggggcggg    540
gccttcggag gccttcgagc ccgcggcaac tagcgcacca cacaagggt cgaggcaggt    600
ctgcaggggg agtgggatgc cacttaggcc tctgagatct gagaggatgc aaggaggacg    660
caagaggatg agcctgcgta ccgaagtggg gacgaggcca tagcaggac tccccgtgg    720
tctgagaggg ggcaggccag agctggctgg ctacgcggcc tgcgagctaa ccggggactg    780
cagtaaaaaa gctgtgaagc tgagctttat gactgtttta attttcagt gattcgttat    840
tgggctaagt taatgacact cttgtaaata tataattatt ttgagatact caaatttgaa    900
agcaatacca actaaaaagc atgtgtggga ttttgttacg gtgtctttct ttttttttt    960
ttttttgata tgcattctaa tgaagtactc tgggtggttc agactcgcct tcgtaatgta    1020
cagtgtggcc tgtgcgtgtt gttttgtgtt tcctgcattg atttttgcct ccgtttattc    1080
tctattgcag tctaaaagtt ggttttaatt ggttgcccac aggattgact tggcctctac    1140
ttcttgttaa ggaaattcat ctcttgtttt atcag                                1175
```

<210>13

<211>416

<212>DNA

<213>Artificial Sequence

<220>

<223>Homo sapiens HYP0 promoter

<400>13

```
tcttttacac gtttggtttt atggtccgtt ggaggttttt tcacgtgag ggtgaagatt    60
cttagtttcc tcaattagca taaaagctgc tccaaggctg ctgagtcttc agttccattc    120
cattcaaaat gttaattcct gagacacgca acctacgccc tactgtcctt tgtcccacgg    180
```

cttctcgaga ttatctctgt gttttaatag ataaaacgcc ttctctcccc acccgtgggg	240
tgcaaagcac agcgcattggg gctggtggcg cccgagagca acacacaacg catgcgccag	300
tccgcaggtg tgggcggagg agaaatcgcg tcggcggcag gggatgacgt aaaaaggccg	360
cgctgtactg cggcttgtgc cgcttcgcga agaaggtttc ctggcctgtt gcagcc	416

<210>14

<211>564

<212>DNA

<213>Artificial Sequence

<220>

<223>Homo sapiens DKC1 promoter

<400>14

gcacactact cctattggcc atctgctcta gctggatggg aactgcaagc agcaaaggcc	60
tttcaacctc tccgagcttc aacatgttca cttatgcata aggactgata caacagcaac	120
ctgaaagggt cgggtgtgaga aagcaaagaa agaggtactg ttacaggagc gttcagccgt	180
gagcccaggc gcaggcgcgg agctaaaccc ggaggtgaag caactacagc tggggccacg	240
cccacagact cgggctgtc ggggcgggcg gcactcgcgc ttccagcctg ggccaaccag	300
gctgcttgag gctcgaagcg catgcgcggt ctgcacctcc gtactggccg agccagcaaa	360
tcgcattgcg cagacgacca gcgggcgcct cggattccgc ccccgggatg gccccgcctc	420
ctccgcgcc gcggcaaggc acgcacaggg cagtgcgcgg gtgggtgggt cctagcagcg	480
cggcctgacg ggaccaaggc ggcgggagtc tgcggtcggt ccctcggctg tggaccgggc	540
ggcacgcacg cgggtgcagg taac	564

<210>15

<211>509

<212>DNA

<213>Artificial Sequence

<220>

<223>Homo sapiens VPS72 promoter

<400>15

acaaaaatta gttgggcatg gtggcgggcg cctgtaatcc cagctactcg ggggactgag	60
gcaggagaat cgcttcagcc cgggggacgg aggttgcagt gagccgagat cgtgccactg	120
