



(12) 发明专利

(10) 授权公告号 CN 101330830 B

(45) 授权公告日 2016. 01. 20

(21) 申请号 200680045545. 8

(22) 申请日 2006. 10. 18

(30) 优先权数据

60/728, 131 2005. 10. 18 US

60/765, 993 2006. 02. 06 US

(85) PCT国际申请进入国家阶段日

2008. 06. 04

(86) PCT国际申请的申请数据

PCT/US2006/040379 2006. 10. 18

(87) PCT国际申请的公布数据

W02007/047583 EN 2007. 04. 26

(73) 专利权人 国家犹太健康中心

地址 美国科罗拉多州

专利权人 科罗拉多大学董事会

(72) 发明人 J·卡姆比尔 Y·里费利

S·约翰逊 B·C·特纳

(74) 专利代理机构 北京市金杜律师事务所

11256

代理人 陈文平

(51) Int. Cl.

C12N 5/10(2006. 01)

(56) 对比文件

W0 99/53028 A, 1999. 10. 21, 全文.

CN 1408850 A, 2003. 04. 09, 说明书第 1 页.

W0 98/10058 A1, 1998. 03. 12, 全文.

刘珺. 鼠 bcl-2alpha 基因在哺乳动物细胞中的表达. 《中国优秀硕士学位论文全文数据库(医药卫生科技辑)》. 2004, (第 2 期), E059-35.

Nestor D. DeoCampo, et al. Cooperation of bcl-2 and myc in the neoplastic transformation of normal rat liver epithelial cells is related to the down-regulation of gap junction-mediated intercellular communication.

《Carcinogenesis》. 2000, 第 21 卷 (第 8 期), 1501-1506.

Anne Wilson, et al. c-Myc controls the balance between hematopoietic stem cell self-renewal and differentiation. 《Genes Dev.》. 2004, 第 18 卷 2747-2763.

审查员 毛颖

权利要求书2页 说明书33页

序列表45页 附图13页

(54) 发明名称

条件无限增殖化长期干细胞和制备和使用所述细胞的方法

(57) 摘要

公开了使干细胞(包括成年干细胞和胚胎干细胞)条件无限增殖化的方法,由所述方法产生的细胞,利用所述细胞的治疗和实验室或研究方法,和利用所述细胞鉴定与细胞分化和发育或治疗疾病有关的化合物的方法。还公开了一种急性髓细胞白血病(AML)小鼠模型和与所述小鼠模型有关的细胞和方法。

1. 一种条件无限增殖化长期成年干细胞,其包含:

(a) 至少一种促进细胞存活和增殖的第一重组蛋白质,或编码它的重组核酸分子,其中所述蛋白质的表达或功能是可诱导的,和其中所述第一重组蛋白质是 MYC 或 Notch-1 的细胞内部分 (ICN-1);和

(b) 至少一种抑制所述干细胞的细胞凋亡的第二重组蛋白质,或编码它的重组核酸分子,其中所述第二重组蛋白质是 Bcl-2 或 hTert。

2. 权利要求 1 的条件无限增殖化长期成年干细胞,其中 (a) 和 / 或 (b) 的核酸分子包含在一种载体中。

3. 权利要求 2 的条件无限增殖化长期成年干细胞,其中载体含有四环素效应元件。

4. 权利要求 2 的条件无限增殖化长期成年干细胞,其中 (a) 和 / 或 (b) 的核酸分子包含在一种载体中,该载体在 (a) 或 (b) 的核酸分子的侧翼包含编码重组酶的识别底物序列的核酸序列。

5. 权利要求 1 至 4 任一项的条件无限增殖化长期成年干细胞,其中所述第一重组蛋白质或编码它的重组核酸分子选自 MYC-ER 和 Notch-1 的 ER- 调节的细胞内部分 (ICN-1-ER)。

6. 权利要求 1 至 4 任一项的条件无限增殖化长期成年干细胞,其中所述抑制干细胞凋亡的第二重组蛋白质是 Bcl-2。

7. 权利要求 1 的条件无限增殖化长期成年干细胞,其中:

(a) 所述第一重组蛋白质是 Tat-MYC;和

(b) 所述第二重组蛋白质是 Tat-Bcl-2。

8. 权利要求 1 的条件无限增殖化长期成年干细胞,其中所述第一重组蛋白质是 Tat- 融合蛋白,或者所述第二重组蛋白质是 Tat- 融合蛋白,或所述第一重组蛋白质和所述第二重组蛋白质都是 Tat- 融合蛋白。

9. 一种生产条件无限增殖化成年干细胞的方法,包括:

(a) 向干细胞群体中递送至少一种促进细胞存活和增殖的第一重组蛋白质、或编码它的核酸分子,其中所述蛋白质的表达或功能是可诱导的,和其中所述第一重组蛋白质是 MYC 或 Notch-1 的 ER- 调节的细胞内部分 (ICN-1);

(b) 通过向干细胞群体中递送至少一种抑制干细胞凋亡的第二重组蛋白质或编码它的核酸分子来抑制所述干细胞的细胞凋亡,其中所述第二重组蛋白质是 Bcl-2 或 hTert;和

(c) 在干细胞生长因子组合的存在下并且在使所述第一重组蛋白质的表达或功能保持活性的条件下扩增所述干细胞,以生产条件无限增殖化成年干细胞。

10. 权利要求 9 的方法,其中向干细胞群体中递送第一重组蛋白质包括在第一 Tat- 融合蛋白的存在下培养所述干细胞,其中 Tat 与所述第一重组蛋白质融合;并且其中通过向干细胞中递送第二重组蛋白质来抑制所述干细胞的细胞凋亡包括所述干细胞与第二 Tat- 融合蛋白一起培养,其中 Tat 与所述第二重组蛋白质融合。

11. 权利要求 9 的方法,其中所述第一重组蛋白质、或编码它的重组核酸分子是 MYC-ER,并且其中所述抑制干细胞凋亡的第二重组蛋白质是 Bcl-2。

12. 权利要求 9 的方法,其中所述抑制干细胞凋亡的蛋白质抑制所述干细胞中的促凋亡蛋白质。

13. 一种鉴定调节谱系定向和 / 或细胞分化和发育的化合物的方法,包括:

a) 使条件无限增殖化长期成年干细胞与待测试化合物接触 ;和

b) 检测条件无限增殖化长期成年干细胞的至少一种基因型或表型特征,与没有所述化合物时的条件无限增殖化长期成年干细胞相比,其中在该化合物存在下检测到特征的差异指示该化合物影响干细胞的特性 ;

其中,条件无限增殖化长期成年干细胞包含 :

i. 至少一种促进细胞存活和增殖的第一重组蛋白质,或编码它的重组核酸分子,其中所述蛋白质的表达或功能是可诱导的,和其中所述第一重组蛋白质是 MYC 或 Notch-1 的 ER- 调节的细胞内部分 (ICN-1) ;和

ii. 至少一种抑制所述干细胞的细胞凋亡的第二重组蛋白质,或编码它的重组核酸分子,其中所述第二重组蛋白质是 Bcl-2 或 hTert。

14. 一种引导条件无限增殖化长期造血干细胞的分化的方法,包括使所述条件无限增殖化长期造血干细胞与促红细胞生成素接触,其中所述条件无限增殖化长期造血干细胞包含 :

(a) 至少一种促进细胞存活和增殖的第一重组蛋白质,或编码它的重组核酸分子,其中所述蛋白质的表达或功能是可诱导的,和其中所述第一重组蛋白质是 MYC 或 Notch-1 的 ER- 调节的细胞内部分 (ICN-1) ;和

(b) 至少一种抑制所述干细胞的细胞凋亡的第二重组蛋白质,或编码它的重组核酸分子,其中所述第二重组蛋白质是 Bcl-2 或 hTert。

15. 一种产生至少一种条件无限增殖化成年干细胞的方法,包括 :

a) 使用包含编码 MYC 分子和人雌激素受体的激素结合域的核酸序列的第一载体和包含编码 Bcl-2 的核酸序列的第二载体转导至少一种干细胞 ;

b) 在培养基中培养所述至少一种干细胞 ;

c) 诱导至少一种干细胞中的 MYC 分子的活性 ;和

d) c) 之后,抑制 MYC 分子的活性,

其中,至少一种干细胞能够被无限增殖化,能够被再次无限增殖化,和能够在去无限增殖化后再次无限增殖化,从而产生至少一种条件无限增殖化成年干细胞。

16. 一种产生条件无限增殖化成年干细胞群的方法,包括 :

a) 将如下物质引入一个或多个干细胞 :

i. 包含 MYC 多肽的外源合成的第一多肽 ;和

ii. 包含 Bcl-2 多肽的外源合成的第二多肽,从而产生至少一种条件无限增殖化干细胞 ;和

b) 在培养基中培养一个或多个干细胞,从而产生条件无限增殖化成年干细胞群,其中所述条件无限增殖化成年干细胞群的维持依赖于所述第一多肽和所述第二多肽的存在,且其中所述条件无限增殖化成年干细胞群在不存在至少所述第一多肽的情况下分化。

条件无限增殖化长期干细胞和制备和使用所述细胞的方法

[0001] 发明领域

[0002] 本发明一般地涉及条件无限增殖化长期干细胞,制备所述细胞的方法,和使用所述细胞的方法,包括治疗方法和药物发现方法。

[0003] 发明背景

[0004] 控制不同血细胞的骨髓输出的能力已经成为多种疾病的治疗的一种重要工具。对于血液恶性肿瘤的一些最好的新疗法基于促使白血病细胞在转化事件之前分化为所定向的谱系的化合物的研制。一个这样的例子是急性早幼粒细胞白血病。在用三氧化二砷治疗患者后,推动恶性细胞沿骨髓单核细胞途径,导致这些肿瘤的缓解。另外一个例子是促进在受照射个体中移植的骨髓干细胞(长期重建性造血干细胞,或 1t-HSC)的成功移植。全身施用已知特异性诱导红血细胞发育(促红细胞生成素,或 Epo)或骨髓细胞发育(粒细胞-巨噬细胞集落刺激因子,或 GM-CSF)的细胞因子可以促进分化的血细胞的出现。最后,使用“动员(mobilization)”法大大简化了从供体中收获 1t-HSC,在该方法中,通过全身施用被称为 G-CSF 的细胞因子,诱导这些细胞从它们正常所处的骨髓部位移动到外周血中。然后可从外周血中收获干细胞,从而避免了疼痛和复杂的骨髓活检组织的采集。所有这些方法都依赖于计划和控制 1t-HSC 的生物学行为的能力。

[0005] 相应地,骨髓(干细胞)移植是一种无价的治疗工具,用于已经经历了放射和/或化疗(例如,癌症患者,或已经暴露于高水平放射)的个体的血液和免疫重建,并且也是治疗免疫缺陷和血液恶性肿瘤的一种重要方法。另外,骨髓移植将是一种对抗衰老对免疫系统和其他细胞和组织的不利影响的非常有用的疗法。预计干细胞移植每年可使超过 35000 名儿童和成人受益。

[0006] 骨髓移植的工作原理是产生所有血细胞类型的放射敏感性 1t-HSC 的置换。最近的研究显示骨髓移植可能对于治疗心脏疾病有价值。尽管这种效果的基础是未知的,但是这一发现和其他发现提高了造血干细胞(1t-HSC)可被重新编程产生其他组织的可能性。如果这是事实,则 1t-HSC 可具有更广泛的应用,并且提供了一种替代有争议的胚胎干细胞治疗的方法。

[0007] 骨髓移植临床应用所面临的主要障碍首先是适当组织相容性骨髓供体的鉴定。这通常通过登记有超过 6 百万潜在供体的机构来完成。选择的供体必须经历诱导干细胞动员进入血液以及随后 4 到 5 天的白细胞清除以分离稀少 1t-HSC 的严格处理。这些细胞移植之后必须仔细监视和处理受者,以使由过路淋巴细胞引起的移植物抗宿主反应最小化。

[0008] 在历史上低频率的相关细胞群体使得难以对造血谱系发展受损的分子基础进行阐明,这阻止了对信号传导和下游反应的生化分析。实际上,这是所有血细胞生成研究中的一个主要的限制性因素。另外,长期造血干细胞(LT-HSC)的有限的可获得性也是治疗人类的许多类型癌症和几种免疫缺陷中的一个主要障碍。据本发明人所知,目前还没有可获得的类似 1t-HSC 自发产生的并且可在体外分化为正常谱系、或可使致死性照射的小鼠或亚致死性照射的人类重建的细胞系,也没有记载任何方法可以预期产生所述细胞系。而且,目前没有持续扩增 1t-HSC 的可行的技术,以致于在每次需要时需要从供体中获得这些细胞。

[0009] 也迫切需要治疗血液恶性肿瘤和免疫缺陷的其他方法,和提高移植 1t-HSC 输出的新的细胞因子。另外,目前还没有用于靶标鉴定和药物发现的合适的平台。缺失的因素是代表造血谱系的不同发展阶段的细胞系。最理想地,所述细胞应该保持在特定谱系中进一步分化的能力。对于基因产品的鉴定,因此对于与细胞发育、增殖和存活的调节有关的新的可药用目标来说,这些细胞系是重要的。另外,在新药开发中,所述细胞系对于用于功能缺失研究的小分子和 shRNA 文库以及用于功能增加研究的 cDNA 文库的筛选是重要的。

[0010] 在本领域中目前的药物发现的障碍包括:(a) 从特定发展阶段分离足够数量的细胞;(b) 细胞在体外增殖一段足够长的时间;和(c) 利用条件致癌基因筛选可影响白血病细胞而不影响正常 HSC 或祖细胞的药物的能力。

[0011] 因此,在本领域中特别需要这样一种方法,该方法产生可以广泛扩增、冷冻并在需要时随时再次使用的 1t-HSC 细胞系,而不需要后续的从供体中收获。

[0012] 发明概述

[0013] 本发明的一个实施方式涉及一种生产条件无限增殖化成年干细胞的方法。该方法包括下列步骤:(a) 获得扩增的成年干细胞的群体;(b) 用包含促进细胞存活和增殖的原癌基因或其生物活性片段或同源物的核酸分子转染所述干细胞,其中所述原癌基因是可诱导的;(c) 用编码抑制该细胞的细胞凋亡的蛋白质的核酸分子转染所述干细胞;和(d) 在干细胞生长因子组合的存在下并且在使原癌基因保持活性的条件下扩增被转染细胞,以生产条件无限增殖化成年干细胞。在该实施方式的一个方面,(b) 和 / 或 (c) 的核酸分子包含在一种整合载体中。在一个方面,利用选自下组的病毒或病毒载体将 (b) 和 / 或 (c) 的核酸分子转染入细胞中:逆转录病毒载体、慢病毒载体、细小病毒、痘苗病毒、冠状病毒、嵌杯样病毒、乳头瘤病毒、黄病毒、正粘病毒 (orthomixovirus)、披膜病毒、小 RNA 病毒、腺病毒载体、修饰和减毒的疱疹病毒。在一个方面,用直接电穿孔将 (b) 和 / 或 (c) 的核酸分子转染入细胞中。在一个方面,(b) 和 / 或 (c) 的核酸分子包含在一种包含编码药物敏感蛋白的核酸序列的载体中。在一个方面,(b) 和 / 或 (c) 的核酸分子包含在一种载体中,该载体在 (b) 或 (c) 的核酸分子的侧翼包含编码重组酶的识别底物序列的核酸序列。

[0014] 在一个方面,该实施方式包括以下附加步骤:(e) 去除使原癌基因保持活性的 (d) 的条件;和(f) 在包含诱导细胞分化的生长因子的培养基中培养 (e) 的细胞。这种方法可进一步包括:(g) 给 (f) 的细胞施加使原癌基因保持活性的 (d) 的条件,以生产处于细胞分化中间阶段的条件无限增殖化细胞。

[0015] 本发明的另一实施方式涉及一种生产条件无限增殖化成年干细胞的方法,包括:(a) 获得扩增的成年干细胞群体;(b) 在下列物质存在下培养所述干细胞:(1) 干细胞生长因子的组合;(2) 第一 Tat- 融合蛋白,其中 Tat 与由促进细胞存活和增殖的原癌基因或其生物活性片段或同源物编码的蛋白质融合;和(3) 第二 Tat- 融合蛋白,其中 Tat 与抑制干细胞的细胞凋亡的蛋白质融合。

[0016] 本发明的另一实施方式涉及生产条件无限增殖化胚胎干细胞的方法,包括:(a) 获得扩增的胚胎干细胞群体;(b) 用包含促进细胞存活和增殖的原癌基因或其生物活性片段或同源物的核酸分子转染所述干细胞,其中所述原癌基因是可诱导的;(c) 用编码抑制该细胞的细胞凋亡的蛋白质的核酸分子转染所述干细胞;和(d) 在干细胞生长因子组合的存在下并且在使原癌基因保持活性的条件下扩增被转染细胞,以生产条件无限增殖化胚胎

干细胞。

[0017] 本发明的另一实施方式涉及生产条件无限增殖化干细胞的方法,包括:(a) 获得扩增的干细胞群体;(b) 在下列物质存在下培养所述干细胞:(1) 干细胞生长因子的组合;(2) 由促进细胞存活和增殖的原癌基因或其生物活性片段或同源物编码的蛋白质;和(3) 抑制干细胞的细胞凋亡的蛋白质。利用任何适当的递送系统将(2)和(3)的蛋白质递送进干细胞中,递送技术包括但不限于:Tat 融合、适体 (aptamer) 技术、或 CHARIOT™技术。

[0018] 本发明的另一实施方式涉及一种生产条件无限增殖化干细胞的方法,包括:(a) 获得扩增的干细胞群体;(b) 将由促进细胞存活和增殖的原癌基因或其生物活性片段或同源物编码的蛋白质或其编码核酸分子递送进该细胞中,其中该原癌基因是可诱导的;(c) 通过将抑制该细胞的细胞凋亡的蛋白质、编码抑制该细胞的细胞凋亡的蛋白质的核酸分子、或抑制该细胞中的促凋亡蛋白的核酸分子或蛋白质递送进细胞中,抑制所述干细胞的细胞凋亡;和(d) 在干细胞生长因子组合的存在下并且在使该原癌基因保持活性的条件下扩增所述细胞,以生产条件无限增殖化成年干细胞。

[0019] 在上述任何实施方式中,原癌基因可选自,但不限于:MYC-ER 和 ICN-I-ER。在任何上述实施方式中,抑制细胞凋亡的蛋白质可选自,但不限于抑制细胞凋亡的 Bcl-2 家族的成员,例如 Bcl-2、Bcl-X、Bcl-w、BclXL、Mcl-1、Dad-1 或 hTERT。当原癌基因为 MYC-ER 或 ICN-I-ER 时,使原癌基因保持活性的条件可包括他莫昔芬或其激动剂的存在。在一个方面,用下列物质转染或递送(作为蛋白质)所述细胞:MYC-ER 和 Bcl-2;MYC-ER 和 hTERT;ICN-1-ER 和 Bcl-2;ICN-1-ER 和 hTERT;或 MYC-ER 和 ICN-1-ER。

[0020] 在上述任何实施方式中,扩增步骤可在介质中进行,所述介质包括但不限于:(1) 白介素-6(IL-6)、IL-3 和干细胞因子(SCF);(2) 包含干细胞因子(SCF)、血小板生成素(TPO)、胰岛素样生长因子 2(IGF-2) 和成纤维细胞生长因子 1(FGF-1) 的无血清培养基。

[0021] 在任何上述实施方式中,成年干细胞包括但不限于:造血干细胞、肠干细胞、成骨细胞干细胞、间充质干细胞、神经干细胞、上皮干细胞、心肌细胞祖干细胞、皮肤干细胞、骨骼肌干细胞、和肝干细胞。在一个方面,间充质干细胞可选自肺间充质干细胞和骨髓基质细胞。在一个方面,上皮干细胞选自:肺上皮干细胞、乳房上皮干细胞、血管上皮干细胞和肠上皮干细胞。在一个方面,皮肤干细胞选自:表皮干细胞和囊泡干细胞(毛囊干细胞)。在一个方面,神经细胞选自神经元多巴胺能干细胞和运动神经元干细胞。在一个方面,干细胞来自新鲜的或冷藏的脐带血。在一个方面,干细胞是从正常人或粒细胞集落刺激因子(G-CSF)治疗患者的外周血中获得的造血祖细胞。

[0022] 在上述任何实施方式中,该方法可进一步包括遗传修饰干细胞以矫正细胞中的遗传缺陷,遗传修饰干细胞以沉默化基因表达,和/或遗传修饰干细胞以过量表达基因。

[0023] 在上述任何实施方式中,该方法可进一步包括保存细胞。在一个方面,该方法进一步包括从保存物中回收细胞并培养细胞。

[0024] 本发明的另一实施方式涉及通过上述或本文别处描述的任何方法生产的细胞。

[0025] 本发明的另一实施方式涉及一种给个体提供成年干细胞或其分化细胞的方法,包括:(a) 提供一种通过上述或本文别处描述的任何方法生产的条件无限增殖化成年干细胞的来源;(b) 去除使(a)的干细胞条件无限增殖化的条件,和(c) 将所述干细胞或其分化细胞施用于个体。在一个方案中,预先从(c)的个体中获得细胞。在一个方案中,从先前冷冻

的所述细胞的贮存液中获得细胞。在一个方面,从个体中新鲜地获得细胞,并用上述或本文别处描述的任何方法进行条件无限增殖化。在一个方面,所述个体患有癌症。在另一方面,该个体患有白血病。在另一方面,该个体患有一种免疫缺陷病。在另一方面,该个体患有一种贫血障碍。在另一方面,该个体正经历整形外科手术。在另一方面,该个体正经历选择性整容外科手术。在另一方面,该个体正经历移植外科手术。在一个方面,该个体需要干细胞或其分化细胞,选自:造血干细胞、肠干细胞、成骨细胞干细胞、间充质干细胞、神经干细胞、上皮干细胞、心肌细胞祖干细胞、皮肤干细胞、骨骼肌干细胞、和肝干细胞。在另一方面,该个体需要提高的免疫细胞功能。在另一方面,该个体具有一种可被干细胞矫正的遗传缺陷。

[0026] 本发明的另一实施方式涉及一种鉴定调节谱系定向和 / 或细胞分化和发育的化合物的方法,包括:(a) 接触通过上述或本文别处描述的任何方法生产的成年干细胞;和(b) 检测(a)的干细胞中至少一种基因型或表型特征,与没有化合物时相比,其中在化合物存在下检测到特征的差异显示该化合物影响干细胞的特性。

[0027] 本发明的另一实施方式涉及一种研究谱系定向和 / 或细胞分化和发育的方法,包括评价通过上述或本文别处描述的任何方法生产的成年干细胞或其分化细胞,以检测该细胞的至少一种基因型或表型特征。

[0028] 本发明的另一实施方式涉及通过上述或本文别处描述的任何方法生产的细胞在治疗其中干细胞移植是有益的疾病或病症的药物中的用途。

[0029] 本发明的另一实施方式涉及一种急性髓细胞白血病(AML)的小鼠模型,包括用一种方法产生的小鼠,所述方法包括:(a) 致死性照射小鼠;(b) 将通过上述或本文别处描述的任何方法生产的条件无限增殖化长期干细胞和来自 Rag^{-/-}小鼠的完整骨髓细胞转移入该小鼠中;和(c) 定期注射他莫昔芬或其激动剂到小鼠中直到小鼠发展出 AML 的临床体征。在一个方案中,该细胞用 MYC-ER 和 Bcl-2 转染或被递送了这些物质(作为蛋白质)。

[0030] 本发明的另一实施方式涉及从上述 AML 小鼠模型中获得的肿瘤细胞。

[0031] 本发明的另一实施方式涉及 AML 小鼠模型在人蛋白质特异性候选药物的临床前试验中、在鉴定、发展和 / 或测试用于诊断、研究或治疗 AML 的化合物中、或在鉴定、发展和 / 或测试用于诊断、研究或治疗 AML 的靶标中的用途。

附图说明

[0032] 图 1 是显示转导细胞的骨髓移植和在体内用 4OHT 活化 MYC 功能后的死亡曲线的图。

[0033] 图 2 是显示在体外转导后,源于年轻和老年小鼠的 HSC 的散射特征和 GFP 表达水平的散布图。点图代表与 IL-3、IL-6 和 SCF 培养 3 天后 HSC 的前向(FSC)和侧向(SSC)散射特征的流式细胞分析数据。这两个标准与细胞大小(FSC)和粒度(SSC)相关。

[0034] 图 3 是显示源于受照射受体的细胞系的表型比较的散布图,该受体用 BCL-2、MYC-ER 和 EGFP- 转导的来自年老(> 60% ID 所有组成成分)和年轻 3-83 μ δ 转基因小鼠的造血干细胞重建。显示了在开始培养 3(年轻)和 3(衰老)个月 after 代表性克隆的表型。

[0035] 图 4 是显示在撤除他莫昔芬后,衰老 LT-HSC 系(ABM46)在体外自发分化的散布图(干细胞和 B 谱系标记物表达通过流式细胞术分析)。

[0036] 图 5 是显示在过继转移入受照射的年轻受体 6 周后,对源于 Lt-HSC 系的造血细

胞区室的分析的散布图。该图中显示了来自三只小鼠的数据,一只小鼠接受了衰老 HSC 系 ABM42,两只小鼠接受了衰老 HSC 系 ABM46。

[0037] 图 6 是显示在用 ABM42 和 ABM46 细胞系重建的小鼠中 B- 细胞区室的发育受到影响的散布图。该图中显示了来自三只小鼠的数据,一只小鼠接受了衰老 HSC 系 ABM42,两只小鼠接受了衰老 HSC 系 ABM46。

[0038] 图 7 是显示在用 ABM42 和 ABM46 细胞系重建的小鼠中 T- 细胞发育的散布图。该图中显示了来自三只小鼠的数据,一只小鼠接受了衰老 HSC 系 ABM42,两只小鼠接受了衰老 HSC 系 ABM46。

[0039] 图 8 是显示细胞系的表型比较的散布图和图,所述细胞系源于从用 BCL-2 和 MYC-ER 逆转录病毒转导并且在体外连续培养中保持超过 90 天的年轻 C57/BL6 小鼠中获得的 HSC。各图显示了对于病毒表达标记物 (GFP 和 Thy1. 1) 以及四种确定小鼠中长期 HSC 所需的标记物 Sca-1、c-kit、CD34 和 Flk-2 的表达的流式细胞分析结果。四种细胞系包含保持 1t-HSC 表型 (Sca-1+、c-kit+、CD34-、flk-2-) 的亚群。

[0040] 图 9 是显示源自从年轻 C57/BL6 小鼠中获得的 HSC 的细胞系的表型比较的散布图和图,所述小鼠被不同癌基因组合逆转录病毒转导并且保持体外连续培养超过 90 天 (pMIG-MYC 和 pMIT-Bcl-2(上图)、pMIG-MYC. ER 和 pMIG-hTERT(中图) 和 pMIG-ICN. 1. ER 和 pMIT-Bcl-2(下图))。

[0041] 图 10 是显示源自从年轻 C57/BL6 小鼠中获得的 HSC 的细胞系的表型比较的散布图和图,所述小鼠被不同癌基因组合逆转录病毒转导并且保持体外连续培养超过 90 天 (pMIG-ICN. 1. ER 和 pMIT-Bcl-2(上图)、pMIG-ICN. 1 和 pMIT-Bcl-2(第二行的图)、或 pMIG-ICN. 1 和 pMIG-Bcl-2(第三行的图)、或 pMIG-hTERT 和 pMIT-Bcl-2(下图))。

[0042] 图 11 是显示源自从年轻 C57/BL6 小鼠中获得的 HSC 的细胞系的表型比较的散布图和图,所述小鼠被不同癌基因组合逆转录病毒转导并且保持体外连续培养超过 90 天 (pMIG-MYC 和 pMIG-ICN. 1(上图)、pMIG-MYC. ER 和 pMIG-ICN. 1(中图)、或 pMIG-ICN. 1. ER 和 pMIG-MYC(下图))。

[0043] 图 12 是显示源自从年轻 C57/BL6 小鼠中获得的 HSC 的细胞系的 T 细胞和 B 细胞区室的体内重建的散布图,所述小鼠被不同癌基因组合逆转录病毒转导并且保持体外连续培养超过 90 天。

[0044] 图 13 是显示重组酶的识别底物序列 (RSS) 的应用的示意图,用于确保在移植前从本发明的条件无限增殖化长期干细胞中切除重组 DNA。

[0045] 图 14 是显示检测自本发明条件无限增殖化长期干细胞分化的 NK 和红细胞谱系细胞的图。

[0046] 发明详述

[0047] 本发明提供了一种解决方案,该方案解决了能够产生、保持和操作源于长期干细胞 (特别是长期造血干细胞 (Lt-HSCs)) 的稳定细胞系的问题,所述细胞在置于适当条件下时可产生正常由这些细胞产生的所有细胞谱系。本发明一般的涉及产生条件无限增殖化长期干细胞的方法,和用所述方法产生的干细胞,和利用所述干细胞的方法。更特别地,利用长期造血干细胞作为一种示例性干细胞群体,发明人已经建立了一种产生条件无限增殖化 (例如,可逆地无限增殖化或在当除去所述条件时为可逆的特定条件下无限增殖化) 干

细胞的有效的方法,所述干细胞能够在体外和体内分化为正常细胞谱系,并且能够重建需要所述细胞的患者。实际上,本发明可消除对骨髓供体的需要,因为本发明提供了在操作(例如,化疗、放疗,等等)之前从患者中收获干细胞、扩增所述细胞和将它们返回患者的能力。而且,所述干细胞可如需要或期望的那样广泛扩增、保存(例如,冷冻),然后恢复和再次扩增、操作、和/或重复利用。例如,可操作所述干细胞以矫正遗传缺陷或给患者提供益处(治疗性或预防性益处),或分化为一种期望的细胞类型。最后,所述细胞可用于多种实验中,用于鉴定与细胞发育、增殖和存活的调节有关的新靶标,和鉴定和开发用于改善或治疗将受益于细胞发育、增殖和/存活的调节的疾病和病症的药物。

[0048] 发明人已经发展了允许长期干细胞条件(在此例如长期造血干细胞(1t-HSC))无限增殖化的新技术。得到的细胞系可在体外无限地指数扩增(繁殖)和/或冷冻保存(储存),并且具有拯救受到致死性照射的小鼠和重建所述动物中的所有血细胞谱系的能力。而且,发明人已经能够在体外通过抑制转化性癌基因的功能而产生分化的血细胞。在此描述的所述细胞和产生它们的方法将允许产生不携带重组DNA的可移植的人干细胞,因此不会给受体带来长期风险。本发明的这些条件无限增殖化1t-HSC可被稳定化在它们的成熟表型,和建立其中在癌基因再激活后保持成熟表型的细胞系。例如,发明人已经能够发展 $CD4^+ \alpha \beta^+$ T细胞,以及树突细胞系。

[0049] 本技术在应用于临床条件时相对于骨髓移植具有下述优点:

- [0050] 1. 需要非常少的1t-HSC建立克隆;
- [0051] 2. 克隆代表了一种可更新的资源,所述资源可被无限地储存,并且可快速获得;
- [0052] 3. 这种治疗的成本应该比传统的骨髓移植更低;
- [0053] 4. 使用1t-HSC克隆可以缓解移植物抗宿主疾病的威胁,和相关的费用;
- [0054] 5. 至少在一些情况中,本技术可以缓解对骨髓供体的需要。

[0055] 另外,本发明提供了条件转化的长期干细胞例如1t-HSC细胞在产生代表分化的谱系(例如,分化的造血谱系,包括造血谱系发育的中间阶段)的细胞中的应用。例如,除了无数的治疗和预防应用以外,这些细胞系将允许鉴定可诱导恶性细胞分化、阻止其生长、或诱导细胞凋亡的新的化合物。这些细胞还将允许筛选新的引导干细胞沿特定途径分化的细胞因子和生长因子。所述细胞系并不存在,并且对于药物发现是重要的。

[0056] 更具体地,为了克服在本领域中关于来源于成人的干细胞(虽然如下所述,本发明并不限于来源于成人的干细胞)长期群体的提供和限制,本发明已经发展了产生代表早期造血干细胞祖细胞的条件转化细胞系的新方法。在此处描述和例示的本技术的一个具体的非限制性例子中,该策略包括来自5-氟尿嘧啶(5-FU)处理的3-83 μ δ 小鼠的骨髓干细胞的转染(例如,通过逆转录病毒转导)。本发明人使用具有编码Bcl-2和绿色荧光蛋白(GFP)(作为报告基因)和MYC-ER和GFP(也作为报告基因)的插入物的pMSCV双顺反子(bisistronic)逆转录病毒载体。选择MYC是因为它在淋巴细胞中具有代替细胞因子产生的存活和增殖信号的能力。通过限制目标细胞,发明人假设将形成干细胞肿瘤。重要的是,在此情况中MYC-ER功能依赖于他莫昔芬,通过从动物或培养物中去除他莫昔芬允许MYC功能终止和转化。在用MYC-ER转导的细胞中,产生融合蛋白,但是融合蛋白保持在细胞质中直到暴露于他莫昔芬。选择Bcl-2是因为它具有抑制细胞凋亡的能力,所述细胞凋亡通常由于暴露于MYC信号而发生,并且更特别地,当通过从细胞中去除他莫昔芬而使

MYC “失活”或移除时发生。这种基因类型新组合（如下面更详细讨论的那样，本发明不限于这些特定的基因）是成功产生本发明的条件无限增殖化干细胞的部分原因，并且如下文所述，可容易地扩展到其他类似的基因组合上。

[0057] 上述转导的干细胞的受体产生肿瘤（在 40HT 的存在下），收获来自骨髓、脾和淋巴结的肿瘤细胞并置于具有他莫昔芬和一种干细胞生长因子混合物的培养物中。发明人已经发现，在没有适当的干细胞生长因子组合时，本方法生产的干细胞将停止生长并且在很短的时间内死亡。因此，在用上述基因组合转染细胞后使用干细胞生长因子“混合物”（即，适当的或适合的干细胞生长因子的组合）是本发明方法的第二个重要方面。这种混合物具有促进并保持干细胞生长的一般特性，不限于特定的生长因子的组合，选择所述因子的参数在下面进行更详细的讨论。

[0058] 本发明方法产生的干细胞可在培养物中扩增，并且同质地是（例如）sca1 阳性的，是内皮糖蛋白和 ckit 阳性的，和 CD34、Flt3、B220、CD19 和 mIgM 阴性的，这指示了 1t-HSC 的表型，这是本领域公知的。这些细胞可被冷冻（冷藏保存，或储存），然后在冷冻后容易地复苏和培养。重要的是，这些复苏的细胞在表型上是同质的，并且显示 1t-HSC 的表型（例如，均质的 GFP 亮细胞是 Sca1、内皮糖蛋白和 ckit 阳性的，CD34、Flt3、B220、CD19 和 mIgM 阴性的）。这种表型很好地对应于在小鼠中提供所有长期重建的长期再生多能性干细胞的公开的特征（Reya 等人，2003，Nature 423 :409-14）。

[0059] 发明人已经进一步发展了本方法，以使其可以完全在体外完成（如上所述，起始步骤部分在体内进行）。发明人也证实了具有与上述类似特征的其他基因组合也导致 1t-HSC 的条件无限增殖化。而且，细胞系可通过移去他莫昔芬并提供适当的生长因子在体外分化为造血谱系，在他莫昔芬被阻止的受体动物中体内分化为所有造血谱系。另外，细胞可分化为发育的中间水平，该中间水平具有稳定的表型并保持在应用或移去适当的信号（在此描述的）后沿着它们的定向途径进一步分化的能力。所述细胞对于多种治疗应用是无价的。所有这些实验在下文和实施例中详细描述。

[0060] 本发明的方法和细胞系不仅提供了一种详细研究与成年干细胞向不同细胞谱系定向相关的分子、生化和细胞事件、以及研究干细胞向不同细胞谱系的分化和发育的机会，还提供了独特的治疗和药物发现工具。

[0061] 例如，本发明的干细胞系提供了可扩增干细胞的独特来源，其可用于多种移植、治疗和预防策略，包括治疗癌症，特别是放射治疗的癌症。例如，在目前对于白血病的治疗中，获得骨髓供体的限制和所述供体提供干细胞的限制严重限制了在放射治疗后重建患者的选择。本发明通过提供一种方法解决了这一问题，该方法用于产生可持续扩增和可更新的自体干细胞或组织相容性干细胞供应，所述细胞可根据需要储存或恢复。所述技术可最终免除对骨髓供体的需要。另外，多种免疫缺陷病和贫血疾病（例如，再生障碍性贫血或溶血性贫血）也将大大从本技术中受益，因为本发明提供了根据个体需要再生个体的造血细胞的能力。而且，衰老过程与造血室的几种重要改变有关，尤其包括越来越不能实现产生性的免疫应答。来源于衰老小鼠的造血干细胞已经显示具有较高水平的对于 DNA 修复问题的 mRNA。这可能最后影响它们自我更新、进行分化、进行增殖、和响应于骨髓细胞因子存活的能力。因此，衰老的个体也可从本发明中受益，因为可提供健康造血细胞的持续供应以矫正或缓解所述缺陷。

[0062] 本发明的技术不限于骨髓干细胞,而是可应用于实际上任何类型的干细胞,并且可延伸到来源于成人的细胞之外,扩展到胚胎干细胞。

[0063] 在一个实施例中,本发明的另一种应用涉及持续可扩增和可更新的毛囊干细胞的生产。源于该谱系的条件无限增殖化干细胞的发育可用于烧伤患者的整形手术中,用于任何经历化疗和/或放射治疗导致头发生长不可逆损失的患者,还可用于经历了影响颅骨的任何外科手术后的患者。而且,所述细胞可用于选择性方法中,所述方法包括诱导被遗传性秃发症影响的个体的头发生长。类似地,本发明的干细胞应用于皮肤对于烧伤患者的伤口愈合和治疗、以及对于创伤和其他患者的整形外科手术、以及选择性外科手术(包括但不限于整容外科手术)是无价的。另外可对所述细胞进行遗传操纵以矫正年轻和衰老个体的先天或后天遗传缺陷。本领域技术人员将理解,基于本公开内容,可从本发明在多种其他干细胞群体的应用中获得益处,这些干细胞包括但不限于来源于肺、乳房和肠上皮的干细胞,和来源于神经和心脏组织的干细胞,仅以这些为例。其他干细胞类型在本文其它部分提到。

[0064] 另外,本发明给个体提供了在个体的一生中根据需要获得自体干细胞和其分化细胞的可扩增供应的独特机会。例如,已知当机体衰老时,免疫功能和免疫记忆恶化。但是,利用本发明提供的技术,将有可能利用能够分化为造血谱系所有细胞的新的自体干细胞再生个体,从而给衰老个体提供“年轻的”免疫系统。另外,如果需要(例如,在个体发展了癌症或免疫缺陷疾病或对于实际上任何类型的新生自体细胞有其他需要的事件中),用本发明方法产生的干细胞可在个体的一生中作为治疗方案的一部分储存和利用。