cactccagtc tggcgacaga gcgaggcgag actccgtccc tacggaaaaa aaaaaccata	180
tatatatata aaataaaaaat aaataaaata aaatgtgttg acttaggggt atgggggtct	240
tcgtaccacc aagtggactc attgagcaag ggagatcgct gggtagtaga gcaagcggcc	300
tagagcatcc gcgcaggagt aggagaggct gaagagtaaa gagggtttag aggggtgggac	360

tgaggagtctc tgcacactct gttactttctg cctgtaaacg cccgacttcc gccgctgggtg	420
gccacccgca ggtagtgatg tcgagcgctg agctcccaaa accgagctgg tgaggggctg	480
caggtggcgg cgcagtctcg gtaggcggt	509

<210>16

<211>654

<212>DNA

<213>Artificial Sequence

<220>

<223>Homo sapiens IT6B4BP promoter

<400>16

tctgtccctc aaggcacage tggactcacc cctttcccca aacctgccct tcctccggct	60
tcatttccat caatacctcc atcatcaacc ctcccgcgag acactcctgg cccctcctct	120
ccctcatgcc tcacaaccga ccagccggag gtctaggtcg atgacagctc ctaaaaagct	180
cctgaatgaa taatgaatga atgaacgcga gcaggctagg cgtggggcca ggcggggctg	240
cgcccagacc gctcgcgacc atagagtcgc ccggaggccg gaggtagagg ggctggatgc	300
gtggcgggga ggcgcgggct ctcccgaag tctccctgga cggaagtga aacggaaacc	360
tttttaggga gtccaaggta cagtcgccgc gtgcggagct tgttactggt tacttggtaa	420
gctggtgtga ggggaacctg ggagggtcag ctccggctct gggtcgggag gggtaggggc	480
cagaggattc agggccggag gttctggtgg gggcccagtg ggcgggaccc gaggacggag	540
gggccgggag gccgagaggg gcggggctgc ggccgggcct gagggacgga ggccgggata	600
cttgggaaag gatccgcccgc cttgaactc ccgcctccgc cgcccctagg cctc	654

<210>17

<211>567

<212>DNA

<213>Artificial Sequence

<220>

<223>Homo sapiens UCRQ promoter

<400>17

gtcacctttt tgttccctcc cccgcctccc gcattcggcc gcttcctgac tgggattcca	60
cagaaaagcc gagggctgag gagaagtgtg agcgccctccg cctgtccact gtcccccaaa	120
gtcagttcaa tccccgacgt cctccgctag gtccaccccc accggcccgg gcagggcctc	180
caaggcacct cccacctacg ggtaaccag tcagcccaact tctttctggg acaaaggcgt	240
catcccttag agacagtagg aaaatggtat ctcccgaag ttacctcacg acctccaaga	300

gcggcttcca accttgccgg aaatgacgaa cgagtcaacc ggatcgggtga ctgtggaggg	360
cgagctgagc cctgtgcgtg agtgggggtct ggttgtgcag tgttcgtgga ccctgggagg	420
ctaggggcgc cccgctgggc tgggaaagga taaggagtgc aggggcagga gtctggggtt	480
gggatggac ccccgcgggg actgcggcgc ttcgcgaaag cgagccaagc gcctgtccac	540
cctcggtcct gcagggcgc cgccaca	567

<210>18

<211>22

<212>DNA

<213>Artificial Sequence

<220>

<223>ME5primer

<400>18

acgcgtgcaa ggcatggaaa aa

22

<210>19

<211>35

<212>DNA

<213>Artificial Sequence

<220>

<223>MP3 primer

<400>19

acgcgtagat ctgaattcta cccgggcgac gcagt

35

<210>20

<211>21

<212>DNA

<213>Artificial Sequence

<220>

<223>CMV5 primer

<400>20

acgcgttgac attgattatt g

21

<210>21

<211>24

<212>DNA

<213>Artificial Sequence

<220>

<223>KMD1 primer

<400>21

tctagagcca aaacaaactc ccat

24

<210>22

<211>24

<212>DNA

<213>Artificial Sequence

<220>

<223>KMD4 primer

<400>22

gctagcggcc tccgcgccgg gttt

24

<210>23

<211>36

<212>DNA

<213>Artificial Sequence

<220>

<223>KMD5 primer

<400>23

acgcgtagat ctgaattcgt ctaacaaaaa agccaa

36

<210>24

<211>41

<212>DNA

<213>Artificial Sequence

<220>

<223>EEF1A1F primer

<400>24
acgcgtgtaa gccagcaatg gtagaggga gattctgcac g 41

<210>25
<211>38
<212>DNA
<213>Artificial Sequence

<220>
<223>EEF1A1R primer

<400>25
ggatcctttt ggcttttagg ggtagttttc acgacacc 38

<210>26
<211>30
<212>DNA
<213>Artificial Sequence

<220>
<223>BApF primer

<400>26
acgcgtgaga tgtccacacc taggatgtcc 30

<210>27
<211>27
<212>DNA
<213>Artificial Sequence

<220>
<223>BApR primer

<400>27
ggatccgggtg agctgcgaga atagccg 27

<210>28
<211>26
<212>DNA

<213>Artificial Sequence

<220>

<223>GAPDHF primer

<400>28

acgcgtttca tccaagcgtg taaggg 26

<210>29

<211>28

<212>DNA

<213>Artificial Sequence

<220>

<223>GAPDHR primer

<400>29

gtttaaacgg tgtctgagcg atgtggct 28

<210>30

<211>31

<212>DNA

<213>Artificial Sequence

<220>

<223>RPLF primer

<400>30

acgcgtaggc ccacctaggg tactttcctt t 31

<210>31

<211>31

<212>DNA

<213>Artificial Sequence

<220>

<223>RPLR primer

<400>31

ggatccggcg acaccaggat cttcagtggc t 31

<210>32

<211>31

<212>DNA

<213>Artificial Sequence

<220>

<223>LENG8F1 primer

<400>32

acgcgtagaa ttgtttgaac ccaggaggcg g 31

<210>33

<211>33

<212>DNA

<213>Artificial Sequence

<220>

<223>LENG8R1 primer

<400>33

gtttaaacia agtagaagac gacggcgcac gcg 33

<210>34

<211>33

<212>DNA

<213>Artificial Sequence

<220>

<223>LENG8F2 primer

<400>34

gtttaaacce acaccagaa ctcttcagat cct 33

<210>35

<211>31

<212>DNA

<213>Artificial Sequence

<220>

<223>LENG8R2 primer

<400>35

gaattccttg accttggggt ataaggggtg g 31

<210>36

<211>30

<212>DNA

<213>Artificial Sequence

<220>

<223>SNX3F primer

<400>36

gaattcaatc cagacgcgtg tctggtgcaa 30

<210>37

<211>22

<212>DNA

<213>Artificial Sequence

<220>

<223>SNX3R primer

<400>37

ggatccttcg ctgtagctgc tg 22

<210>38

<211>37

<212>DNA

<213>Artificial Sequence

<220>

<223>CN0T3F1 primer

<400>38

acgcgtgtag ctctccccc agaccaattg ttttaag 37

<210>39

<211>37

<212>DNA

<213>Artificial Sequence

<220>

<223>CN0T3R1 primer

<400>39