[0065] 本发明还提供了基因治疗的独特机会。具体地,通过以下方法现在可以矫正基因缺陷或者可以将有益的基因修饰导入体细胞中:操作利用本发明方法进行条件无限增殖化及扩增的从个体中获得的自体干细胞。然后可将所述干细胞再导入它们的来源个体中。

[0066] 通过本发明方法产生的干细胞也可用于多种药物发现实验中。因为现在可以生产实际上无限供应的易于储存、恢复、扩增和操作的同质的干细胞,所述干细胞可在试验中作为干细胞使用或分化为多种细胞谱系使用,用于测试多种化合物对细胞分化、基因表达和细胞过程的影响。所述细胞可在与化合物接触之前被操作,例如通过遗传操作。来自于具有遗传缺陷的个体的干细胞可用所述实验进行评价,以鉴定治疗性化合物(例如,癌症治疗剂)和评价基因替代疗法。实际上,本发明的技术提供了一种针对特定个体的细胞以鉴定药物候选物和治疗候选物和“适合”个体细胞的策略的机会。所述实验的一个例子在下面有更详细的描述。

[0067] 关于在本发明之前对于谱系定向和细胞分化和发育的研究和发现,无法获得和不能产生足够数量的期望的细胞群体来进行期望的实验严重阻碍了所述研究。例如,为了鉴定或筛选在特定祖细胞系分化中的中间产物,必须获得足够数量的细胞来提供有意义的和可重复的结果。祖细胞系也应保持在已经定向的谱系中进一步分化的能力,因此产生了这些目前还不存在的新工具,也没有关于产生这些细胞所需要的技术的其他描述。利用本发明之前可获得的技术,这是不可能的。本发明通过提供可扩增的和基本上无限供应的可用于多种实验的同质干细胞解决了这一问题。这一技术将大大提高细胞分化和发现领域的研究能力。

[0068] 如上所述,本发明的条件无限增殖化 1t-HSC 的方法可适用于来源于其他组织的其他干细胞。例如,如果需要,通过改变基因传递和生长因子,本发明可应用于如下所述的

多种不同的干细胞。所述细胞也可在体外扩增,并在癌基因灭活后继续分化,如此处对于造血干细胞所述。这些细胞然后可用于治疗用途,包括组织修复和组织再生/工程化。相应地,可通过在此描述的或本领域技术人员根据本公开内容认为适合的任何方法(包括多种病毒介导的方法),将 MYC-ER 和 Bcl-2 基因的组合或在此描述的任何其他组合转染入细胞中,所述细胞包括但不限于:间充质干细胞(包括但不限于肺间充质细胞,骨髓基质细胞),神经干细胞(包括但不限于神经元多巴胺能干细胞和运动神经元干细胞),上皮干细胞(包括但不限于肺上皮干细胞、乳房上皮干细胞、和肠上皮干细胞),心肌细胞祖干细胞,皮肤干细胞(包括但不限于表皮干细胞和囊干细胞),骨骼肌干细胞,内皮干细胞(例如,肺内皮干细胞),和肝干细胞,以产生可在体外扩增并且在癌基因灭活后继续分化的条件无限增殖化细胞系。除了所述细胞系的治疗潜力以外,这些细胞系可进一步在体外(或离体)修饰,以矫正先天的遗传缺陷,以及用于研究早期谱系定向和分化的分子基础。这些细胞可能是潜在相关的治疗目标的新资源,这些细胞系也可用于筛选防止或诱导分化的小分子,和鉴定用于多种治疗(包括但不限于,癌症治疗和免疫缺陷治疗)的新的化合物和分子靶标。

[0069] 一般定义

[0070] 根据本发明,本文提到分离的核酸分子是指已经从其自然环境中分离出的核酸分子(即,已经进行了人工操作),其自然环境是在其中天然发现该核酸分子的基因组或染色体。因此,“分离的”并不必然反映核酸分子被纯化的程度,而是表示该分子不包括在其中天然发现该核酸分子的全基因组或全染色体。分离的核酸分子可包括基因。包括基因的分离的核酸分子不是包括所述基因的染色体的片段,而是包括与该基因相关的编码区和调节区,但没有在相同染色体上天然发现的其他基因。分离的核酸分子还可包括特定的核酸分子序列,该序列的侧面(即,在该序列的 5' 和 / 或 3' 端)为在天然状态下通常不在该特定核酸序列侧面的其他核酸(即,异源序列)。分离的核酸分子可包括 DNA、RNA(例如 mRNA)、或 DNA 或 RNA 的衍生物(例如, cDNA、siRNA、shRNA)。尽管“核酸分子”主要指物理的核酸分子,“核酸序列”主要指核酸分子上的核苷酸序列,但是这两个词可互换使用,特别是对于能够编码蛋白质或蛋白质结构域的核酸分子或核酸序列来说。

[0071] 优选地,本发明的分离的核酸分子利用重组 DNA 技术(例如,聚合酶链反应(PCR)扩增、克隆)或化学合成法来生产。分离的核酸分子包括天然核酸分子和其同源物,包括但不限于天然等位变体和修饰的核酸分子,其中已插入、缺失、置换和 / 或倒置核苷酸,以使所述修饰提供期望的效果(例如,如在此描述的,提供可诱导的原癌基因)。

[0072] 核酸分子同源物可利用多种本领域技术人员已知的方法生产(见例如 Sambrook 等人, *Molecular Cloning: A Laboratory Manual*, ColdSpring Harbor Labs Press(1989))。例如,可利用多种技术修饰核酸分子,包括但不限于经典的诱变技术和重组 DNA 技术,例如定点诱变,用于诱导突变的核酸分子的化学处理,核酸片段的限制性酶切,核酸片段的连接,核酸序列的所选区域的 PCR 扩增和 / 或诱变,寡核苷酸混合物的合成和混合物的连接以“建立”核酸分子及其组合的混合物。可通过对由核酸编码的蛋白质的功能进行筛选和 / 或通过与野生型基因杂交,从修饰的核酸混合物中选择核酸分子同源物。

[0073] 本发明的核酸分子或多核苷酸的最小的大小是足以编码用于本发明的蛋白质的大小,该蛋白质例如是由原癌基因编码的蛋白质或其功能片段(即,具有全长蛋白质的生物学活性并足以用于本发明的方法的部分),或抗细胞凋亡蛋白质或其功能片段(即,具有

全长蛋白质的生物学活性并足以用于本发明的方法的部分)。其他的可用于本发明的核酸分子可包括具有足以形成探针或寡核苷酸引物的最小大小的核酸分子,所述探针或引物能够与编码天然蛋白质的核酸分子的互补序列形成稳定的杂合体(即,在中度、高度或极高严格条件下),其长度一般至少为5个核苷酸,优选范围是5到大约50或大约500个核苷酸或更长,包括在全部数字增量之间的任何长度(例如,5、6、7、8、9、10、... 33、34、... 256、257、... 500)。除了实用性的限制之外,对于本发明的核酸分子的最大大小没有限制,核酸分子可包括足以用于本发明在此描述的任何实施方式的序列。

[0074] 如此处所用的,严格杂交条件指标准杂交条件,在这种条件下,核酸分子用于鉴定类似的核酸分子。所述标准条件在(例如)Sambrook等人, *Molecular Cloning :A Laboratory Manual*, Cold Spring Harbor Labs Press, 1989 中公开。Sambrook等人(如上)全文引入本文作为参考(特别见,第9.31-9.62页)。另外,例如Meinkoth等人, 1984, *Anal. Biochem.* 138, 267-284 公开了为获得允许不同核苷酸错配程度的杂交而计算适当的杂交和清洗条件的公式;Meinkoth等人(如上)全文引入本文作为参考。

[0075] 更特别地,此处所指的中度严格杂交和清洗条件指允许分离与用于杂交反应的探针的核酸分子具有至少大约70%核酸序列同一性的核酸分子的条件(即,允许大约30%或更少的核苷酸错配的条件)。此处所指的高度严格杂交和清洗条件指允许分离与用于杂交反应的探针的核酸分子具有至少大约80%核酸序列同一性的核酸分子的条件(即,允许大约20%或更少的核苷酸错配的条件)。此处所述的极高度严格杂交和清洗条件指允许分离与用于杂交反应的探针的核酸分子具有至少大约90%核酸序列同一性的核酸分子的条件(即,允许大约10%或更少的核苷酸错配的条件)。如上所述,本领域技术人员可利用Meinkoth等人(如上)中的公式计算获得特定核苷酸错配水平所需的适当杂交和清洗条件。所述条件将根据是形成DNA:RNA杂合体还是形成DNA:DNA杂合体而改变。计算的DNA:DNA杂合体的解链温度比DNA:RNA杂合体低10℃。在特定的实施方式中,用于DNA:DNA杂合体的严格杂交条件包括在6X SSC(0.9M Na⁺)的离子强度下、在大约20℃到大约35℃之间(较低的严格度)杂交、更优选在大约28℃到大约40℃之间(更严格)、更优选在大约35℃到大约45℃之间(甚至更严格)的温度下杂交,并使用适当的清洗条件。在特别的实施方式中,DNA:RNA杂合体的严格杂交条件包括在6XSSC(0.9M Na⁺)的离子强度下、在大约30℃到45℃之间(较低的严格度)、更优选在大约38℃到大约50℃之间(更严格)、更优选在大约45℃到大约55℃之间(甚至更严格)的温度下杂交,并使用类似的严格清洗条件。这些数值是基于对具有大于大约100个核苷酸、0%甲酰胺和大约40%的G+C含量的分子的解链温度的计算。或者,T_m可如Sambrook等人(同上)第9.31到9.62页所述根据经验计算。通常清洗条件应该尽可能严格,并且应该适合于选择的杂交条件。例如,杂交条件可包括盐和温度条件的组合,该温度大约比计算的特定杂合体的T_m低大约20-25℃,清洗条件一般包括盐和温度条件的组合,该温度大约比计算的特定杂合体的T_m低大约12-20℃。适用于DNA:DNA杂合体的杂交条件的一个例子包括在6X SSC(50%甲酰胺)中在大约42℃下杂交2-24小时,然后是包括在室温下在大约2XSSC中一次或多次清洗的清洗步骤,紧接着在更高的温度和更低的离子强度下再进行清洗(例如,在大约37℃下在大约0.1X-0.5X SSC中至少清洗一次,紧接着在大约68℃下在大约0.1X-0.5X SSC中至少清洗一次)。

[0076] 在本发明的一个实施方式中,可生产任何此处描述的氨基酸序列,包括所述序列

的截短形式（片段或部分）和同源物，其在给定氨基酸序列的 C- 和 / 或 N- 末端每侧具有从至少一个到大约 20 个附加的异源氨基酸。得到的蛋白质或多肽可被称为“基本由给定的氨基酸序列组成”。根据本发明，如果利用给定氨基酸序列的来源生物体的标准密码子将天然存在的序列中的所述核苷酸进行翻译，异源氨基酸是在给定氨基酸序列侧翼未天然发现的（即，在体内未天然发现）序列，或者在存在于该基因中时，是不被位于编码给定氨基酸序列的天然存在的核酸序列侧翼的核苷酸所编码的序列。类似地，“基本由 组成”，当在此用于核酸序列时，是指编码给定氨基酸序列的核酸序列，在编码给定氨基酸序列的核酸序列的 5' 和 / 或 3' 末端的侧翼可以有至少一个到大约 60 个另外的异源核苷酸。如在天然基因中存在，异源核苷酸在编码给定氨基酸序列的核酸序列侧翼未被天然发现（即，在体内未天然发现）。

[0077] 根据本发明，重组载体（通常称为重组核酸分子，特别是当它包含根据本发明的目标核酸序列时）是一种工程化的（即，人工生产的）核酸分子，其被用作操作选择的核酸序列和将所述核酸序列导入宿主细胞的工具。重组载体因此适用于对选择的核酸序列的克隆、测序和 / 或其他操作，例如通过表达选择的核酸序列和 / 或将其递送入宿主细胞中。所述载体典型地包含异源核酸序列，即，天然地或通常发现不与待克隆或递送的核酸序列相邻的核酸序列，虽然载体也可包含调节性核酸序列（例如，启动子、非翻译区），该调节性核酸序列被天然发现与本发明的核酸分子相邻，或可用于表达本发明的核酸分子（下面更详细地讨论）。载体可以是 RNA 或 DNA，原核的或真核的，典型地是质粒或病毒载体。载体可作为染色体外元件（例如，质粒）保持或可被整合入宿主细胞的染色体中。完整的载体可保持在宿主细胞内适当位置，或在某种条件下，质粒 DNA 可被除去，留下本发明的核酸分子。在其他条件下，将载体设计为在选择的时间从宿主细胞的基因组上切除（移去）（下面更详细地讨论）。整合的核酸分子可在染色体启动子的控制下，在天然或质粒启动子的控制下，或在几种启动子组合的控制下。核酸分子的单个或多个拷贝可被整合入染色体中。本发明的重组载体可包含至少一个选择性标记。

[0078] 根据本发明，“有效连接”指以一种方式将核酸分子与表达控制序列（例如，转录控制序列和 / 或翻译控制序列）连接，使该分子在转染（例如，转化、转导、转染、缀合或导致）入宿主细胞中时可以表达。转录控制序列是控制转录的启动、延伸或终止的序列。特别重要的转录控制序列是控制转录启动的序列，例如启动子、增强子、操纵子和阻遏子序列。适合的转录控制序列包括可在该重组核酸分子待导入的宿主细胞或生物体内起作用的任何转录控制序列。

[0079] 根据本发明，术语“转染”是指将外源核酸分子（即，重组核酸分子）插入细胞内的任何方法。术语“转导”是一种特定类型的转染，其中遗传物质从一种来源转移到另一种来源，例如通过病毒（例如，逆转录病毒）或转导噬菌体。当该术语用于指核酸分子导入微生物细胞例如细菌和酵母中时，术语“转化”可与“转染”互换使用。在微生物系统中，术语“转化”用于描述由于微生物获得外源核酸而导致的遗传改变，并且与“转染”基本上同义。但是，在动物细胞中，转化有第二种含义，可指细胞在（例如）成为癌性之后在培养中的生长性质改变。因此，为避免混淆，“转染”在此优选用于外源核酸向动物细胞中的导入。因此，术语“转染”在此一般包括动物细胞的转染或转导，和微生物细胞的转化或转导，这些术语涉及外源核酸向细胞内的导入。转染技术包括但不限于转化、转导、粒子轰击、扩散、主动

转运、超声浴、电穿孔术、显微注射法、脂质转染法、吸附、感染和原生质融合。

[0080] 如此处所用，本发明中提到的分离蛋白质或多肽包括全长蛋白质、融合蛋白、嵌合蛋白或所述蛋白质的任何片段（截短形式、部分）或同源物。更特别地，根据本发明的分离的蛋白是从其天然环境中分离出（即，已经进行了人工操作）的蛋白质（包括多肽或肽），可包括但不限于纯化的蛋白质、部分纯化的蛋白质、重组产生的蛋白质、膜结合蛋白质、和与其他蛋白质相关的分离的蛋白质。因此，“分离的”不反映蛋白质被纯化的程度。优选地，本发明的分离蛋白是重组生产的。另外，作为特定蛋白质（Bcl-2）的例子，“人 Bcl-2 蛋白”或“源于”人 Bcl-2 蛋白的蛋白质是指来自人（*Homo sapiens*）的 Bcl-2 蛋白质（包括天然存在的 Bcl-2 蛋白质的同源物或部分）或指根据天然存在的人 Bcl-2 蛋白质的结构（例如，序列）和（也许）功能的知识以其他方式生产的 Bcl-2 蛋白。换句话说，人 Bcl-2 蛋白包括具有与天然存在的人 Bcl-2 蛋白基本类似的结构和功能或为此处描述的天然存在的人 Bcl-2 蛋白的生物活性（即，具有生物活性）同源物的任何 Bcl-2 蛋白。因此，人 Bcl-2 蛋白可包括纯化的、部分纯化的、重组的、突变的 / 修饰的和合成的蛋白质。根据本发明，术语“修饰”和“突变”可互换使用，特别是对于此处描述的蛋白质的氨基酸序列（或核酸序列）的修饰 / 突变。

[0081] 如此处所用的术语“同源物”是指由于对天然存在的蛋白质或肽的修饰（包括小修饰）而与天然存在的蛋白质或肽（即，“原型”或“野生型”蛋白质）不同，但保持了天然存在形式的基本蛋白质和侧链结构的蛋白质或肽。所述改变包括但不限于：一个或几个（即，1、2、3、4、5、6、7、8、9 或 10 个）氨基酸侧链的改变；一个或几个氨基酸的改变，包括缺失（例如，蛋白质或截短形式的蛋白质或肽）、插入和 / 或置换；一个或几个原子的立体化学的改变；和 / 或小的衍生化，包括但不限于：甲基化、糖基化、磷酸化、乙酰化、肉豆蔻酰化、异戊烯化、棕榈酸化、酰胺化和 / 或添加糖基磷脂酰肌醇。同源物可具有与天然存在的蛋白质或肽相比提高的、降低的或基本类似的性质。同源物可包括蛋白质的激动剂或蛋白质的拮抗剂。

[0082] 同源物可以是天然等位变异或天然突变的结果。编码蛋白质的核酸的天然存在的等位变体是基本上存在于与编码所述蛋白质的基因相同的基因组位点处的基因，但由于例如突变或重组引起的天然变异，具有类似的但不相同的序列。等位变体典型地编码与所对比的基因编码的蛋白质具有类似活性的蛋白质。一组等位变体可编码相同的蛋白质，但由于遗传密码的简并性而具有不同的核酸序列。等位变体也可包含在基因的 5' 或 3' 非翻译区（例如，在调节控制区）的改变。等位变体对于本领域技术人员是公知的。

[0083] 同源物可利用本领域已知的生产蛋白质的技术生产，包括但不限于对分离的天然存在的蛋白质的直接修饰、直接的蛋白质合成、或对编码蛋白质的核酸序列的修饰，例如，利用经典的或重组 DNA 技术实现随机或靶向诱变。

[0084] 在一个实施方式种，给定蛋白质的同源物包含一种氨基酸序列、基本由该序列组成、或由该序列组成，所述氨基酸序列与参考蛋白质的氨基酸序列至少大约 45%、或至少大约 50%、或至少大约 55%、或至少大约 60%、或至少大约 65%、或至少大约 70%、或至少大约 75%、或至少大约 80%、或至少大约 85%、或至少大约 90%、或至少大约 95% 相同、或至少大约 95% 相同、或至少大约 96% 相同、或至少大约 97% 相同、或至少大约 98% 相同、或至少大约 99% 相同（或在 45% 到 99% 之间的全部整数增量的任何百分同一性）。在一个

实施方式中,同源物包含一种氨基酸序列、基本上由该氨基酸序列组成、或由该氨基酸序列组成,所述氨基酸序列与参考蛋白质的天然存在的氨基酸序列少于 100%相同,小于大约 99%相同,小于大约 98%相同,小于大约 97%相同,小于大约 96%相同,小于大约 95%相同,等等,以 1%的增量,直到小于大约 70%相同。

[0085] 如此处所用,除非另外说明,百分(%)同一性是指对同源性的评价,其通过下述方法进行:(1)具有标准缺省参数的 BLAST 2.0BasicBLAST 同源性检索,其使用 blastp 进行氨基酸检索,使用 blastn 进行核酸检索,其中默认对低复杂性区域过滤查询序列(描述于 Altschul, S. F., Madden, T. L., Schaaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997) " Gapped BLAST and PSI-BLAST: a new generation of protein database search programs." Nucleic Acids Res. 25 :3389-3402,全文在此引入作为参考);(2)BLAST 2 比对(使用下述的参数);(3)和/或具有标准缺省参数的 PSI-BLAST(位点特异性迭代 BLAST)。应该注意到,由于在 BLAST 2.0BASIC BLAST 和 BLAST 2 之间的标准参数有一些区别,两种特定序列可能利用 BLAST 2 程序确定为具有显著同源性,而在 BLAST 2.0BASIC BLAST 中使用一种序列作为查询序列进行检索可能不能鉴定出极其匹配的第二序列。另外,PSI-BLAST 提供了一种自动的、易于使用的“分布图”检索模式,其是一种寻找序列同源物的灵敏方法。该程序首先进行缺口 BLAST 数据库检索。PSI-BLAST 程序使用来自任何显著比对的信息,所述信息被返回以构建一种位点特异性评分矩阵,其替代用于下一轮数据库检索的查询序列。因此,可以理解百分同一性可以利用这些程序中的任何一种来确定。

[0086] 两种特定序列可利用如 Tatusova 和 Madden, (1999), " Blast 2sequences-a new tool for comparing protein and nucleotide sequences", FEMS Microbiol Lett. 174 : 247-250 所述的 BLAST 2 程序进行相互比对,该文献在此全文引入作为参考。BLAST 2 序列比对以 blastp 或 blastn 进行,利用 BLAST 2.0 算法在两个序列之间进行缺口 BLAST 检索(BLAST 2.0),允许在获得的比对中引入缺口(gaps)(缺失和插入)。在此为了清楚起见,利用下述的标准缺省参数进行 BLAST 2 序列比对:

[0087] 对于 blastn,使用 0BLOSUM62 矩阵:

[0088] 匹配赠分 = 1

[0089] 错配罚分 = -2

[0090] 开放缺口 (5) 和延伸缺口 (2) 罚分

[0091] 缺口 x_ 下降 (50)期望 (10) 字长 (11) 过滤 (开)

[0092] 对于 blastp,使用 0BLOSUM62 矩阵:

[0093] 开放缺口 (11) 和延伸缺口 (1) 罚分

[0094] 缺口 x_ 下降 (50)期望 (10) 字长 (3) 过滤 (开)

[0095] 根据本发明,分离的蛋白质,包括其生物活性同源物或片段,具有野生型或天然蛋白质的至少一种生物活性特征。一般地,蛋白质的生物活性或生物作用指归因于蛋白质的天然存在形式的蛋白质抽表现出的或执行的任何功能,其在体内(即,在蛋白质的天然生理环境中)或在体外(即,在实验室条件下)测量或观察。导致蛋白质表达降低或蛋白质活性降低的修饰、活动或相互作用可被称为蛋白质的失活(完全或部分)、下调、作用降低、或作用或活性降低。类似地,导致蛋白质表达升高或蛋白活性升高的修饰、活动或相互作用

可被称为蛋白质的扩增、过量生产、活化、增强、上调或作用增强。

[0096] 本发明的条件无限增殖化方法

[0097] 本发明的一个实施方式涉及生产条件无限增殖化成年干细胞、优选长期干细胞的方法。该方法一般包括下述步骤：(a) 获得扩增的成年干细胞群体；(b) 用包含促进细胞存活和增殖的原癌基因的载体转染（转导）干细胞，其中原癌基因是可调节的（可诱导、可控制）；(c) 用编码抑制细胞凋亡的蛋白质的载体转染干细胞；和 (d) 在干细胞生长因子组合的存在下并在使原癌基因为活性的条件下扩增转染的细胞。在一个实施方式中，载体是一种整合载体。通过这种方法生产的细胞可被培养、扩增、储存、恢复、用于治疗方法中、用于研究和发现方法中、遗传操作、通过消除使原癌基因保持活性的条件来诱导分化，和 / 或用于此处所述的或基于该公开内容对于本领域技术人员显而易见的任何其他方法。步骤 (b) 和 (c) 可以任何顺序执行。

[0098] 根据本发明，“条件无限增殖化”是指以一种可逆的方式无限增殖化的细胞（例如，能够以依赖于细胞因子的方式无限生长而不分化，同时保持在适当条件下分化为大量不同谱系的能力和潜能），所述细胞在一组特定的条件下被无限增殖化，当条件被消除或改变（或增加其他条件）时，细胞不再无限增殖并且可分化为其他细胞型。“条件无限增殖化”可与“可逆无限增殖化”可互换使用。例如，关于本发明的方法，当将干细胞置于允许原癌基因活化的条件下时（例如，在 MYC-ER 的例子中的他莫昔芬或其激动剂），促进细胞存活和增殖的可调节原癌基因的存在使细胞保持无限增殖化的表型。换句话说，细胞在培养物中无限地生长和扩增，并且在这些特定条件下保持未分化的状态。当这些条件被消除（例如，在 MYC-ER 的例子中除去他莫昔芬）时，干细胞不再无限增殖并且可根据适当的环境（例如，适当的生长因子组合）分化为多种细胞谱系。

[0099] 如此处所用的“干细胞”是指本领域一般理解的术语。例如，干细胞，不管它们的来源如何，是能够长时间自我分裂和更新的细胞，是非特化的（非分化的），并且可产生（分化为）特定的细胞类型（即，它们是多种不同特定细胞类型的祖细胞或前体细胞）。“长期”，当与干细胞联用时，是指干细胞在长时间内（例如，数月，例如至少 3 个月，到几年）根据干细胞的特定类型通过分裂为相同的非特定细胞类型而自我更新的能力。如此处所述，源于不同动物种类的多种长期干细胞例如长期造血干细胞 (lt-HSC) 的表型特征是本领域已知的。例如，鼠 lt-HSC 可根据下列细胞表面标志物表型的存在而被鉴定：c-kit⁺、Sca-1⁺、CD34⁻、f1k2⁻（见实施例）。成年干细胞包括可获自任何非胚胎组织或来源的干细胞，并且典型地产生它们所处的组织的细胞类型。术语“成年干细胞”可与术语“体干细胞”互换使用。胚胎干细胞是从任何胚胎组织或来源获得的干细胞。

[0100] 在本发明的一个实施方式中，用于本发明的干细胞可包括获自任何来源的任何成年干细胞。在本发明的另一实施方式中，干细胞可包括胚胎干细胞。本发明中有用的干细胞包括但不限于：造血干细胞，间充质干细胞（包括但不限于，肺间充质干细胞，骨髓基质细胞），神经干细胞，上皮干细胞（包括但不限于肺上皮干细胞、乳房上皮干细胞、血管上皮干细胞、和肠上皮干细胞），肠干细胞，心肌细胞祖干细胞，皮肤干细胞（包括但不限于表皮干细胞和囊泡干细胞（毛囊干细胞）），骨骼肌干细胞，成骨细胞前体干细胞，和肝干细胞。

[0101] 造血干细胞产生所有类型的血细胞，包括但不限于红细胞（红血球）、B 淋巴细胞、T 淋巴细胞、自然杀伤细胞、嗜中性粒细胞、嗜碱性粒细胞、嗜酸性粒细胞、单核细胞、巨噬细

胞、和血小板。

[0102] 间充质干细胞（包括骨髓基质细胞）产生多种细胞类型，包括但不限于骨细胞、软骨细胞、脂肪细胞、肺细胞，和其他种类的结缔组织细胞，例如腱中的结缔组织细胞。

[0103] 脑中的神经干细胞产生三种主要的细胞类型：神经细胞（神经元）和两类肺神经元细胞：星形胶质细胞和少突胶质细胞。

[0104] 不同组织的内层（lining）的上皮干细胞产生形成组织上皮的几种细胞类型。

[0105] 皮肤干细胞存在于表皮的底层和毛囊的底部。表皮干细胞产生角质化细胞，其迁移到皮肤表面形成保护层，囊泡干细胞可产生毛囊和表皮。本领域技术人员也熟知其他来源的成年干细胞。

[0106] 胚胎干细胞可产生身体的所有组织和细胞。

[0107] 获得所述干细胞和提供起始培养条件例如液体培养基或半固体培养基的方法是本领域已知的。开始时通过将干细胞来源与一种在组织源或培养物中扩增或富集所述细胞的适当试剂接触，在体内或在体外扩增细胞。例如，在造血干细胞的例子中，供体个体可用富集造血干细胞和促进所述细胞增殖而不分化的试剂例如 5- 氟尿嘧啶处理。其他适合于扩增期望的干细胞类型的试剂是本领域公知的。可选地，并且优选地，成年干细胞从组织来源中分离，然后在体外通过暴露于适当试剂而扩增或富集。例如，对于造血干细胞，生产成年造血祖细胞的扩增培养物的方法在 Van Parijs 等人，(1999 ;Immunity, 11, 763-70) 中描述。通过获得动物的细胞样品的任何适当方法从个体中获得细胞，包括但不限于体液（例如，血液）的骨髓收集、收集脐带血、组织穿刺和组织解剖，特别包括但不限于皮肤、肠、角膜、脊髓、脑组织、头皮、胃、乳房、肺（例如，包括灌洗和支气管检查）、骨髓、羊水、胎盘和卵黄囊的细针抽吸物的任何活组织检查。

[0108] 在一个实施方式中，用于本发明的细胞也可获自新鲜的或冷冻保藏（储存）的脐带血、可源于胚胎干细胞（ES）体外定向分化的造血祖细胞群、从正常或粒细胞集落刺激因子（G-CSF）治疗的患者的外周血获得的造血干细胞（HSC），所述患者已被诱导将其 1t-HSCs 动员到外周循环中。

[0109] 一旦获得扩增的干细胞群体（可以获得、提供或产生），即用以下物质连续（以任何顺序）或同时转染细胞：(1) 包含促进细胞存活和增殖的原癌基因的载体，其中原癌基因是可调节的（可诱导，可控制），和 (2) 编码抑制细胞的细胞凋亡的蛋白质的载体。优选地，所述载体是整合性载体，此处定义为具有整合入细胞基因组的能力的任何载体（例如，逆转录病毒载体）。在下面对多种载体和转染方法进行更详细的描述。所述原癌基因是可调节的（可诱导或可控制），因此可以根据需要活化和灭活该原癌基因（即，开启或关闭），以使干细胞保持无限增殖状态或使其分化为期望的细胞类型。可以选择或设计原癌基因，以用任何适当的方法进行调节，包括响应于任何条件，例如化合物或试剂、温度或任何其他适当条件的存在或不存在。作为例子，此处描述的原癌基因 MYC-ER（雌激素受体（ER）- 调节的 MYC）和 ICN-1-ER（Notch-1 的 ER- 调节的细胞内部分）在他莫昔芬存在下都是可诱导的。应该注意到所述基因也可被工程化以响应于其他二聚化药物，例如 FK1012、改变形式的雷帕霉素，或可从包含四环素响应元件的载体中表达。后一种情况调节蛋白质的表达，不调节细胞中存在的多肽的功能。这一平台技术的其他类似修饰对于本领域技术人员是显而易见的。

[0110] 在本发明的方法中有用的原癌基因是促进细胞存活和增殖的任何原癌基因。优选的用于本发明的方法的原癌基因包括但不限于 MYC、ICN-1、hTERT(人端粒酶的逆转录酶成分)、NMYC、S-MYC、L-MYC、Akt(十四烷基化的)。另外,其他适用的基因和本发明的方法或修饰基因以获得期望的结果的方法包括但不限于使用下游信号效应物例如丙酮酸脱氢酶激酶 1(PDK-1);雷帕霉素的哺乳动物靶标(mTOR);shRNA 引起的磷酸酶和张力蛋白同源物(PTEN)的丧失;Bcl-3、细胞周期蛋白 D1、细胞周期蛋白 D3、Bcl-10、Bcl-6、BCR-ABL(与 ABL 融合的断裂点簇集区)及其各种突变形式、stat5 和 stat3 的组成型活性形式、AML1-ETO(急性髓细胞白血病 1 和矮小相关转录因子 1 的融合物)、MLL-ENL(混合谱系白血病和十一十九白血病(eleven nineteen leukemia))、Hox 基因、活化形式的白介素-3(IL-3)受体 β 链、和其他细胞因子受体链(表皮生长因子受体(EGFR)、c-kit、血小板来源的生长因子受体(PDGFR)等),以及 wnt(所有哺乳动物形式)、 β -联蛋白、shh 蛋白(shh-1 和所有哺乳动物形式)、bmi-1 和 c-jun(所有哺乳动物形式)。并且,本发明包括 shRNA 诱导(或抑制)细胞周期蛋白激酶抑制剂的丧失,包括但不限于 pi 6、pi 9、p21 和 p27。在一个实施方式中,本发明包括任何或所述原癌基因(例如,MYC-ER 或 ICN-1-ER)或其他基因的可调节同源物的应用。实施例描述了使用 MYC-ER 或 ICN-1-ER 利用本发明的方法成功地生产条件无限增殖化 1t-HSC。

[0111] 编码人 MYC 的核酸序列在此处表示为 SEQ ID NO:1,其编码此处表示为 SEQ ID NO:2 的氨基酸序列。编码 hTERT 的核酸序列在此处表示为 SEQ ID NO:3,其编码此处表示为 SEQ ID NO:4 的氨基酸序列。编码人 ICN-1 的核酸序列在此处表示为 SEQ ID NO:11,其编码此处表示为 SEQ ID NO:12 的氨基酸序列。ICN-1 是 Notch-1 的一部分,具体地,是 Notch-1 的氨基酸 1757-2555(见 Aster 等人, Mol Cell Biol. 2000Oct;20(20):7505-15,在此全文引入作为参考)。MYC-ER 的核苷酸和氨基酸序列是本领域已知的,并且 MYC-ER 蛋白在 Soloman 等人, Oncogene. 1995Nov 2;11(9):1893-7 中描述,在此全文引入作为参考。ICN-1-ER 是本发明人创造的,编码此蛋白质的核酸序列在此处表示为 SEQ ID NO:13,其编码 SEQ ID NO:14 表示的氨基酸序列。

[0112] 类似地,一种优选的抗细胞凋亡基因是 Bcl-2,尽管本发明也包括其他编码抑制细胞凋亡并且特别在原癌基因在干细胞中失活时维持细胞存活的蛋白质的基因。编码 Bcl-2 α 的核酸序列在此处表示为 SEQ ID NO:5,其编码 SEQ ID NO:6 的氨基酸序列。Bcl-2 β 在此处表示为 SEQ ID NO:7,其编码 SEQ ID NO:8 的氨基酸序列。“抗细胞凋亡”基因在此处被定义为编码甚至在能够诱导细胞凋亡的条件存在下仍能够抑制(减少、阻止、降低)细胞中与细胞凋亡相关的过程或促进(提高、增强、刺激、允许)细胞存活的蛋白质的任何基因。与细胞凋亡相关的蛋白质,和编码所述蛋白质的基因,是本领域公知的。所述其他基因包括但不限于可能在成年干细胞的条件转化情况中起重要作用的 Bcl-2 家族中的任何基因。这些基因包括但不限于 Bcl-2 家族的其他促存活成员,例如,Bcl-X、Bcl-w、BclXL、Mcl-1、Dad-1、或 hTERT(人端粒酶的逆转录酶成分,已经显示其抑制增殖)。所述基因在被调节的癌基因存在下异常过量表达,如本文的实施例中描述的 Bcl-2 那样。另外,本发明的这一方面包括对促凋亡 bcl-2 家族的 BH3-only 成员(例如,Bim、PUMA、NOXA、Bax、Bak、BclXS、Bad、Bar 等)进行 shRNA 介导的基因下调(knockdown)(或利用任何其他方法的破坏或抑制),以及破坏胱天蛋白酶 3,9,10、MLL-1(和所有哺乳动物形式)、En1-1(无胚

乳-1 (Endospermless-1)) 和所有哺乳动物形式、Apaf-1 和其他形成凋亡体一部分的成分。

[0113] 上述每个基因的核酸序列或其编码区是本领域已知的并且是公众可获得的, 包括人基因。类似地, 由这些基因编码的蛋白质的氨基酸序列是本领域已知并且公众可获得的。

[0114] 本发明人已经利用本发明的方法并利用原癌基因和抗细胞凋亡基因的不同组合制备了几种不同的长期条件无限增殖化干细胞, 包括下述组合: MYC-ER 和 Bcl-2; MYC-ER 和 hTERT (人端粒酶的逆转录酶成分); ICN-1-ER 和 Bcl-2; ICN-1-ER 和 hTERT; 和 MYC-ER 和 ICN-1-ER。

[0115] 应注意对于用于本方法的原癌基因或编码抗细胞凋亡蛋白的基因, 不需要在此处描述的构建体中使用整个基因, 因为本发明包括该基因的任何部分或编码期望的功能蛋白质产物、其功能部分或其功能同源物的核酸序列 (例如, cDNA)。相应地, 此处提到的用于转染干细胞的基因或转基因被理解为示例性的, 并且包括任何编码完整基因的核酸分子、基因的全部编码区、或基因的部分或其同源物的应用, 只要所述核酸序列编码适于本发明应用的功能性蛋白质即可。

[0116] 在本发明的一个实施方式中, 本方法另外包括针对编码促凋亡蛋白的 RNA 的 shRNA 或 siRNA 的应用, 所述促凋亡蛋白例如为 Bcl-2 家族的促细胞凋亡成员, 即 BH3-only 类型的成员 (Bim、Bax、Bak、Puma、Noxa, 等等)。预期在被调节的癌基因的背景中破坏促细胞凋亡基因将导致某些干细胞群的更有效的无限增殖化。RNA 干扰 (RNAi) 是双链 RNA 和 (在哺乳动物系统中) 短干扰 RNA (siRNA) 或短发夹 RNA (shRNA) 用于抑制互补基因的表达或使其沉默化的过程。在目标细胞中, siRNA 是未缠绕的并且与 RNA 诱导的沉默复合物 (RISC) 结合, RISC 随后被引导到与 siRNA 互补的 mRNA 序列上, 由此 RISC 剪切 mRNA。shRNA 在载体中被转染入目标细胞中, 在那里它被转录, 然后被 DICER 酶加工形成 siRNA 样分子, 该分子活化 RISC, 然后 RISC 与 siRNA 一起被引导到与 shRNA 互补的 mRNA 序列上, 由此 RISC 剪切 mRNA。

[0117] 可用任何适当的转染细胞特别是哺乳动物细胞的方法, 包括利用多种技术的组合, 用包含原癌基因和编码抗细胞凋亡蛋白的载体转染干细胞。本发明人已经发现在产生此处所述的条件无限增殖化长期干细胞的表达基因 (或构建体) 之间有特别的协同作用。实施例已经证实了逆转录载体的应用, 其他方法包括但不限于使用慢病毒载体、细小病毒、痘苗病毒、冠状病毒、嵌杯样病毒、乳头瘤病毒、黄病毒、正粘病毒、披膜病毒、小 RNA 病毒、腺病毒载体、修饰的和减毒的疱疹病毒。任何所述病毒可进一步用特定的表面表达的分子修饰, 所述分子将它们靶向于 HSC 或其他干细胞, 例如膜结合的 SCF, 或其他干细胞特异性生长因子配体。转染哺乳动物细胞的其他方法包括但不限于对哺乳动物表达载体进行直接电穿孔, 例如利用 NUCLEOFECTOR™ 技术 (AMAXA Biosystems)。这种技术是用于大多数初级细胞和难转染的细胞系的高效非病毒基因转移方法, 其是对于长期已知的电穿孔方法的改进, 基于利用电流和溶液的细胞类型特定组合将聚阴离子大分子直接转移到细胞核中。另外, 适当的转染方法可包括本领域已知的任何细菌、酵母或其他人工基因递送方法。