ggatcctcca tccttc cage caggagccaa taccgac 37

<210>40

<211>31

<212>DNA

<213>Artificial Sequence

<220>

<223>CN0TF2 primer

<400>40

agatcttggg gctggctctt tgcagatag c 31

<210>41

<211>31

<212>DNA

<213>Artificial Sequence

<220>

<223>CN0TR2 primer

<400>41

ggatcccttc cctgccctac agacgcactc t 31

<210>42

<211>31

<212>DNA

<213>Artificial Sequence

<220>

<223>CPNE1F1 primer

<400>42

acgcgtgtcc atttaatcct caaaaaactt a 31

<210>43

<211>34

<212>DNA

<213>Artificial Sequence

<220>

<223>CPNE1R1 primer

<400>43

ggatcctttt tactgcagtc cccgttatta gctc 34

<210>44

<211>31

<212>DNA

<213>Artificial Sequence

<220>

<223>CPNE1F2 primer

<400>44

agatctagct gtgaagctga gctttatgac t 31

<210>45

<211>31

<212>DNA

<213>Artificial Sequence

<220>

<223>CPNE1R2 primer

<400>45

ggatccctga taaaacaaga gatgaatttc c 31

<210>46

<211>31

<212>DNA

<213>Artificial Sequence

<220>

<223>HYPOF primer

<400>46

acgcgttctt ttacacgttt ggttttatgg t 31

<210>47

<211>31

<212>DNA

<213>Artificial Sequence

<220>

<223>HYPOR primer

<400>47

ggatccggct gcaacaggcc aggaaacctt c 31

<210>48

<211>25

<212>DNA

<213>Artificial Sequence

<220>

<223>DKC1F primer

<400>48

acgcgtgcac actactccta ttggc 25

<210>49

<211>25

<212>DNA

<213>Artificial Sequence

<220>

<223>DKC1R primer

<400>49

gaattcggtta ccctgcaccg cgtgc 25

<210>50

<211>25

<212>DNA

<213>Artificial Sequence

<220>

<223>VPS72F primer

<400>50

acgcgtacaa aaattagttg ggcat 25

<210>51

<211>25

<212>DNA

<213>Artificial Sequence

<220>

<223>VPS72R primer

<400>51

gaattcaccg cctaccgaga ctgcg 25

<210>52

<211>27

<212>DNA

<213>Artificial Sequence

<220>

<223>ITGB4BPF primer

<400>52

acgcgttctg tccctcaagg cacagct 27

<210>53

<211>40

<212>DNA

<213>Artificial Sequence

<220>

<223>ITGB4BPR primer

<400>53

gtttaaacga ggcctagggg cggcggaggc gggagttcaa 40

<210>54

<211>27

<212>DNA

<213>Artificial Sequence

<220>

<223>UQCRQF primer

<400>54

acgcgtgtca cctttttgtt ccctccc 27

<210>55

<211>28

<212>DNA

<213>Artificial Sequence

<220>

<223>UQCRQR primer

<400>55

gtttaaactg tggcggcggc cctgcagg 28

<210>56

<211>22

<212>DNA

<213>Artificial Sequence

<220>

<223>RGpA F primer

<400>56

ggatcctttt ccctctgccaa aa 22

<210>57

<211>25

<212>DNA

<213>Artificial Sequence

<220>

<223>RGpA R primer

<400>57

actagtataa gagaagaggg acagc

25

<210>58

<211>38

<212>DNA

<213>Artificial Sequence

<220>

<223>RPL F50 primer

<400>58

acgcgtacgc gcgcagacag accgcctata taagccat

38

<210>59

<211>37

<212>DNA

<213>Artificial Sequence

<220>

<223>RPL F100 primer

<400>59

acgcgttgac gtctgacaga gcgtccaccc gtcttcg

37

<210>60

<211>38

<212>DNA

<213>Artificial Sequence

<220>

<223>RPL F200 primer

<400>60

acgcgtctgg ccgcccgcgg ccctgggtacc cggtcacc 38

<210>61

<211>39

<212>DNA

<213>Artificial Sequence

<220>

<223>RPL F500 primer

<400>61

acgcgtgtct cccctccgg cctcccgggt tgacaaagg 39

<210>62

<211>39

<212>DNA

<213>Artificial Sequence

<220>

<223>RPL F1000 primer

<400>62

acgcgtgtgc gctcgagcag gatttctctc cgtccttcc 39

<210>63

<211>36

<212>DNA

<213>Artificial Sequence

<220>

<223>RPL R TSS primer

<400>63

ggatccgcgc tcctccgcct gcgcatggct tatata 36

<210>64

<211>31

<212>DNA

<213>Artificial Sequence

<220>

<223>LENG F50 primer

<400>64

acgcgtgtga cgtcaggacg ccgcggtcag g 31

<210>65

<211>31

<212>DNA

<213>Artificial Sequence

<220>

<223>LENG F100 primer

<400>65

acgcgttggc gttcattggc tgtgcagggc c 31

<210>66

<211>31

<212>DNA

<213>Artificial Sequence

<220>

<223>LENG F200 primer

<400>66

acgcgtttgt cccctcgggg ccaccgtccc c 31

<210>67

<211>31

<212>DNA

<213>Artificial Sequence

<220>

<223>LENG F1000 primer

<400>67

acgcgtttgt atcagagtcc tggacggaaa c 31

<210>68	
<211>33	
<212>DNA	
<213>Artificial Sequence	
<220>	
<223>LENG R TSS primer	
<400>68	
gtttaaacct ctggtcttct ttggttcga cgt	33
<210>69	
<211>42	
<212>DNA	
<213>Artificial Sequence	
<220>	
<223>SCV3LB primer	
<400>69	
ggatccctcg agcgataaaa taaaagattt tatttagtct cc	42
<210>70	
<211>32	
<212>DNA	
<213>Artificial Sequence	
<220>	
<223>SCV3LRI primer	
<400>70	
gaattcgctcg actgaaagac ccccgctgac gg	32
<210>71	
<211>24	
<212>DNA	
<213>Artificial Sequence	
<220>	
<223>3' LTR-1 primer	

<400>71
gctagcccct gtgccttatt tgaa 24

<210>72
<211>36
<212>DNA
<213>Artificial Sequence

<220>
<223>NEWMCSF primer

<400>72
acgcgtttta accgcggaat tcggatccac atcgtg 36

<210>73
<211>36
<212>DNA
<213>Artificial Sequence
<220>
<223>NEWMCSR primer

<400>73
ctcgagatct aggcctcacg atgtggatcc gaattc 36

<210>74
<211>30
<212>DNA
<213>Artificial Sequence

<220>
<223>eGFP5 primer

<400>74
acgcgtggat ccatggtgag caagggcgag 30

<210>75
<211>30
<212>DNA

<213>Artificial Sequence

<220>

<223>eGFP3 primer

<400>75

ctcgagagat ctttacttgt acagctcgtc 30

<210>76

<211>27

<212>DNA

<213>Artificial Sequence

<220>

<223>GP91F primer

<400>76

ggatccatgg ggaactgggc tgtgaat 27

<210>77

<211>35

<212>DNA