[0118] 在适当的生长因子的存在下扩增转染的干细胞或培养干细胞和外源融合蛋白 (例如, 在以下所述该方法的变化形式中描述的 Tat- 融合蛋白) 的步骤可包括使用任何适当的培养条件, 包括在此特别描述的条件。适当的干细胞生长因子的组合可包括任何允许本发明的转染 (例如, 转导) 细胞在培养中生长、存活和增殖的干细胞因子。虽然在此

描述了特定的组合,并且虽然这是本方法的一个重要步骤,但这一步骤可以简单地描述为提供适合干细胞生长、增殖和存活的生长因子的任何组合,并且包括本领域已知的任何组合。因此,本发明不限于特定的组合。一种优选的生长因子组合包括:白介素-6(IL-6)、IL-3 和干细胞因子(SCF)。另外一种优选的生长因子组合包括在无血清培养基中的干细胞因子(SCF)、血小板生成素(TPO)、胰岛素样生长因子 2(IGF-2) 和成纤维细胞生长因子 1(FGF-1)。这后一种组合最近由 Zhang 和 Lodich(2005; Murine hematopoietic stem cells change their surface phenotype during ex vivo expansion, Blood 105, 4314-20) 描述。预期用编码此处所述的组合蛋白质(例如,如实施例中描述的 MYC-ER 和 Bcl-2) 的核酸分子转染的干细胞在生长因子混合物中也被条件无限增殖化,如同在上述例子中描述的混合物(使用 IL-3、IL-6 和 SCF) 一样。其他用于本发明的生长因子包括但不限于血管生成素样蛋白(例如,Agpt12、Angpt13、Angpt15、Angpt17 等)、增殖蛋白-2(PLF2)、糖原合酶激酶-3 抑制剂、wnt 和 Notch 信号途径诱导剂、Flt3L 和相关细胞因子、成纤维细胞生长因子 2(FGF2) 和相关细胞因子、wnt-1 和 wnt 途径的其他活化剂、sbh 蛋白(shh-1) 和该途径的其他活化剂。其他合适的生长因子组合适用本发明的方法,并且对于本领域技术人员来说是显而易见的。实际上,利用本发明方法产生的细胞系可容易地用于筛选能够用于在体外在中性或定向条件下扩增长期干细胞或其衍生的任何祖细胞的其他细胞因子和生长因子。

[0119] 根据本发明,适合培养动物细胞的培养基可包括为培养动物细胞、特别是哺乳动物细胞而开发的任何可获得的培养基,或可用动物细胞生长必需的适当成分例如可同化的碳、氮和微量营养素在实验室中制备的培养基。所述培养基包括基础培养基,其是适于动物细胞生长的任何基础培养基,包括但不限于 Iscove 改良的 Dulbecco 培养基(IMDM)、Dulbecco 改良的 Eagles 培养基(DMEM)、 α MEM(Gibco)、RPMI 1640、或任何其他适合的商业上可获得的培养基。基础培养基中加入了可同化的碳源、氮源和微量营养素源,包括但不限于,血清源、生长因子、氨基酸、抗生素、维生素、还原剂和/或糖源。应注意包含基础培养基和动物细胞生长需要的许多附加成分的完全培养基是商业上可获得的,并且一些培养基可用于特定类型的细胞培养。另外,许多无血清培养基是可获得的,并且可能特别适于根据本发明的干细胞的培养。

[0120] 细胞和组合物

[0121] 本发明的另一实施方式涉及根据此处所述的本发明的方法生产的细胞、细胞系或细胞群体。本发明还包括包含所述细胞、细胞系或细胞群体的组合物。用于治疗方法时,所述组合物可包含药学上可接受的载体,包括药学上可接受的赋形剂和/或递送载体,用于将细胞、细胞系或细胞群体递送给患者。如此处所用的药学上可接受的载体是指适于将在本发明的方法中有用的治疗组合物递送到适当的体内部位的任何物质。

[0122] 为生产中间发育阶段的细胞谱系而对本发明方法的修改

[0123] 本发明的另一实施方式涉及修改此处所述的新方法以生产捕获造血谱系的中间发育阶段的细胞系。根据本发明,发育或分化的“中间”阶段指细胞发育或分化的多能阶段,其位于产生“中间”细胞的干细胞发育或分化阶段的下游,但位于细胞分化的终点或端点的上游。例如,前 B 细胞是造血干细胞的中间阶段,其仍然可分化为成熟 B 细胞。本领域技术人员可理解发育或分化的中间阶段。

[0124] 更特别地,对于许多治疗和发现或研究应用,以及细胞系的保存,期望细胞系具有一种稳定的表型并且当转染细胞的癌基因被关闭时保持沿着其定向途径进一步分化的能力。相应地,本发明包括生产没有完全分化(没有最终分化)、而是处于分化中间阶段的细胞的附加步骤。在该实施方式的一种非限制性例子中,在去除保持原癌基因或其他促进细胞存活和增殖的基因的活性的条件(例如,在他莫昔芬依赖性原癌基因的例子中的 4-OHT)后,或通过应用关闭(失活)原癌基因/癌基因的适当条件,利用上述方法生产的长期干细胞在体外随机分化。这一步骤可在使培养物保持在中性细胞因子生长条件(例如,IL-3、IL-6 和 SCF)下的同时进行,或者通过将能够特异性引导向某种谱系分化的细胞因子(例如,对于淋巴谱系的 IL-7 和 Notch 配体,对于树突细胞的 GM-CSF 和 IL-4,对于骨髓单核细胞的 G-CSF,等等)替换为对于分化为中性的细胞因子(不引导或驱动细胞的分化)而进行。一旦培养物开始显示与特定谱系一致的分化标记物,再次给培养基施加活化原癌基因的条件(例如,4-OHT)或使培养基暴露于以其他方式再活化原癌基因的条件,以稳定表型和产生具有稳定的中间分化表型的细胞系。

[0125] 以本方法为例,本发明人通过从培养基中除去 4-OHT 并在分化后再添加 4OHT,已经在体外由 ABM42 细胞(通过本发明的方法生产的 1t-HSC,见实施例)产生了 CD4⁺, $\alpha\beta$ + T 细胞。本发明人还通过在 GM-CSF、IL-4 和 FLT3L 中培养 ABM46 细胞(见实施例)、然后在分化后将培养物放回存在 4-OHT 的条件下而产生了树突细胞系。

[0126] 另一种生产所述细胞系的方法包括将 ct1t-HSC 细胞导入小鼠中,以允许在注射 4-OHT 之后在体内分化,并且阻滞或稳定表型。该方法在实施例 8 中详细描述。简言之,例如,将通过本方法产生的 1t-HSC 注射到无免疫应答的动物(例如,无免疫应答的小鼠)中。利用注射活化剂(例如,4-OHT)再活化 1t-HSC 中的癌基因,然后收集细胞,然后在体外培养细胞以使该细胞分化,然后根据需要保存或使用。这种方法,和如上所述的其他方法,可用于鼠和人 ct1t-HSC 细胞系,例如利用 NOD/SCID 小鼠作为受体,或新生 Rag-1^{-/-}小鼠,将对其进行肝内注射。

[0127] 本发明的方法对胚胎干细胞的应用

[0128] 本发明的另一实施方式涉及将条件无限增殖化干细胞的方法应用于胚胎干细胞(ES)。所述方法可用于产生更容易源自 ES 细胞的细胞系,例如神经元谱系细胞,包括神经元干细胞。

[0129] 在本实施方式中,本发明的方法,包括用原癌基因和抑制细胞凋亡的基因(例如,MYC-ER 和 Bcl-2)转导细胞(在此例中,ES 细胞),可应用于 ES 细胞以进一步控制这些细胞的定向分化。在此实施方式中,所述细胞可用于产生转基因小鼠,例如,并且另外,任何 ES 细胞和其衍生的相关祖细胞群体可通过暴露于活化剂而进行原癌基因的活化,因此可以产生新的条件转化的干细胞系(不同组织类型)或对于感兴趣的组织类型的成熟细胞系。另外,转导的 ES 细胞在体外的定向分化也可用于如上所述捕获分化的中间状态。以此方式使用 ES 细胞或来源于 ES 的细胞提供了一种在不同疾病背景下用于药物发现和目标鉴定的新平台。

[0130] 例如,神经元干细胞可应用于本发明的该实施方式中,以及利用本发明的方法使 ES 细胞定向分化为神经元途径中。来自于海马的神经元干细胞的分离和转导已经在以前对于小鼠进行了描述。神经球(neurosphere)的培养条件使这些细胞能够增殖,导致它们

对于本发明的病毒介导的基因（例如，MYC-ER 和 Bcl-2）转导是易感的，从而产生条件转化的神经元干细胞系。移植后它们在体外和在体内的分化可根据病毒编码的报告基因和先前确定的神经元分化的标记物进行监测。另外，在条件转化的神经干细胞移植后给小鼠施用活化剂（例如，4-OHT）可导致神经恶性肿瘤的发展（成神经细胞瘤、成胶质细胞瘤，等等）。这些肿瘤将提供一种临床前研究和靶标鉴定的新方法。

[0131] 已被例如 MYC-ER 和 Bcl-2 转导的 ES 细胞的定向分化可在先前确定的生长培养基以及细胞因子的存在下进行。在培养过程中任何时间添加活化剂（例如，4-OHT）将使细胞能够稳定在一种中间表型，并且导致仍然保持进一步分化能力的细胞系的产生。例如，通常通过添加视黄酸和 FGF8 来实现来自 ES 细胞的多巴胺能神经元的产生。这种类型的神经元对于修复在阿尔茨海默氏病患者中观察到的脑损伤是理想的。但是，完全分化的神经元细胞的移植可能阻碍它们成功的移植和植入。定向于多巴胺能神经元途径、但在移植后仍保持进一步分化能力的条件转化的细胞，如本文所预计的，可大大提高植入和成功移植的机会。对于来自 ES 细胞的运动神经元的产生，可建议一种类似的方案，即向培养物中添加视黄酸和 sbh 蛋白激动剂。这些神经元细胞可帮助修复脊髓损伤。再次，完全分化的细胞将不用于本实施方式中，而是使用保持分化能力的定向祖细胞（通过本发明方法产生）。

[0132] 用于去除转基因的条件无限增殖化方法的变化或修改

[0133] 在本发明的一个实施方式中，为了避免将包含例如此处描述的转基因（例如，MYC-ER）的干细胞引入人和 / 或小鼠中的风险，设计重组构建体，以使这些 DNA 片段可被切除。这一实施方式可利用任何适当的方法实现：首先，根据本发明方法建立长期干细胞，然后将细胞（或患者）暴露于重组 DNA 将被去除、切除或完全沉默化的条件下。

[0134] 例如，在本发明的一个方案中，使用一种细菌重组酶方法。在本发明的这一方案中，优选地使用两种不同的重组酶以允许控制两种基因中的哪一种在任何时间点被切除。所述重组酶的两个例子是 Cre 和 F1p 重组酶，它们是本领域公知的。简言之，将一种重组酶的识别底物序列（RSS' s）导入逆转录病毒构建体中，使它们位于癌基因以及报告基因（例如，GFP 或 Thy1.1）的开放阅读框的侧面。在此例子中，细胞在包含 Tat-Cre 融合蛋白（即，与 Cre 融合的 HIV 或其他逆转录病毒 Tat 蛋白）的培养基中培养。这种重组蛋白以前已经被描述并且显示能够被动进入细胞，并且介导基因组 DNA 的依赖于 loxP 位点的重组。其他的基因（核酸分子）切除方法也是本领域技术人员公知的，并且能够容易地应用于本发明中。实施例 5 和 13 说明了本发明的这一实施方式。

[0135] 在本发明的另一实施方式中，为提供另一种避免将包含例如此处描述的转基因的干细胞引入人和 / 或其他动物（例如，小鼠）中的风险的方法，而不是用原癌基因或抗细胞凋亡蛋白的重组构建体的组合转染干细胞，利用 Tat- 融合蛋白进行本发明，作为一种使蛋白质进入细胞内部而不必将转基因引入细胞中的方法。例如，编码 tat- 原癌基因或 tat- 抗细胞凋亡基因（例如，Tat-MYC-ER 或 Tat-Bcl-2）的重组构建体可用于使干细胞条件无限增殖化。在发明的本实施方式中，目标干细胞在适当的培养条件下、在包含纯化的由选择的特定基因组合（例如，MYC-ER 和 Bcl-2）编码的重组 Tat- 融合蛋白的培养基中培养。在本发明的本实施方式中，原癌基因的产物或类似基因的产物可以为可诱导的，如上面的实施方式一样。可选地，或另外，可以只通过提供或从培养物中去除蛋白质而调节蛋白质的作用。尽管用本方法产生的细胞系将继续依赖于外源 Tat- 融合蛋白的添加，但是它们中将不会

导入特定的外源核苷酸序列。预期不含外源癌基因序列可以改善本发明方法的临床使用。人免疫缺陷病毒-1 (HIV-1) Tat 是 Tat 蛋白的一个例子,其他逆转录病毒 Tat 蛋白是本领域公知的。作为一种非限制性的例子,编码 HIV-1 Tat 的核酸序列在此处表示为 SEQ ID NO :9,其编码此处表示为 SEQ ID NO :10 的氨基酸序列。

[0136] 在另一实施方式中,为提供另一种避免将包含例如此处描述的转基因的干细胞引入人和 / 或其他动物 (例如,小鼠) 中的风险的方法,而不是用原癌基因或抗细胞凋亡蛋白的重组构建体的组合转染干细胞,通过利用适体技术将蛋白质 (例如, MYC 和 Bcl-2) 引入细胞中而实施本发明。适体是根据它们以高亲和力和特异性结合预定的特定目标分子的能力而从随机化组合核酸文库中选择的合成核酸的短链 (通常是 RNA,也可以是 DNA)。适体具有一种确定的三维结构,并且能够区别在结构上具有极小差别的化合物。相应地,适体可与用于本发明的蛋白质或编码蛋白质的非整合性 cDNA 连接,例如,用于将蛋白质或 DNA 递送到细胞中。另外,适体可容易地用于将 siRNA 递送到细胞中,例如,当根据本发明破坏促凋亡蛋白时。适体技术例如在 Davidson, 2006, *Nature Biotechnol.* 24 (8) :951-952 ;和 McNamara 等人, 2006, *Nature Biotechnol.* 24 (8) :1005-1015 中描述。此外,预期不含外源癌基因序列可改善本发明方法的临床使用。

[0137] 在另一实施方式中,为提供另一种避免将包含例如此处所述的转基因的干细胞引入人和 / 或其他动物 (例如,小鼠) 中的风险的方法,而不是用原癌基因或抗细胞凋亡蛋白的重组构建体的组合转染干细胞,通过利用 CHARIOT™ 技术 (Krackeler Scientific, Inc., Albany, NY) 将原癌基因和 / 或抗细胞凋亡蛋白引入细胞内实施本发明。利用这种技术,在 CHARIOT™ 肽和目标蛋白质之间形成了非共价键。这防止了蛋白质降解并在转染过程中保持其自然性质。在递送到细胞中后,复合体解离并且 CHARIOT™ 被转运到细胞核中,而被递送的蛋白质具有生物活性并且自由进入它的细胞目标中。有效的递送可在存在或不存在血清的情况下发生,并且不依赖于内体途径,内体途径可在细胞内化过程中修饰大分子。该递送系统也绕开了转录-翻译过程。因此,本发明中有用的蛋白质可被递送到细胞中并且释放,以使细胞无限增殖化,而不需要将原癌基因或癌基因导入到细胞中。如上所述,预期不含外源致癌基因序列可以改善本发明方法的临床使用。

[0138] 作为通过在此描述的多种病毒途径控制原癌基因插入到宿主细胞基因组中的可能性并由此避免转化事件的另一种替代 (或另外的) 方法,可将药物敏感性 (药物易感性) 盒引入将要使用的病毒构建体中,使其将在每种被转导的细胞和其分化的子代中表达。药物敏感盒或药物易感盒是一种编码蛋白质的核酸序列,所述蛋白质使细胞对于特定药物的存在易感或敏感,以致在暴露于药物后,细胞活性被抑制并且优选地经历细胞凋亡。对于其中特定血细胞群体的水平在没有明显原因 (例如,感染、创伤、应激,等等) 的情况下增长的患者,可以给予一段时间的已引入对其敏感性的药物,以除去这些细胞并且缓解任何可能的附加并发症,所述并发症涉及其中遗传插入可能已经无意中引起致癌突变的细胞。相应地,作为一种非限制性例子,可以将编码 HPRT 的 cDNA 的盒引入本发明方法使用的构建体中,以使转导的细胞对于 6- 硫代鸟嘌呤易感。另一非限制性例子是来自单纯疱疹病毒家族成员 (HSV-TK) 的胸苷激酶 cDNA 的引入,以使转导的细胞对于相关抑制剂例如更昔洛韦、阿昔洛韦和任何相关衍生物易感。另外,任何其他这样的药物敏感性盒以及它们相关的激动剂在本文中都是有用的。

[0139] 根据本发明将核酸或蛋白质引入到细胞中的其他方法对于本领域技术人员来说是显而易见的。本发明优选那些最小化或消除了将重组 DNA 引入到宿主细胞基因组中的风险的方法,上面描述了许多这样的例子。

[0140] 本发明的条件无限增殖化细胞的使用方法

[0141] 本发明的另一实施方式包括通过本发明方法产生的任何干细胞群体,包括混合群体和克隆群体,以及本发明的干细胞在此处描述的任何方法中的应用,包括分化为期望的细胞类型,和任何移植、细胞替换、疾病治疗、遗传工程、药物发现、和此处描述的细胞发育和分化研究的方法。

[0142] 由于现在可以生产实际上无限供应的易于保存、恢复、扩增和操作的均质干细胞,所述干细胞可用作干细胞或分化为不同细胞谱系并在试验中用于检测不同化合物对细胞分化、基因表达和 / 细胞过程的影响。因此,本方法的一个实施方式涉及一种鉴定影响细胞分化、基因表达和 / 细胞过程的化合物的方法。该方法一般包括以下步骤:使通过本发明的方法生产的干细胞与测试化合物接触,并测定特定结果,特别是期望的结果,例如基因表达、生物学活性、细胞分化、细胞生长、细胞增殖等等(见下文),与不存在该化合物的结果进行比较,以确定该测试化合物对于干细胞是否具有期望的效果。该方法可用于测试细胞分化、细胞活性或基因表达的实际上任何方面。在一个方面,干细胞在与化合物接触之前被操作,例如,进行遗传操作。例如,为了鉴定治疗化合物(例如,癌症治疗剂)和评价基因取代治疗,可利用所述试验评价来自具有遗传缺陷的个体的干细胞。实际上,本发明的技术提供了针对特定个体的细胞以鉴定药物候选物和治疗候选物和“适合”个体细胞的策略的机会。而且,如上所述,所述试验还可用于鉴定适于在培养物中保持本发明的干细胞的其他生长因子或培养条件。所述试验的一个例子在实施例 7 中进行了详细的描述,尽管本发明不限于该试验。

[0143] 本发明的另一实施方式涉及一种研究细胞谱系定向和 / 或来自干细胞的细胞分化和发育的方法,其一般包括培养本发明的条件无限增殖化干细胞,和在不同条件下和在存在及不存在可影响细胞谱系定向或分化的化合物或试剂的情况下评价所述细胞的与细胞发育和分化有关的遗传和生物标记物。如上所述,在本发明之前,由于无法获得以及不能产生足够数量的期望的细胞群体来进行期望的试验,严重阻碍了这样的研究。例如,为了鉴定或筛选特定组细胞系分化的中间体,必须获得足够数量的细胞才能提供有意义的和可再现的结果。利用本发明之前可获得的技术,这是不可能的。但是,本发明通过提供可扩增的和基本上无限供应的可用于多种试验的均质干细胞解决了这一问题。这一技术将大大提高在细胞分化和发现领域的研究能力。在一个方案中,使本发明的条件无限增殖化干细胞扩增,然后将其一部分在不存在使细胞保持条件无限增殖化状态的条件(例如,根据此处所述的示例性的方法,不存在他莫昔芬)的情况下培养。可评价细胞在基因表达、细胞表面标记物、生物分子分泌或任何其他基因型或表型标记物方面的改变,以研究细胞分化和谱系定向的过程。可向培养物中加入生长因子或其他因子,例如,以驱使沿着特定细胞谱系途径分化,并且可以在存在或不存在这些因子的情况下评价细胞的变化。此外,可以利用细胞评价影响(调节)细胞分化和发育的培养条件、体内条件、因子和试剂。

[0144] 在本发明任何试验中检测细胞基因型或表型特征改变的多种方法是本领域已知的。可用于测定或检测基因序列或表达的方法的例子包括但不限于:多聚酶链反应(PCR)、

逆转录酶-PCR(RT-PCR)、原位PCR、定量PCR(q-PCR)、原位杂交、Southern印迹法、Northern印记法、序列分析、微阵列分析、报告基因的检测、或其他DNA/RNA杂交平台。测定蛋白质水平的方法包括但不限于:Western印迹法、免疫印迹法、酶联免疫吸附试验(ELISA)、放射免疫试验(RIA)、免疫沉淀法、表面等离子体共振、化学发光、荧光偏振、磷光、免疫组织化学分析、基质辅助激光解吸/电离飞行时间(MALDI-TOF)质谱法、微细胞分析法、微阵列、显微镜检查、荧光激活细胞分选(FACS)、流式细胞术、和基于蛋白质性质的试验,包括但不限于DNA结合、配体结合、与其他蛋白质配偶体相互作用、细胞信号传导、酶活性、和可溶性因子或蛋白质的分泌。

[0145] 在药物筛选试验中,术语“测试化合物”,“假定的抑制化合物”或“假定的调节化合物”是指在特定过程中具有未知的或先前未意识到的调节活性的化合物。同样,鉴定化合物的方法中的术语“鉴定”意为包括所有的化合物,其用于特定目的(例如,调节细胞分化)的化合物的可用性通过本发明的方法确定,优选在存在和不存在所述化合物的情况下确定。利用本发明的方法筛选的化合物包括已知的有机化合物,例如抗体、肽文库的产物、和化学组合文库的产物。化合物也可利用合理的药物设计来鉴定。所述方法对于本领域技术人员来说是已知的,并且可包括使用三维成像软件程序。例如,可用于设计或选择在本发明中有用的模拟物或其他治疗化合物的多种药物设计方法在Maulik等人,1997,Molecular Biotechnology:Therapeutic Applications and Strategies,Wiley-Liss,Inc.中公开,其全文在此引入作为参考。

[0146] 在任何上述的试验中,本发明的细胞、细胞裂解物、核酸分子或蛋白质暴露于或接触假定的调节化合物(例如通过混合)的条件,是任何适当的培养或测定条件,其可包括使用有效的培养基,在该培养基中可以培养细胞(例如,如上所述)或者可以在存在及不存在假定的调节化合物的情况下评价细胞裂解物。本发明的细胞可在多种容器中培养,包括但不限于,组织培养瓶、试管、微量滴定皿和培养皿。培养在适合细胞的温度、pH和二氧化碳含量下进行。所述培养条件也在本领域技术人员的能力范围内,特别适合培养本发明的无限增殖化干细胞的条件在本文别处进行了详细描述。细胞与假定调节化合物在一定条件下接触,所述条件考虑每个接触容器中的细胞数目、对细胞应用的假定调节化合物的浓度、假定调节化合物与细胞的孵育时间、和给细胞应用的化合物的浓度。有效方案的确定可以由本领域技术人员根据例如容器大小、容器中液体的体积、已知适于培养在试验中使用的特定细胞类型的条件、和所测试的假定调节化合物的化学组成(即,大小、电荷,等等)等变量来实现。

[0147] 在本发明的一个实施方式中,本发明的细胞和方法对于旨在评价来源于人脐带血的ctlt-HSC、CD34+细胞、或分离自外周血的成年CD34+细胞的多能性的方法是有用的。所述方法在实施例11中描述。

[0148] 本发明的另一种实施方式涉及ctlt-HSC细胞系作为产生急性髓细胞白血病(AML)的新模型的平台的应用。更特别地,本发明人已经利用本发明的ctlt-HSC产生了一种急性髓细胞白血病的小鼠模型。这是由类似于HSC的细胞(基于它们的表面标记物表达)组成的白血病。为产生促进小鼠中的白血病的ctlt-HSC,将 10^3 - 10^5 个ctlt-HSC与 10^5 个Rag-1^{-/-}全骨髓细胞一起转移入受到致死性照射的受体小鼠中。小鼠每周给予4-OHT以保持癌基因活性,并且监视与白血病相关的临床指征,这是本领域已知的。从这些动物中回收

肿瘤,并且可在缺乏 4-OHT 的培养物中增殖。这些细胞保持了它们的 HSC 样表型,表明它们的增殖、存活和阻滞的分化不再强烈依赖 MYC 的高活性。白血病细胞系在再次移植给受照射的受体小鼠后也引起疾病。这些工具为研究 AML 和相关疾病的生物学和探索新的治疗途径提供了新平台。而且,将 ctlt-HSC 细胞系引入到用 4-OHT 治疗的小鼠中将为治疗提供一种良好的内部阳性对照:撤除 4-OHT。体内肿瘤建立之后产生的次级细胞系也可用于理解 AML 的药物抗性形式的相关治疗目标。

[0149] 本发明的其他实施方式涉及通过本发明方法产生的干细胞以及由这些干细胞分化的细胞在多种治疗和健康相关方法中的应用。这些方法一般包括以下步骤:获得通过本发明方法产生的条件无限增殖化干细胞的群体、培养物或细胞系,去除使所述细胞条件无限增殖化的条件,然后在一种治疗方案中使用该细胞。例如,可将该细胞直接施用于需要该细胞的个体,或者可使细胞在体外分化为期望的细胞类型,然后施用于个体。另外,在去除使细胞无限增殖化的条件之前或紧接着之后,可在体外遗传修饰细胞,以表达或沉默化一种基因或多种基因,作为在受控环境下的基因治疗新方法。细胞然后可作为干细胞施用于个体,或者首先在体外分化为期望的细胞谱系。

[0150] 为获得干细胞,在一个实施方式中,从待治疗的个体获得干细胞,然后根据本发明的方法进行条件无限增殖化。这些细胞可被广泛地扩增、保存(例如,冷冻或冷藏保存),然后根据需要恢复并再次扩增、操作和/或重复使用。在另一实施方式中,通过获取以前保存的来自待治疗个体的条件无限增殖化干细胞的来源而获得干细胞。在另一实施方式中,从以前产生的一组人干细胞系中获得干细胞,按照目前用于鉴定“匹配”供体的标准,该组人干细胞系覆盖了该群体中高百分比的部分。在一个实施方式中,细胞获自新鲜的或冷藏保存的脐带血、可由 ES 细胞体外定向分化产生的造血祖细胞群体、从正常患者或 G-CSF 治疗的患者的外周血获得的 HSC, G-CSF 治疗的患者已经被诱导使其 lt-HSC 动员到外周循环中。干细胞的其他来源对于本领域技术人员是显而易见的。根据本文前面描述的方法培养细胞,并可在适当时间去除控制无限增殖化的条件。另外,在将细胞给予个体之前,可以操作细胞以切除负责条件无限增殖化的基因或构建体(即,原癌基因和/或抗细胞凋亡编码基因),或者如果利用可溶性融合蛋白在培养基中保持细胞,如以上对于 Tat- 融合蛋白所述,则可从培养物中逐渐或立即去除这些融合蛋白。

[0151] 因此,本发明包括向个体(可包括任何动物)中递送通过本发明的方法产生的干细胞(包括包含所述干细胞的组合物),或由这些细胞分化的细胞。由于这些方法中使用的干细胞是在体外产生的,即使干细胞最初是从患者中分离的,整个细胞施用过程也基本上是一种离体(ex vivo)施用方案。离体施用是指在患者之外进行部分调节步骤,这样产生来自个体的条件无限增殖化干细胞(除基本正常的干细胞之外,可包括产生遗传修饰的干细胞),并将所述细胞或由所述细胞分化的细胞返回到患者中。可通过任何适当的施用方法,将根据本发明产生的干细胞或由这些细胞分化的细胞返回到个体中,或施用于个体。所述施用可以是全身施用、粘膜施用和/或邻近目标部位的位置施用。优选的施用途径对于本领域技术人员来说是显而易见的,取决于所预防或治疗的疾病的类型或施用原因。优选的施用方法包括但不限于静脉内施用、腹膜内施用、肌肉内施用、结节内(intranodal)施用、冠状动脉内施用、动脉内施用(例如,施用于颈动脉中)、皮下施用、经皮递送、气管内施用、皮下施用、关节内施用、心室内施用、脊柱内施用、肺部施用、导管渗透、和直接注射入组织

中（例如，肝脏插管）。

[0152] 细胞可与载体或药学上可接受的赋形剂一起施用。载体典型地是提高治疗组合物在被治疗个体中的半衰期的化合物。适合的载体包括但不限于多聚控释制剂、生物可降解植入物、脂质体、油、酯和乙二醇。如此处所用的药学上可接受的赋形剂指任何适于将通过本发明的方法产生的细胞递送到适当的体内部位的物质。优选的药学上可接受赋形剂能够使细胞保持一种形式，在细胞到达体内目标组织或部位后，细胞能够以有益于个体的方式发挥作用。

[0153] 根据本发明，有效的施用方案包括适当的剂量参数和施用模式，其导致有效数量的功能细胞递送到患者中以提供对于患者的短期或长期的益处。有效的剂量参数可利用本领域中对于特定病症或疾病的方法标准来确定。所述方法包括，例如，确定存活率、副作用（即，毒性）和疾病的发展或退化。

[0154] 根据本发明的干细胞或其分化细胞的适当单剂量是在一段适当的时期内施用一次或多次时能够给患者提供有益数量的细胞的剂量。例如，本发明的干细胞的一种优选的单剂量是每个个体每次施用大约 0.5×10^4 到大约 5.5×10^8 、或从大约 0.5×10^5 到大约 5.5×10^7 、或从大约 0.5×10^6 到大约 5.5×10^{10} 个干细胞，从大约 1×10^8 到大约 5.5×10^{10} 的剂量是更优选的。本发明包括在 0.5×10^4 到大约 5.5×10^{10} 之间以 10^2 个细胞的增量增加的任何剂量。更高或更低的剂量是本领域技术人员已知的，取决于待施用的干细胞或分化细胞的类型，还取决于施用途径。本领域技术人员知道，给动物施用的剂量取决于疾病或病症的程度和个体患者对治疗的反应。因此，在本发明的范围内适当的剂量包括治疗特定疾病所需要的任何剂量。

[0155] 如此处所用的词语“防止疾病”指减轻疾病的症状、减少疾病的发生、和 / 或降低疾病的严重性。保护动物（个体、受试者）是指当施用于动物时，根据本发明生产的细胞防止疾病发生和 / 或治疗或缓解疾病症状、体征或病因的能力。同样，防止动物的疾病包括防止疾病发生（预防性治疗）和治疗患有某种疾病或经历疾病的起始症状的动物（治疗性治疗）。术语“疾病”指相对于哺乳动物的正常健康状态的任何偏离，包括出现疾病症状的状态，以及发生偏离（例如，感染、基因突变、遗传缺陷，等等）但尚未表现出症状的情况。

[0156] 如上所述，可以给个体施用本发明的干细胞以治疗或预防多种疾病。例如，本发明的干细胞系提供了可扩增的干细胞的独特来源，可用于多种移植和治疗策略，包括癌症的治疗，特别是，放射治疗的癌症。另外，多种免疫缺陷障碍和贫血障碍（例如，再生障碍性贫血或溶血性贫血）也将从该技术中大大受益，因为本发明提供了根据个体需要再生个体的造血细胞的能力。本发明的另一应用涉及持续可扩增和可更新的毛囊干细胞的产生，其用于，例如，烧伤患者的整形手术，用于任何经历化疗和 / 或放射治疗导致头发生长不可逆损失的患者，以及进行了任何影响颅骨的外科手术或在包括在被遗传型秃顶影响的个体中诱导头发生长的选择性处理中的患者。类似地，本发明干细胞对于皮肤的应用对于以下治疗是无价的：烧伤患者的创伤愈合和治疗，以及创伤和其他患者的整形外科手术，以及选择性外科手术，包括但不限于整容外科手术。所述细胞可另外进行遗传操作以矫正年轻和年老个体中的先天或后天遗传缺陷。本领域技术人员将理解，基于本公开内容，对多种其他干细胞群体使用本发明可以获得益处，所述细胞群体包括但不限于来源于肺、乳房和肠上皮的干细胞和来源于神经和心脏组织的干细胞，仅举这几个例子。

[0157] 另外,如上所述,本发明给个体提供了在个体一生中根据需要获得可扩增的自体干细胞和其分化细胞供应的独特机会。本方法产生的所述干细胞可以保存,并在个体一生中在需要时(例如,在个体发生癌症或免疫缺陷疾病时)作为治疗方案的一部分使用。

[0158] 通过以下方法,现在可以矫正遗传缺陷或将有益的基因修饰导入体细胞中:操作利用本发明方法条件无限增殖化并且扩增的从个体中获得的自体干细胞。然后可将干细胞再引入到它们的来源个体中。

[0159] 本发明的其他应用包括使用干细胞系修复由于 COPD、IPE、气肿、哮喘和吸烟而发生的肺损伤。另外,所述细胞可用于治疗心脏中的血管损伤,并且有助于致死性照射后的自身免疫疾病(例如,SLE、糖尿病、RA)。

[0160] 在本发明的方法中,可根据本发明方法产生的细胞和包含这些细胞的组合物施用于任何动物,包括脊椎动物纲、哺乳动物纲的任何成员,包括但不限于,灵长类、啮齿类、家畜和驯养宠物。优选的待治疗的哺乳动物是人。

[0161] 本发明的不同方面在下面的实施例和附图中进行了更详细的描述。但是,本发明不限于这些实施例和对发明的举例说明。

[0162] 实施例

[0163] 实施例 1

[0164] 下述实施例描述了使长期造血干细胞(Lt-HSC)可逆无限增殖化的方法的发展。

[0165] 在历史上低频率的相关细胞群体使得难以对造血谱系发展受损的分子基础进行阐明,这阻止了对信号传导和下游反应的生化分析。实际上,这是所有血细胞生成研究中的一个主要的限制性因素。另外,LT-HSC 的有限的可获得性也是治疗人类的许多类型癌症和几种免疫缺陷中的一个主要障碍。

[0166] 在克服该限制的努力中,本发明人发展了一种产生条件转化的细胞系的方法,所述细胞系代表早期造血干细胞祖细胞。最初的策略包括来自 5FU 治疗的年轻和免疫学衰老 3-83 小鼠的骨髓干细胞的逆转录病毒转导。本发明人使用具有编码 Bcl-2 和 GFP 及 MYC-ER 和 GFP 的插入的 pMSCV 双顺反子逆转录病毒载体 [Van Parijs, L., Y. Refaeli, A. K. Abbas, 和 D. Baltimore. (1999) Autoimmunity as a consequence of retrovirus-mediated expression of C-FLIP in lymphocytes. *Immunity*, 11, 763-70]。选择这些基因是因为本发明人知道 MYC 具有取代淋巴细胞中来源于细胞因子的存活和增殖信号的能力。通过限制靶细胞,发明人设想可以形成干细胞肿瘤。重要的是,在此情况下 MYC-ER 功能依赖于他莫昔芬,通过从动物或培养物中除去他莫昔芬使 MYC 功能和转化终止。在用 MYC-ER 转导的细胞中,产生融合蛋白,但其保持在细胞质中直到暴露于他莫昔芬。

[0167] 更具体地,用两种逆转录病毒(编码 MYC-ER 和 Bcl-2)转导来自 5FU 治疗的小鼠的干细胞群体,并将其转移入受到致死性照射(1200 拉德)的受体小鼠中。十天后,开始每周腹膜内注射 1mg/小鼠的在油中乳化的 4-羟泰米芬(4OHT),以活化 MYC 功能(图 1)。在四周内,年轻的(而不是年老的)转导干细胞的接受体发展出肿瘤。从骨髓、脾和淋巴结中收获肿瘤,并在体外在不添加细胞因子的情况下与他莫昔芬一起培养。这些细胞生长大约 10 天,然后停止生长,并且细胞最终死亡,发明人怀疑细胞正在分化并且认为这可能是由于细胞生长需要细胞因子。参见图 1,曲线代表了移植后死亡率的动力学和体内 MYC 功能的活化。小鼠均死于白血病。虽然 MYC 的过量表达可代替细胞因子-依赖的增殖和存活功能,

但看起来并不涉及来源于细胞因子的分化信号。

[0168] 当患病时,使小鼠安乐死,收获脾和淋巴结细胞并置于具有他莫昔芬和干细胞生长因子混合物(IL-6、IL-3和干细胞因子(SCF))的培养物中。平行地,通过流式细胞术分析细胞(图2)。参见图2,点图代表了在与IL-3、IL-6和SCF培养三天之后HSC的前向(FSC)和侧向(SSC)散射特性的流式细胞分析数据。这两个标准与细胞大小(FSC)和粒度(SSC)有关。这两个群体有类似的分布。柱状图代表了在每种细胞群体中表达的GFP水平。这反映了用编码MYC-ER和Bcl-2的cDNA的逆转录病毒进行体外逆转录病毒转导的效率。

[0169] 在所有情况中,离体GFP⁺细胞>90% Sca-1⁺并且谱系标记物阴性。培养几天后,细胞开始生长并且冷冻大约400个系用于后续研究。增殖之后,这些细胞保持EGFP的表达并且一致地为SCA1阳性和CD34、Flk2和谱系标记物阴性(图3)。在来源于年轻小鼠的标记物和来源于年老小鼠的标记物之间,标记物表达的唯一差别是年轻小鼠中c-kit的表达提高。不被理论所束缚,本发明人相信这可能是由于在分析标记物之前年老细胞系在c-kit配体中的培养时间较长(3个月相对于3周)。最终,发明人发现了这些细胞系可在冷冻后容易地恢复并保持它们最初的表型。重要的是,这些细胞系在表型上是同质的并且显示出提供小鼠中所有长期重建的LT-HSC的表型(Reya, T., Duncan, A. W., Ailles, L., Domen, J., Scherer, D. C, Willert, K., Hintz, L., Nusse, R., 和 Weissman, LL. (2003). A role for Wnt signaling in self-renewal of hematopoietic stem cells. Nature 423, 409-14)。

[0170] 最近,发明人融化了10个如上所述生产的来源于骨髓的细胞系,通过在细胞因子混合物和40HT中培养能够容易地恢复10个细胞系中的9个。发明人对这些肿瘤进行了表型确定,结果非常有希望。具体来说,基于前向和90°光散射,每个细胞系包含两种不同的细胞群体。这9个细胞系仅仅在这些群体的比例上有差异。这些群体中细胞大小较大的均是GFP发光和Sca1、内皮糖蛋白和ckit阳性,但为Flt3、B220、CD19和mIgM阴性。CD34也显示为阴性,尽管这需要证实(图3)。这种表型完全对应于长期再生多能性干细胞公开的特征(Reya等人,上文)。发明人在最近从白血病获得的细胞系上观察到相同的起始表型,所述白血病是从获自年轻供体小鼠的转导HSC发展而来的(图3)。

[0171] 为了测试这些细胞分化的能力,在分析表型标记物之前,使用他莫昔芬及不使用他莫昔芬并且在IL-3、IL-6和SCF的存在下培养代表性的细胞系以终止MYC-ER功能7天。如图4所示,显著比例的细胞获得B谱系标记物,包括B220(~12%)、CD19(~10%)和mIgM(~10%)。另外,发明人已经能够通过从培养物中去除40HT在体外产生下述谱系:CD4⁺ab T-细胞,骨髓细胞(Mac-1⁺), ter-119⁺红系祖细胞, NK1.1表达细胞, 嗜中性粒细胞(Gr-1⁺细胞)。进一步的试验将评价这些细胞产生其他谱系的能力,以及改变细胞因子方案对分化的影响。虽然没有进行比较,但本发明人预期,与年老动物相比,年轻动物的分化在B细胞生产中更有效。据本发明人所知,这是可在体外被诱导分化的条件无限增殖化造血干细胞的第一个例子。

[0172] 实施例2

[0173] 下面的实施例描述了向致死性照射的受体中过继转移LT-HSC细胞系的结果。

[0174] 如果实施例1所述的HSC细胞系是分析年老动物中缺陷B细胞淋巴细胞生成的基础的适当对象,它们应该在体内再现缺陷。发明人已经通过将LT-HSC细胞系过继转移入致死性照射的受体中开始解决这一问题。在开始的试验中,来自年老动物的细胞系(>60%

ID) 与 RAG2⁺骨髓一起转移,并且受体不用他莫昔芬处理以使 MYC-ER 沉默化。六周后收获受体的骨髓和脾细胞,并且分析 GFP⁺ 细胞(源于 HSC 系的 GFP 标记细胞)的恢复和表型(图 5)。

[0175] 在来自图 5 中表示的三个小鼠的数据中,一个小鼠接受衰老 HSC 系 ABM42,两只小鼠接受衰老 HSC 系 ABM46。根据转移的细胞系,淋巴散射门中 30 到 70% 的细胞是 GFP⁺。如图 5 所示,检测的两种细胞系(ABM46 和 ABM42)都产生 B(CD19⁺) 和 T(TCR⁺, CD4⁺, CD8⁺) 细胞、巨噬细胞(CD11b⁺) 和粒细胞(GR1⁺)。这些谱系的比例存在一定的受体间差异。但是,重要的是,两种测试细胞系产生成熟 CD4 和 CD8 单阳性 T 细胞(图 7),但 B 细胞发育没有超过祖细胞阶段(图 6)。虽然 B220⁺, CD19⁺ 细胞发育,但是它们没有发展到 mIg⁺ 阶段。这完全是根据实验结果预测的结果,所述实验包括利用来自免疫学老年小鼠的 BM HSC 的自重建和继承性重建(Johnson, S. A., S.J. Rozzo 和 J. C. Cambier, Aging-dependent exclusion of antigen-inexperienced cells from the peripheral B cell repertoire. J Immunol, 2002. 168(10): p. 5014-23)。换句话说,当来自免疫学老年小鼠的完整骨髓用于移植时,可观察到相同的发育阻滞。

[0176] 本发明人已经发现,这一系统可进一步采取一个步骤,成功地由原始 HSC 系的过继受体的骨髓重建 LT-HSC 系(数据未显示)。这只需通过在干细胞细胞因子加他莫昔芬中培养骨髓细胞以再活化 MYC 来完成。这些细胞正在生长并显示原始的表型。

[0177] 实施例 3

[0178] 下列实施例描述了一种利用完全在体外进行的方法使 HSC 可逆地无限增殖化的方法。

[0179] 除了此前描述的产生条件无限增殖化长期 HSC 细胞系的方法以外,本发明人已经能够完全在体外进行这一程序。上述方法依赖于将转导的 HSC 引入小鼠中,并且在体内诱导它们的转化。在体外进行这一程序的优点是本方法的每一方面都在受控的环境下进行。

[0180] 本方法首先包括用 5-氟尿嘧啶(5-FU) 处理供体小鼠以富集 HSC 并且诱导这些细胞增殖。收集来自小鼠胫骨和股骨的 5FU 富集的造血干细胞,然后以每孔 $1.8-2.0 \times 10^6$ 细胞的密度置于 24 孔组织培养板上的 DMEM 培养基中,该培养基包含 15% 的热灭活的胎牛血清和 IL-3、IL-6 和 SCF。对细胞进行三轮旋转感染(spin infection),以用编码 MYC-ER 和 Bcl-2 的逆转录病毒载体转导细胞。简言之,用 pMIG-MYC-ER 或 pMIT-Bcl2 转染细胞。收集包含病毒的上清液,并补充 4 μ g/ml 的 1,5-二甲基-1,5-二氮十一亚甲基聚甲溴化物(polybrene) 和 10mM 的 HEPES,并且通过 0.45 μ m 的滤器。两种不同的病毒上清液以 1:1 的比例混合并且添加到孔中。然后以 2000rpm 离心细胞 1 小时。在每次旋转感染的最后替换病毒上清液。最后一轮感染 24 小时之后,通过流式细胞分析确定转导水平以确定转导效率。转导的细胞然后在包含 IL-3、IL-6、SCF 和 10nM 4OHT 的 DMEM 培养基中孵育。培养基每 3 天更换一次,特别强调保证细胞因子和 4OHT 的新鲜供应。根据需要,细胞缓慢通过。

[0181] 利用这种体外方法,发明人已经能够产生具有下列基因组合的条件无限增殖化细胞系:MYC-ER 和 Bcl-2;MYC-ER 和 hTERT(人端粒酶的逆转录酶成分);ICN-1-ER(Notch-1 的细胞内部分的 ER-调节的活性成分)和 Bcl-2;ICN-1-ER 和 hTERT;以及 MYC-ER 和 ICN-1-ER。图 8-11 所示的数据显示大多数这些细胞系的起始特征。他们产生了由 c-kit⁺, Sca-1⁺, CD34⁻, flk2⁻ 细胞组成的细胞系,这是与正常长期造血干细胞表现的表型一致的表

型。图 8-11 所示的数据来自于对逆转录病毒编码的报告基因 (GFP 和 thy 1.1) 以及干细胞的四种标记物 :c-kit、sca-1、CD34 和 flk-2 的流式细胞术分析。图 8-11 所示的细胞系已经在测定表型之前培养了 5 周。至此这些细胞已经在连续培养中扩增并分开超过 35 天。

[0182] 参见图 8, 此图显示了源于从年轻 C57/BL6 小鼠中获得的 HSC 的细胞系的表型比较, 所述细胞系用 BCL-2 和 MYC-ER 逆转录病毒转导并且在体外连续培养中保持超过 90 天。显示的是典型克隆在 90 天持续培养后 3(年轻的)个月时的表型。

[0183] 参见图 9, 此图显示了源于从年轻 C57/BL6 小鼠中获得的 HSC 的细胞系的表型比较, 所述细胞系用不同的癌基因组合逆转录病毒转导并且在体外连续培养中保持超过 90 天。5FU 富集的 HSC 用 pMIG-MYC 和 pMIT-Bcl-2(上图)、pMIG-MYC. ER 和 pMIG-hTERT(中图)或 pMIG-ICN. 1. ER 和 pMIT-Bcl-2 逆转录病毒转导。细胞保持在补充了 15% 的胎牛血清和 IL-6、IL-3 和 SCF 混合物的 DMEM 中。显示的是典型克隆在 90 天连续培养后 3(年轻的)个月时的表型。该图显示了对于病毒表达标记物 (GFP 和 Thy1.1) 以及四种定义小鼠的长期 HSC 所需的标记物 (Sca-1、c-kit、CD34 和 Flk-2) 的表达的流式细胞术分析结果。这四种细胞系包含保持 1t-HSC 的表型的亚群 (Sca-1⁺、c-kit⁺、CD34⁻、flk-2⁻)。

[0184] 参见图 10, 此图显示了源于从年轻 C57/BL6 小鼠中获得的 HSC 的细胞系的表型比较, 所述细胞系用癌基因的不同组合逆转录病毒转导并且在体外连续培养中保持超过 90 天。5FU 富集的 HSC 用 pMIG-ICN. 1. ER 和 pMIT-Bcl-2(上图)、pMIG-ICN. 1 和 pMIT-Bcl-2(第二行的图)、或 pMIG-ICN. 1 和 pMIG-Bcl-2(第三行的图)、或 pMIG-hTERT 和 pMIT-Bcl-2(下图) 逆转录病毒转导。这些细胞保持在补充了 15% 的胎牛血清和 IL-6、IL-3 和 SCF 混合物的 DMEM 中。显示的是典型克隆在 90 天连续培养后 3(年轻的)个月时的表型。该图显示了对于病毒表达标记物 (GFP 和 Thy1.1) 以及四种定义小鼠的长期 HSC 所需的标记物 (Sca-1、c-kit、CD34 和 Flk-2) 的表达的流式细胞术分析结果。这四种细胞系包含保持 1t-HSC 的表型的亚群 (Sca-1⁺、c-kit⁺、CD34⁻、flk-2⁻)。

[0185] 参见图 11, 此图显示了源于从年轻 C57/BL6 小鼠中获得的 HSC 的细胞系的表型比较, 所述细胞系用癌基因的不同组合逆转录病毒转导并且在体外连续培养中保持超过 90 天。5FU 富集的 HSC 用 pMIG-MYC 和 pMIG-ICN. 1(上图)、pMIG-MYC. ER 和 pMIG-ICN. 1(中图)或 pMIG-ICN. 1. ER 和 pMIG-MYC 逆转录病毒转导。这些细胞保持在补充了 15% 的胎牛血清和 IL-6、IL-3 和 SCF 混合物的 DMEM 中。显示的是典型克隆在 90 天连续培养后 3(年轻的)个月时的表型。该图显示了对于病毒表达标记物 (GFP 和 Thy1.1) 以及四种定义小鼠的长期 HSC 所需的标记物 (Sca-1、c-kit、CD34 和 Flk-2) 的表达的流式细胞术分析结果。这四种细胞系包含保持 1t-HSC 的表型的亚群 (Sca-1⁺、c-kit⁺、CD34⁻、flk-2⁻)。

[0186] 这些细胞系也被用于体内重建细胞区室。参见图 12, 此图显示了由源于从年轻 C57/BL6 小鼠中获得的 HSC 的细胞系在体内重建 T 细胞和 B 细胞区室的结果, 所述细胞系用癌基因的不同组合逆转录病毒转导并且在体外连续培养中保持超过 90 天。简言之, 5FU 富集的 HSC 用 pMIG-ICN. 1-ER 和 pMIG-hTERT(上图)、pMIG-MYC. ER 和 pMIG-hTERT(中图)或 pMIG-MYC-ER 和 pMIT-Bcl-2(下图) 逆转录病毒转导。这些细胞系保持在补充了 15% 的胎牛血清和 IL-6、IL-3 和 SCF 混合物的 DMEM 中。利用来自 Rag2^{-/-} 小鼠的骨髓干细胞和体外产生的 LT-HSC 系重建受到致死性照射的年轻 C57/BL16 小鼠。六周后, 收获骨髓, 并用一组特定谱系标记物染色。成熟 CD4 和 B220 阳性 /GFP 阳性细胞的发育可容易地观察到。来

自四个代表性小鼠的数据显示在此图中。在每个组中,大约 30% 的小鼠保持 GFP 标记物。

[0187] 实施例 4

[0188] 下面的实施例描述了用于使来源于人脐带血和骨髓的 HSC 在体外可逆无限增殖化的方法的延伸。

[0189] 本技术的另外一个应用是通过它们的条件无限增殖化能够在体外扩增人长期造血干细胞。发明人因此将前面实施例描述的体外方法作了一些改变用于人细胞。首先,逆转录病毒优选地用双嗜性被膜包装,以保证人细胞的有效转导。另外,细胞源是按照发明人机构的审查委员会 (Institutional Review Boards) 的所有规定和制度,从脐带血库匿名获得的人脐带血。产生的细胞将表达报告基因,其最终可用于通过高速细胞分选法分离纯的群体。发明人已经注意到许多由鼠 1t-HSC 细胞系产生的成熟细胞丧失了表面标记物的表达,这可能是由于谱系确定和分化后逆转录病毒基因组的甲基化作用。发明人期望在人细胞中看见类似的表现,在这种情况下,可根据这些细胞中报告基因的存在,结合细胞群体的细胞表面标记物,来监测 1t-HSC 及其在移植受体中的分布。

[0190] 实施例 5

[0191] 下面的实施例描述了一种从条件无限增殖化 HSC 细胞中连续切除编码 MYC-ER 和 Bcl-2 的 DNA 片段的方法。

[0192] 为了避免将包含编码 MYC, ER 和 Bcl-2 的转基因的 HSC 引入人和 / 或小鼠中的风险,将利用细菌重组酶方法切除这两个 DNA 片段。为了允许控制两个基因中的哪一个在任何时间点切除,使用两种不同的重组酶。所述重组酶的两个例子是 Cre 和 F1p 重组酶。简言之,将一种重组酶的识别底物序列 (RSS) 导入逆转录病毒构建体中,以使它们位于癌基因的开放阅读框的侧面,并且导入报告基因 (GFP 或 Thy1. 1)。在此例中,细胞在包含 Tat-Cre 融合蛋白的培养基中孵育。重组蛋白先前已被描述并且显示能够被动进入细胞,并介导基因组 DNA 的依赖于 loxP 位点的重组。

[0193] 这种方法将允许获得大量能产生许多 HSC 用于体外和体内分化的条件。首先,给细胞逐渐停止它们在条件转化过程中已经适应的高水平的增殖和存活信号。第二,使细胞重新适应依赖正常细胞因子行使其自身稳定功能和分化。第三,报告基因表达的连续丧失将允许确定每个被研究基因的缺失状态和程度。相应地,表达两种报告基因 (GFP 和 Thy1. 1) 的细胞包含两种序列 (分别为 MYC 和 Bcl-2),表达 Thy1. 1 而不表达 GFP 的细胞已经成功缺失了 MYC 编码序列,但仍含有 Bcl-2 基因,最后,不表达 GFP 或 Thy1. 1 的细胞已经缺失了这两种等位基因。图 13 以示意表的形式显示了这种方法。

[0194] 另外,通过从表达人 MYC 转基因的小鼠株获得 5FU 富集的 BM-HSC 在小鼠中测试本方法,所述 MYC 转基因可通过去除四环素和含有被称为 tTA (四环素反式作用蛋白) 的细菌蛋白质而诱导。人 MYC cDNA 被克隆到四环素调节转录元件 (TRE) 的下游。TRE-MYC 小鼠用 5FU 处理,并用于收获 BM-HSC。这些细胞在体外用表达 Bcl-2 和 tTA 的逆转录病毒 (pMIT-Bcl2 和 pMIG-tTA) 转导。细胞在多西环素的连续存在下培养,以使 MYC 转基因保持沉默化。一旦对细胞进行流式细胞术分析,它们可移植回小鼠中,所述小鼠将不给予包含多西环素的饮食 (这是通常在体内使用的四环素的一种更稳定的形式)。

[0195] 一旦产生 1t-HSC 细胞系,则检测在多西环素存在下体外培养它们的效果,平行地用在不含 4OHT 的情况下培养的包含 MYC, ER 的细胞系进行试验。用 western 印迹法和细胞

内染色法全程监测 MYC 的蛋白质水平。

[0196] 实施例 6

[0197] 下面的实施例描述了从液体组织培养基中去除 40HT 后许多造血谱系的体外产生。

[0198] 传统的用于确定 HSC 的潜能的方法涉及使用具有确定细胞因子的半固体培养基（甲基纤维素）来增强 HSC 向特定谱系的分化。发明人对于确定利用本发明方法在体外产生的细胞群体的多能性感兴趣。为了检查这种组织，将此处描述的 ABM42 和 ABM46 细胞系保持在包含 IL-3、IL-6 和 SCF、但不含 40HT 的培养基中。除了发明人能够在重建小鼠中检测的谱系（即，淋巴、骨髓和粒细胞谱系）之外，也可检测表达 NK1.1 或 ter-119 的 GFP⁺ 细胞（图 14）。NK1.1 细胞可以是 NK- 细胞，或 NK-T 细胞。Ter-119 表达细胞属于红细胞谱系。这些发现显示了这些细胞系能够产生正常造血系统的所有成分并且这些细胞对于产生大量用于被动治疗的特定成分是有用的。另外，它们对于研究血细胞生成中的早期事件和鉴定用于治疗性干预遗传病症或引起正常化疗的并发症或甚至是感染疾病的新疗法具有极大的应用和重要性。

[0199] 实施例 7

[0200] 下面的实施例描述了一种高通量筛选诱导或抑制条件转化的长期 HSC 分化的小分子或生物剂的方法。

[0201] 下面是一种筛选诱导或抑制 HSC 分化的小分子或生物剂的一般方法。以前，无法获得大量的干细胞这一事实阻碍了这些类型的大规模筛选。由于本发明人目前具有使长期 HSC 无限增殖化的能力，现在可以进行这样的技术了。

[0202] 例如，一种这样的方法是已由 Schneider 等人修改的骨髓分化读出 (read-out) (Schneider, T. 和 Issekutz, A. C. (1996) Quantitation of eosinophil and neutrophil infiltration into rat lung by basic assays for eosinophil peroxidase and myeloperoxidase. Application in a Brown Norway rat model of allergic pulmonary inflammation. J Immunol Methods 198, 1-14)。简言之，将条件转化的长期 HSC 以不同浓度的细胞数（通常为 2×10^4 – 5×10^4 细胞 / 孔）置于 96 孔平底板中。在添加 40HT 以使细胞保持未分化状态或不添加 40HT 以诱导分化的完全培养基 (DMEM+15% 热灭活的胎牛血清, $1 \times$ 青霉素 / 链霉素, $1 \times$ L- 谷氨酰胺和 $1 \times$ 非必需氨基酸, 补充了 IL-3、IL-6 和 SCF) 中进行筛选。这些条件已经显示产生 Mac-1⁺ 细胞，与骨髓分化模式一致。可以添加额外的细胞因子以引导沿特定途径分化，尽管该系统也可用于筛选一组细胞因子的特定功能。在此例中，加入没有补充 IL-3、IL-6 和 SCF、而是补充了将要测试或用于引导分化的给定细胞因子（例如，CSF-1、G-CSF、GM-CSF、EPO、TEPO，等等）的完全培养基。

[0203] 如果需要并且根据待测试的试剂或分子而确定，用 96 孔板滴定小分子、生物剂或阳性对照物质（例如，三氧化二砷），并与 1t-HSC 孵育 24 到 72 小时或更长时间。孵育后，细胞用 PBS 清洗并且重悬浮于 PBS 中在 -80°C 下过夜储存以裂解细胞。然后在室温下融化细胞，将培养板在 3,000rpm 下离心 10 分钟。然后将上清液转移入新的 96 孔板中并与四甲基联苯胺 (TMB) 混合 40 分钟。反应用 4N H_2SO_4 终止，在 450nm 下读取 O.D.。这种类型的高通量试验可用于测试小分子或生物剂诱导或阻止条件转化的长期 HSC 分化为广泛种类的细胞类型的能力。这些筛选的结果然后可进一步测试诱导或抑制 HSC 体内分化的能力。这

种测试形式的变化对于本领域技术人员是显而易见的,并包括在本发明中。

[0204] 实施例 8

[0205] 下面的实施例描述了本发明的方法产生中间造血谱系细胞系的用途。

[0206] 下面的方案可用于在将条件无限增殖化 1t-HSC 细胞系移植入致死性照射的小鼠中后诱导代表造血谱系发育中间阶段的细胞系的发育。首先,将根据本发明方法产生的 10^3 - 10^5 条件转化的 1t-HSC 细胞系转移入几组受到致死性照射的受体小鼠中。移植物还可包括 10^5 Rag-1^{-/-} 作为载体,以确保受照射小鼠在开始时存活。给小鼠每周腹膜内注射 1mg 他莫昔芬,以使来源于条件转化的 1t-HSC 细胞系的部分分化的细胞无限增殖化。注射在开始移植 3 天至 1 周后开始或在移植后 8 天开始,直到小鼠已被条件转化的 1t-HSC 细胞系完全重建。用他莫昔芬治疗 3 天后,或当它们显示出与白血病相关的临床症状时,从来自小鼠的脾和骨髓细胞中收集细胞。细胞在具有 4-OHT 的标准骨髓培养条件 (DMEM, 15% 胎牛血清,青霉素 / 链霉素, L- 谷氨酰胺, 非必需氨基酸, IL-3、IL-6 和 SCF) 下培养,或在其他细胞因子和用于其他造血细胞类型的介质的存在下培养。将细胞系冷冻和 / 或扩增,细胞系也是通过有限稀释克隆的并且通过原病毒整合的 PCR 扩增确定的单细胞,冷冻,然后通过流式细胞术表征其表面标记物的表达。这些类型的方法用于鼠和人 ct1t-HSC 细胞系,利用 NOD/SCID 小鼠作为受体,或使用新生 Rag-1^{-/-} 小鼠,将对其进行肝内注射。

[0207] 实施例 9

[0208] 下述实施例描述了使用本发明的方法和采用体外产生成熟 D4+ α β T 细胞的方案发展代表 T 细胞发育中间阶段的细胞系。

[0209] 在本试验中,将根据本发明方法产生的条件无限增殖化的 1t-HSC 细胞系在正常细胞因子混合物的存在下接种于培养板,其中补充了 IL-7 并且不含他莫昔芬。在表达 jagged (一种 Notch-1 配体) 的 OP-9 基质细胞层上建立平行培养。在为了监测 T 细胞发育的指征而开始培养之后,每 48 小时对 T 细胞谱系标记物进行细胞染色。将显示 T- 谱系定向和发育指征的孔转换为包含他莫昔芬的培养基,以稳定表型和建立细胞系。如实施例 8 所述将得到的细胞系扩增、克隆和表征。为了确定它们的 T 细胞受体所有成分的使用,将 T 细胞系针对各个 TCR-V β 等位基因进行特异性染色。适当时,将一些成熟 T- 细胞系或代表祖细胞群体的细胞系移植入 Rag-1^{-/-} 小鼠中,以评价它们遵循体内正常耐受和自身稳定的机制的能力,以及它们进一步体内分化的能力。最后,在体外和体内评价它们对抗原刺激应答的能力。

[0210] 实施例 10

[0211] 下面的实施例描述了使用本发明方法和采用用于使 HSC 定向分化为骨髓细胞谱系的方案发展中间发育细胞系和髓细胞白血病模型。

[0212] 在本试验中,将根据本发明方法产生的条件无限增殖化的 1t-HSC 细胞系在正常细胞因子混合物的存在下接种于培养板,其中补充了 G-CSF 并且不含他莫昔芬。在为了监测骨髓发育的指征而开始培养之后,每 48 小时对骨髓谱系标记物进行细胞染色。将显示骨髓谱系定向和发育指征的孔转换为包含他莫昔芬的培养基,以稳定表型和建立细胞系。如实施例 8 所述将得到的细胞系扩增、克隆和表征。将一些得到的细胞系移植回小鼠中以监测它们再生 Op/Op 小鼠 (天然缺乏巨噬细胞的突变小鼠) 的能力。还将这些细胞系移植入将始终保持在他莫昔芬条件下的野生型小鼠中,以确定这些细胞系是否还将产生类似于人

AML、CML 和 APL 的髓细胞白血病。这些新的肿瘤提供了用于临床前治疗的新模型。

[0213] 实施例 11

[0214] 下面的实施例描述了人成年 ctlt-HSC 细胞系的产生以及利用 NOD/SCID 或 RAG^{-/-} 异种移植模型在体内检测它们的多能性。

[0215] 在此试验中,用编码 MYC-ER、Bcl-2 和 GFP(用于以后检测移植的细胞)的逆转录病毒载体在体外转导 CD34⁺ 细胞(来自被动员的血或脐带血),所述病毒载体利用双嗜性被膜包裹(根据本发明的方法)。如上面利用鼠 HSC 所述,通过在 40HT 和生长因子存在下体外增殖来选择 lt-HSC。所选细胞的多能性通过以下方法评价:将 lt-HSC 系移植入亚致死性照射的 NOD/SCID 或 NOD/SCID/ β -2M^{-/-} 或 Rag-1^{-/-} 或 Rag-2^{-/-} 小鼠中,在 6-12 周后通过免疫荧光流式细胞术分析所有血细胞谱系。更特别地,在利用本发明方法产生 ctlt-HSC 细胞系之后,可利用两种不同的和补充的方法检查它们的多能性。在第一种方法中,将不同数量的克隆 ctlt-HSC 细胞系引入亚致死性照射的 NOD/SCID 小鼠或 NOD/SCID/B-2M^{-/-} 小鼠中。在此例中,在小鼠实施亚致死照射方案(0.3Gy)之后,静脉内转移来源于人 ctlt-HSC 细胞系的 10³-10⁵ 细胞。在移植 12 周之后分析小鼠的重建。第二,通过直接注射将来源于人 ctlt-HSC 细胞系的 10³-10⁵ 细胞引入新生 Rag-1^{-/-} 或 Rag-2^{-/-} 小鼠的肝中。还在移植 6-12 周后分析这些异种移植物的适当的重建。

[0216] 实施例 12

[0217] 下面的实施例描述了在移植后利用条件法消除 ctlt-HSC 的 MYC 和 Bcl-2 表达。

[0218] 在此试验中,重新改造用于转化干细胞的病毒(病毒载体),使其在 MYC-ER、Bcl-2 和 GFP 开放阅读框(ORFS)侧翼包含两个 loxp 位点。当细胞被移植时,用 Cre 或 CRE-TAT 融合蛋白的调节形式删除癌基因编码序列,从而消除受体中由插入片段引起恶性肿瘤的风险。该方法首先在小鼠中发展,然后应用于人 lt-HSC。

[0219] 在第二种方法中,利用来自 TRE-MYC 小鼠的 lt-HSC 产生具有编码 Bcl-2 或 rtTA 的逆转录病毒的细胞系。将其移植入小鼠中。给小鼠饲以多西环素,以消除 MYC 的表达。可以使用从 TRE-MYCxTRE-Bcl-2 双基因(bigenic)小鼠获得的 lt-HSC,其可利用 pMIG-rtTA 逆转录病毒转导以消除 MYC 和 Bcl-2 的表达。

[0220] 实施例 13

[0221] 下面的实施例描述了利用 HIV-1Tat 与 MYC 和/或 Bcl-2 的融合蛋白在不对 lt-HSC 进行遗传修饰的情况下获得条件转化。

[0222] 利用已经建立的方案产生和纯化 MYC-Tat 和 Bcl-2 融合蛋白。通过处理细胞检测融合蛋白,其中可容易地测定过量表达的 Bcl-2(例如,活化的 T 细胞, B- 细胞淋巴瘤细胞系,其具有对 BCMA-Fc 的抗性,等等)或 MYC(例如,无反应性 B- 细胞,幼稚 T 细胞,活化的 T 细胞)的效果。利用 MYC-Tat 和 Bcl-2-Tat 蛋白的组合使 lt-HSC 在移植之前增殖。这种方法在小鼠系统中容易地发展并测试,然后应用于人。

[0223] 美国临时专利申请 NO. 60/728, 131 和美国临时专利申请 No. 60/765, 993 的全部公开内容在此引入作为参考。

[0224] 尽管本发明的多个实施方式已经详细描述,但是这些实施方式的修饰和改变对于本领域技术人员来说是显而易见的。然而,应该清楚的理解,所述修饰和改变在本发明的范围内,如下面权利要求中阐明的那样。

序列表

<110> 国家犹太医学研究中心
科罗拉多大学董事会

<120> 条件无限增殖化长期干细胞和制备和使用所述细胞的方法

<130>2879-117PCT

<150>60/728, 131

<151>2005-10-18

<150>60/765, 993

<151>2006-0206

<160>14

<170>PatentIn version 3.3

<210>1

<211>2377

<212>DNA

<213> 人

<400>1

acccccgagc tgtgctgctc gcggccgcca ccgccgggcc ccggccgtcc ctggctcccc	60
tcctgcctcg agaagggcag ggcttctcag aggcttggcg ggaaaaagaa cggagggagg	120
gategcgctg agtataaaag ccggttttcg gggctttatc taactcgctg tagtaattcc	180
agcgagaggc agagggagcg agcgggcggc cggctagggt ggaagagccg ggcgagcaga	240
gctgcgctgc gggcgctcctg ggaagggaga tccggagcga atagggggct tcgcctctgg	300
cccagccctc ccgctgatcc cccagccagc ggtccgcaac ccttgccgca tccacgaaac	360
tttgcccata gcagcgggcg ggcactttgc actggaactt acaacacccg agcaaggacg	420
cgactctccc gacgcgggga ggctattctg cccatttggg gacacttccc cgccgctgcc	480
aggacceget tctctgaaag gctctccttg cagctgctta gacgctggat ttttttcggg	540
tagtggaana ccagcagcct cccgcgacga tgccctcaa cgtagcttc accaacagga	600
actatgacct cgactacgac tcgggtgcagc cgtatttcta ctgcgacgag gaggagaact	660
tetaccagca gcagcagcag agcgagctgc agccccggc gccagcgag gatatctgga	720
agaaattcga gctgctgccc accccgcccc tgtcccctag ccgccgtcc gggctctgct	780
cgccctccta cgttgcggtc acacccttct cccttcgggg agacaacgac ggcggtggcg	840

ggagcttctc	cacggccgac	cagctggaga	tggtgaccga	gctgctggga	ggagacatgg	900
tgaaccagag	tttcatctgc	gacccggacg	acgagacctt	catcaaaaac	atcatcatcc	960
aggactgtat	gtggagecgc	ttctcggccg	ccgccaagct	cgtctcagag	aagctggcct	1020
cctaccagge	tgcgcgcaaa	gacageggca	gcccgaaccc	cgcccgcggc	cacagcgtct	1080
gctccacctc	cagcttgtac	ctgcaggatc	tgagcgccgc	cgccctcagag	tgcctcgacc	1140
cctcgggtgt	cttcccctac	cctctcaacg	acagcagctc	gcccgaagtc	tgcgcctcgc	1200
aagactccag	cgccttctct	ccgtcctcgg	attctctgct	ctcctcgacg	gagtcctccc	1260
cgcagggcag	ccccgagccc	ctgggtgctc	atgaggagac	accgcccacc	accagcagcg	1320
actctgagga	ggaacaagaa	gatgaggaag	aaatcgatgt	tgtttctgtg	gaaaagaggc	1380
aggctcctgg	caaaaggtea	gagtctggat	caccttctgc	tgagggccac	agcaaacctc	1440
ctcacagccc	actggctctc	aagagggtgc	acgtctccac	acatcagcac	aactacgcag	1500
cgcctccctc	cactcggaag	gactatcctg	ctgccaagag	ggtcaagttg	gacagtgtca	1560
gagtcctgag	acagatcagc	aacaaccgaa	aatgcaccag	ccccaggtec	tcggacaccg	1620
aggagaatgt	caagaggcga	acacacaacg	tcttgagcgc	ccagaggagg	aacgagctaa	1680
aacggagctt	ttttgccttg	cgtgaccaga	tcccggagtt	ggaaaacaat	gaaaaggccc	1740
ccaaggtagt	tatccttaaa	aaagccacag	catacatcct	gtccgtccaa	gcagaggagc	1800
aaaagctcat	ttctgaagag	gacttggtgc	ggaaacgacg	agaacagttg	aaacacaaac	1860
ttgaacagct	acggaactct	tgtgcgtaag	gaaaagtaag	gaaaacgatt	ccttctaaca	1920
gaaatgtcct	gagcaatcac	ctatgaactt	gtttcaaatg	catgatcaaa	tgcaacctca	1980
caaccttggc	tgagtcttga	gactgaaaga	tttagccata	atgtaaactg	cctcaaattg	2040
gactttgggc	ataaaagaac	ttttttatgc	ttaccatctt	ttttttttct	ttaacagatt	2100
tgtattttaag	aattgttttt	aaaaaatttt	aagattttaca	caatgtttct	ctgtaaatat	2160
tgccattaaa	tgtaaataac	tttaataaaa	cgtttatagc	agttacacag	aatttcaatc	2220
ctagtatata	gtacctagta	ttataggtac	tataaacctt	aatttttttt	atttaagtac	2280
attttgcctt	ttaaagttag	tttttttcta	ttgttttttag	aaaaaataaa	ataactggca	2340
aatatatcat	tgagccaaaa	aaaaaaaaaa	aaaaaaa			2377

<210>2

<211>454

<212>PRT

<213> 人

<400>2

Met Asp Phe Phe Arg Val Val Glu Asn Gln Gln Pro Pro Ala Thr Met

1 5 10 15

Pro Leu Asn Val Ser Phe Thr Asn Arg Asn Tyr Asp Leu Asp Tyr Asp

20 25 30

Ser yal Gln Pro Tyr Phe Tyr Cys Asp Glu Glu Glu Asn Phe Tyr Gln

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

Asn Asn Arg Lys Cys Thr Ser Pro Arg Ser Ser Asp Thr Glu Glu Asn
 355 360 365
 Val Lys Arg ArgThr Hi s Asn Val Leu Glu Arg Gln Arg Arg Asn Glu
 370 375 380
 Leu Lys Arg Ser Phe Phe Ala Leu Arg Asp Gln Ile Pro Glu Leu Glu
 385 390 395 400
 Asn Asn Glu Lys Ala Pro Lys Val Val Ile Leu Lys Lys Ala Thr Ala
 405 410 415
 Tyr Ile Leu Ser Val Gln Ala Glu Glu Gln Lys Leu Ile Ser Glu Glu
 420 425 430
 Asp Leu Leu Arg Lys Arg Arg Glu Gln Leu Lys His Lys Leu Glu Gln
 435 440 445
 Leu Arg Asn Ser Cys Ala
 450

<210>3

<211>51552

<212>DNA

<213> 人

<220>

<221>misc_feature

<222>(31450).. (31450)