<213>Artificial Sequence

<220>

<223>GP91R primer

<400>77

ggatccctcg agttagaagt tttccttggt gaaaa 35

<210>78

<211>37

<212>DNA

<213>Artificial Sequence

<220>

<223>L1F primer

<400>78

gctcttccgc tcacgtgtga tcaatttaaa tttcgaa

37

<210>79

<211>37

<212>DNA

<213>Artificial Sequence

<220>

<223>L1R primer

<400>79

agcggaagag cttcgaaatt taaattgatc acacgtg

37

<210>80

<211>63

<212>DNA

<213>Artificial Sequence

<220>

<223>L2F primer

<400>80

aggcctggtc accggccatt atggccacgt gatcatttaa atttgaagca tttatcaggg
tta

60

63

<210>81

<211>27

<212>DNA

<213>Artificial Sequence

<220>

<223>L2R primer

<400>81

tattcgcgcg tttcggtgat gaatatt

27

<210>82

<211>39

<212>DNA

<213>Artificial Sequence

<220>

<223>BApF primer

<400>82

gtcgacatta atgccggtga gtgagcggcg cggggccaa

39

<210>83

<211>32

<212>DNA

<213>Artificial Sequence

<220>

<223>BApR primer

<400>83

ggatccggtg gcgcgtcgcg ccgctgggtt tt 32

<210>84

<211>27

<212>DNA

<213>Artificial Sequence

<220>

<223>NeoF primer

<400>84

agatctatgg gatcgccat tgaacaa 27

<210>85

<211>29

<212>DNA

<213>Artificial Sequence

<220>

<223>pAR primer

<400>85

catatgtcat aatcagccat accacattt 29

<210>86

<211>27

<212>DNA

<213>Artificial Sequence

<220>

<223>SVpAF primer

<400>86

ctcgagatgg gatcgccat tgaacaa 27

<210>87

<211>28

<212>DNA

<213>Artificial Sequence

<220>

<223>SVpAR primer

<400>87

catatgagta atcagccata ccacattt 28

U3- 失活的反转录病毒载体

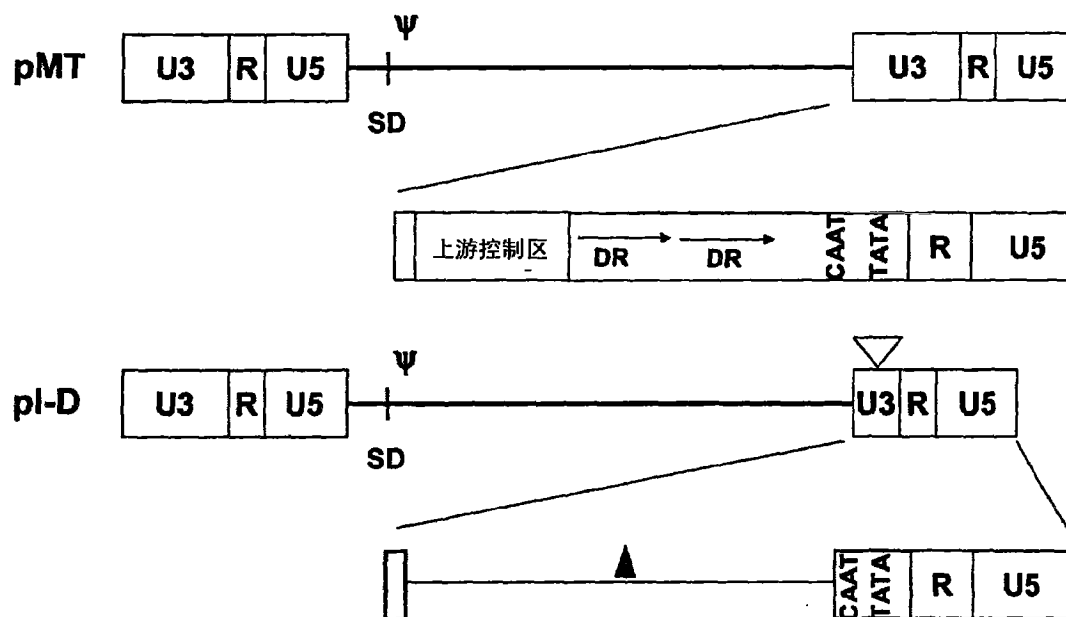


图 1

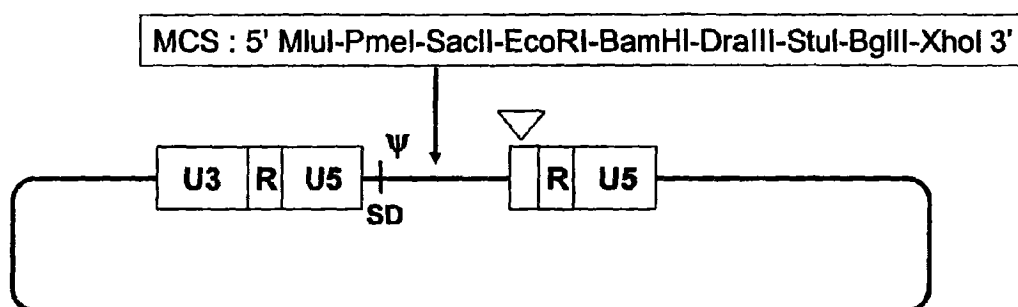
PI-ND

图 2

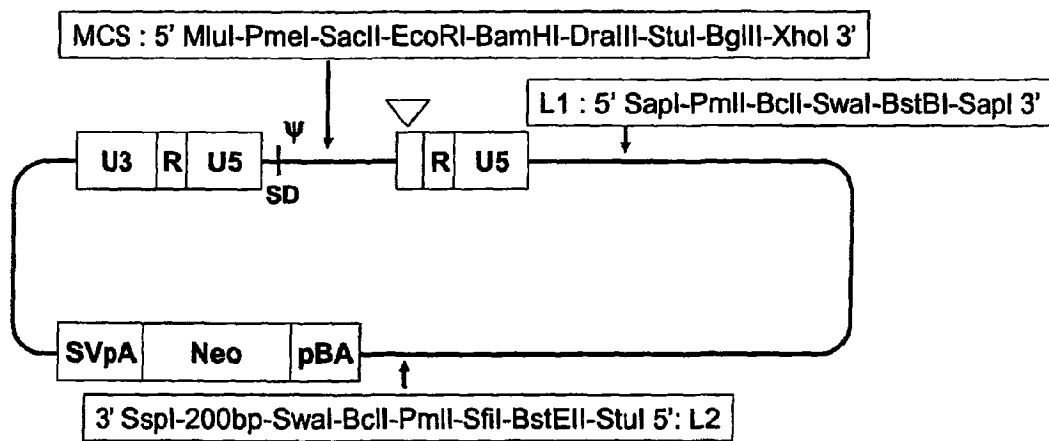
pl-LND-n

图 3

gp91 表达

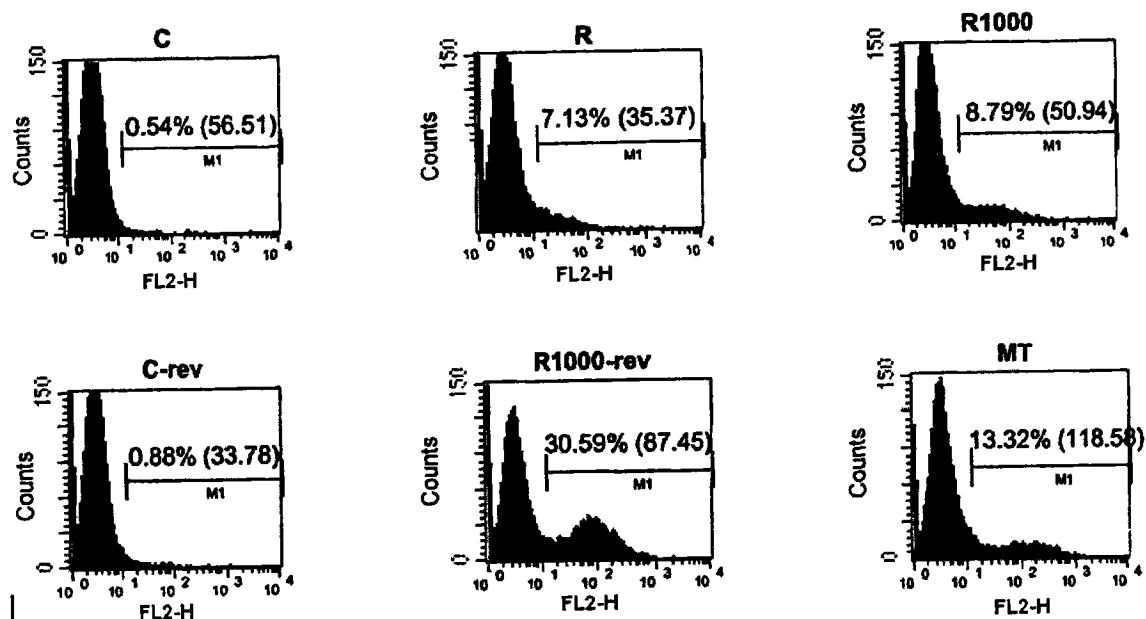


图 4

GAPDH 启动子

```

TTCA'FCCAAGCGTGTAAGGGTCCCCGTCTTGACTCCCTAGTGTCCTGCTGCCCCACAGTC
CAGTCCTGGGAACCAGCACCGATCACCTCCCATCGGGCCAATCTCAGT'CCCTTCCCCCCT
ACGTGGGGGCCACACGCTCGGTGCGTGCCAGTTGAACCAGGCGGCTGCGGAAAAAA
AAAAGCGGGGAGAAAGTAGGGCCCCGGCTACTAGCGGTTTACGGGCGCACGTAGCTCA
GGCCTCAAGACCTTGGGCTGGGACTGGCTGAGCCTGGCGGGAGGCGGGGTTCGAGTCA
CCGCCTGCCGCGCGCCCCCGGTTTCTATAAATTGAGCCCGCAGCCTCCCGCTTCGCTCT
          BRE          TATA box          +1
CTGCTCCTCCTGTTTCGACAGTCAGCCGCATCTTCTTTTTCGTCGCCAGGTGAAGACGGGC
GGAGAGAAACCCGGGAGGCTAGGGACGGCCTGAAGGCGGCAGGGCGGGCGCAGGC
CGGATGTGTTCCGCGCCGCTGCGGGGTGGCCCCGGCGGCCTCCGCATTGCAGGGCGG
GCGGAGGACGTGATCGGCGCGGGCTGGGCATGGAGGCCTGGTGGGGAGGGGAGGG
GAGGCGTGTGTGTCGGCCGGGGCCACTAGGCGCTCACTGTTCTCTCCCTCCGCGCAGCC
GAGCCACATCGCTCAGACACC
  
```

图 5

RPL10 启动子

GTGCGCTCGAGCAGGATTTCTCCCGTCTCTCCCTGTCAAAGGACGGGAAGACTTTGTTA