<223>n 是 a, c, g, 或 t

<400>3

acttgagccc aagagttcaa ggctacggtg agccatgatt gcaacaccac acgccagcct	60
tggtgacaga atgagaccct gtctcaaaaa aaaaaaaaaa aattgaaata atataaagca	120
tcttctctgg ccacagtgga acaaaaccag aaatcaacaa caagaggaat tttgaaaact	180
atacaaacac atgaaaatta aacaatatac ttctgaatga ccagtgagtc aatgaagaaa	240
ttaaaaagga aattgaaaaa tttattttaag caaatgataa cggaacata acctctcaaa	300
accacaggta tacagcaaaa gcagtgctaa gaaggaagtt tatagctata agcagctaca	360
tcaaaaaagt agaaaagcca ggcgcagtgg ctcatgcctg taatcccagc actttgggag	420
gccaaaggcg gcagatcgcc tgaggtcagg agttcgagac cagcctgacc aacacagaga	480
aaccttgctg ctactaaaaa taaaaatta gctgggcatg gtggcacatg cctgtaatcc	540
cagctactcg ggaggctgag gcaggataac cgcttgaacc caggaggtgg aggttgcggt	600
gagccgggat tgcgccattg gactccagcc tgggtaacaa gaggtaaacc ctgtctcaag	660
aaaaaaaaaa aagtagaaaa acttaaaaat acaacctaat gatgcacctt aaagaactag	720
aaaagcaaga gcaaaactaaa cctaaaattg gtaaaagaaa agaaataata aagatcagag	780

cagaaataaa	tgaaactgaa	agataacaat	acaaaagatc	aacaaaatta	aaagttggtt	840
ttttgaaaag	ataaacaaaa	ttgacaaacc	tttggccaga	ctaagaaaaa	aggaaagaag	900
acctaataaa	ataaagtcag	agatgaaaaa	agagacatta	caactgatac	cacagaaatt	960
caaaggatca	ctagaggcta	ctatgagcaa	ctgtacacta	ataaattgaa	aaacctagaa	1020
aaaatagata	aattcctaga	tgcatacaac	ctaccaagat	tgaacatga	agaaatccaa	1080
agcccaaaca	gaccaataac	aataatggga	ttaaagccat	aataaaaaag	ctcctagcaa	1140
agagaagccc	aggacceaat	ggcttccttg	ctggatttta	ccaatcattt	aaagaagaat	1200
gaattccaat	cctactcaaa	ctattctgaa	aaatagagga	aagaatactt	ccaaactcat	1260
tctacatggc	cagtattacc	ctgattccaa	aaccagacaa	aaacacatca	aaaacaaaca	1320
aacaaaaaaa	cagaaagaaa	gaaaactaca	ggccaatatc	cctgatgaat	actgatacaa	1380
aaatcctcaa	caaaacacta	gcaaaccaaa	ttaaacaaca	ccttcgaaag	atcattcatt	1440
gtgatcaagt	gggattttatt	ccagggatgg	aaggatgggt	caacatatgc	aaatcaatca	1500
atgtgataca	tcateccaac	aaaatgaagt	acaaaaacta	tatgattatt	tcactttatg	1560
cagaaaaagc	atttgataaa	attctgcacc	cttcattgata	aaaaccctca	aaaaaccagg	1620
tatacaagaa	acatacaggc	caggcacagt	ggctcacacc	tgcgatccca	gcactctggg	1680
aggccaaggt	gggatgattg	cttgggcccc	ggagtttgag	actagcctgg	gcaacaaaat	1740
gagacctggt	ctacaaaaaa	ctttttttaa	aaattagcca	ggcatgatgg	catatgcctg	1800
tagtcccagc	tagtctggag	gctgaggtgg	gagaatcact	taagcctagg	aggctgaggc	1860
tgcagtgagc	catgaacatg	tcactgtact	ccagcctaga	caacagaaca	agaccccact	1920
gaataagaag	aaggagaagg	agaagggaga	agggaggagg	aaggaggagg	gaggagaagg	1980
aggaggtgga	ggagaagtgg	aaggggaagg	ggaagggaaa	gaggaagaag	aagaaacata	2040
tttcaacata	ataaaagccc	tatatgacag	accgaggtag	tattatgagg	aaaaactgaa	2100
agcctttcct	ctaagatctg	gaaaatgaca	agggccccact	ttcaccactg	tgattcaaca	2160
tagtactaga	agtcctagct	agagcaatca	gataagagaa	agaaataaaa	ggcatccaaa	2220
ctggaaagga	agaagtcaaa	ttatcctggt	tgcagatgat	atgatcttat	atctggaaaa	2280
gacttaagac	accactaaaa	aactattaga	gctgaaattt	ggtacagcag	gatacaaaat	2340
caatgtacaa	aaatcagtag	tattttctata	ttccaacagc	aaacaatctg	aaaaagaaac	2400
caaaaaagca	gctacaaata	aaattaacaa	gctaggaatt	aaccaaagaa	gtgaaagatc	2460
tctacaatga	aaactataaa	atgttgataa	aagaaattga	agagggcaca	aaaaaaagaaa	2520
agatatcca	tgttcataga	ttggaagaat	aaatactgtt	aaaatgtcca	tactacccaa	2580
agcaatttac	aaattcaatg	caatccctat	taaaatacta	atgacgttct	tcacagaaat	2640
agaagaaaca	attctaagat	ttgtacagaa	ccacaaaaga	cccagaatag	ccaaaagctat	2700
cctgacaaaa	aagaacaaaa	ctggaagcat	cacattacct	gacttcaaat	tatactacaa	2760
agctatagta	acccaaacta	catggtactg	gcataaaaaac	agatgagaca	tggaccagag	2820
gaacagaata	gagaatccag	aaacaaatcc	atgcatctac	agtgaactca	tttttgacaa	2880
aggtgccaaag	aacatacttt	ggggaaaaga	taatctcttc	aataaatggg	gctggaggaa	2940
ctggatatcc	atatgcaaaa	taacaatact	agaactctgt	ctctcaccat	atacaaaaagc	3000
aaatcaaaat	ggatgaaagg	cttaaatcta	aaacctcaaa	ctttgcaact	actaaaagaa	3060
aacaccggag	aaactctcca	ggacattgga	gtgggcaaag	acttcttgag	taattccctg	3120

caggcacagg	caaccaaagc	aaaaacagac	aatgggagtc	atatcaagtt	aaaaagcttc	3180
tgcccagcaa	aggaaacaat	caacaaagag	aagagacaac	ccacagaatg	ggagaatata	3240
tttgcaaact	attcatctaa	caaggaatta	ataaccagta	tatataagga	gctcaaacta	3300
ctctataaga	aaaacaccta	ataagctgat	tttcaaaaat	aagcaaaaaga	tctgggtaga	3360
cattttctca	aataagtcac	acaaatggca	aacaggcagc	tgaaaatgtg	ctcaacacca	3420
ctgatcatca	gagaaatgca	aatcaaaaact	actatgagag	atcatctcat	cccagttaaa	3480
atggctttta	ttcaaaaagac	aggcaataac	aatggccagt	gaggatgtgg	ataaaaaggaa	3540
acccttggac	actgttgggtg	ggaatggaaa	ttgctaccac	tatggagaac	agtttgaaag	3600
ttcctcaaaa	aactaaaaat	aaagctacca	tacagcaatc	ccattgctag	gtatatactc	3660
caaaaaaggg	aatcagtgtg	tcaacaagct	atctccactc	ccacatttac	tgcagcactg	3720
ttcatagcag	ccaaggtttg	gaagcaacct	cagtgtccat	caacagacga	atggaaaaag	3780
aaaatgtggt	gcacatacac	aatggagtag	tacgcagcca	taaaaaagaa	tgagatcctg	3840
tcagttgcaa	cagcatgggg	ggcactgggc	agtatgttaa	gtgaaataag	ccaggcacag	3900
aaagacaaac	ttttcatgtt	ctcccttact	tgtggggagca	aaaattaaaa	caattgacat	3960
agaaatagag	gagaatgggtg	gttctagagg	ggtggggggac	agggtgacta	gagtcaacaa	4020
taattttattg	tatgttttaa	aataactaaa	agagtataat	tgggttgttt	gtaacacaaa	4080
gaaaggataa	atgcttgaag	gtgacagata	ccccatttac	cctgatgtga	ttattacaca	4140
ttgtatgcct	gtatcaaaaat	atctcatgta	tgctatagat	ataaaccccta	ctatattaaa	4200
aattaaaatt	ttaatggcca	ggcacgggtg	ctcatgtccg	taatcccagc	actttgggag	4260
gccgagcg	gtggatcacc	tgaggctcagg	agtttgaaac	cagtctggcc	accatgatga	4320
aaccctgtct	ctactaaaga	tacaaaaatt	agccaggcgt	ggtggcacat	acctgtagtc	4380
ccaactactc	aggaggctga	gacaggagaa	ttgcttgaac	ctgggaggcg	gaggttgacg	4440
tgagccgaga	tcatgccact	gcactgcagc	ctgggtgaca	gagcaagact	ccatctcaaa	4500
acaaaaacaa	aaaaaagaag	attaaaattg	taatttttat	gtaccgtata	aataataact	4560
ctactatatt	agaagttaaa	aattaaaaca	attataaaaag	gtaattaacc	acttaatact	4620
aaataagaac	aatgtatgtg	gggtttctag	cttctgaaga	agtaaaagtt	atggccacga	4680
tggcagaaat	gtgaggaggg	aacagtggaa	gttactgttg	ttagacgctc	atactctctg	4740
taagtgactt	aattttaacc	aaagacaggc	tgggagaagt	taaagaggca	ttctataaagc	4800
cctaaaacaa	ctgctaataa	tggtgaaagg	taatctctat	taattaccaa	taattacaga	4860
tatctctaaa	atcgagctgc	agaattggca	cgtctgatca	caccgtcctc	tcattcacgg	4920
tgtttttttt	cttgtgtgct	tggagatttt	cgattgtgtg	ttcgtgtttg	gttaaaactta	4980
atctgtatga	atcctgaaac	gaaaaatggt	ggtgatttcc	tccagaagaa	ttagagtacc	5040
tggcaggaag	caggtggctc	tgtggacctg	agccacttca	atcttcaagg	gtctctggcc	5100
aagaccag	tgcaaggcag	aggcctgatg	acccgaggac	aggaaaagctc	ggatgggaag	5160
gggcgatgag	aagcctgcct	cgttgggtgag	cagcgcatga	agtgccctta	tttacgcttt	5220
gcaaagattg	ctctggatac	catctggaaa	aggcgccag	cgggaatgca	aggagtacga	5280
agcctcctgc	tcaaaccag	gccagcagct	atggcgccca	cccgggcgtg	tgccagaggg	5340
agaggagtca	aggcacctcg	aagtatggct	taaatctttt	tttcacctga	agcagtgacc	5400
aagggtgtatt	ctgagggaag	cttgagttag	gtgccttctt	taaaacagaa	agtcattggaa	5460

gcacccttct	caagggaaaa	ccagacgccc	gctctgcggt	catttacctc	tttctctct	5520
ccctctcttg	ccctcgcggt	ttctgatecg	gacagagtga	cccccggtga	gcttctccga	5580
gcccgtgctg	aggaccctct	tgcaaagggc	tccacagacc	cccgccctgg	agagaggagt	5640
ctgagcctgg	cttaataaca	aactgggatg	tggctggggg	cggacagcga	cggcgggatt	5700
caaagactta	attccatgag	taaattcaac	ctttccacat	ccgaatggat	ttggatttta	5760
tcttaatat	ttcttaaat	tcataaata	acattcagga	ctgcagaaat	ccaaaggcgt	5820
aaaacaggaa	ctgagctatg	tttgccaagg	tccaaggact	taataacat	gttcagaggg	5880
atTTTTgc	ctaagtactt	tttattgggt	ttcataaggt	ggcttagggg	gcaagggaaa	5940
gtacacgagg	agaggcctgg	gcggcagggc	tatgagcacg	gcagggccac	cggggagaga	6000
gtccccggcc	tgggaggtg	acagcaggac	cactgaccgt	cctccctggg	agctgccaca	6060
ttgggcaacg	cgaaggcggc	cacgtgcgt	gtgactcagg	acccataacc	ggcttcctgg	6120
gccacccac	actaaccag	gaagtcacgg	agctctgaac	ccgtggaaac	gaacatgacc	6180
cttgccctgc	tgcttccctg	ggtgggtcaa	gggtaatgaa	gtggtgtgca	ggaaatggcc	6240
atgtaaatta	cacgactctg	ctgatgggga	ccgttccttc	catcattatt	catcttcacc	6300
cccaaggact	gaatgattcc	agcaacttct	tgggtgtga	caagccatga	caaaactcag	6360
tacaaacacc	actcttttac	taggcccaca	gagcacggsc	cacaccctg	atatattaag	6420
agtccaggag	agatgaggct	gctttcagcc	accaggctgg	ggtgacaaca	gcggctgaac	6480
agtctgttcc	tctagactag	tagaccctgg	caggcactcc	cccagattct	agggcctggg	6540
tgctgcttcc	cgagggcgcc	atctgccctg	gagactcagc	ctgggggtgcc	acactgaggc	6600
cagccctgtc	tccacaccct	ccgcctccag	gcctcagctt	ctccagcagc	ttcctaaacc	6660
ctgggtgggc	cgtgttccag	cgctactgtc	tcacctgtcc	cactgtgtct	tgtctcagcg	6720
acgtagctcg	cacggttcct	cctcacatgg	ggtgtctgtc	tccttcccca	acactcacat	6780
gcgttgaagg	gaggagattc	tgcgcctccc	agactggctc	ctctgagcct	gaacctggct	6840
cgtggccccc	gatgcagggt	cctggcgctc	ggctgcacgc	tgacctccat	ttccaggcgc	6900
tccccgtctc	ctgtcatctg	ccggggcctg	ccggtgtgtt	cttctgtttc	tgtgctcctt	6960
tccacgtcca	gctgcgtgtg	tctctgcccg	ctagggctctc	ggggttttta	taggcatagg	7020
acgggggcgt	ggtgggccag	ggcgctcttg	ggaaatgcaa	catttgggtg	tgaaagtagg	7080
agtgcctgtc	ctcacctagg	tccacgggca	caggcctggg	gatggagccc	ccgccaggga	7140
cccgccttc	tctgcccagc	actttcctgc	ccccctccct	ctggaacaca	gagtggcagt	7200
ttccacaagc	actaagcatc	ctcttcccaa	aagaccagc	attggcaccc	ctggacattt	7260
gccccacagc	cctgggaatt	cacgtgacta	cgcacatcat	gtacacactc	ccgtccacga	7320
ccgacccccg	ctgttttatt	ttaatagcta	caaagcaggg	aaatccctgc	taaaatgtcc	7380
tttaacaaac	tggttaaaca	aacgggtcca	tccgcacggt	ggacagtccc	tcacagtga	7440
gaggaacatg	ccgtttataa	agcctgcagg	catctcaagg	gaattacgct	gagtcaaaac	7500
tgccacctcc	atgggatacg	tacgcaacat	gctcaaaaag	aaagaatttc	accccatggc	7560
aggggagtg	ttaggggggt	taaggacggt	gggggcggca	gctgggggct	actgcacgca	7620
ctttttacta	aagccagttt	cctggttctg	atggtatttg	ctcagttatg	ggagactaac	7680
cataggggag	tggggatggg	ggaacccgga	ggctgtgcca	tctttgccat	gcccagtggt	7740
cctgggcagg	ataatgctct	agagatgccc	acgtcctgat	tccccaaac	ctgtggacag	7800

aacccgcccc	gccccagggc	ctttgcaggt	gtgatctccg	tgaggaccct	gaggtctggg	7860
atccttcggg	actacctgca	gccccaaaa	gtaatccagg	ggttctggga	agaggcgggc	7920
aggagggtca	gaggggggca	gcctcaggac	gatggaggca	gtcagtctga	ggctgaaaag	7980
ggagggaggg	cctcgagccc	aggcctgcaa	gcgcctccag	aagctggaaa	aagcggggaa	8040
gggaccctcc	acggagcctg	cagcaggaag	gcacggctgg	cccttagccc	accaggggccc	8100
atcgtggacc	tccggcctcc	gtgccatagg	agggcactcg	cgttgccctt	ctagcatgaa	8160
gtgtgtgggg	atttgcagaa	gcaacaggaa	acccatgcac	tgtgaatcta	ggattatttc	8220
aaaacaaagg	tttacagaaa	catccaagga	cagggtgaa	gtgcctccgg	gcaagggcag	8280
ggcaggcacg	agtgatttta	tttagctatt	ttattttatt	tacttacttt	ctgagacaga	8340
gttatgctct	tgttgcccag	gctggagtgc	agcggcatga	tcttggctca	ctgcaacctc	8400
cgtctcctgg	gttcaagcaa	ttctcgtgcc	tcagcctccc	aagtagctgg	gatttcaggc	8460
gtgcaccacc	acaccggct	aattttgtat	ttttagtaga	gatgggcttt	caccatgttg	8520
gtcaagctga	tctcaaaatc	ctgacctcag	gtgatccgcc	cacctcagcc	tcccaaagtg	8580
ctgggattac	aggcatgagc	cactgcacct	ggcctattta	accattttaa	aacttccctg	8640
ggctcaagtc	acaccactg	gtaaggagtt	catggagtgc	aatttcccct	ttactcagga	8700
gttaccctcc	tttgatat	tctgtaattc	tctgtagact	ggggatacac	cgtctcttga	8760
catattcaca	gtttctgtga	ccacctgtta	tcccatggga	cccactgcag	gggcagctgg	8820
gaggctgcag	gcttcaggtc	ccagtggggg	tgccatctgc	cagtagaaac	ctgatgtaga	8880
atcagggcgc	aagtgtggac	actgtcctga	atctcaatgt	ctcagtgtgt	gctgaaacat	8940
gtagaaatta	aagtcacatc	ctcctactct	actgggattg	agcccccttc	ctatcccccc	9000
ccaggggcag	aggagtccct	ctcactcctg	tggaggaagg	aatgatactt	tgttattttt	9060
cactgctggg	actgaatcca	ctgtttcatt	tgttggtttg	tttggtttgt	tttgagaggc	9120
ggtttcactc	ttgttgctca	ggctggaggg	agtgcattgg	cgcgatcttg	gcttactgca	9180
gcctctgcct	cccaggttca	agtgattctc	ctgcttccgc	ctcccatttg	gctgggatta	9240
caggcacccg	ccaccatgcc	cagctaattt	tttgtatttt	tagtagagac	gggggtgggt	9300
ggggttcacc	atgttggcca	ggctgggtctc	gaacttctga	cctcagatga	tccacctgcc	9360
tctgcctcct	aaagtgtctg	gattacaggt	gtgagccacc	atgccagct	cagaattttac	9420
tctgttttaga	aacatctggg	tctgaggtag	gaagctcacc	ccactcaagt	gttgtggtgt	9480
tttaagccaa	tgatagaatt	tttttattgt	tgttagaaca	ctcttgatgt	tttactactgt	9540
gatgactaag	acatcatcag	cttttcaaag	acacactaac	tgcacccata	atactggggg	9600
gtcttctggg	tatcagcaat	cttcattgaa	tgccgggagg	cgtttcctcg	ccatgcacat	9660
ggtgttaatt	actccagcat	aatcttctgc	ttccatttct	tctcttccct	cttttaaaat	9720
tgtgttttct	atgttggctt	ctctgcagag	aaccagtgtg	agctacaact	taacttttgt	9780
tggaaacaaat	tttccaaacc	gcccccttgc	cctagtggca	gagacaattc	acaaacacag	9840
ccctttaaaa	aggcttaggg	atcactaagg	ggatttctag	aagagcgacc	tgtaatccta	9900
agtattttaca	agacgaggct	aacctccagc	gagcgtgaca	gcccaggggag	ggtgcgaggc	9960
ctgttcaaat	gctagctcca	taaataaagc	aatttccctc	ggcagtttct	gaaagtagga	10020
aaggttacat	ttaaggttgc	gtttgttagc	atttcagtgt	ttgccgacct	cagctacagc	10080
atccctgcaa	ggcctcggga	gaccagaag	tttctcgecc	ccttagatcc	aaacttgagc	10140

aacccggagt	ctggattcct	gggaagtcct	cagctgtcct	gcggttgtgc	cggggcccca	10200
ggtctggagg	ggaccagtgg	ccgtgtggct	tctactgtcg	ggctggaagt	cgggcctcct	10260
agctctgcag	tccgaggctt	ggagccaggt	gcctggaccc	cgaggctgcc	ctccaccctg	10320
tgcgggcggg	atgtgaccag	atgtttggcct	catctgccag	acagagtgcc	ggggcccgagg	10380
gtcaaggccg	ttgtggctgg	tgtgaggcgc	ccggtgcgcg	gccagcagga	gcgcctggct	10440
ccatttccca	ccctttctcg	acgggaccgc	cccgggtgggt	gattaacaga	tttgggggtgg	10500
tttgctcatg	gtggggaccc	ctcgccgcct	gagaacctgc	aaagagaaat	gacgggcctg	10560
tgtcaaggag	cccaagtcgc	ggggaagtgt	tgcaggagg	cactccggga	ggtcccgcgt	10620
gcccgtccag	ggagcaatgc	gtcctcgggt	tcgtccccag	ccgcgtctac	gcgcctccgt	10680
cctccccctt	acgtccggca	ttcgtggtgc	ccggagcccc	acgccccgcg	tccggacctg	10740
gaggcagccc	tgggtctccg	gatcaggcca	gcggccaaag	ggtcgccgca	cgcacctgtt	10800
cccagggcct	ccacatcatg	gcccctccct	cgggttaccc	cacagcctag	gccgattcga	10860
cctctctccg	ctggggccct	cgctggcgtc	cctgcaccct	gggagcgcg	gcggcgcgcg	10920
ggcggggaag	cgcggcccag	acccccgggt	ccgcccggag	cagctgcgct	gtcggggcca	10980
ggccgggctc	ccagtggatt	cgcgggcaca	gacgcccagg	accgcgtcc	ccacgtggcg	11040
gagggactgg	ggacccgggc	acccgtcctg	ccccttcacc	ttccagctcc	gcctcctccg	11100
cgcggacccc	gccccgtccc	gaccctccc	gggtccccgg	cccagcccc	tccgggcctt	11160
cccagcccc	ccccttcctt	tccgcggccc	cgcctctccc	tcgcggcgcg	agtttcaggc	11220
agcgctgcgt	cctgctgcgc	acgtgggaag	ccctggcccc	ggccaccccc	gcgatgccgc	11280
gcgctccccg	ctgcccagcc	gtgcgctccc	tgtctgcgag	ccactaccgc	gaggtgctgc	11340
cgctggccac	gttcgtgcgg	cgcctggggc	cccagggtcg	gcggctgggtg	cagcgcgggg	11400
acccggcggc	tttcgcgcg	ctgggtggccc	agtgcctgg	gtgcgtgccc	tgggacgcac	11460
ggccgcccc	cgcgcgcgc	tccttcgcgc	aggtgggcct	ccccggggtc	ggcgtccggc	11520
tggggttgag	ggcggccggg	gggaaccagc	gacatgcgga	gagcagcgca	ggcgactcag	11580
ggcgcttccc	ccgcaggtgt	cctgcctgaa	ggagctgggtg	gcccagatgc	tgcagaggct	11640
gtgcgagcgc	ggcgcggaaga	acgtgctggc	cttcggcttc	gcgctgctgg	acggggccccg	11700
cgggggcccc	cccagggcct	tcaccaccag	cgtgcgcagc	tacctgcccc	acacggtgac	11760
cgacgcactg	cgggggagcg	gggcgtgggg	gctgctgctg	cgccgcgtgg	gcgacgacgt	11820
gctggttcac	ctgctggcac	gctgcgcgct	ctttgtgctg	gtggctcccc	gctgcgccta	11880
ccaggtgtgc	gggcccgcgc	tgtaccagct	cggcgctgcc	actcaggccc	ggcccccgcc	11940
acacgctagt	ggaccccgaa	ggcgtctggg	atgcgaacgg	gcctggaacc	atagcgtcag	12000
ggaggccggg	gtccccctgg	gcctgccagc	cccgggtgcg	aggaggcgcg	ggggcagtgc	12060
cagccgaagt	ctgccgttgc	ccaagaggcc	caggcgtggc	gctgccccctg	agccggagcg	12120
gacgcccgtt	gggcaggggt	cctgggcccc	cccgggcagg	acgcgtggac	cgagtgaccg	12180
tggttttctgt	gtggtgtcac	ctgccagacc	cgcggaagaa	gccacctctt	tggagggtgc	12240
gctctctggc	acgcgccact	cccaccatc	cgtgggcccgc	cagcaccacg	caggcccccc	12300
atccacateg	cggccaccac	gtccctggga	cacgccttgt	cccccggtgt	acgccgagac	12360
caagcacttc	ctctactcct	caggcgacaa	ggagcagctg	cggccctcct	tcctactcag	12420
ctctctgagg	cccagcctga	ctggcgctcg	gaggctcgtg	gagaccatct	ttctgggttc	12480

caggccctgg	atgccaggga	ctccccgcag	gttgccccgc	ctgccccagc	gctactggca	12540
aatgcggccc	ctgtttcttg	agctgcttgg	gaaccacgcg	cagtgtccct	acgggggtgct	12600
cctcaagacg	cactgccccg	tgcgagctgc	ggtcacccca	gcagccgggtg	tctgtgccccg	12660
ggagaagccc	cagggtctctg	tggcggcccc	cgaggaggag	gacacagacc	cccgctgcct	12720
ggtgcagctg	ctccgccagc	acagcagccc	ctggcagggtg	tacggcttcg	tgcgggcctg	12780
cctgcgccgg	ctggtgcccc	caggcctctg	gggctccagg	cacaacgaac	gccgcttcct	12840
caggaacacc	aagaagtcca	tctccctggg	gaagcatgcc	aagctctcgc	tgcaggagct	12900
gacgtggaag	atgagcgtgc	gggactgcgc	ttagctgcgc	aggagcccag	gtgaggaggt	12960
ggtggccgtc	gagggccccag	gccccagagc	tgaatgcagt	aggggctcag	aaaagggggc	13020
aggcagagcc	ctggtcctcc	tgtctccatc	gtcacgtggg	cacacgtggc	ttttcgctca	13080
ggacgtcgag	tggacacggt	gatctctgcc	tctgctctcc	ctcctgtcca	gtttgcataa	13140
acttacgagg	ttcaccttca	cgttttgatg	gacacgcggt	ttccaggcgc	cgaggccaga	13200
gcagtgaaca	gaggaggctg	ggcgcggcag	tggagccggg	ttgccggcaa	tggggagaag	13260
tgtctggaag	cacagacgct	ctggcgaggg	tgcctgcagg	ttacctataa	tcctcttcgc	13320
aatttcaagg	gtgggaatga	gagggtggga	cgagaacccc	ctcttcctgg	gggtgggagg	13380
taagggtttt	gcaggtgcac	gtggtcagcc	aatatgcagg	tttgtgttta	agatttaatt	13440
gtgtgttgac	ggccagggtg	ggtggctcac	gccggtaatc	ccagcacttt	gggaagctga	13500
ggcagggtga	tcacctgagg	tcaggagttt	gagaccagcc	tgaccaacat	ggtgaaaccc	13560
tatctgtact	aaaaatacaa	aaattagctg	ggcatgggtg	tgtgtgcctg	taatcccagc	13620
tacttgggag	gctgaggcag	gagaatcact	tgaaccagg	aggcggaggc	tgcagtgagc	13680
tgagattgtg	ccattgtact	ccagcctggg	cgacaagagt	gaaactctgt	ctttaaaaaa	13740
aaaaagtgtt	cgttgattgt	gccaggacag	ggtagaggga	gggagataag	actgttctcc	13800
agcacagatc	ctggtcccat	ctttaggtat	gaagagggcc	acatgggagc	agaggacagc	13860
agatggctcc	acctgctgag	gaagggacag	tgtttgtggg	tgttcagggg	atggtgctgc	13920
tgggccctgc	cgtgtcccca	ccctgttttt	ctggatttga	tgttgaggaa	cctccgctcc	13980
agcccccttt	tggctcccag	tgtctcccag	ccctaccgtg	gcagctagaa	gaagtcccga	14040
tttccacccc	tccccacaaa	ctcccaagac	atgtaagact	tccggccatg	cagacaagga	14100
gggtgacctt	cttggggctc	ttttttttct	ttttttcttt	ttatgggtgg	aaaagtcata	14160
taacatgaga	ttggcactcc	taacaccgtt	ttctgtgtac	agtgacagaat	tgctaactcg	14220
gcgggtgttta	cagcagggtt	cttgaaatgc	tgcgtcttgc	gtgactggaa	gtccctaccc	14280
atcgaacggc	agctgcctca	cacctgctgc	ggctcagggtg	gaccacgccg	agtcagataa	14340
gcgtcatgca	accagttttt	gctttttgtg	ctccagcttc	cttcgttgag	gagagtttga	14400
gttctctgat	caggactctg	cctgtcattg	ctgttctctg	acttcagatg	aggtcacaaat	14460
ctgcccctgg	cttatgcagg	gagtgaggcg	tgggtccccg	gtgtccctgt	cacgtgcagg	14520
gtgagtgagg	cgttgccccc	aggtgtccct	gtcacgtgta	gggtgagtga	ggcgcggccc	14580
ccgggtgtcc	ctgtcccgtg	cagcgtgatt	gagggtgtgg	ccccgggtgt	ccctgtcacg	14640
tgtagggtga	gtgaggcgcc	atccccgggt	gtccctgtca	cgtgtagggt	gagtgaggcg	14700
tgggtccccg	gtgtccctgt	cccgtgcagg	gtgagtgagg	cactgtcccc	gggtgtccct	14760
gtcacgtgca	gggtgagtga	ggcgcgggtc	ccgggtgtcc	ctctcagggt	taggggtgagt	14820

gaggecgggc	cccaggggtgt	ccctgtcacg	tgtaggggtga	gtgaggcacc	gtccctgggt	14880
gtccctccca	ggtatagggt	gagttaggca	ctgtccccgg	gtgtccctgt	cacgtgcagg	14940
gtgagtgagg	cgcggccccc	gggtgtccct	ctcaggtgca	gggtgagtga	ggcgtgtcc	15000
ctgggtgtcc	ctgtctctgt	tagggtagt	gaggctctgt	ccccaggtgt	ccttggcgtt	15060
tgtcacttg	agcttgctcc	tgaatgtttg	ctctttctat	agccacagct	gcgccggttg	15120
cccattgcct	gggtagatgg	tgcaggcgca	gtgttggtcc	ccaagcctat	cttttctgat	15180
gctcggtct	tcttgggtac	ctctccgttc	cattttgcta	cggggacacg	ggactgcagg	15240
ctctcgctc	ccgcgtgcca	ggcactgcag	ccacagcttc	aggtccgctt	gcctctgttg	15300
ggcctggctt	gtcaccacg	tgcgcgccac	atgcatgctg	ccaatactcc	tctcccagct	15360
tgtctcatgc	cgaggctgga	ctctgggctg	cctgtgtctg	ctgccacgtg	ttgtggaga	15420
catcccagaa	agggttctct	gtgccctgaa	ggaaagcaag	tcaccccagc	cccctcactt	15480
gtcctgtttt	ctcccaagct	gccccctgc	ttggccccct	tgggtgggtg	gcaacgcttg	15540
tcaccttatt	ctgggcacct	gccgtcatt	gcttaggctg	ggctctgcct	ccagtcgccc	15600
cctcacatgg	attgacgtcc	agccacaggt	tggagtgtct	ctgtctgtct	cctgctctga	15660
gaccacgtg	gagggccggt	gtctccgcca	gccttcgtca	gacttccctc	ttgggtctta	15720
gttttgaatt	tcactgattt	acctctgacg	ttcttatctc	tccattgtat	gctttttctt	15780
ggtttattct	ttcattccct	ttctagcttc	ttagtttagt	catgcctttc	cctctaagtg	15840
ctgccttacc	tgcaccctgt	gttttgatgt	gaagtaatct	caacatcagc	cactttcaag	15900
tgtttctaaa	atacttcaaa	gtgttaatac	ttcttttaag	tattcttatt	ctgtgatttt	15960
tttctttgtg	cacgctgtgt	tttgacgtga	aatcattttg	atatcagtga	cttttaagta	16020
ttcttttagct	tattctgtga	tttctttgag	cagtgagtta	tttgaacact	gtttatgttc	16080
aagatatgta	gagtatcaag	atacgtagag	tatttttaagt	tatcatttta	ttattgattt	16140
ctaactcagt	tgtgtagtgg	tctgtataat	accaattatt	tgaagtttgc	ggagccttgc	16200
tttgtgatct	agtgtgtgca	tggtttccag	aactgtccat	tgtaaatttg	acatcctgtc	16260
aatagtgggc	atgcatgttc	actatatcca	gcttattaag	gtccagtgca	aagcttctgt	16320
ctccttctag	atgcatgaaa	ttccaagaag	gaggccatag	tccctcacct	gggggatggg	16380
tctgttcatt	tcttctcggt	tggtagcatt	tatgtgaggc	attgttaggt	gcatgcacgt	16440
ggtagaattt	ttatcttcct	gatgagtga	tcttttggag	acttctatgt	ctctagtaat	16500
ctagtaattc	tttttttaaa	ttgctcttag	tactgccaca	ctgggcttct	tttgattagt	16560
attttcttgc	tgtgtctgtt	ttctgccttt	aatttatata	tatatatata	tttttttttt	16620
ttttgagaca	gagtcttgggt	ctgtcgccca	gggtgagtgc	agtgggtgtga	tcacaggtca	16680
gtgtaacttt	taccttctgg	cctgagccgt	cctctcacct	cagcctcctg	agtagctgga	16740
actgcagaca	cgcaccgcta	cacctggcta	atttttaaat	tttttctgga	gacagggctct	16800
tgtgtgtttg	cccaggctgg	tctcaaactc	ttggactcaa	gggatccatc	tacctgggct	16860
tcccaaagtg	ctgaattaca	ggcatgagcc	accatgtctg	gcctaatttt	caacactttt	16920
atattcttat	agtgtgggtg	tgtcctgtta	acagcatgta	ggtgaatttc	caatccagtc	16980
tgacagtctg	tgtttaactg	gataacctga	tttattttca	tttttttgtc	actagagacc	17040
cgcctgggtg	actctgattc	tccacttgcc	tgttgcatgt	cctcgttccc	ttgtttctca	17100
ccacctcttg	ggttgccatg	tgcgtttcct	gccgagtgtg	tgttgatcct	ctcgttgcct	17160

cctggtcact	gggcatttgc	ttttatttct	ctttgcttag	tgttaccccc	tgatcttttt	17220
attgtcgttg	tttgcttttg	tttattgaga	cagtctcact	ctgtcaccca	ggctggagtg	17280
taatggcaca	atctcggtc	actgcaacct	ctgcctcctc	ggttcaagca	gttctcattc	17340
ctcaacctca	tgagtagctg	ggattacagg	cgtccaccac	cacgcctggc	taatttttgt	17400
attttttagta	gagataggct	ttcaccatgt	tggccaggct	ggtctcaaac	tcctgacctc	17460
aagtgatctg	cccgccttgg	cctccacag	tgctgggatt	acaggtgcaa	gccaccgtgc	17520
ccggcatacc	ttgatctttt	aaaatgaagt	ctgaaacatt	gctacccttg	tcctgagcaa	17580
taagaccctt	agtgtatttt	agctctggcc	accccccagc	ctgtgtgctg	ttttccctgc	17640
tgacttagtt	ctatctcagg	catcttgaca	ccccacaaag	ctaagcatta	ttaatatgtt	17700
tttccgtgtt	gagtgtttct	gtagctttgc	ccccgccttg	cttttccctc	tttgttcccc	17760
gtctgtcttc	tgtctcaggc	ccgccgtctg	gggtcccttt	ccttgtcctt	tgcgtggttc	17820
ttctgtcttg	ttattgctgg	taaaccacag	ctttacctgt	gctggcctcc	atggcatcta	17880
gcgacgtccg	gggacctctg	cttatgatgc	acagatgaag	atgtggagac	tcacgaggag	17940
ggcggtcate	ttggcccgtg	agtgtctgga	gcaccacgtg	gccagcgttc	cttagccagt	18000
gagtgcacgc	aacgtccgct	cggcctgggt	tcagcctgga	aaaccacagg	catgtcgggg	18060
tctggtggct	ccgcggtgtc	gagtttgaaa	tcgcgcaaac	ctgcggtgtg	gcgccagctc	18120
tgacggtgct	gcctggcggg	ggagtgtctg	cttcctccct	tctgcttggg	aaccaggaca	18180
aaggatgagg	ctccgagccg	ttgtcgccca	acaggagcat	gacgtgagcc	atgtggataa	18240
ttttaaaatt	tctaggctgg	gcgcggtggc	tcacgcctgt	aatcccagca	ctttgggagg	18300
ccaaggcggg	tggatcacga	ggtcaggagg	tcgagaccat	cctggccaac	atgatgaaac	18360
cccatctgta	ctaaaaacac	aaaaattagc	tgggcgtggt	ggcgggtgcc	tgtaatccca	18420
gctactcggg	aggctgaggc	aggagaattg	cttgaacctg	ggagttggaa	gttgcagtga	18480
gccgacattg	caccactgca	ctccagcctg	gcaacacagc	gagactctgt	ctcaaaaaaa	18540
aaaaaaaaaa	aaaaaaaaaa	aattctagta	gccacattaa	aaaagtaaaa	aagaaaaggt	18600
gaaattaatg	taataataga	ttttactgaa	gcccagcatg	tccacacctc	atcatttttag	18660
ggtgttattg	gtgggagcat	cactcacagg	acatttgaca	ttttttgagc	tttgtctgcg	18720
ggatcccgtg	tgtaggtccc	gtgcgtggcc	atctcggcct	ggacctgctg	ggcttcccat	18780
ggccatggct	gttgtaccag	atggtgcagg	tccgggatga	ggtcgccagg	ccctcagtga	18840
gctggatgtg	cagtgtccgg	atggtgcacg	tctgggatga	ggtcgccagg	ccctgctgtg	18900
agctggatgt	gtggtgtctg	gatggtgcag	gtcaggggtg	aggtctccag	gccctcgggt	18960
agctggagggt	atggagtccg	gatgatgcag	gtccgggggtg	aggtcgccag	gccctgctgt	19020
gagctggatg	tgtggtgtct	ggatggtgca	ggtcaggggt	gaggtctcca	ggccctcgggt	19080
aagctggagg	tatggagtcc	ggatgatgca	ggtcgggggt	gaggtcgcca	ggccctgctg	19140
tgagctggat	gtgtggtgtc	tggatggtgc	aggtctgggg	tgaggtcacc	aggccctgcg	19200
gtgagctggg	tgtgcggtgt	ctggatggtg	caggtctgga	gtgaggtcgc	cagacgggtgc	19260
cagaccatgc	ggtgagctgg	atatgcggtg	tccggatggt	gcaggtctgg	ggtgaggttg	19320
ccaggccctg	ctgtgagttg	gatgtgggggt	gtccggatgc	tgcaggtccg	gtgtgaggtc	19380
accaggccct	gctgtgagct	ggatgtgtgg	tgtctggatg	gtgcaggtct	ggggtgaagg	19440
tcgccaggcc	cctgcttgtg	agctggatgt	gtggtgtctg	gatggtgcag	gtctggagtg	19500

aggtcgccag	gccctcggtg	agctggatgt	gcagtgtcca	gatgggtgcag	gtccgggggtg	19560
aggtcgccag	accctgcggt	gagctggatg	tgcggtgtct	ggatgggtgca	ggtctggagt	19620
gaggtcgcca	ggccctcggt	gagctggatg	tatggagtcc	ggatgggtgcc	ggtccgggggt	19680
gaggtcgcca	gaccctgctg	tgagctggat	gtgcggtgtc	tggatgggtac	aggtctggag	19740
tgaggtcgcc	agaccctgct	gtgagctgga	tatgcggtgt	ccggatgggtg	caggtcaggg	19800
gtgaggctct	caggccctcg	gtgagctgga	ggtatggagt	ccggatgatg	caggtcggg	19860
gtgaggtcgc	caggccctgc	tgtgaactgg	atgtgcggcg	tctggatggg	gcaggctctg	19920
ggtgtggtcg	ccaggccctc	ggtgagctgg	aggtatggag	tccggatgat	gcaggtcgg	19980
ggtgaggtcg	ccaggccctg	ctgtgagctg	gatgtgcggc	gtctggatgg	tgcaggctctg	20040
gggtgtggtc	gccaggccct	cggtagctg	gaggtatgga	gtccggatga	tgcaggtcgg	20100
gggtgaggtt	gccaggccct	gctgtgagct	ggatgtgctg	tatccggatg	gtgcagtcgg	20160
gggtgaggtc	gccaggccct	gctgtgagct	ggatgtgctg	tatccggatg	gtgcaggctct	20220
ggggtgaggt	caccaggccc	tgcggtgagc	tggttgtgcg	gtgtccgggt	gctgcaggtc	20280
cggggtgagt	tcgccaggcc	ctcggtgagc	tggatgtgcg	gtgtccccgt	gtccggatgg	20340
tgcagggtcca	gggtgaggtc	gctaggccct	tgggtgggctg	gatgtgccgt	gtccggatgg	20400
tgcaggctctg	gggtgaggtc	gccaggccct	tggtagctg	gatgtgcggg	gtctgcatgg	20460
tgcaggctctg	gggtgaggtc	gccaggccct	tgggtgggctg	gatgtgtggg	gtccggatgg	20520
tgcaggtcgg	gcgtgaggtc	gccaggccct	gctgtgagct	ggatgtgcgg	tgtctggatg	20580
gtgcagggtcc	ggggtgaggt	agccaaggcc	tccggtgagc	tggatgtggg	gtgtccggat	20640
ggtgcaggtc	cggggtgagg	tcgccaggcc	ctgcggttag	ctggatatgc	ggtgtccgga	20700
tgggtcagggt	cggggtgag	gtcaccaggc	cctgcggtta	gctggatgtg	cgggtgtctgg	20760
atgggtgcagg	tccggggtga	ggtcgccagg	ccctgctgtg	agctggatgt	gctgtatccg	20820
gatgggtgcag	gtccgggggtg	aggtcgccag	gccctgcagt	gagctggatg	tgctgtatcc	20880
ggatgggtgca	ggtctggcgt	gaggtcgcca	ggccctgcgg	ttagctggat	atgcggtgtc	20940
ggatgggtgca	ggtccgggggt	gaggtcacca	ggccctgcgg	ttagctggat	gtgcggtgtc	21000
cggatgggtg	aggtctgggg	tgaggtcgcc	aggccctgct	gtgagctgga	tgtgctgtat	21060
ccggatgggtg	caggtcggg	gtgaggtcgc	caggccctgc	ggtgagctgg	atgtgctgta	21120
tccggatgggt	gcaggctctgg	cgtgaggtcg	ccaggccctg	cggtagctg	gatgtgcagt	21180
gtacggatgg	tgcaggtcgg	gggtgaggtc	gccaggccct	gcggtgggct	gtatgtgtgt	21240
tgtctggatg	gtgcagggtcc	ggggtgagtt	cgccaggccc	tgcggtgagc	tggatgtgtg	21300
gtgtctggat	gctgcaggtc	cggggtgagt	tcgccaggcc	ctcggtgagc	tggatatgcg	21360
gtgtccccgt	gtccgaatgg	tgcagggtcca	gggtgaggtc	gccaggccct	tgggtgggctg	21420
gatgtgccgt	gtccggatgg	tgcaggctctg	gggtgaggtc	gccaggccct	tggtagctg	21480
gatgtgcggg	gtccggatgg	tgcaggtcgg	gggtgaggtc	accaggccct	cggtagctg	21540
gatgtggcat	gtccttctcg	ttaaaggggt	tggctgtgtt	ccggccgcag	agcaccgtct	21600
gcgtgaggag	atcctggcca	agttcctgca	ctggctgatg	agtgtgtacg	tcgtcgagct	21660
gtcagggtct	ttcttttatg	tcacggagac	cacgtttcaa	aagaacaggc	tctttttcta	21720
ccggaagagt	gtctggagca	agttgcaaag	cattggaatc	aggtactgta	tccccacgcc	21780
aggcctctgc	ttctcgaagt	cctggaacac	cagcccggcc	tcagcatgcg	cctgtctcca	21840

cttgccctgtg	cttccctggc	tgtgcagctc	tgggctggga	gccaggggcc	ccgtcacagg	21900
cctgggtccaa	gtggattctg	tgcaaggctc	tgactgcctg	gagctcacgt	tctcttactt	21960
gtaaaatcag	gagtttgtgc	caagtggctc	ctagggtttg	taaagcagaa	gggatttaaa	22020
ttagatggaa	acactaccac	tagcctcctt	gcctttccct	gggatgtggg	tctgattctc	22080
tctctctttt	ttttttcttt	tttgagatgg	agtctcactc	tgttgcccag	gctggagtgc	22140
agtggcataa	tcttggtctc	ctgcaacctc	cacctcctgg	gtttaagcga	ttcaccagcc	22200
tcagcctcct	aagtagctgg	gattacaggc	acctgccacc	acgcctggct	aatttttgta	22260
cttttaggag	agacgggggt	tcacatgtt	ggccaggtcg	gtctcgaact	catgacctca	22320
ggtgatccac	ccaccttggc	ctcccaaagt	gctgggttta	caggctaagc	caccgtgccc	22380
agccccgat	tctcttttaa	ttcatgctgt	tctgtatgaa	tcttcaatct	attggattta	22440
ggtcatgaga	ggataaaatc	ccaccactt	ggcgactcac	tgcagggagc	acctgtgcag	22500
ggagcacctg	gggataggag	agttccacca	tgagctaact	tctaggtggc	tgcatttgaa	22560
tggctgtgag	attttgtctg	caatgttcgg	ctgatgagag	tgtgagattg	tgacagattc	22620
aagctggatt	tgcatacgtg	agggacggga	gcgctggctc	gggagatgcc	agcctggctg	22680
agcccaggcc	atggtattag	cttctccgtg	tcccgccag	gctgactgtg	gagggttta	22740
gtcagaagat	cagggttcc	ccagctcccc	tgcacactcg	agtccttggg	gggccttgtg	22800
acaccccatg	ccccaaatca	ggatgtctgc	agagggagct	ggcagcagac	ctcgtcagag	22860
gtaacacagc	ctctgggctg	gggaccccg	cgtgggtgctg	gggccatttc	cttgcatctg	22920
ggggagggtc	agggttttcc	ctgtgggaac	aagttaatac	acaatgcacc	ttacttagac	22980
tttacacgta	tttaatggtg	tgcgacccaa	catggctcatt	tgaccagtat	tttgaaaga	23040
atttaattgg	ggtgaccgga	aggagcagac	agacgtgggtg	gtccccaaga	tgctccttgt	23100
cactactggg	actgtttgtc	tgcctggggg	gccttggagg	ccccctctcc	ctggacaggg	23160
taccgtgcct	tttctactct	gctgggcctg	cggcctgcgg	tcagggcacc	agctccggag	23220
cacccgcggc	cccagtgtcc	acggagtgcc	aggctgtcag	ccacagatgc	ccagggtccag	23280
gtgtggccgc	tccagcccc	gtgcccccat	gggttggttt	gggggaaaag	gccaagggca	23340
gaggtgtcag	gagactggtg	ggctcatgag	agctgattct	gctccttggc	tgagctgccc	23400
tgagcagcct	ctccgcctt	ctccatctga	agggatgtgg	ctctttctac	ctgggggtcc	23460
tgcctggggc	cagccttggg	ctaccccagt	ggctgtacca	gaggggacagg	catcctgtgt	23520
ggaggggcat	gggttcacgt	ggccccagat	gcagcctggg	accaggctcc	ctgggtgctga	23580
tgggtgggaca	gtcacccctg	gggttgaccg	ccggactggg	cgtccccagg	gttgactata	23640
ggaccaggtg	tccaggtgcc	ctgcaagtag	aggggctctc	agaggcgtct	ggctggcatg	23700
ggtggacgtg	gccccgggca	tggccttcag	cgtgtgctgc	cgtgggtgcc	ctgagccctc	23760
actgagtcgg	tgggggcttg	tggcttcccg	tgagcttccc	cctagtctgt	tgtctggctg	23820
agcaagcctc	ctgaggggct	ctctatttga	gacagcactt	gaagagggtg	cagctgcggg	23880
agctgtcgga	agcagaggtc	aggcagcatc	gggaagccag	gcccgccttg	ctgacgtcca	23940
gactccgctt	catccccaag	cctgacgggc	tgcggccgat	tgtgaacatg	gactacgtcg	24000
tgggagccag	aacgttccgc	agagaaaaga	gggtggctgt	gcttttggtt	aacttccttt	24060
ttaaacagaa	gtgcgtttga	gccccacatt	tggatatcagc	ttagatgaag	ggccccgagg	24120
aggggccacg	ggacacagcc	agggccatgg	cacggcgcca	acccatttgt	gcgcacagtg	24180

aggtggccga	ggtgccggtg	cctccagaaa	agcagcgtgg	gggtgtaggg	ggagctcctg	24240
gggcagggac	aggctctgag	gaccacaaga	agcagccggg	ccagggcctg	gatgcagcac	24300
ggccccaggt	cctggatccg	tgctctgctg	tggtgcgag	cctccgtgcg	cttccgctta	24360
cggggcccg	ggaccaggcc	acgactgcc	ggagcccacc	gggctctgag	gaccttgac	24420
cttgccccac	ggctctgca	ccccaccct	gtggctgcgg	tggtgcgggt	gaccccgta	24480
tctgaggaga	gtgtggggtg	aggtggacag	aggtgtggca	tgaggatccc	gtgtgcaaca	24540
cacatgcggc	caggaaccgc	tttcaaacag	ggtctgagga	agctgggagg	ggttctaggt	24600
ccgggtctg	ggtggctggg	gacactgggg	aggggctgct	tctcccctgg	gtccctatgg	24660
tggggtgggc	acttggccgg	atccactttc	ctgactgtct	cccatgctgt	ccccgccagg	24720
ccgagcgtct	cacctcgagg	gtgaaggcac	tgttcagcgt	gctcaactac	gagcgggcgc	24780
ggcgccccgg	cctcctgggc	gcctctgtgc	tgggcctgga	cgatatccac	agggcctggc	24840
gcaccttcgt	gctgcgtgtg	cgggcccagg	acccgccgcc	tgagctgtac	tttgtcaagg	24900
tgggtgccgg	ggacccccgt	gagcagccct	gctggaccct	gggagtggct	gcctgattgg	24960
cacctcatgt	tgggtggagg	aggtactcct	gggtgggccc	caggagtgct	aggtgaccct	25020
gtcactgttg	aggacacacc	tggcacctag	ggtggaggcc	ttcagccctt	cctgcagcac	25080
atggggccga	ctgtgcaccc	tgactgcccg	ggctcctatt	ccaaggagg	gtccactgg	25140
attccagttt	ccgtcagaga	aggaaccgca	acggctcagc	caccaggccc	cgggtgccttg	25200
caccccagtc	ctgagccagg	ggtctcctgt	cctgaggctc	agagagggga	cacagcccgc	25260
cctgcccttg	gggtctggag	tggtgggggt	cagagagaga	gtgggggaca	ccgccaggcc	25320
aggccctgag	ggcagagggt	atgtctgagt	ttctgcgtgg	ccactgtcag	tctcctcgcc	25380
tccactcaca	caggtggatg	tgacgggcgc	gtacgacacc	atcccccagg	acaggctcac	25440
ggaggtcatt	gccagcatca	tcaaacccca	gaacacgtac	tgcgtgcgtc	ggtatgccgt	25500
ggtccagaag	gccgcccatt	ggcacgtccg	caaggccttc	aagagccacg	taagggtcac	25560
gtgtgatagt	cgtgtccagg	atgtgtgtct	ctgggatatt	aatgtgtcta	gaatgcagtc	25620
gtgtctgtga	tgcgtttctg	tggtggagggt	acttccatga	tttacacatc	tgtgatattg	25680
gtgtgtggca	cgtgtgtgtc	gtgggtgcatg	tatctgtggc	gtgcatatct	gtggtgtgtg	25740
tgtgtgtggc	acgtgtgtgt	ccatgggtgtg	tgtgcctgtg	gtgtgcatgt	gtgtgtgtct	25800
gtgacacgtg	catgttcatg	ctgtgtgtctg	catgtctgtg	atgtgcctat	ttgtggtgtg	25860
tgtgtgcatg	tgtccgtgac	atatgcgtgt	ctatggcatg	ggtgtgtgtg	gccccttggc	25920
cttactcctt	cctcctccag	gcatgggtccg	caccattgtc	ctcacgctct	cgggtgctgg	25980
tttggggagc	tccacattca	gggtcctcac	ttctagcatg	ggtgccccctg	tcctgtcaca	26040
gggctgggcc	ttggagactg	taagccagggt	ttgagaggag	agtagggatg	ctggtggtac	26100
cttcttgga	ccctggcacc	cccaggaccc	cagtctggcc	tatgccggct	ccatgagata	26160
taggaaggct	gattcaggcc	tcgtcccccg	ggacacactc	ctcccagagc	ggccgggggc	26220
cttggggctc	ggcaggggtg	aaaggggccc	tgggcttggg	ttcccaccca	gtggtcatga	26280
gcacgttgga	ggggttaagcc	ctcaaagtcg	tgccaggccg	gggtgcagag	gtgaagaagt	26340
atccctggag	cttcggctctg	gggagaggca	catgtggaaa	cccacaagga	cctctttctc	26400
tgacttcttg	agcttgtggg	attggttttc	atgtgtggga	taggtgggga	tctgtgggat	26460
tggtttttat	gagtggggta	acacagagtt	caaggcgagc	tttcttcctg	tagtgggtct	26520

gcaggtgctc	caacagcttt	attgaggaga	ccatatcttc	ctttgaacta	tggtcggggt	26580
tatagtaagt	caggggtgtg	gaggcctccc	ctgggctccc	tgttctgttt	cttccactct	26640
ggggctcgtg	ggtgcctgct	gtgggtgtgtg	gccgggtgggc	agggcctcca	ggcctccttg	26700
tgttcattgg	cctggatgtg	gccctggcta	cgtcccgtcc	ttggaattcc	cctgcgagtt	26760
ggaggcttcc	tttctttctt	ttttcttttc	tttttttttt	tttttgataa	cagagtctcg	26820
ctcttttttg	cccaggctgg	agtgggtttg	cgtgatcttg	gctcactgca	acctgtgctt	26880
cctgagttca	agcaattctc	ttgcctcagc	ctcccaagta	gctggaatta	taggcgcccc	26940
ccaccatgct	gactaatttt	tgtaatttta	gtagagacga	ggtttctcca	tgttggccag	27000
gctggtctcg	aactcctgac	ctcaggatgat	cctcccacct	cggcctccca	aagtgtctggg	27060
atgacagggtg	tgaaccgccc	cgcgccggcc	agactcgctt	cctgcagctt	ccgtgagatc	27120
tgcagcgata	gctgcctgca	gccttgggtgc	tgacaacctc	cgttttccct	ctccaggctct	27180
cgttaggggt	ctttccattt	catgactctc	ttcacagaag	agtttcacgt	gtgctgattt	27240
cccggctgtt	tcctgcgtaa	ttgggtgtctg	ctgtttatcg	atggcctcct	tccatttcct	27300
ttaggctttg	tttattgttg	ttttccggc	tccttgaagg	aaaagtctcg	attatggatg	27360
tttgaacttt	cttttctaaa	caagcatctg	aagttgccgt	tttccctcta	aagcagggat	27420
cccaggcccc	ctggctgttg	agtggcaccc	gtctggggcc	tgttaggaac	ccggcgcccc	27480
gcgggaggct	aggtgggggtg	tggggagcca	gcgttcccgc	ctgagccccg	cccctctcag	27540
atcagcagtg	gcatgcgggtg	ctcagaggcg	cacacacct	actgagaact	gtgcgtgaga	27600
ggggctctaga	ttctgtgctc	cttatgggaa	tctaatgcct	gatgatctga	ggtggaaccg	27660
tttgctccca	aaaccatccc	cttccccact	gctgtcctgt	ggaaaaatcg	tcttccacga	27720
aaccagtccc	tgggtaccaca	atgggtgggg	accctgtgct	aaagacctgc	ttcagcagcc	27780
tctcgtcagt	gttgatata	tggcttttct	gtgttgagtc	cagaataatt	acggatttct	27840
gtgatgcttt	ccgcgcacct	cagacccatg	ggctatttgt	gggcgtgttg	cctgctcctg	27900
ggttgggaag	ggtgcaggcc	ccatgtacct	tcctgttact	gccttccagg	ttggttctca	27960
gggttgaatc	gtactcgatg	tggtttttagc	ccacggccct	gccgccagct	cctgggggct	28020
ggggaacatg	ctgaagcaca	gagtcaccgt	gcgcgtcttt	tgatgcctca	caagctcgag	28080
gcctcctgtg	tcctgtttag	tgtgtgtcac	gtgcctgtc	acatcctgtc	ttggggacgc	28140
aggggcttag	caggtcccgt	agtaaataac	aagcgtcctg	ggggagtctg	cagaatagga	28200
ggtgggggtg	ccggtctctc	tcccgcgtct	tcagactctt	ctcctgcctg	tgtgtggct	28260
gcacctgcat	ccctgcaatc	cctccagcac	tgggctggag	aggcccggga	gctcgagtgc	28320
cacttgtgcc	acgtgactgt	ggatggcagt	cggtcacggg	ggtctgatgt	gtggtgactg	28380
tggatggcgg	ttggtcacag	gggtctgatg	tgtggtgact	gtggatggcg	gtcgtggggt	28440
ctgatgtggt	gactgtggat	ggcggtcgtg	gggtctgatg	tgtggtgact	gtggatggcg	28500
gtcgtggggt	ctgatgtggt	gactgtggat	ggcggtcgtg	gggtctgatg	tgggtgactgt	28560
ggatggcggt	cgtgggggtct	gatgtgggtga	ctgtggatgg	cagtcgtggg	gtctgatgtg	28620
tgggtgactgt	ggatggcggt	cgtgggggtct	gatgtgggtga	ctgtggatgg	cagtcgtggg	28680
gtctgatgtg	tgggtgactgt	ggatggcggt	cgtgggggtct	gatgtgtggt	gactgtggat	28740
ggcggtcgtg	gggtctgatg	tgtggtgact	gtggatggcg	gtcgtggggt	ctgatgtgtg	28800
gtgactgtgg	atggcggtcg	tggggctctga	tgtggtgact	gtggatggcg	gtcgtggggt	28860

ctgatgtgtg	gtgactgtgg	atggtgatcg	gtcacagggg	tctgatgtgt	ggtgactgtg	28920
gatggcggtc	gtgggtctcg	atgtgtggtg	actgtggatg	gtgatcggtc	acaggggtct	28980
gatgtgtggt	gactgtggat	ggcggtcgtg	gggtctgatg	tgtggtgact	gtggatggcg	29040
gttggtcccc	gggtctgat	gtgtggtgac	tgtggatggc	gatcggtcac	aggggtctga	29100
tgtgtggtga	ctgtggatgg	cggtcgtggg	gtctgatgtg	tggtgactgt	ggatggcggt	29160
cgtgggggtct	gatgtgtggt	gactgtggat	ggcggtcgtg	gggtctgatg	tggtgactgt	29220
ggatggcggt	cgtgggggtct	gatgtggtga	ctgtggatgg	cggtcgtggg	gtctgatgtg	29280
tggtgactgt	ggatggcggt	tggccccggg	ggtctgatgt	gtggtgactg	tggatggcgg	29340
tcgtggggtc	tgatgtggtg	actgtggatg	gcagtcgtgg	ggtctgatgt	gtggtgactg	29400
tggatggcgg	tcgtggggtc	tgatgtgtgg	tgactgtgga	tggcggtcgt	gggtctgat	29460
gtgtggtgac	tgtggatggc	ggtcgtgggg	tctgatgtgt	ggtgactgtg	gatggcggtc	29520
gtgggggtctg	atgtggtgac	tgtggatggc	ggtcgtgggg	tctgatgtgt	ggtgactgtg	29580
gatggtgatc	ggtcacaggg	gtctgatgtg	tggtgactgt	ggatggcggt	cgtgggggtct	29640
gatgtgtggt	gactgtggat	ggcggtcgtg	gggtctgatg	tggtgactgt	ggatggcggt	29700
cgtgggggtct	gatgtgtggt	gactgtggat	ggcggtcgta	gggtctgatg	tgtggtgact	29760
gtggatggca	gtcggtcaca	gggtctgat	gtgtggtgac	tgtggatggc	ggtcgtgggg	29820
tctgatgtgt	ggtgactgtg	gatggcggtc	gtgggggtctg	atgtgtggtg	actgtggatg	29880
gcggtcgtgg	ggtctgatgt	gtggtgactg	tggatggcgg	tcgtggggtc	tgatgtggtg	29940
actgtggatg	gtgatcggtc	acaggggtct	gatgtgtggt	agctgcaggt	ggagtcacag	30000
gtgtgtctgt	agctactttg	cgtcctcggc	cccccgccc	ccgtttccca	aacagaagct	30060
tccagggcgc	tctctgggct	tcattcccgc	atcgggcttg	gccgcaggtc	cacacgtcct	30120
gatcggaaga	aacaagtgcc	cagctctggc	cggggcaggc	cacatttgtg	gctcatgccc	30180
tctcctctgc	cggcagggtct	ctaccttgac	agacctccag	ccgtacatgc	gacagtctgt	30240
ggctcacctg	caggagacca	gcccgttgag	ggatgccgtc	gtcatcgagc	aggtctgggc	30300
actgccctgc	agggttgggc	acggactccc	agcagtgggt	cctccccctgg	gcaatcactg	30360
ggctcatgac	cggacagact	gttggccctg	gggggcagtg	gggggaatga	gctgtgatgg	30420
gggcatgatg	agctgtgtgc	cttggcgaaa	tctgagctgg	gccatgccag	gctgcgacag	30480
ctgctgcatt	caggcacctg	ctcacgtttg	actgcgcggc	ctctctccag	ttccgcagtg	30540
cctttgttca	tgatttgcta	aatgtcttct	ctgccagttt	tgatcttgag	gccaaaaggaa	30600
aggtgtcccc	ctccttttagg	agggcaggcc	atgttttagc	cgtgtcctgc	ccagctggcc	30660
cctcagtgtc	gggtctgagg	ccaaaggaaa	cgtgtccccc	ttcttaggag	gacgggccgt	30720
gtttgagcca	cgccccgctg	agcgggcctc	tcagtgtctg	gtctgtccac	gtggccctgt	30780
ggccctttgc	agatgtggtc	tgtccacgtg	gcccgttggc	tctttgcaga	tgcctgttag	30840
cacttgctcg	gctctagggg	acagtcgtgt	ccaccgcatg	aggctcagag	acctctgggc	30900
gaatttcctt	ggctcccagg	gtgggggtgg	aggtggcctg	ggctgtctggg	acccagaccc	30960
tgtgccccgc	agctgggcag	caactcctgg	atcacatatg	ccatccgggc	cacgggtgggc	31020
tgtgtgggtg	tgagcccagc	tggaccaca	ggtggcccag	aggagacgtt	ctgtgtcaca	31080
cactctgcct	aagcccatgt	gtgtctgcag	agactcggcc	cggccagccc	acgatggccc	31140
tgcattccag	cccagccccg	cacttcatca	caaacactga	ccccaaaagg	gacggagggt	31200

cttggccacg	tggctcctgcc	tgtctcagca	cccaccggct	cactcccatg	tgtctcccgt	31260
ctgttttgcg	agagctcctc	cctgaatgag	gccagcagtg	gcctcttcga	cgtcttctta	31320
cgtttcatgt	gccaccacgc	cgtgcgcate	aggggcaagt	gagtcagggtg	gccagggtgcc	31380
attgccctgc	gggttggttg	gcgggctggc	agggcttctg	ctcacctctc	tcctgcccct	31440
tccccactgn	ccttctgccc	ggggccacca	gagtctcctt	ttctggcccc	cgccccctcc	31500
ggctcctggg	ctgcaggctc	ccgaggcccc	ggaaacatgg	ctcggcttgc	ggcagccgga	31560
gcggagcagg	tgccacacga	ggcctggaaa	tggcaagcgg	ggtgtggagt	tgctcctgcg	31620
tggaggacga	ggggcggggg	gtgtgtctgg	gtcagggtgtg	cgccgagcgt	ttgagcctgc	31680
agcttgtcag	ctccaagtta	ctactgacgc	tggacacccg	gctctcacac	gcttgtatct	31740
ctctctcccc	atacaaaagg	attttatccg	attctcattc	ctgtccctgt	cgtgtgaccc	31800
ccgcgagggc	gcgggctctt	ctctctgtga	ctagatttcc	catctggaaa	gtgcgggggtt	31860
gaccgtgtag	tttgctcctc	tcgggggggc	tgtgggtggc	atggggcagg	cggcctggga	31920
gagctgccgt	cacacagcca	ctgggtgagc	cacactcacg	gtggtagagc	cacagtgcct	31980
ggtgccacat	cacgtcctct	ggatttttaag	taaaaccaca	cacctcccgg	caggcatctg	32040
cctgcgaccc	tgtgtgtgcc	tggggagagt	ggtagcacgg	aggaaattcg	tgcacactca	32100
aggtcatcag	caaggtcctc	cgcagtcagg	tggaacgtgg	aggcctctct	ctgggatcgt	32160
ctccagcgga	taaaggactg	tgcacagctt	cggaaagctt	tatttaaaaa	tataactatt	32220
aattatttga	ttataagtaa	tcactaatgg	tatcagcaat	tataatatatt	attaaagtat	32280
aattagaaat	attaagtagt	acacacgttc	tggaaaaaca	caaattgcac	atggcagcag	32340
agtgaatttt	ggccgaggga	cacgtgtgca	catgtgtgta	agcggccccc	aggcccacag	32400
aattcgctga	caaagtcacc	tccccagaga	agccaccacg	ggcctccttc	gtggtcgtga	32460
attttattaa	gatggatcaa	gtcacgtacc	gtccacgtgt	ggcagggctt	tggggaatgt	32520
gaggtgatga	ctgcgtcctc	atgccctgac	agacaggagg	tgactgtgtc	tgtcctgtcc	32580
ctaggacacg	gacaggcccc	aagctctagt	ccccatcgtg	gtccagtttg	gcctctgaat	32640
aaaaacgtct	tcaaaacctg	ttgccccaaa	aactaagaac	agagagagtt	tcccatccca	32700
tgtgtctaca	ggggcgatc	tgcttgcgtt	gactcgctgg	gctggccgga	ctcctagagt	32760
tgggtcggtgt	gcttctgtgc	aaaaagtgc	gtcctcttgc	ccatcactgt	gatatctgca	32820
ccagcaagga	aagcctcttt	tcttttcttt	cttttttttt	ttttgagacg	gaacgtcact	32880
gttgtctgcc	tgggcttgag	tgcagtggcg	cgatctcaac	tcactgcaac	ctccgcctcc	32940
cgggttccag	catttctcct	gcctcagcct	cccagcagc	tgagattaca	ggcaccacc	33000
ccctgcgcct	ggctaatttt	tgtattttta	gtagagaggg	gtttttgcca	tgttggccag	33060
gctggctctg	aactcctgac	ctcagggtgat	ccaccacact	cggcctccca	aagtgtctgg	33120
attacagggtg	tgagccatca	cgcccagccg	gaaagcctct	ttttaagggtg	accacctata	33180
gcgcttcccc	aaaataacag	gtcttgtttt	tgcagtaggc	tgcaagcgtc	tcttagcaac	33240
aggagtggcg	tcctgtgggc	tctggggatg	gctgagggtc	gcgtggcagc	catgccttct	33300
gtgtgcacct	ttaggttcca	cggggctatt	ctgctctcac	tgtttgtctg	aaaacgcacc	33360
cttggcatcc	ttgtttggag	agtttctgct	tctcgttggg	catgctgaaa	ctaggggcaa	33420
ggttgtatcc	gttggcgcg	agcggctaca	tgtagggtca	tgagtctttc	accgtggaca	33480
aattccttga	aaaaaaaaaa	aggagtccgg	ttaagcattc	attccgggtc	aagtgtctgg	33540

ttctgtgaat	aaactctaag	atttaagaaa	ccttaatgaa	agaaaacctt	gatgattcag	33600
agcaaggatg	tggtcacacc	tgtggctgga	tctgtttcag	ccgccccagt	gcatggtgag	33660
agtggggagc	agggattgtt	tgttcagagg	tctcatctgg	tatgtttctg	aggtgtttgc	33720
cggctgaatg	gtagacgtgt	cgtttgtgtg	tatgaggttc	tgtgtctgtg	tgtggctcgg	33780
tttgagtgtg	cgcatgtcca	gcacatgccc	tgcccgtctc	tcacctgtgt	cttccccgcc	33840
caggtcctac	gtccagtgcc	aggggagccc	gcagggtctc	atcctctcca	cgctgctctg	33900
cagcctgtgc	tacggcgaca	tggagaacaa	gctgtttgcg	gggattcggc	gggacgggtg	33960
aggcctcctc	tccccaggg	gggcttgggt	gggggttgat	ttgcttttga	tgcattcagt	34020
gttaatatcc	ctggtgctct	ggagaccatg	actgctctgt	cttgaggaa	cagacaaggt	34080
tgcagccctt	tcttggatat	aagccgcacg	ggaggggttg	cacagcctga	ggactgcggg	34140
ctccacgcag	gctctgtcca	gcggccatgt	ccagaggcct	cagggtcag	caggcgggag	34200
ggccgctgcc	ctgcatgatg	agcatgtgaa	ttcaacaccg	aggaagcaca	ccagcttctg	34260
tcacgtcacc	caggttccgt	tagggtcctt	ggggagatgg	ggctgggtgca	gcctgaggcc	34320
ccacatctcc	cagcaggccc	tcgacagggt	gcctggactg	ggcgccctctt	cagccccattg	34380
cccatccac	ttgcatgggg	tctacacca	aggacgcaca	cacctaaata	tcgtgccaac	34440
ctaattgtgt	tcaactcagc	tggcttttat	tgacagcagt	tacttttttt	tttttaatac	34500
tttaagtctt	agggtagatg	tgcacgacgt	gcaggttagt	tacatatgta	tacatgtgcc	34560
atgttgggtg	gctgcacca	ttaactcacc	atttacatta	ggtatatctc	ctaattgctat	34620
ccctccccc	tccccccacc	ccatgacagg	ccctgggtgtg	tgatgttccc	caccctgtgt	34680
ccaagtgttc	tcattgttca	gttcccacct	gtgagtgaga	acatgtgggtg	tttggttttc	34740
tttccttgca	atagtttgct	cagagtgatg	gtttccagct	tcgtccatgt	ccctacaaag	34800
gacatgaact	cacccttttt	tatgactgca	tagtattccg	tgggtgtatat	gtgccacatt	34860
ttcttaatcc	agtctatcat	cgatggacat	ttgggttgggt	tgcaagtctt	tgctactgtg	34920
aatagtgccg	caataaacat	acgtgtgcat	gtgtcttttat	agcagcatga	tttataatcc	34980
tttgggtata	taccagtaaa	tgggatggct	gggtcaaatg	gtattttctag	ttctagatcc	35040
ttgaggaatc	accacactgt	cttccacaat	ggttgaacta	gtttacactc	ccaccaacag	35100
tgtaaaagtg	ttctgggtgct	ggagaggatg	tggacagcag	ttattttttt	atgaaaatag	35160
tatcactgaa	caagcagaca	gttagtgaag	gatgcgtcag	gaagcctgca	ggccacacag	35220
ccatttctct	cgaagactcc	gggtttttcc	tgtgcatctt	ttgaaactct	agctccaatt	35280
atagcatgta	cagtggatca	aggttcttct	tcattaaggt	tcaagttcta	gattgaaata	35340
agtttatgta	acagaaacaa	aaatttcttg	tacacacaac	ttgctctggg	atttggagga	35400
aagtgtcctc	gagctggcgg	cacactgggtc	agccctcttg	gacaggatac	ctctggccca	35460
tggatcatggg	gcgctgggct	tgggcctgag	ggtcacacag	tgcacatgct	ccagcttccct	35520
gtggatagga	tctgggtctc	ggatcatgct	gaggaccaca	gctgccatgc	tggtaaaggg	35580
caccacgtgg	ctcagagggg	gcgaggttcc	cagccccagc	tttcttaccg	tcttcagtta	35640
tttttcccta	agagtctgag	aagtggggcc	gcgcctgatg	gccttcgttc	gtcttcagct	35700
ggcacagaat	tgcacaagct	gatggtaaac	actgagtact	tataatgaat	gaggaattgc	35760
tgtagcagtt	aactgtagag	agctcgtctg	ttggaaagaa	atttaagttt	ttcatttaac	35820
cgttttgag	aatgttactt	tatttatggc	tgtgtaaatt	gtttgacatt	cagtcctctg	35880

tagacagata	ctacgtaaaa	agtgtaaagt	taaccttgct	gtgtattttc	cettatttta	35940
ggctgtcct	gcgtttgggtg	gatgatttct	tgttgggtgac	acctcacctc	accacacgca	36000
aaaccttct	caggtgaggc	ccgtgccgtg	tgtctgtggg	gacctccaca	gcctgtgggc	36060
tttgcagttg	agccccccgt	gtcctgcccc	tggcaccgca	gcgttgtctc	tgccaagtcc	36120
tctctctctg	ccggtgctgg	atccgcaaga	gcagaggcgc	ttggccgtgc	accaggcct	36180
gggggcgcag	gggcaccttc	gggaggaggt	gggtaccgtg	caggccctgg	tcctgcagag	36240
acgcaccag	gttacacacg	tggtgagtg	aggcggtgac	ctggctcctg	ctgctctttg	36300
gaaagtcaag	agtggcggt	cctggggccc	cagttagacc	cccaggagct	gtgcacaggg	36360
cctgcagggc	cgaggcgga	gcctcctccc	cagggtgcac	ctgagcctgc	ggagagcagg	36420
agctgctgag	tgagctggcc	cacagcgttc	gctgcgggtca	cgttcctgcg	tggggttgtt	36480
tgggatcggt	gggagaattt	ggatttgctg	agtgtctgtg	tcttgaacca	cggagatggc	36540
taggagtggg	tttcagagtt	gatttttgtg	aatcaaacta	aaatcaggca	caggggacct	36600
ggcctcagca	caggggattg	tccaatgtgg	tccccctcaa	gggcgccccca	cagagccggt	36660
gggcttgttt	taaagtgcga	tttgacgagg	gacgagaaac	cttgaaagct	gtaaagggaa	36720
ccctcagaaa	atgtggccgc	caggggtggt	ttcaggtgct	ttgctgggct	gtgtttgtga	36780
aaaccatttt	ggaccgcgcc	tccaagtcca	ccctccaggt	ccaccctcca	gggccgcctt	36840
gggctggggg	tatgcctggc	gttccttggtg	ccgcagcccg	gagcacagca	ggctgtgcac	36900
atttaaatcc	actaagattc	actcgggggg	agcccagggtc	ccaagcaact	gagggctcag	36960
gagtcctgag	gctgctgagg	ggacagagca	gacggggaac	gctgcttctg	tgtggcaagt	37020
tcctgagggt	gctggccagg	gaggtggctc	agagtgtatg	ttgggggtccc	accgggggca	37080
gaactctgtc	tctgatgagt	cggcagccat	gtaacaggaa	gggggtggcca	cagggagctg	37140
ggaatgcacc	aggggagctg	cgcagctggc	cgagggtccca	gggccaggcc	acaggaaggg	37200
cagggggacg	cccggggcca	cagcagaggc	cgcaggaagg	gaaggggatg	cccaggccag	37260
agcagaggct	accgggcaca	ggggggctcc	ctgagctggg	tgagcgaggc	tcatgactcg	37320
gcgagggaac	ctccttgacg	tgaagctgac	gactggtgtt	gcccagctca	cagcccagcc	37380
aggtcccgcg	cctgagcagg	aactcagaac	cctccccctt	gtctaaagca	cagcagatgc	37440
cttcagggca	tctaggagaa	aacaggcaaa	gtcgttgaga	aacgtcttaa	aagaaggtgg	37500
gatggtggca	atttcttgct	cagatttttag	tctgcccccg	accacagatg	agtctataac	37560
gggatttggtg	tgttgccatg	gggacacatg	agatggacca	tcacagaggc	cactggggct	37620
gcacctccca	tctgagtcct	ggctgtcccc	ggtccaggcc	aggttcttgc	atgtcacct	37680
acctgtcctg	cccgggagac	agggaagca	ccccgaagtc	tggagcaggg	ctgggtccag	37740
gtcctcaga	gtccttgcca	ggcccagcac	cctgtctcaa	atcaccactt	ctctgggggt	37800
ttccaaagca	tttaacaagg	gtgtcaggtt	acctcctggg	tgacggcccc	gcatcctggg	37860
gctgacattg	cccctctgcc	ttaggacctt	ggtccgaggt	gtccctgagt	atggctgcgt	37920
ggtgaacttg	cggaagacag	tggtgaactt	ccctgtagaa	gacgaggccc	tgggtggcac	37980
ggcttttgtt	cagatgccgg	cccacggcct	attccccctg	tgcggcctgc	tgctggatac	38040
ccggaccctg	gaggtgcaga	gcgactactc	caggtgagcg	cacctggccg	gaagtggagc	38100
ctgtgccccg	ctggggcagg	tgctgctgca	gggccgttgc	gtccacctct	gcttccgtgt	38160
ggggcaggcg	actgccaatc	ccaaagggtc	agaggccaca	gggtgccccct	cgccccatct	38220

ggggctgagc	agaaatgcat	ctttctgtgg	gagtgagggt	gctcacaacg	ggagcagttt	38280
tctgtgctat	tttggtaaaa	ggaaatggtg	caccagacct	gggtgcaactg	aggtgtcttc	38340
agaaagcagt	ctggatccga	acccaagacg	cccgggccct	gctgggcgtg	agtctctcaa	38400
acccaacac	aggggcccctg	ctgggcatga	gtccctctga	acccgagacc	ctggggccct	38460
gctgggcgtg	agtctctccg	aacccagaga	cttcagggcc	cttttgggcg	tgagtctctc	38520
cgctgtgagc	cccacactcc	aaggetcate	cacagtctac	aggatgccat	gagttcatga	38580
tcacgtgtga	cccatcaggg	gacagggcca	tggtgtgggg	gggtctctta	caaaattctg	38640
gggtcttggt	tccccagagc	ccgagagctc	aaggccccgt	ctcaggctca	gacacaaatg	38700
aattgaagat	ggacacagat	gcagaaatct	gtgctgtttc	ttttatgaat	aaaaagtatc	38760
aacattccag	gcagggcaag	gtggctcaca	cctataatcc	cagcactttg	ggaggccgag	38820
gtgggtggat	cacttgaggc	caggagtttg	aggccaacct	aaccaacata	gtgaaattcc	38880
atttctactt	aaaaaataca	aaaattagcc	tggcctgggtg	gcacacgcct	gtagtccccg	38940
ctatgcggga	ggctgaggca	ggagaatcat	ttgaaccag	gaggcagagg	ttgcagtgag	39000
ccgagatcac	accactgcac	tccagcctgg	gcaacagagt	gagacttcat	cttaaaaaaa	39060
aaaaaaaaag	tatcagcatt	ccaaaacct	agtggaacag	tgttttttta	ttctgtcctt	39120
cgataatatt	tactggtgct	gtgctagagg	ccggaactgg	gggtgccttc	ctctgaaagg	39180
cacaccttca	tgggaagaga	aataagtgg	gaatggttgt	taaaccagag	gtttaaactg	39240
gggtcctgtc	gttctgagtt	aacagtccag	atctggactt	tgccctcttc	cagaatgtct	39300
cctggggttt	gcttcatggg	ggagcagcag	gtgtggacac	cctcgtgatg	ggggagcagc	39360
aggtgcagac	gccctcatga	tgggggagtg	gcagggtgcag	acacccttgt	gcatggtgcc	39420
cagcatgtcc	ctgttgagc	tccctcccca	caaggatgcc	ggtctcctgt	gctccccaca	39480
gtccctgctt	ccctctcaca	gccttacctg	gtcctggcct	ccactggctt	tgtctgcatg	39540
atttccacat	ttcctgggct	cccagcacct	cttcgcctct	cccaggcacc	tctgcagtgc	39600
tggccatacc	agtcagctgt	gaactgtcca	ctgcttattt	tgctccccat	gaaatgtatt	39660
ttttaggaca	ggcacccttg	gttccagcct	ctggcacagc	atcagtgaat	gttattgaag	39720
gacaaaggac	agacaaacaa	atcaggaaaa	tgggttctct	ctaaacacat	tgcaaaagcca	39780
cagaggctag	tgcaggatgg	gtgggcatca	ggtcatcaga	tgtgggtcca	atgccagaat	39840
attctgtgct	cccaaaggcc	acttggctcag	agtgtgtgct	tgcagagggtg	gctctaaaag	39900
ctcagcagtg	gaggcagtg	ttcgccatac	tcagggtgaa	ctcacatcct	ctgtgtctga	39960
agtatacagc	agaggcttga	agggcactctg	ggagaagaaa	acaggcaaaa	tgattaagaa	40020
aagtgaaaaa	ggaaaagtgg	taagatggga	attttcttgt	ccagatttta	gtctcccaaa	40080
ccacagctca	gatggtagaa	tgtggtcaga	actgatggac	agaacaatag	aacaaaacgg	40140
aagccctatc	tctcagaaac	gtgtgttaat	gtggtatgtg	gcacagctga	tggaaaagag	40200
agtgtgtgtg	taattttttt	ttctgagaaa	actgactgga	agcaaataag	ttgtgtcttt	40260
acagcatata	ccagagcaga	ttctaggttag	aagaggagac	acatgcaaac	aacaccagca	40320
acagaaataa	aacaaaagac	tcaaagggaa	gggaggtgaa	cgttccctgg	tttggtgttg	40380
gggaaggaca	cacagggagg	cggatgaaac	cagttaggca	acgggcattg	ctttcactgc	40440
agagaaactc	agcttgccctg	agccacagtg	aaaatggcca	ttccctggag	cgtttgtgca	40500
cgtgatttat	ttaaggcgcc	ctgtgaggtc	ctgcacattc	atcctctcac	tttgttctcc	40560