 CCCCACCGCGCCCCACCTGCAGAATGGTGGACAGATACCTCCAGATCCCACTTCCCC
 AGGACGCCCCGCTGCTCTGCGCACCTCTCCCCGGATGCTGCCCCGTGGGCGGGTGGGGG
 CCGCCCTGCTTCCCCACGACCCCCAGACGCACCCGGAGGGACTC
 TTGAGCACAGTGGAGTGGGAAGGCGGAGGTGGGGCGGTGCCCAGGCGAGAGCGGCTC
 ATGGGAGGCGGCGCCCGAGACGCAGCTGGTGGGACGGTGGGGTTCAGGGTGGGCGG
 AGCGGGGCTAGAGATGCCCCGGGGTTCCCAGGCCATGAGTCTCCGTGGAGATTTCTCC
 TCGACCTCTTCCCCGCGGCAATGTGCGAACCCTGGGTCTCCAGGAAACGGGGATACGG
 GGCATGGCTCCCAGCAAGGCCTGGTCCAGCCTCTCCGGTAGGGGAATGGGTCTCCCCCT
 CCGGCCCTCCGGGTGACAAAGGAACGCGGGCCAGATCCCCGTATGGCGCTTACCCG
 CCGGGGCTCTAGCCTAGAAGGAGGCAACGAGCGCGTGTCCGAGACCCGTGCAAGCTC
 AGGGACACTCTCGCGGTGCGCCGGGAGGCCACCTAGGGTACTTTCCCTTTTTCCTACTCTC
 AGAAATATACGTCTGTACAGTTAACGGCAAAGCCTAGGGCAAGAGTTCTACGCCCAAG
 ATGGCCAGCCGGAAGCGGGCTTCTCGCGACCATGTGGCGAAGCCCCATTCTGTCAGCTGG
 CCGCCCCGCGGCCCTGGTACCCGGTCACCTCTCTGATCTGCGCATGTGCTGGGCTACGCCC
 GGGCGCAAGCGCCAAGAGCGGCTGCGTCTATGGTCATGACGTCTGACAGAGCGTCCAC
 CCGTCTTCGACAGGACTCTATGGTTCTTACGCGCGCAGACAGACCGCCTATATAAGCCAT
 TATA box
 CCGCAGGCGGAGGAGCCGCTCTTTCCCTTCGGTGTGGTGAGTAAGCGCAGTTGTCGTCT
 +1
 CTTGCGGTGCCGTTGCTGGTTCTCACACCTTTTAGGTCTGTTCTCGTCTTCCGTTCCGACT
 CTCTCTTTTTCGTTGCAGCCACTGAAGATCCTGGTGTCCG

图 6

LENG8 启动子

TTGTATCAGAGTCCTGGACCGAAACAGATGGCACTCAAAGGTGGCGCGCAGTTCAGA
 GAAATGCCTATGTACGGATTGTGGTCCAATGCCCTCAGCCTGACCTCAGGGACCTTCGGGG
 GTCTGCTCCGCGCCACCCCTTACACATCTGTGACCCCAACACTTCCACCCAGCGCCA
 CTGCCAACAGCTACACCCATCCCCCTCCAACCGCGTCAGCTTCCAGCCTCGGTCCATCT
 GAACTCGCCGTGCCCCCTCCCCCTGCGCCCTTCCAGATTCAATTTGCTAGGGAAGCCCCGT
 CTTCGGGTGGAGCTGTTCCCTCATCCCCCTTCTTTATCATTTCTCTCCCAGGGCTTCCACA
 TCACCGTGCTGTGGACAATCCCGGAATCCTGTACGCCAGTTTACATTTAGGAACAGT
 AATGGCTCCCACTGACTCAGTCAAAACAAGGCTGCGGCCGGGCACGGTGGCTCACGCC
 CGTAATCCCAGCACTTTGGGAAGCCGAGACGGAGGGAACACGAGGTGAGGAGTTCGAG
 AACAGCCTGGCCGACA'TGGTGAAACCCCGTCTTTACAAAAACAAAAA'TTAGCCGGG
 CATAGTGGCGCGCGCTGTAATCCCAGCTACTCCGGAGGCTGAGGCAGAATTGTTTGAA
 CCCAGGAGGCGGAGGTTGCAGTGAGCAGAGATCACGCCACTGTACTCTATCGTGGGCG
 ACGACAGAGCAAGAGCAAGACTCCGTCTCCGAGAACAACAACAACAGCAACAAGAA
 AACAACAATAAAAAAATAAGGCTGCGTGGGAGGCAGAAAGAGCTAATGCGGCCACG
 CTTGTCCCCTCGGGGCCACCGTCCCCACCCAGACTTCCGGTCTGCCCTTAAATGTTTATG
 CGTAAGTGCGTGGGCAGGAAGGCGGGCTCAAGCGCAGCTCGTGGCGTTTATTGGCTGTG
 CAGGGCCGAGGGAGGCGGTGCAAGGCCGCCGCGTGACGTGAGGACGCCCGCGGTGAGG
 ACGTCGAAGCCAAAGAAGACCAGAGCCAGCCGGGTGGCACAGCGGTGTCGTGGCCGT
 (+1) +1
 GTTGCTGATCGCCTGGGTGGTGTGTTGGCGTGTCCCTGCAGCGAAGGATCCTGGTTGGTAA
 GGGGAGCGGCGGGCGAGCAGGCGGGCGGGGATAGCATCTCCTTTTGGTCTTGGCCCCC
 GCGAGCCCCGAGGCCCTTCTCGGCCGTGCGAGCAGCAGACGCCGCGCGGAGCGTGCAC