taaccacctg	agaggttagag	gaggaaaggc	tccaggggag	cagccgccct	tggtcaccca	40620
gctggcaaa	ggcatgcatg	attgcagcct	ggcctcctgc	tccggggccc	ttgctctgcc	40680
cgaggacccc	acacaagtca	gacccatagg	ctcaggggtga	gccggagccc	aaggctcgtgt	40740
tggggatggc	tgtgaaagaa	gaaatggacg	tctgatgcac	acttggaag	gtcctaccag	40800
cagcgtcaaa	gaaatgcatg	tgaaactgac	agcgagaccc	atccctcaaa	gaaacgcacg	40860
tgaaactgat	ggcgagacct	gtccccatcc	ctcatgctgg	ctccttttct	gggcttgcca	40920
agagccagca	tcaggttgag	gcaagctgga	aagacttttc	tggaaagcag	cttgtttgca	40980
tggaagtcct	cacaatgtcc	tgtgtcttcc	cagtaattcc	acttctgaag	tgaccagaca	41040
ttatcacggg	tcttatttac	catttccagt	gttccaggca	gggggacttg	ccacagcaag	41100
tcacgaacct	gccc aaatac	agggctaagg	agatattatg	catcacaaaa	cttgctctgc	41160
cattaaacat	ttttcaaaga	atttttgaag	aatgtttaat	ggcacaaaac	gtttattttca	41220
atgtagcagt	gttcaaagct	ggatgtaaaa	gaacacaccc	caggagcctg	ccgtgaatgt	41280
catgtgtgtt	catctttgga	catggacata	catgggcagt	gagtgggtgg	gaggccctgg	41340
aggacatcgg	tgggatgcct	ccatcctgcc	cctctggaga	caccatgtgt	gccacgtgca	41400
ctcactggag	ccctgttttag	ctgggtgccac	ctggctcttc	catccctgag	attcaaacac	41460
agttagattc	cccacgccc	actcagtgtt	ctcccacaaa	aaacctgagt	cacacctgtg	41520
ttcactcgag	ggacgcccgg	gagccagggc	tccacagttt	attatgtgtt	tttggctgag	41580
ttatgtgcag	atctcatcag	ggcagatgat	gagtgcacaa	acacggccgt	gcgaggtttg	41640
gatacactca	acatcactag	ccaggtcctg	gtggagtittg	gtcatgcaga	gtctggatgg	41700
catgtagcat	ttggagtcca	tggagtgage	accagccccc	ctcgggctgc	agcgcatgcc	41760
ccaggcagga	caaggaagcg	ggaggaaggc	aggaggctct	ttggagcaag	ctttgcagga	41820
gggggctggg	tgtggggcag	gcacctgtgt	ctgacattcc	cccctgtgtc	tcagctatgc	41880
ccggacctcc	atcagagcca	gtctcacctt	caaccgcggc	ttcaaggctg	ggaggaacat	41940
gcgtcgcaaa	ctctttgggg	tcttgcggt	gaagtgtcac	agcctgtttc	tggatttgca	42000
ggtgagcagg	ctgatggtea	gcacagagtt	cagagttcag	gagggtgtgtg	cgcaagtatg	42060
tgtgtgtgtg	tgtgcgcgcg	tgcctgcaag	gctgatgggtg	actggctgca	cgtaagagtg	42120
cacatgtacg	catatacacg	tgagcacata	catgtgtgca	tgtgtgtaca	tgaaggcatg	42180
gcagtgtgtg	cacaggtgtg	caagggcaca	agtgtgtgca	catgcgaatg	cacacctgac	42240
atgcatgtgt	gttcgtgcac	agtcgtgtgg	gcattcacgt	gagggtgcatg	cgtgtgggtg	42300
tgcagtgtga	gtagcatgtg	tgcacataac	atgtattgag	gggtcctcgt	gttcaccccg	42360
ctaggctctc	agcaccagtg	ccactcctta	caggatgaga	cgggggtccc	ggccttggtg	42420
ggctgaggct	ctgaagctgc	agccctgagg	gcattgtccc	atctgggcat	ccgcgtccac	42480
tccctctcct	gtgggcttct	gtgtccactc	ccccctcct	gtgggcattt	acatccactc	42540
cactccctct	ctcctgtggg	catccgcgtc	cactccccct	ctctgtgggc	atctgcgtcc	42600
acctccctc	tctgtgggca	tttgcgtcca	ctccctctcc	tggttccttc	ctgtcttgge	42660
cgagcctcgg	gggcaggcag	atgacacaga	gtcttgactc	gcccagggtg	gttcgcagct	42720
gccgggtgag	ggccaggccg	gatttcactg	ggaagaggga	tagtttcttg	tcaaaatgtt	42780
cctctttctt	gttccatctg	aatggatgat	aaagcaaaaa	gtaaaaactt	aaaatcccag	42840
agaggtttct	accgtttctc	actctttctt	ggcgactcta	ggtgaacagc	ctccagacgg	42900

tgtgcaccaa	catctacaag	atcctcctgc	tgcaggcgta	caggtgagcc	gccaccaagg	42960
ggtgcaggcc	cagcctccag	ggaccctccg	cgctctgctc	acctctgacc	cggggcttca	43020
ccttggaaact	cctgggtttt	aggggcaagg	aatgtcttac	gttttcagtg	gtgctgctgc	43080
ctgtgcacag	ttctgttccg	gtggctctgt	gcaaagcacc	tgttctccat	ctctgggtag	43140
tggtaggagc	cgggtgtggc	ccaggtgtcc	ccactgtgcc	tgtgcactgg	ccgtgggacg	43200
tcatggaggc	catcccagg	cagcaggggc	atggggtaaa	gagatgttta	tggggagtct	43260
tagcagagga	ggctgggaag	gtgtctgaac	agtagatggg	agatcagatg	cccggaggat	43320
ttggggctctc	agcaaagagg	gccgaggtgg	gtgcaggtga	gggtcgctgg	ccccaccccc	43380
gggaaggtgc	agcagagctg	tggctcccca	cacagcccgg	ccagcacctg	tgctctgggc	43440
atggctgtgc	tcctggaacg	ttccctgtcc	tggctgggtca	gggggtgccc	ctgccaagaa	43500
tcgacaactt	tatcacagag	ggaagggcca	atctgtggag	gccacagggc	cagcttctgc	43560
ctggagtcag	ggcaggtggt	ggcacaagcc	tggggctgt	accaaagggc	agtcgggcac	43620
cacaggcccc	ggcctccacc	tcaacaggcc	tcccagagcca	ctgggagctg	aatgccagga	43680
ggccgaagcc	ctcgcccat	gagggctgag	aaggagtgtg	agcatttgtg	ttaccagg	43740
ccgaggctgc	gcgaattacc	gtgcacactt	gatgtgaaat	gaggtcgctg	tctatcgtgg	43800
aaaccagca	agggctcag	ggagagtttt	ccattacaag	gtcgtaccat	gaaaatggtt	43860
tttaaccga	gtgcttgcgc	cttcatgctc	tggcaggag	ggcagagcca	cagctgcatg	43920
ttaccgcctt	tgcaccagct	ccagaggctt	gggaccaggc	tgtctcagtt	ccagggtgcg	43980
tccggctcag	accgccctcc	tctctgcctt	ctctctctgc	ctcaaactct	ccctcgtttg	44040
catctccctg	acgcgtgcct	gggccctcgt	gcaagctgct	tgactccttt	ccggaaaccc	44100
ttgggggtgtg	ctggatacag	gtgccactga	ggactggagg	tgtctgacac	tgtggttgac	44160
cccagggtcc	agctggcgctg	cttggggcct	ccttgggcca	tgatgaggtc	agaggagttt	44220
tcccaggtga	aaactcctgg	gaaactccca	gggcatgtg	acctgccacc	tgctcctccc	44280
atattcagct	cagtcttgtc	ctcatttccc	caccagggtc	tctagctccg	aggagctccc	44340
gtagagggcc	tgggctcagg	gcagggcggc	tgagtttccc	cacccatgtg	gggacccttg	44400
ggtagtcgct	tgattgggta	gccctgagga	ggccgagatg	cgatgggcca	cgggccgttt	44460
ccaaacacag	agtcaggcac	gtggaaggcc	caggaatccc	cttccctcga	ggcaggagtg	44520
ggagaacgga	gagctgggccc	ccgatttcac	ggcagccagg	ctgcagtggg	cgaggctgtg	44580
gtgggtccacg	tggcgctggg	ggcggggctt	gattcaaatc	cgctggggct	cggccttcct	44640
ggcccgtgct	ggccgcgcct	ccacacgggc	ttgggggtgga	cgccccgacc	tctagcaggt	44700
ggctattttct	ccctttggaa	gagagccctt	cacccatgct	aggtgtttcc	ctcctgggtc	44760
aggagcgtgg	ccgtgtggca	accccgggac	cttaggctta	tttatttgtt	taaaaacatt	44820
ctgggccttg	cttccgttgt	tgctaaatgg	ggaaaagaca	tcccacctca	gcagagttac	44880
tgagaggctg	aaaccggggt	gctggcttga	ctgggtgat	ctcaggctcat	tccagaagtg	44940
gctcaggaag	tcagtgagac	caggtacatg	gggggctcag	gcagtgggtg	agatgaggta	45000
cacggggggc	tcaggcagtg	ggtgaggcca	ggtacatggg	gggctcaggc	actgggtgag	45060
atgaggtaca	cggggggctc	aggcagagg	tcagaccagg	tacacggggg	ctctgatcac	45120
acgcacatat	gagcacatgt	gcacatgtgc	tgtttcatgg	tagccaggtc	tgtgcacacc	45180
tgccccaag	tcccaggaag	ctgagaggcc	aaagatggag	gctgacagg	ctggcgcggt	45240

ggctcacacc	tgtagtccca	gcactttggg	aggccgaggc	gagaggatcc	cttgagccca	45300
ggagtttaag	accagcctga	gcaacatagt	agaaccccat	ctctatgaaa	aataaaaaca	45360
aaaattagct	gaacatggtg	gtgtgcgcct	gtagtcccaa	tacttgggag	gctgaagtgg	45420
gaggatcact	tgagcccagg	aggtggaagc	tgcagtgagc	tgagattgca	ccactgtact	45480
gcagcctggg	tgacagagtg	agagcccate	tcaacaacaa	caaagaagac	tgacaaatgc	45540
agttttcttg	aaagaaacat	ttagtaggaa	cttaacctac	acacagaagc	caagtcgggtg	45600
tctcggtgtc	agtgagatga	gatgatgggt	cctcacacca	tcaccccaga	cccagggttt	45660
atgcaccaca	ggggcggtg	gctcagaagg	gatgcgcagg	acgttgatat	acgatgacat	45720
caaggttgct	tgacgaaggg	caggattcat	gataagtacc	tgctggtaca	caaggaacaa	45780
tggataaact	ggaacacctt	gaggccttcc	cggaacaggg	gctaatacaga	agccagcatg	45840
gggggctggc	atccaggatg	gagctgcttc	agcctccaca	tgcgtgttca	tacagatggt	45900
gcacagaaac	gcagtgtacc	tgtgcacaca	cagacacgca	gctactcgca	cacacaagca	45960
cacacacaga	catgcatgca	tgcattcggt	tgtgtgcacc	tgtgccccatg	aggaaaccca	46020
tgcattgtga	ttcatgcacg	cacacaggca	cgggtggggc	catgcccaca	cccacgagca	46080
ccgtctgatt	aggaggcctt	tcctctgacg	ctgtccgcca	tcctctcagg	tttcacgcat	46140
gtgtgctgca	gctcccattt	catcagcaag	tttggaagaa	ccccacattt	ttcctgcgcg	46200
tcattctctga	cacggcctcc	ctctgtctact	ccatcctgaa	agccaagaac	gcaggatatgt	46260
gcagggtgct	ggcctcagtg	gcagcagtg	ctgcctgctg	gtgttagtgt	gtcaggagac	46320
tgagtgaatc	tgggcttagg	aagtctctac	cccttttcgc	atcaggaagt	ggtttaaccc	46380
aaccactgtc	aggctcgtct	gcccgccttc	tcgtgggggtg	agcagagcac	ctgatggaag	46440
ggacaggagc	tgtctgggag	ctgccatcct	tcccaccttg	ctctgcctgg	ggaagcgtg	46500
gggggcctgg	tctctcctgt	ttgccccatg	gtgggatttg	gggggcctgg	cctctcctgt	46560
ttgccctgtg	gtgggattgg	gctgtctccc	gtccatggca	cttagggccc	ttgtgcaaac	46620
ccaggccaag	ggcttaggag	gaggccaggc	ccaggctacc	ccacccctct	caggagcaga	46680
ggccgcgtat	caccacgaca	gagccccgcg	ccgtcctctg	cttcccagtc	accgtcctct	46740
gcccctggac	actttgtcca	gcatcaggga	ggtttctgat	ccgtctgaaa	ttcaagccat	46800
gtcgaacctg	cggctctgag	cttaacagct	tctactttct	gttctttctg	tgttgtggaa	46860
atttcacctg	gagaagccga	agaaaacatt	tctgtcgtga	ctcctgcggt	gcttgggtcg	46920
ggacagccag	agatggagcc	accccgagca	ccgtcgggtg	tgggcagctt	tccggtgtct	46980
cctgggaggg	gagctgggct	gggcctgtga	ctcctcagcc	tctgttttcc	cccagggatg	47040
tcgttggggg	ccaagggcgc	cgccggccct	ctgccctccg	aggccgtgca	gtggctgtgc	47100
caccaagcat	tcctgtctca	gctgactcga	caccgtgtca	cctacgtgcc	actcctgggg	47160
tcactcagga	caggcaagtg	tgggtggagg	ccagtgcggg	ccccacctgc	ccaggggtca	47220
tccttgaacg	ccctgtgtgg	ggcgagcagc	ctcagatgct	gctgaagtgc	agacgcccc	47280
gggcctgacc	ctgggggcct	ggagccacgc	tggcagccct	atgtgattaa	acgttggtgt	47340
ccccaggcca	cggagcctgg	cagggtcccc	aacttcttga	acccctgctt	cccatctcag	47400
gggcgatggc	tccccacgct	tgggagcctt	ctgacccctg	acctgtgtcc	tctcacagcc	47460
tcttccctgg	ctgctgccct	gagctcctgg	ggctcctgagc	aagtctctct	cccggccccg	47520
cgtccagcgc	tcactgggct	gcctgtctgc	tcgccccggt	ggaggggtgt	ctgtcccttc	47580

actgaggttc	ccaccagcca	gggccacgag	gtgcaggccc	tgcctgcccc	gccaccacaca	47640
cgctctagga	gggtttgagg	atgccacctc	tggcctcttc	tggaacggag	tctgattttg	47700
gccccgcage	ccagacgcag	ctgagtcgga	agctccccggg	gacgacgctg	actgccctgg	47760
aggccgcage	caacccggca	ctgccctcag	acttcaagac	catcctggac	tgatggccac	47820
ccgcccacag	ccaggccgag	agcagacacc	agcagccctg	tcacgccggg	ctctacgtcc	47880
cagggaggga	ggggcgggcc	acaccagggc	ccgcaccgct	gggagtctga	ggcctgagtg	47940
agtgtttggc	cgaggcctgc	atgtccggct	gaaggctgag	tgtccggctg	aggcctgagc	48000
gagtgtccag	ccaagggtg	agtgtccagc	acacctgccg	tcttcaacttc	cccacaggct	48060
ggcgctcggc	tccaccccag	ggccagcttt	tcctcaccag	gagccccggct	tccactcccc	48120
acataggaat	agtccatccc	cagattcgcc	attgttcacc	cctcgccctg	ccctcctttg	48180
ccttccaccc	ccaccatcca	ggtggagacc	ctgagaagga	ccctggggagc	tctgggaatt	48240
tggagtgacc	aaaggtgtgc	cctgtacaca	ggcgaggacc	ctgcacctgg	atgggggtcc	48300
ctgtgggtca	aattgggggg	aggtgctgtg	ggagtaaaat	actgaatata	tgagtttttc	48360
agttttgaaa	aaaatctcat	gtttgaatcc	taatgtgcac	tgcatagaca	ccactgtatg	48420
caattacaga	agcctgtgag	tgaacggggg	ggtggtcagt	gcgggccccat	ggcctggctg	48480
tgcattttacg	gaagtctatg	agtgaatggg	gttgtgggtca	gtgcggggccc	atggcctggc	48540
tgggcctggg	aggtttctga	tgtgtgagg	caggagggga	aggagggtag	gggatagaca	48600
gtgggagccc	ccaccctgga	agacataaca	gtaagtccag	gcccgaaggg	cagcagggat	48660
gctggggggc	cagcttgggc	ggcggggatg	atggaggggc	tggccagggt	ggcagggatg	48720
atggggggccc	cagctggggg	ggcaggggtg	atgggggggg	ctgggtctggg	tggcggggaa	48780
gatggggaag	cctggctggg	ccccctcctc	ccctgcctcc	cacctgcagc	cgtggatccg	48840
gatgtgcttc	cctgggtgcac	atcctctggg	ccatcagctt	tcatggagggt	ggggggcagg	48900
ggcatgacac	catcctgtat	aaaatccagg	attcctcctc	ctgaacgccc	caactcaggt	48960
tgaagtcac	attccgcctc	tggccattct	cttaagagta	gaccaggatt	ctgatctctg	49020
aagggtgggt	aggggtgggg	agtggagggt	gtggacacag	gaggcttcag	ggtggggctg	49080
gtgatgtctt	ctcatcctct	tatcatctcc	cagtctcatt	tctcatcctc	ttatcatctc	49140
ccagtctcat	ctgtcttctt	cttatctccc	agtctcatct	gtcatcctct	taccatctcc	49200
cagtctcatt	tcttatectc	ttatctccta	gtctcatcca	gacttacctc	ccagggcggg	49260
tgccaggctc	gcagtggagc	tggacatacg	tccttcctca	ggcagaagga	actggaagga	49320
ttgcagagaa	caggaggggc	ggctcagagg	gacgcagctt	tgggggtgaag	aaacagcccc	49380
tcctcagaag	ttggcttggg	ccacacgaaa	ccgaggggccc	tgcgtgagtg	gctccagagc	49440
cttccagcag	gtcccttggtg	gggccttatg	gtatggccgg	gtcctactga	gtgcaccttg	49500
gacagggctt	ctggtttgag	tgcagcccg	acgtgcctgg	tgtcgggggtg	ggggcttatg	49560
gccactggat	atggcgctcat	ttattgtctg	tgttcagag	aatgtctgag	tgaccgagcc	49620
taatgtgtat	ggtgggccc	agtccacaga	ctgtgtcgta	aatgcactct	ggtgcctgga	49680
gccccgtat	aggagctgtg	aggaaggagg	ggctcttggc	agccggcctg	ggggcgccct	49740
tgccttgcaa	actggaagg	agcggccccg	ggcgccgtgg	gcggacgacc	tcaagtgaga	49800
ggttggacag	aacaggggcg	ggacttccca	ggagcagagg	ccgctgctca	ggcacacctg	49860
ggtttgaatc	acagaccaac	aggtcaggcc	attgttcagc	tatccatctt	ctacaaagct	49920

```

ccagattcct gtttctccgg gtgttttttg ttgaaatttt actcaggatt acttatattt 49980
tttgctaaag tattagaccc ttaaaaaagg tatttgcttt gatatggctt aactcactaa 50040
gcacctactt tatttgcttg tttttattta ttattattat tattattaga gatgggtgtct 50100
actctgtcac ccaggttggt agtgcagtgg cacagtcatt gctcgctgta gccgcaaacc 50160
cccaggetca agtgatectc eggcctcagc ttcccagagt gctgggatta caggtgtgag 50220
ccactgccct tgcctggcac ttttaaaaac cactatgtaa ggtcaggctc agtggcttcc 50280
acacctgtca tcccagtagt ttgggaagcc gaggcagaag gattgtctga ggccaggagt 50340
ttgagaccag catgggtaac ataggagagc cccatctcta caaaaaatgc aaaaagtatt 50400
ccgggcgtgg ggtccagcat ctgtagtccc agctgctcgg gaggtgagt gggaggatcg 50460
cttgagcccc ggaggtcatt gctgcagtga gctgtgattg taccatcgca ctccagcctg 50520
ggcaacagag tgagaccctg tctcaaaaaa aaaaaaaaaa aaagaaggag aaggagaaga 50580
gaagaagaag gaagaaggaa agagaagaag aaggagaaga gaagaaagaa ggagaaggag 50640
gcctgctagg tgctaggtag actgtcaaat ctccagcaa aatgaaaata acaaagtttt 50700
aaagggaag aaaaaccca gctctttgga ctcccttagg cctgaacttc atctcaagca 50760
gcttccttcc acagacaagc gtgtatggag cgagtgagtt caaagcagaa agggaggaga 50820
agcaggcaag ggtggaggct gtgggtgaca ccagccagga cccctgaaag ggagtgggtg 50880
ttttcctgcc tcagccccc gctcctgccg gtcctgcacc tgctgtaacc gtcgatgttg 50940
gtgccagggtg cccacctggg aaggatgctg tgcagggggc ttgccaaact ttggtgggtt 51000
tcagaagccc caggcacttg tggcaggcac aattacagcc cctccccaaa gatgcccacg 51060
tccttctcct ggaacctgtg aatgtgtcac ccgcaaggca gaggtggtg aaggctgcag 51120
gtggaatcac ggctgccagt cagccgatct taaggtcatt ctggattatc tgggtggcct 51180
gatatggcca caagggtccc tagaagttag agagggaggc aggggagagt cagagagggg 51240
acgtgagaag gaccactggc cactgctggc tttgagatgg aggaggggggt ccccagccaa 51300
ggaatggggg cagccgctcc atgctggaaa agcaagcaat cctccccggt cctgagggca 51360
cacggccctg cccacgcctc gatttcaggc cagtgggacc tgtttcagct ttccggcctc 51420
cagagctgta agatgatgcg tttgtgttca gccactaagc tgcagtgatt cgtcacagca 51480
gcaaatggaa tagcagtaca gggaaatgaa tacagggaca gttctcagag tgactctcag 51540
cccacccctg gg 51552

```

<210>4

<211>1132

<212>PRT

<213> 人

<400>4

Met Pro Arg Ala Pro Arg Cys Arg Ala Val Arg Ser Leu Leu Arg Ser

1

5

10

15

His Tyr Arg Glu Val Leu Pro Leu Ala Thr Phe Val Arg Arg Leu Gly

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

Asp Lys Glu Gln Leu Arg Pro Ser Phe Leu Leu Ser Ser Leu Arg Pro		
340	345	350
Ser Leu Thr Gly Ala Arg Arg Leu Val Glu Thr Ile Phe Leu Gly Ser		
355	360	365
Arg Pro Trp Met Pro Gly Thr Pro Arg Arg Leu Pro Arg Leu Pro Gln		
370	375	380
Arg Tyr Trp Gln Met Arg Pro Leu Phe Leu Glu Leu Leu Gly Asn His		
385	390	395
Ala Gln Cys Pro Tyr Gly Val Leu Leu Lys Thr His Cys Pro Leu Arg		
405	410	415
Ala Ala Val Thr Pro Ala Ala Gly Val Cys Ala Arg Glu Lys Pro Gln		
420	425	430
Gly Ser Val Ala Ala Pro Glu Glu Glu Asp Thr Asp Pro Arg Arg Leu		
435	440	445
Val Gln Leu Leu Arg Gln His Ser Ser Pro Trp Gln Val Tyr Gly Phe		
450	455	460
Val Arg Ala Cys Leu Arg Arg Leu Val Pro Pro Gly Leu Trp Gly Ser		
465	470	475
Arg His Asn Glu Arg Arg Phe Leu Arg Asn Thr Lys Lys Phe Ile Ser		
485	490	495
Leu Gly Lys His Ala Lys Leu Ser Leu Gln Glu Leu Thr Trp Lys Met		
500	505	510
Ser Val Arg Asp Cys Ala Trp Leu Arg Arg Ser Pro Gly Val Gly Cys		
515	520	525
Val Pro Ala Ala Glu His Arg Leu Arg Glu Glu Ile Leu Ala Lys Phe		
530	535	540
Leu His Trp Leu Met Ser Val Tyr Val Val Glu Leu Leu Arg Ser Phe		
545	550	555
Phe Tyr Val Thr Glu Thr Thr Phe Gln Lys Asn Arg Leu Phe Phe Tyr		
565	570	575
Arg Lys Ser Val Trp Ser Lys Leu Gln Ser Ile Gly Ile Arg Gln His		
580	585	590
Leu Lys Arg Val Gln Leu Arg Glu Leu Ser Glu Ala Glu Val Arg Gln		
595	600	605
His Arg Glu Ala Arg Pro Ala Leu Leu Thr Ser Arg Leu Arg Phe Ile		
610	615	620
Pro Lys Pro Asp Gly Leu Arg Pro Ile Val Asn Met Asp Tyr Val Val		
625	630	635
Gly Ala Arg Thr Phe Arg Arg Glu Lys Arg Ala Glu Arg Leu Thr Ser		

					645						650						655	
Arg	Val	Lys	Ala	Leu	Phe	Ser	Val	Leu	Asn	Tyr	Glu	Arg	Ala	Arg	Arg			
					660						665						670	
Pro	Gly	Leu	Leu	Gly	Ala	Ser	Val	Leu	Gly	Leu	Asp	Asp	Ile	His	Arg			
					675						680						685	
Ala	Trp	Arg	Thr	Phe	Val	Leu	Arg	Val	Arg	Ala	Gln	Asp	Pro	Pro	Pro			
					690						695						700	
Glu	Leu	Tyr	Phe	Val	Lys	Val	Asp	Val	Thr	Gly	Ala	Tyr	Asp	Thr	Ile			
705						710						715						720
Pro	Gln	Asp	Arg	Leu	Thr	Glu	Val	Ile	Ala	Ser	Ile	Ile	Lys	Pro	Gln			
					725						730						735	
Asn	Thr	Tyr	Cys	Val	Arg	Arg	Tyr	Ala	Val	Val	Gln	Lys	Ala	Ala	His			
					740						745						750	
Gly	His	Val	Arg	Lys	Ala	Phe	Lys	Ser	His	Val	Ser	Thr	Leu	Thr	Asp			
					755						760						765	
Leu	Gln	Pro	Tyr	Met	Arg	Gln	Phe	Val	Ala	His	Leu	Gln	Glu	Thr	Ser			
					770						775						780	
Pro	Leu	Arg	Asp	Ala	Val	Val	Ile	Glu	Gln	Ser	Ser	Ser	Leu	Asn	Glu			
785						790						795						800
Ala	Ser	Ser	Gly	Leu	Phe	Asp	Val	Phe	Leu	Arg	Phe	Met	Cys	His	His			
					805						810						815	
Ala	Val	Arg	Ile	Arg	Gly	Lys	Ser	Tyr	Val	Gln	Cys	Gln	Gly	Ile	Pro			
					820						825						830	
Gln	Gly	Ser	Ile	Leu	Ser	Thr	Leu	Leu	Cys	Ser	Leu	Cys	Tyr	Gly	Asp			
					835						840						845	
Met	Glu	Asn	Lys	Leu	Phe	Ala	Gly	Ile	Arg	Arg	Asp	Gly	Leu	Leu	Leu			
					850						855						860	
Arg	Leu	Val	Asp	Asp	Phe	Leu	Leu	Val	Thr	Pro	His	Leu	Thr	His	Ala			
865						870						875						880
Lys	Thr	Phe	Leu	Arg	Thr	Leu	Val	Arg	Gly	Val	Pro	Glu	Tyr	Gly	Cys			
					885						890						895	
Val	Val	Asn	Leu	Arg	Lys	Thr	Val	Val	Asn	Phe	Pro	Val	Glu	Asp	Glu			
					900						905						910	
Ala	Leu	Gly	Gly	Thr	Ala	Phe	Val	Gln	Met	Pro	Ala	His	Gly	Leu	Phe			
					915						920						925	
Pro	Trp	Cys	Gly	Leu	Leu	Leu	Asp	Thr	Arg	Thr	Leu	Glu	Val	Gln	Ser			
					930						935						940	
Asp	Tyr	Ser	Ser	Tyr	Ala	Arg	Thr	Ser	Ile	Arg	Ala	Ser	Leu	Thr	Phe			
945						950						955						960

Asn Arg Gly Phe Lys Ala Gly Arg Asn Met Arg Arg Lys Leu Phe Gly		
965	970	975
Val Leu Arg Leu Lys Cys His Ser Leu Phe Leu Asp Leu Gln Val Asn		
980	985	990
Ser Leu Gln Thr Val Cys Thr Asn Ile Tyr Lys Ile Leu Leu Leu Gln		
995	1000	1005
Ala Tyr Arg Phe His Ala Cys Val Leu Gln Leu Pro Phe His Gln		
1010	1015	1020
Gln Val Trp Lys Asn Pro Thr Phe Phe Leu Arg Val Ile Ser Asp		
1025	1030	1035
Thr Ala Ser Leu Cys Tyr Ser Ile Leu Lys Ala Lys Asn Ala Gly		
1040	1045	1050
Met Set Leu Gly Ala Lys Gly Ala Ala Gly Pro Leu Pro Ser Glu		
1055	1060	1065
Ala Val Gln Trp Leu Cys His Gln Ala Phe Leu Leu Lys Leu Thr		
1070	1075	1080
Arg His Arg Val Thr Tyr Val Pro Leu Leu Gly Set Leu Arg Thr		
1085	1090	1095
Ala Gln Thr Gln Leu Set Arg Lys Leu Pro Gly Thr Thr Leu Thr		
1100	1105	1110
Ala Leu Glu Ala Ala Ala Asn Pro Ala Leu Pro Set Asp Phe Lys		
1115	1120	1125
Thr Ile Leu Asp		
1130		

<210>5

<211>6492

<212>DNA

<213> 人

<400>5

tttctgtgaa gcagaagtct gggaatcgat ctggaaatcc tcctaatttt tactccctct	60
ccccgcgact cctgattcat tgggaagttt caaatcagct ataactggag agtgctgaag	120
attgatggga tcgttgcctt atgcatttgt ttgtgtttta caaaaaggaa acttgacaga	180
ggatcatgct gtacttaaaa aataacaacat cacagaggaa gtagactgat attaacaata	240
cttactaata ataacgtgcc tcatgaaata aagatccgaa aggaattgga ataaaaattt	300
cctgcatctc atgccaaggg ggaaacacca gaatcaagtg ttccgcgtga ttgaagacac	360
cccctcgtcc aagaatgcaa agcacatcca ataaaatagc tggattataa ctcctcttct	420
ttctctgggg gccgtggggg gggagctggg gcgagaggtg ccgttggccc ccgttgcctt	480

tcctctggga	aggatggcgc	acgctgggag	aacagggtac	gataaccggg	agatagtgat	540
gaagtacatc	cattataagc	tgtegcagag	gggctacgag	tgggatgcgg	gagatgtggg	600
cgccgcgccc	ccgggggccc	ccccgcacc	gggcatcttc	tcctcccagc	ccgggcacac	660
gccccatcca	gcccgatecc	gggacccggg	cgccaggacc	tcgccgctgc	agaccccggc	720
tgcccccggc	gcccgcgcgg	ggcctgcgct	cagcccgggtg	ccacctgtgg	tccacctgac	780
cctccgccag	gcccggcgacg	acttctcccc	ccgctaccgc	cgcgacttcg	ccgagatgtc	840
cagccagctg	cacctgacgc	ccttcaccgc	gcggggacgc	tttgccacgg	tgggtggagga	900
gctcttcagg	gacgggggtga	actgggggag	gattgtggcc	ttctttgagt	tcgggtggggt	960
catgtgtgtg	gagagcgtca	accggggagat	gtcgccccctg	gtggacaaca	tcgccctgtg	1020
gatgactgag	tacctgaacc	ggcacctgca	cacctggatc	caggataacg	gaggctggga	1080
tgcctttgtg	gaactgtacg	gccccagcat	gcggcctctg	tttgatttct	cctggctgtc	1140
tctgaagact	ctgctcagtt	tggccctggg	gggagcttgc	atcacccctg	gtgcctatct	1200
gggccacaag	tgaagtcaac	atgcctgccc	caaacaata	tgcaaaagg	tcactaaagc	1260
agtagaaata	atatgcattg	tcagtgatgt	accatgaaac	aaagctgcag	gctgtttaag	1320
aaaaaataac	acacatataa	acatcacaca	cacagacaga	cacacacaca	cacaacaatt	1380
aacagtcttc	aggcaaaacg	tcgaatcagc	tatttactgc	caaagggaaa	tatcatttat	1440
tttttacatt	attaagaaaa	aaagatttat	ttatttaaga	cagtcccatc	aaaactcctg	1500
tcttttgaaa	tccgaccact	aattgcccaag	caccgcttcg	tgtggctcca	cctggatgtt	1560
ctgtgcctgt	aaacatagat	tcgctttcca	tgttgttggc	cggatcacca	tctgaagagc	1620
agacggatgg	aaaaaggacc	tgatcattgg	ggaagctggc	tttctggctg	ctggaggctg	1680
gggagaagg	gttcattcac	ttgcatttct	ttgccctggg	ggctgtgata	ttaacagagg	1740
gagggttcc	gtggggggaa	gtccatgcct	ccctggcctg	aagaagagac	tctttgcata	1800
tgactcacat	gatgcatacc	tgggtgggagg	aaaagagttg	ggaacttcag	atggacctag	1860
taccactga	gatttccacg	ccgaaggaca	gcgatgggaa	aaatgccctt	aatcatagg	1920
aaagtatttt	tttaagctac	caattgtgcc	gagaaaagca	ttttagcaat	ttatacaata	1980
tcatccagta	ccttaagccc	tgattgtgta	tattcatata	ttttggatac	gcacccccca	2040
actcccaata	ctggctctgt	ctgagtaaga	aacagaatcc	tctggaactt	gaggaagtga	2100
acatttcgg	gacttccgca	tcaggaaggc	tagagttacc	cagagcatca	ggccgccaca	2160
agtgcctgct	tttaggagac	cgaagtccgc	agaacctgcc	tgtgtcccag	cttggaggcc	2220
tggtcctgga	actgagccgg	ggccctcact	ggcctcctcc	agggatgatc	aacagggcag	2280
tgtggtctcc	gaatgtctgg	aagctgatgg	agctcagaat	tccactgtca	agaaagagca	2340
gtagaggggt	gtggctgggc	ctgtcaccct	ggggccctcc	aggtaggccc	gttttcacgt	2400
ggagcatggg	agccacgacc	cttcttaaga	catgtatcac	tgtagaggga	aggaacagag	2460
gccctgggcc	cttcctatca	gaaggacatg	gtgaaggctg	ggaacgtgag	gagaggcaat	2520
ggccacggcc	cattttggct	gtagcacatg	gcacgttggc	tgtgtggcct	tggcccacct	2580
gtgagtttaa	agcaaggctt	taaatgactt	tggagagggt	cacaaatcct	aaaaagaagca	2640
ttgaagttag	gtgtcatgga	ttaattgacc	cctgtctatg	gaattacatg	taaaacatta	2700
tcttgtcact	gtagtttgg	tttatttgaa	aacctgacaa	aaaaaaagtt	ccagggtgtg	2760
aatatggggg	ttatctgtac	atcctggggc	attaaaaaaa	aatcaatgg	tggggaacta	2820

taaagaagta	acaaaagaag	tgacatcttc	agcaaataaa	ctaggaaatt	tttttttctt	2880
ccagtttaga	atcagccttg	aaacattgat	ggaataactc	tgtggcatta	ttgcattata	2940
taccatttat	ctgtattaac	tttggaatgt	actctgttca	atgtttaatg	ctgtggttga	3000
tatttcgaaa	gctgctttta	aaaaatacat	gcctctcagc	gtttttttgt	ttttaattgt	3060
atttagttat	ggcctataca	ctatttgtga	gcaaagggtga	tcgttttctg	tttgagattt	3120
ttatctcttg	attcttcaaa	agcattctga	gaagggtgaga	taagccctga	gtctcagcta	3180
cctaagaaaa	acctggatgt	cactggccac	tgaggagcct	tgtttcaacc	aagtcattgt	3240
catttccacg	tcaacagaat	tgtttattgt	gacagttata	tctgttgctc	ctttgacctt	3300
gtttcttgaa	ggtttctctg	tccctgggca	attccgcatt	taattcatgg	tattcaggat	3360
tacatgcatg	tttggttaaa	cccatgagat	tcattcagtt	aaaaatccag	atggcaaatg	3420
accagcagat	tcaaacttat	ggtgggtttga	cctttagaga	gttgctttac	gtggcctgtt	3480
tcaacacaga	cccaccaga	gccctcctgc	cctccttccg	cgggggcttt	ctcatggctg	3540
tccttcaggg	tcttctgaa	atgcagtggg	gcttacgctc	caccaagaaa	gcaggaaacc	3600
tgtggtatga	agccagacct	ccccggcggg	cctcagggaa	cagaatgatc	agacctttga	3660
atgattctaa	tttttaagca	aaatattatt	ttatgaaagg	tttacattgt	caaagtgatg	3720
aatatggaat	atccaatcct	gtgctgctat	cctgccaaaa	tcattttaat	ggagtcagtt	3780
tgcagtatgc	tccacgtggg	aagatcctcc	aagctgcttt	agaagtaaca	atgaagaacg	3840
tggacgtttt	taatataaag	cctgttttgt	cttttgttgt	tgttcaaacg	ggattcacag	3900
agtatttgaa	aaatgtatat	atattaagag	gtcacggggg	ctaattgctg	gctggctgcc	3960
ttttgctgtg	gggttttgtt	acctggtttt	aataacagta	aatgtgcca	gcctcttggc	4020
cccagaactg	tacagtattg	tggctgcact	tgctctaaga	gtagttgatg	ttgcattttc	4080
cttattgtta	aaaacatgtt	agaagcaatg	aatgtatata	aaagcctcaa	ctagtcattt	4140
ttttctcttc	ttcttttttt	tcattatata	taattatttt	gcagttgggc	aacagagaac	4200
catccctatt	ttgtattgaa	gagggattca	catctgcata	ttaactgctc	tttatgaatg	4260
aaaaaacagt	cctctgtatg	tactcctctt	tacactggcc	agggtcagag	ttaaatagag	4320
tatatgcaat	ttccaaattg	gggacaaggg	ctctaaaaaa	agccccaaaa	ggagaagaac	4380
atctgagaac	ctcctcggcc	ctcccagtc	ctcgctgcac	aaatactccg	caagagaggc	4440
cagaatgaca	gctgacaggg	tctatggcca	tcgggtcgct	tccgaagatt	tggcaggggc	4500
agaaaactct	ggcaggctta	agattttgaa	taaagtcaca	gaattaagga	agcacctcaa	4560
tttagttcaa	acaagacgcc	aacattctct	ccacagctca	cttacctctc	tgtgttcaga	4620
tgtggccttc	catttatatg	tgatctttgt	tttattagta	aatgcttata	atctaaagat	4680
gtagctctgg	cccagtggga	aaaattagga	agtgattata	aatcgagagg	agttataata	4740
atcaagatta	aatgtaaata	atcagggcaa	tcccaacaca	tgtctagctt	tcacctccag	4800
gatctattga	gtgaacagaa	ttgcaaatag	tctctatttg	taattgaact	tatcctaaaa	4860
caaatagtgt	ataaatgtga	acttaaacct	taattaattc	caactgtact	tttaaggcag	4920
tggctgtttt	tagactttct	tatcacttat	agtttagta	gtacacctac	tctatcagag	4980
aaaaacagga	aaggctcgaa	atacaagcca	ttctaaggaa	attagggagt	cagttgaaat	5040
tctattctga	tcttattctg	tgggtgtctt	tgcagcccag	acaaatgtgg	ttacacactt	5100
tttaagaaat	acaattctac	attgtcaagc	ttatgaaggt	tccaatcaga	tctttattgt	5160

tattcaattt ggatctttca gggatTTTT ttttaaatta ttatgggaca aaggacattt	5220
gttggagggg tgggaggag gaagaatttt taaatgtaaa acattcccaa gtttggatca	5280
gggagtTgga agttttcaga ataaccagaa ctaagggtat gaaggacctg tattggggtc	5340
gatgtgatgc ctctgcgaag aaccttTgtg gacaaatgag aaacattttg aagtttTgtg	5400
tacgaccttt agattccaga gacatcagca tggetcaaag tgcagctccg tttggcagtg	5460
caatggTata aatttcaagc tggatatgtc taatgggtat ttaaacaata aatgtgcagt	5520
tttaactaac aggatatTTa atgacaacct tctggttggT agggacatct gtttctaaat	5580
gtttattatg tacaatacag aaaaaattt tataaaatta agcaatgtga aactgaattg	5640
gagagtGata atacaagtcc tttagtctta ccagtgaaT cattctgttc catgtctttg	5700
gacaaccatg acctTggaca atcatgaaat atgcatctca ctggatgcaa agaaaatcag	5760
atggagcatg aatggTactg taccggTtca tctggactgc ccagaaaaa taacttcaag	5820
caaacatcct atcaacaaca aggtTgttct gcataccaag ctgagcacag aagatgggaa	5880
cactggTgga ggatggaaag gctcgctcaa tcaagaaaat tctgagacta ttaataaata	5940
agactgtagt gtagatactg agtaaTcca tgcacctaaa cctttTggaa aatctgccgt	6000
gggccctcca gatagctcat ttcattaagt tttccctcc aaggtagaat ttgcaagagt	6060
gacagtggat tgcatttctt ttggggaagc tttctttTgg tggttttgtt tattatacct	6120
tcttaagtTt tcaaccaagg tttgctttTg ttttgagtta ctggggTtat ttttgtttta	6180
aataaaaaata agtgtacaat aagtgttttt gtattgaaag cttttgttat caagattttc	6240
atacttttac ctTccatggc tctttttaag attgatactt ttaagaggTg gctgatattc	6300
tgcaacactg tacacataaa aaatacggta aggatacttt acatggTtaa ggtaaagtaa	6360
gtctccagTt ggccaccatt agctataatg gcactttgtt tgtgtTgtg gaaaaagtca	6420
cattgccatt aaactttcct tgtctgtcta gttaatatTg tgaagaaaaa taaagtacag	6480
tgtgagatac tg	6492