 AGGGTGTGGCGCGCAGGGGCAGCCACTGCGCCTGCGCACCGGGCCTGGGGCCGCGCG
 TTCGGGCACTAGCGCGCGTGGCCCGTGGTCTTCTACTTTCCACACCCAGAACTCTTCAGA
 TCCTTGACCCCACTGGCTTTTCAGTCAGCCTCOCCTTTTCTGCCAGCTTCTCTTGAGTCC
 ATCTACTTTTCTTCCCCACTTTGTGACGTGTTTTTAGCTCCCCCTTAAAGTCTCCCTAACTCA
 TTCTTTTTCTCATAGGCAGTGAAAAAGCAGTCTGGCTCCCGAGGTCCACCCCTTATACCC
 CAAGGTCCAG

图 7

SNX3 启动子

AATCCAGACGCGTGTCTGCTGCAACGCTCGGGTTTATGGCAAATCATCTCAGGCATTT
 GCTTAACCTTCTCCAGAAAGGCATTTTCAGGGGTTTCACAGTGAGACGGTGACAGGTTG
 GCACAGAGTTAGTAGGGGCAGTTTGTTCGATTTGCGGGCAAATCTCTAAGATCTCTCC
 GTTTAACTTTGCCCCGCAATTCCCAAAGCCGCTAAAGCCGTTTCCGGCGCTCTACCCCG
 CCGCAGGCCGAGGCTGGCCGAGAGACAGGAAGCGCCAGCTCTGGGCGTCTGGGTC
 CTCGCCTCCTCGGCCGAGCCCCGCGGCGGCGGCTCGCGGTGCATGTGTTGGGCGCTGTA
 GTCCGGCCGGAACCTGTTTGCGALCCCCGAGTCCCATGACACCGCTTCTCCTCACACCCC
 +1
 AGTCCGCAGTGCCCCCTCCCCAGCCTCGGCCGGGCTCCCGGGAGCCGGGCGTGGCGTTC
 CAGCTAGTGAGCCGTTTCTCCCTCGGGCTCGGAGGCGGAAGCTTGAGGGGCGCGGGGA
 GGACCTTCGCGTGCGGGGTGAACGCCCGCTCTACGTGCTCGTTCTCTTCGCGACCGCTG
 CGCGCGAGCCCCGTGTCCCCACGGCGGGCAGCAGCGCGCGCGCGGCGGCTGAACGC
 GGAGGGGGCGGAGGGAGCCCCGCGGCGGCGGCAGCAGCTACAGCGAA

图 8

ITGB4BP 启动子

TCTGTCCCTCAAGGCACAGCTGGACTCATCCCTTTCCCCAAACCTGCCCTTCTCCGGCT
 TCATTTCCATCAATACCTCCATCATCAACCCTTCCGCGAGACACTCCTGGCCCCCTCCTCTC
 CCTCATGCCTCACAACCGACCAGCCGGAGGTCTAGGTGATGACAGCTCCTAAAAAGCT
 CCTGAATGAATAATGAATGAATGAACGCGAGCAGGCTAGGCGTGGGGCCAGGCGGGT
 CGCGCCCAGACCGCTCGCGACCATAGAGTCCGCGGAGGCGCGGAGGAGAGGGGCTGG
 ATGCGTGGCGGGGAGCGCCGGGCTCTCCCGGAAGTCTCCCTGGACGGAAGTGGAAACG
 +1
 GAAACCTTTTTAGGGAGTCCAAGGTACAGTCGCCGCGTGCGGAGCTTGTTACTGGTTACT
 (+1)
 TGGTAAGCTCGTGTGAGGGGAACCTGGGAGGGTCAGCTCCGCTCCTGGGTCCGGAGCGG
 GTGGGGGCCAGAGGATTCAGGGCCGGAGGTCTGTTGGGGGCCAGTGGGCGGGACCC
 GAGGACGGAGGGGGCCGGAGGCCGAGAGGGGCGGGTCCGCGCGGGGCTGAGGGA
 CGGAGGCCGGGATACTTGGGAAAGGATCCGCCGGCCTTGAACTCCCGCCTCCGCCGCCC
 CTAGGCCTC

图 9

UQCRQ 启动子

```
GTCACCTTTTGTTCCTCCCCCGCCTCCCGCATTCGGCCGCTTCCTGACTGGGATTCCAC
AGAAAAGCCGAGGGCTGAGGAGAAGTGTGAGCGCCTCCGCCTGTCCACTGTCCCCCAA
AGTCAGTTCAATCCCCGACGTCCTCCGCTAGGCTCCACCCCAACGGCCCGGGCAGGGCC
TCCAAGGCACCTCCACCTACGGGTACCCAGTCAGCCCACTTCTTCTGGGACAAAGG
CGTCATCCCTTAGAGACAGTAGGAAAATGGTATCTCCCGGAAGTTACCTCACGACCTCC
AAGAGCGGCTTCCAACCTTGCCGGAAATGACGAACGAGTCAACCGGATCGGTGACTGT
                                     (+1)                +1
GGAGGGCGAGCTGAGCCCTGTGCGTGAGTGGGGTCTGGTTGTGCAGTGTTCGTGGACCC
TGGGAGGCTACGGGCGCCCCGCTGGGCTGGGAAAGGATAAGGAGTGCAGGGGCAGGA
CTCTCGGGTTGGGGATGGACCCCCGCGGGGACTGCGGCGCTTCGCGAAAGCGAGCCAA
CGGCCTGTCCACCTCGG TCCTGCAGGGCCGCCGCCACA
```

图 10