<210>6

<211>239

<212>PRT

<213>人

<400>6

Met	Ala	His	Ala	Gly	Arg	Thr	Gly	Tyr	Asp	Asn	Arg	Glu	Ile	Val	Met
1				5						10				15	
Lys	Tyr	Ile	His	Tyr	Lys	Leu	Ser	Gln	Arg	Gly	Tyr	Glu	Trp	Asp	Ala
				20						25				30	
Gly	Asp	Val	Gly	Ala	Ala	Pro	Pro	Gly	Ala	Ala	Pro	Ala	Pro	Gly	Ile
				35						40				45	
Phe	Ser	Ser	Gln	Pro	Gly	His	Thr	Pro	His	Pro	Ala	Ala	Ser	Arg	Asp
				50						55				60	

Pro Val Ala Arg Thr Ser Pro Leu Gln Thr Pro Ala Ala Pro Gly Ala	
65	70 75 80
Ala Ala Gly Pro Ala Leu Ser Pro Val Pro Pro Val Val His Leu Thr	
	85 90 95
Leu Arg Gln Ala Gly Asp Asp Phe Ser Arg Arg Tyr Arg Arg Asp Phe	
	100 105 110
Ala Glu Met Ser Ser Gln Leu His Leu Thr Pro Phe Thr Ala Arg Gly	
	115 120 125
Arg Phe Ala Thr Val Val Glu Glu Leu Phe Arg Asp Gly Val Asn Trp	
	130 135 140
Gly Arg Ile yal Ala Phe Phe Glu Phe Gly Gly Val Met Cys Val Glu	
145	150 155 160
Ser Val Asn Arg Glu Met Ser Pro Leu Val Asp Asn Ile Ala Leu Trp	
	165 170 175
Met Thr Glu Tyr Leu Asn Arg His Leu His Thr Trp Ile Gln Asp Asn	
	180 185 190
Gly Gly Trp Asp Ala Phe Val Glu Leu Tyr Gly Pro Ser Met Arg Pro	
	195 200 205
Leu Phe Asp Phe Ser Trp Leu Ser Leu Lys Thr Leu Leu Ser Leu Ala	
	210 215 220
Leu Val Gly Ala Cys Ile Thr Leu Gly Ala Tyr Leu Gly His Lys	
225	230 235

<210>7

<211>1207

<212>DNA

<213>人

<400>7

tttctgtgaa gcagaagtct gggaatcgat ctggaaatcc tcctaatttt tactccctct	60
ccccgcgact cctgattcat tgggaagttt caaatcagct ataactggag agtgctgaag	120
attgatggga tcgttgcctt atgcatttgt tttggtttta caaaaaggaa acttgacaga	180
ggatcatgct gtacttaaaa aataacaacat cacagaggaa gtagactgat attaacaata	240
cttactaata ataacgtgcc tcatgaaata aagatccgaa aggaattgga ataaaaattt	300
cctgcatctc atgccaaggg ggaaacacca gaatcaagtg ttccgcgtga ttgaagacac	360
cccctcgtec aagaatgcaa agcacatcca ataaaatagc tggattataa ctccctctct	420
ttctctgggg gccgtggggg gggagctggg gcgagagggt ccgttggccc ccgttgcttt	480
tcctctggga aggatggcgc acgctgggag aacagggtac gataaccggg agatagtgat	540
gaagtacatc cattataagc tgtcgcagag gggctacgag tgggatgcgg gagatgtggg	600

```

cgccgcgccc ccgggggccc cccccgcacc gggcatcttc tcctcccagc ccgggcacac    660
gccccatcca gccgcatecc gggacccggg cgccaggacc tcgccgctgc agaccccggc    720
tgcccccggc gccgcgcggg ggcttgcgt cagcccgggt ccacctgtgg tccacctgac    780
cctccgccag gccggcgacg acttctcccg ccgtaccgc cgcgacttcg ccgagatgtc    840
cagccagctg cacctgacgc cttcacccgc gcggggacgc tttgccacgg tgggtggagga    900
gctcttcagg gacgggggtga actgggggag gattgtggcc ttctttgagt tcggtggggg    960
catgtgtgtg gagagcgtea accgggagat gtcgccctg gtggacaaca tcgccctgtg   1020
gatgactgag tacctgaacc ggcacctgca cacctggatc caggataacg gaggctgggt   1080
aggtgcactt ggtgatgtga gtctgggctg aggccacagg tccgagatgc gggggttgga   1140
gtgcgggtgg gctcctgggg caatgggagg ctgtggagcc ggcgaaataa aatcagagtt   1200
gttgcta                                           1207

```

<210>8

<211>205

<212>PRT

<213> 人

<400>8

```

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

```

Ser Val Asn Arg Glu Met Ser Pro Leu Val Asp Asn Ile Ala Leu Trp
165 170 175
Met Thr Glu Tyr Leu Ash Arg His Leu His Thr Trp Ile Gln Asp Asn
180 185 190
Gly Gly Trp Val Gly Ala Leu Gly Asp Val Ser Leu Gly
195 200 205

<210>9

<211>2594

<212>DNA

<213> 人类免疫缺陷病毒 1

<400>9

```

atggagccag tagatcctag actagagccc tggaagcatc caggaagtca gcctaaaact 60
gcttgtacca attgctattg taaaaagtggt tgctttcatt gccaaagtttg tttcataaca 120
aaagccttag gcattctccta tggcaggaag aagcggagac agcgacgaag agctcatcag 180
aacagtcaga ctcatcaagc ttctctatca aagcagtaag tagtacatgt aatgcaacct 240
ataccaatag tagcaatagt agcattagta gtagcaataa taatagcaat agttgtgtgg 300
tccatagtaa tcatagaata taggaaaata ttaagacaaa gaaaaataga cagggttaatt 360
gatagactaa tagaaagagc agaagacagt ggcaatgaga gtgaaggaga aatatcagca 420
cttgttgaga tgggggtgga gatggggcac catgctcctt gggatgttga tgatctgtag 480
tgctacagaa aaattgtggg tcacagtcta ttatggggta cctgtgtgga aggaagcaac 540
caccactcta ttttgtgcat cagatgctaa agcatatgat acagaggtag ataattgttg 600
ggccacacat gcctgtgtac ccacagaccc caaccacaa gaagtagtat tggtaaattgt 660
gacagaaaat ttttaacatgt ggaaaaatga catggtagaa cagatgcatg aggatataat 720
cagtttatgg gatcaaagcc taaagccatg tgtaaaatta accccactct gtgttagttt 780
aaagtgcact gatttgaaga atgatactaa taccaatagt agtagcggga gaatgataat 840
ggagaaagga gagataaaaa actgctcttt caatatcagc acaagcataa gaggttaaggt 900
gcagaaagaa tatgcatttt ttataaaact tgatataata ccaatagata atgatactac 960
cagctataag ttgacaagtt gtaacacctc agtcattaca caggcctgtc caaaggtatc 1020
ctttgagcca attcccatac attattgtgc cccggctggg tttgcgattc taaaatgtaa 1080
taataagacg ttcaatggaa caggaccatg tacaaatgtc agcacagtac aatgtacaca 1140
tggaattagg ccagtagtat caactcaact gctgttaaata ggcagtctag cagaagaaga 1200
ggtagtaatt agatctgtca atttcacgga caatgctaaa accataatag tacagctgaa 1260
cacatctgta gaaattaatt gtacaagacc caacaacaat acaagaaaaa gaatccgtat 1320
ccagagagga ccaggagag catttggttac aataggaaaa ataggaaata tgagacaagc 1380
acattgtaac attagtagag caaaatggaa taacacttta aaacagatag ctagcaaatt 1440
aagagaacaa tttggaaata ataaaacaat aatctttaag caatcctcag gaggggaccc 1500
agaaattgta acgcacagtt ttaattgtgg aggggaattt ttctactgta attcaacaca 1560

```



```

actgtttaat agtacttggg ttaatagtagc ttggagtact gaagggtcaa ataactga 1620
aggaagtgc acaatcaccc tccatgcag aataaaacaa attataaaca tgtggcagaa 1680
agtaggaaaa gcaatgtatg cccctcccat cagtggacaa attagatggt catcaaatat 1740
tacagggtcg ctattaacaa gagatgggtg taatagcaac aatgagtcgg agatcttcag 1800
acctggagga ggagatatga gggacaattg gagaagtga ttatataaat ataaagtagt 1860
aaaaattgaa ccattaggag tagcacccac caaggcaaag agaagagtgg tgcagagaga 1920
aaaaagagca gtgggaatag gagctttgtt ccttgggttc ttgggagcag caggaagcac 1980
tatgggcgca gcctcaatga cgctgacggg acaggccaga caattattgt ctggtatagt 2040
gcagcagcag aacaatttgc tgagggtat tgaggcgcaa cagcatctgt tgcaactcac 2100
agtctggggc atcaagcagc tccaggcaag aatcctggct gtggaaagat acctaaagga 2160
tcaacagctc ctggggattt ggggttgctc tggaaaactc atttgcacca ctgctgtgcc 2220
ttggaatgct agttggagta ataatctct ggaacagatt tggaatcaca cgacctggat 2280
ggagtgggac agagaaatta acaattacac aagcttaata cactccttaa ttgaagaatc 2340
gcaaaaccag caagaaaaga atgaacaaga attattggaa ttagataaat gggcaagttt 2400
gtggaattgg tttaacataa caaattggct gtggtatata aaattattca taatgatagt 2460
aggaggcttg gtaggtttta gaatagtttt tgctgtactt tctatagtga atagagttag 2520
gcagggatat tcaccattat cgtttcagac ccacctccca accccgaggg gacccgacag 2580
gcccgaagga atag 2594

```

<210>10

<211>86

<212>PRT

<213> 人类免疫缺陷病毒 1

<400>10

```

Met Glu Pro Val Asp Pro Arg Leu Glu Pro Trp Lys His Pro Gly Ser
1           5           10           15
Gln Pro Lys Thr Ala Cys Thr Asn Cys Tyr Cys Lys Lys Cys Cys Phe
20           25           30
His Cys Gln Val Cys Phe Ile Thr Lys Ala Leu Gly Ile Ser Tyr Gly
35           40           45
Arg Lys Lys Arg Arg Gln Arg Arg Arg Ala His Gln Asn Ser Gln Thr
50           55           60
His Gln Ala Ser Leu Ser Lys Gln ProThr Ser Gln Pro Arg Gly Asp
65           70           75           80
Pro Thr Gly Pro Lys Glu
85

```

<210>11

<211>2397

<212>DNA

<213> 人

<400>11

ctgtcccgcga	agcgccggcg	gcagcatggc	cagctctggt	tccttgaggg	cttcaaagtg	60
tctgaggcca	gcaagaagaa	gcggcgggag	cccctcggcg	aggactccgt	gggcctcaag	120
cccctgaaga	acgcttcaga	cggtgccctc	atggacgaca	accagaatga	gtggggggac	180
gaggacctgg	agaccaagaa	gttcgggttc	gaggagcccg	tggttctgcc	tgacctggac	240
gaccagacag	accaccggca	gtggactcag	cagcacctgg	atgccgctga	cctgcgcatg	300
tctgccatgg	ccccacacc	gccccagggt	gaggttgacg	ccgactgcat	ggacgtcaat	360
gtccgcgggc	ctgatggctt	caccccgctc	atgatgcct	cctgcagcgg	gggcggcctg	420
gagacgggca	acagcgagga	agaggaggac	gcgcgggccc	tcattctccga	cttcattctac	480
cagggcgcca	gcctgcacaa	ccagacagac	cgcacgggcg	agaccgcctt	gcacctggcc	540
gcccgtact	cacgctctga	tgcgcgaag	cgcctgctgg	aggccagcgc	agatgccaac	600
atccaggaca	acatgggccc	caccccgctg	catgcggctg	tgtctgccga	cgcacaaggt	660
gtcttccaga	tcctgatccg	gaaccgagcc	acagacctgg	atgcccgcct	gcatgatggc	720
acgacgccac	tgatcctggc	tgcgcgctg	gccgtggagg	gcatgctgga	ggacctcatc	780
aactcacacg	ccgacgtcaa	cgccttagat	gacctgggca	agtccgccct	gactggggcc	840
gccgccgtga	acaatgtgga	tgcgcgagtt	gtgctcctga	agaacggggc	taacaaagat	900
atgcagaaca	acagggagga	gacacccctg	tttctggccg	cccgggaggg	cagctacgag	960
accgccaagg	tgctgctgga	ccactttgcc	aaccgggaca	tcacggatca	tatggaccgc	1020
ctgccgcgcg	acatcgcaca	ggagcgcgat	catcacgaca	tcgtgaggct	gctggacgag	1080
tacaacctgg	tgcgcagccc	gcagctgcac	ggagccccgc	tggggggcac	gcccaccctg	1140
tcgccccgcg	tctgctcgcc	caacggctac	ctgggcagcc	tcaagcccgg	cgtgcagggc	1200
aagaaggtcc	gcaagcccag	cagcaaaggc	ctggcctgtg	gaagcaagga	ggccaaggac	1260
ctcaaggcac	ggaggaagaa	gtcccaggac	ggcaagggct	gcctgctgga	cagctccggc	1320
atgctctcgc	ccgtggactc	cctggagtca	ccccatggct	acctgtcaga	cgtggcctcg	1380
ccgccactgc	tgccttcccc	gttccagcag	tctccgtccg	tgccccctcaa	ccacctgcct	1440
gggatgcccc	acacccacct	gggcatcggg	cacctgaacg	tggcggccaa	gcccagatg	1500
gcggcgctgg	gtggggggcg	ccggctggcc	tttgagactg	gcccacctcg	tctctccac	1560
ctgcctgtgg	cctctggcac	cagcaccgtc	ctgggctcca	gcagcggagg	ggccctgaat	1620
ttcactgtgg	gcgggtccac	cagtttgaat	ggtcaatgcg	agtggctgtc	ccggctgcag	1680
agcggcatgg	tgccgaacca	atacaacct	ctgcggggga	gtgtggcacc	aggccccctg	1740
agcacacagg	ccccctccct	gcagcatggc	atggtaggcc	cgctgcacag	tagccttget	1800
gccagegccc	tgteccagat	gatgagctac	cagggcctgc	ccagcaccgc	gctggccacc	1860
cagcctcacc	tgggtgcagac	ccagcaggtg	cagccacaaa	acttacagat	gcagcagcag	1920
aacctgcage	cagcaaacad	ccagcagcag	caaagcctgc	agccgccacc	accaccacca	1980

```

cagccgcacc ttggcgtgag ctcagcagcc agcggccacc tgggccggag cttcctgagt    2040
ggagagccga gccaggcaga cgtgcagcca ctgggccccca gcagcctggc ggtgcacact    2100
attctgcccc aggagagccc cgccctgccc acgtcgtctgc catcctcgtt ggtcccaccc    2160
gtgaccgcag cccagtctct gacgcccccc tcgcagcaca gctactcctc gcctgtggac    2220
aacacccccca gccaccagct acaggtgcct gagcaccctt tcttcacccc gtcccctgag    2280
tcccctgacc agtggtccag ctcgtccccc cattccaacg tctccgactg gtccgagggc    2340
gtctccagcc cteccaccag catgcagtcc cagatcgecc gcattccgga ggccttc    2397

```

<210>12

<211>799

<212>PRT

<213>人

<400>12

```

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

```

Leu Glu Ala Ser Ala Asp Ala Asn Ile Gln Asp Asn Met Gly Arg Thr		
195	200	205
Pro Leu His Ala Ala Val Ser Ala Asp Ala Gln Gly Val Phe Gln Ile		
210	215	220
Leu Ile Arg Asn Arg Ala Thr Asp Leu Asp Ala Arg Met His Asp Gly		
225	230	235
Thr Thr Pro Leu Ile Leu Ala Ala Arg Leu Ala Val Glu Gly Met Leu		
	245	250
Glu Asp Leu Ile Asn Ser His Ala Asp Val Asn Ala Val Asp Asp Leu		
	260	270
Gly Lys Ser Ala Leu His Trp Ala Ala Ala Val Asn Asn Val Asp Ala		
	275	285
Ala Val Val Leu Leu Lys Asn Gly Ala Asn Lys Asp Met Gln Asn Asn		
	290	300
Arg Glu Glu Thr Pro Leu Phe Leu Ala Ala Arg Glu Gly Ser Tyr Glu		
305	310	315
Thr Ala Lys Val Leu Leu Asp His Phe Ala Asn Arg Asp Ile Thr Asp		
	325	330
His Met Asp Arg Leu Pro Arg Asp Ile Ala Gln Glu Arg Met His His		
	340	350
Asp Ile Val Arg Leu Leu Asp Glu Tyr Asn Leu Val Arg Ser Pro Gln		
	355	365
Leu His Gly Ala Pro Leu Gly Gly Thr Pro Thr Leu Ser Pro Pro Leu		
	370	380
Cys Ser Pro Asn Gly Tyr Leu Gly Ser Leu Lys Pro Gly Val Gln Gly		
385	390	395
Lys Lys Val Arg Lys Pro Ser Ser Lys Gly Leu Ala Cys Gly Ser Lys		
	405	415
Glu Ala Lys Asp Leu Lys Ala Arg Arg Lys Lys Ser Gln Asp Gly Lys		
	420	430
Gly Cys Leu Leu Asp Ser Ser Gly Met Leu Ser Pro Val Asp Ser Leu		
	435	445
Glu Ser Pro His Gly Tyr Leu Ser Asp Val Ala Ser Pro Pro Leu Leu		
	450	460
Pro Ser Pro Phe Gln Gln Ser Pro Ser Val Pro Leu Asn His Leu Pro		
465	470	480
Gly Met Pro Asp Thr His Leu Gly Ile Gly His Leu Asn Val Ala Ala		
	485	495
Lys Pro Glu Met Ala Ala Leu Gly Gly Gly Gly Arg Leu Ala Phe Glu		

500	505	510
Thr Gly Pro Pro Arg Leu Ser His	Leu Pro Val Ala Ser Gly Thr Ser	
515	520	525
Thr Val Leu Gly Ser Ser Ser	Gly Gly Ala Leu Asn Phe Thr Val Gly	
530	535	540
Gly Ser Thr Ser Leu Asn Gly Gln Cys Glu Trp	Leu Ser Arg Leu Gln	
545	550	555
Ser Gly Met Val Pro Asn Gln Tyr Asn Pro	Leu Arg Gly Ser Val Ala	
565	570	575
Pro Gly Pro Leu Ser Thr Gln Ala Pro Ser	Leu Gln His Gly Met Val	
580	585	590
Gly Pro Leu His Ser Sar Leu Ala Ala Ser	Ala Leu Ser Gln Met Met	
595	600	605
Ser Tyr Gln Gly Leu Pro Ser Thr Arg	Leu Ala Thr Gln Pro His Leu	
610	615	620
Val Gln Thr Gln Gln Val Gln Pro Gln Asn	Leu Gln Met Gln Gln Gln	
625	630	635
Asn Leu Gln Pro Ala Asn Ile Gln Gln Gln	Gln Ser Leu Gln Pro Pro	
645	650	655
Pro Pro Pro Pro Gln Pro His Leu Gly	Val Ser Ser Ala Ala Ser Gly	
660	665	670
His Leu Gly Arg Ser Phe Leu Ser Gly	Glu Pro Ser Gln Ala Asp Val	
675	680	685
Gln Pro Leu Gly Pro Ser Ser Leu Ala	Val His Thr Ile Leu Pro Gln	
690	695	700
Glu Ser Pro Ala Leu Pro Thr Ser Leu	Pro Ser Ser Leu Val Pro Pro	
705	710	715
Val Thr Ala Ala Gln Phe Leu Thr Pro	Pro Ser Gln His Ser Tyr Ser	
725	730	735
Ser Pro Val Asp Asn Thr Pro Ser His	Gln Leu Gln Val Pro Glu His	
740	745	750
Pro Phe Leu Thr Pro Ser Pro Glu Ser	Pro Asp Gln Trp Ser Ser Ser	
755	760	765
Ser Pro His Ser Asn Val Ser Asp Trp	Ser Glu Gly Val Ser Ser Pro	
770	775	780
Pro Thr Ser Met Gln Ser Gln Ile Ala	Arg Ile Pro Glu Ala Phe	
785	790	795

<210>13

<211>814

<212>DNA

<213> 人

<400>13

```

ccttcacctg cctctgcegc ccaggetaca cgggccacca ctgcgagacc aacatcaacg      60
agtgtccag ccagccctgc cgcctacggg gcacctgcca ggacccggac aacgcctacc      120
tctgtttctg cctgaagggg accacaggac ccaactgcga gatcaacctg gatgactgtg      180
ccagcagccc ctgcgactcg ggcacctgtc tggacaagat cgatggctac gagtgtgcct      240
gtgagccggg ctacacaggg agcatgtgta acagcaacat cgatgagtgt gcgggcaacc      300
cctgccacaa cgggggcacc tgcgaggacg gcatcaatgg cttcacctgc cgctgccccg      360
agggtacca cgacccacc tgcctgtctg aggtcaatga gtgcaacagc aaccctgcg      420
tccacggggc ctgccgggac agcctcaacg ggtacaagtg cgactgtgac cctgggtgga      480
gtgggaccaa ctgtgacatc aacaacaacg agtgtgaatc caacccttgt gtcaacggcg      540
gcacctgcaa agacatgacc agtggcatcg tgtgcacctg ccgggagggc ttcagcggtc      600
ccaactgcca gaccaacatc aacgagtgtg cgtccaaccc atgtctgaac aagggcacgt      660
gtattgacga cgttgccggg tacaagtgca actgcctgct gccctacaca ggtgccacgt      720
gtgaggtggt gctggccccg tgtgccccca gcccctgcag aaacggcggg gagtgcaggc      780
aatccgagga ctatgagagt tgtcactatg tcct                                     814

```

<210>14

<211>799

<212>PRT

<213> 人

<400>14

```

Leu Ser Arg Lys Arg Arg Arg Gln His Gly Gln Leu Trp Phe Pro Glu
1           5           10           15
Gly Phe Lys Val Ser Glu Ala Ser Lys Lys Lys Arg Arg Glu Pro Leu
20           25           30
Gly Glu Asp Ser Val Gly Leu Lys Pro Leu Lys Asn Ala Ser Asp Gly
35           40           45
Ala Leu Met Asp Asp Asn Gln Asn Glu Trp Gly Asp Glu Asp Leu Glu
50           55           60
Thr Lys Lys Phe Arg Phe Glu Glu Pro Val Val Leu Pro Asp Leu Asp
65           70           75           80
Asp Gln Thr Asp His Arg Gln Trp Thr Gln Gln His Leu Asp Ala Ala
85           90           95

```

Asp	Leu	Arg	Met	Ser	Ala	Met	Ala	Pro	Thr	Pro	Pro	Gln	Gly	Glu	Val
						100				105				110	
Asp	Ala	Asp	Cys	Met	Asp	Val	Asn	Val	Arg	Gly	Pro	Asp	Gly	Phe	Thr
						115				120				125	
Pro	Leu	Met	Ile	Ala	Ser	Cys	Ser	Gly	Gly	Gly	Leu	Glu	Thr	Gly	Asn
						130				135				140	
Ser	Glu	Glu	Glu	Glu	Asp	Ala	Pro	Ala	Val	Ile	Ser	Asp	Phe	Ile	Tyr
						145				150				155	
Gln	Gly	Ala	Ser	Leu	His	Asn	Gln	Thr	Asp	Arg	Thr	Gly	Glu	Thr	Ala
						165				170				175	
Leu	His	Leu	Ala	Ala	Arg	Tyr	Ser	Arg	Ser	Asp	Ala	Ala	Lys	Arg	Leu
						180				185				190	
Leu	Glu	Ala	Ser	Ala	Asp	Ala	Asn	Ile	Gln	Asp	Asn	Met	Gly	Arg	Thr
						195				200				205	
Pro	Leu	His	Ala	Ala	Val	Ser	Ala	Asp	Ala	Gln	Gly	Val	Phe	Gln	Ile
						210				215				220	
Leu	Ile	Arg	Asn	Arg	Ala	Thr	Asp	Leu	Asp	Ala	Arg	Met	His	Asp	Gly
						225				230				235	
Thr	Thr	Pro	Leu	Ile	Leu	Ala	Ala	Arg	Leu	Ala	Val	Glu	Gly	Met	Leu
						245				250				255	
Glu	Asp	Leu	Ile	Asn	Ser	His	Ala	Asp	Val	Asn	Ala	Val	Asp	Asp	Leu
						260				265				270	
Gly	Lys	Ser	Ala	Leu	His	Trp	Ala	Ala	Ala	Val	Asn	Asn	Val	Asp	Ala
						275				280				285	
Ala	Val	Val	Leu	Leu	Lys	Asn	Gly	Ala	Asn	Lys	Asp	Met	Gln	Asn	Asn
						290				295				300	
Arg	Glu	Glu	Thr	Pro	Leu	Phe	Leu	Ala	Ala	Arg	Glu	Gly	Ser	Tyr	Glu
						305				310				315	
Thr	Ala	Lys	Val	Leu	Leu	Asp	His	Phe	Ala	Asn	Arg	Asp	Ile	Thr	Asp
						325				330				335	
His	Met	Asp	Arg	Leu	Pro	Arg	Asp	Ile	Ala	Gln	Glu	Arg	Met	His	His
						340				345				350	
Asp	Ile	Val	Arg	Leu	Leu	Asp	Glu	Tyr	Asn	Leu	Val	Arg	Ser	Pro	Gln
						355				360				365	
Leu	His	Gly	Ala	Pro	Leu	Gly	Gly	Thr	Pro	Thr	Leu	Ser	Pro	Pro	Leu
						370				375				380	
Cys	Ser	Pro	Asn	Gly	Tyr	Leu	Gly	Ser	Leu	Lys	Pro	Gly	Val	Gln	Gly
						385				390				395	
Lys	Lys	Val	Arg	Lys	Pro	Ser	Ser	Lys	Gly	Leu	Ala	Cys	Gly	Ser	Lys

405	410	415
Glu Ala Lys Asp Leu Lys Ala Arg Arg Lys Lys Ser Gln Asp Gly Lys		
420	425	430
Gly Cys Leu Leu Asp Ser Ser Gly Met Leu Ser Pro Val Asp Ser Leu		
435	440	445
Glu Ser Pro His Gly Tyr Leu Ser Asp Val Ala Ser Pro Pro Leu Leu		
450	455	460
Pro Ser Pro Phe Gln Gln Ser Pro Ser Val Pro Leu Asn His Leu Pro		
465	470	475
Gly Met Pro Asp Thr His Leu Gly Ile Gly His Leu Asn Val Ala Ala		
485	490	495
Lys Pro Glu Met Ala Ala Leu Gly Gly Gly Gly Arg Leu Ala Phe Glu		
500	505	510
Thr Gly Pro Pro Arg Leu Ser His Leu Pro Val Ala Ser Gly Thr Ser		
515	520	525
Thr Val Leu Gly Ser Ser Ser Gly Gly Ala Leu Asn Phe Thr Val Gly		
530	535	540
Gly Ser Thr Ser Leu Asn Gly Gln Cys Glu Trp Leu Ser Arg Leu Gln		
545	550	555
Ser Gly Met Val Pro Asn Gln Tyr Asn Pro Leu Arg Gly Ser Val Ala		
565	570	575
Pro Gly Pro Leu Ser Thr Gln Ala Pro Ser Leu Gln His Gly Met Val		
580	585	590
Gly Pro Leu His Ser Ser Leu Ala Ala Ser Ala Leu Ser Gln Met Met		
595	600	605
Ser Tyr Gln Gly Leu Pro Ser Thr Arg Leu Ala Thr Gln Pro His Leu		
610	615	620
Val Gln Thr Gln Gln Val Gln Pro Gln Asn Leu Gln Met Gln Gln Gln		
625	630	635
Asn Leu Gln Pro Ala Asn Ile Gln Gln Gln Gln Ser Leu Gln Pro Pro		
645	650	655
Pro Pro Pro Pro Gln Pro His Leu Gly Val Ser Ser Ala Ala Ser Gly		
660	665	670
His Leu Gly Arg Ser Phe Leu Ser Gly Glu Pro Ser Gln Ala Asp Val		
675	680	685
Gln Pro Leu Gly Pro Ser Ser Leu Ala Val His Thr Ile Leu Pro Gln		
690	695	700
Glu Ser Pro Ala Leu Pro Thr Ser Leu Pro Ser Ser Leu Val Pro Pro		
705	710	715
		720

Val	Thr	Ala	Ala	Gln	Phe	Leu	Thr	Pro	Pro	Ser	Gln	His	Ser	Tyr	Ser
				725					730					735	
Ser	Pro	Val	Asp	Asn	Thr	Pro	Ser	His	Gln	Leu	Gln	Val	Pro	Glu	His
			740					745					750		
Pro	Phe	Leu	Thr	Pro	Ser	Pro	Glu	Ser	Pro	Asp	Gln	Trp	Ser	Ser	Ser
		755					760				765				
Ser	Pro	His	Ser	Asn	Val	Ser	Asp	Trp	Ser	Glu	Gly	Val	Ser	Ser	Pro
	770					775				780					
Pro	Thr	Ser	Met	Gln	Ser	Gln	Ile	Ala	Arg	Ile	Pro	Glu	Ala	Phe	
785				790					795						

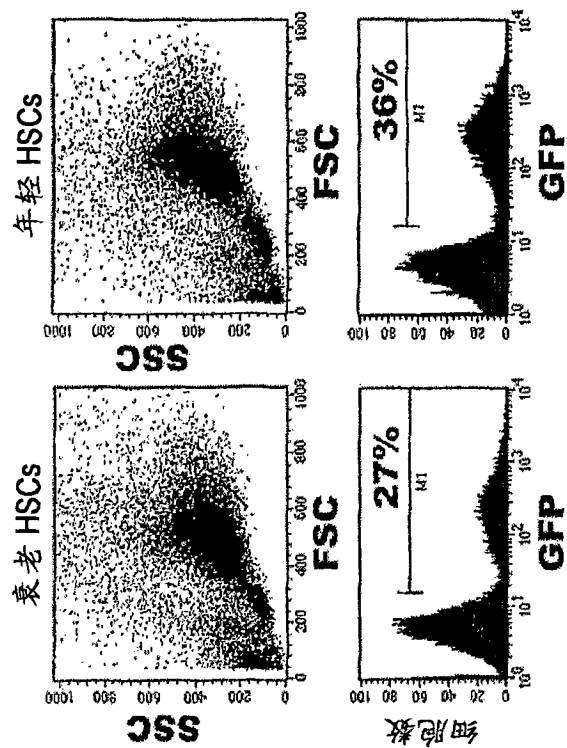


图 2

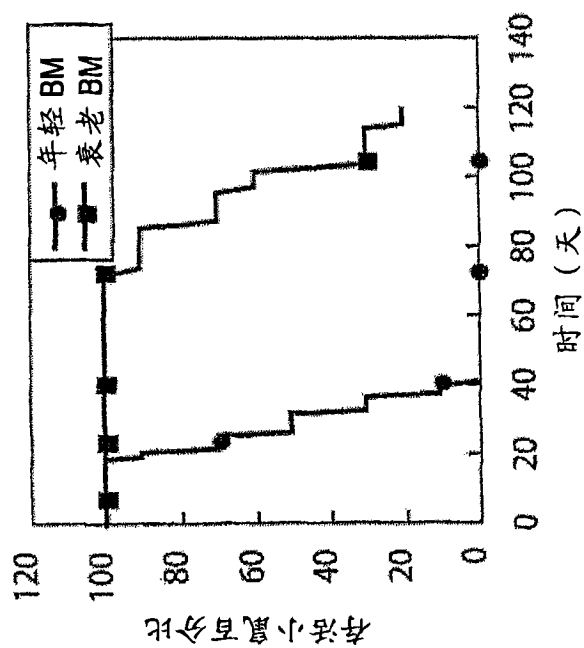
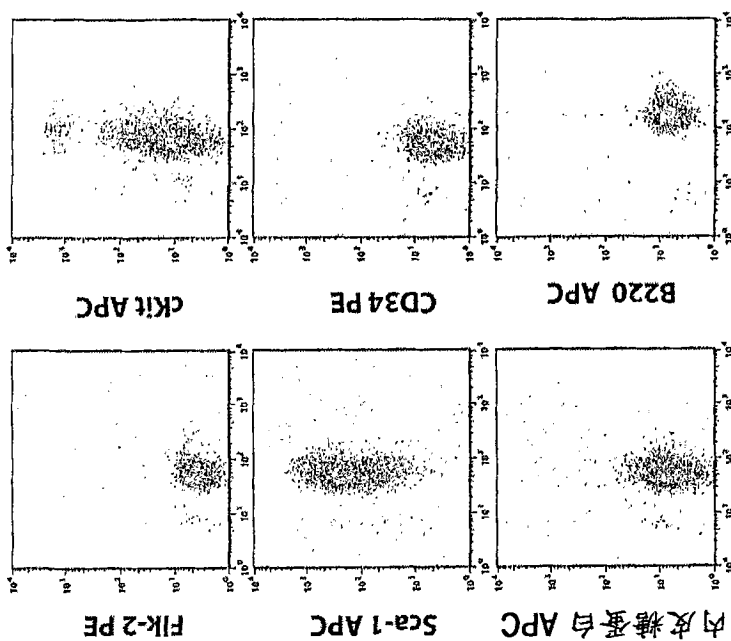


图 1

来源于衰老 HSC 的细胞系



来源于年轻 HSC 的细胞系

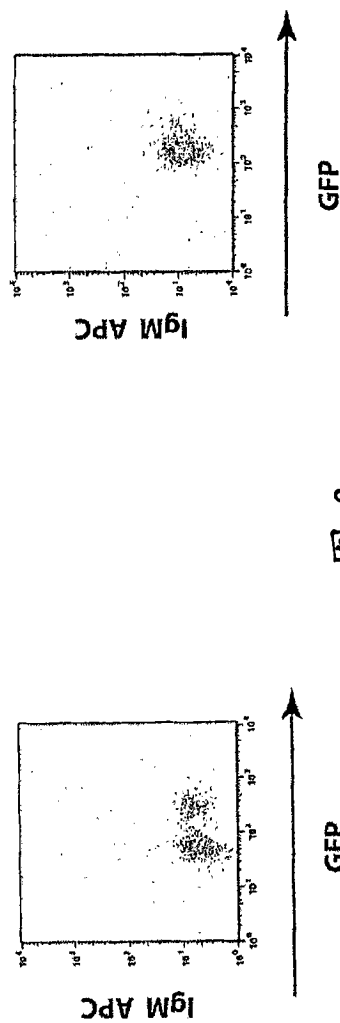
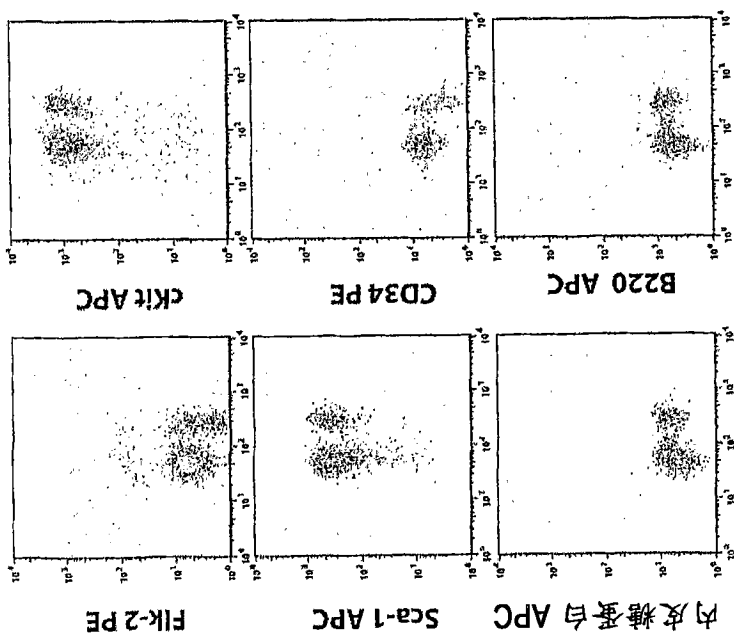
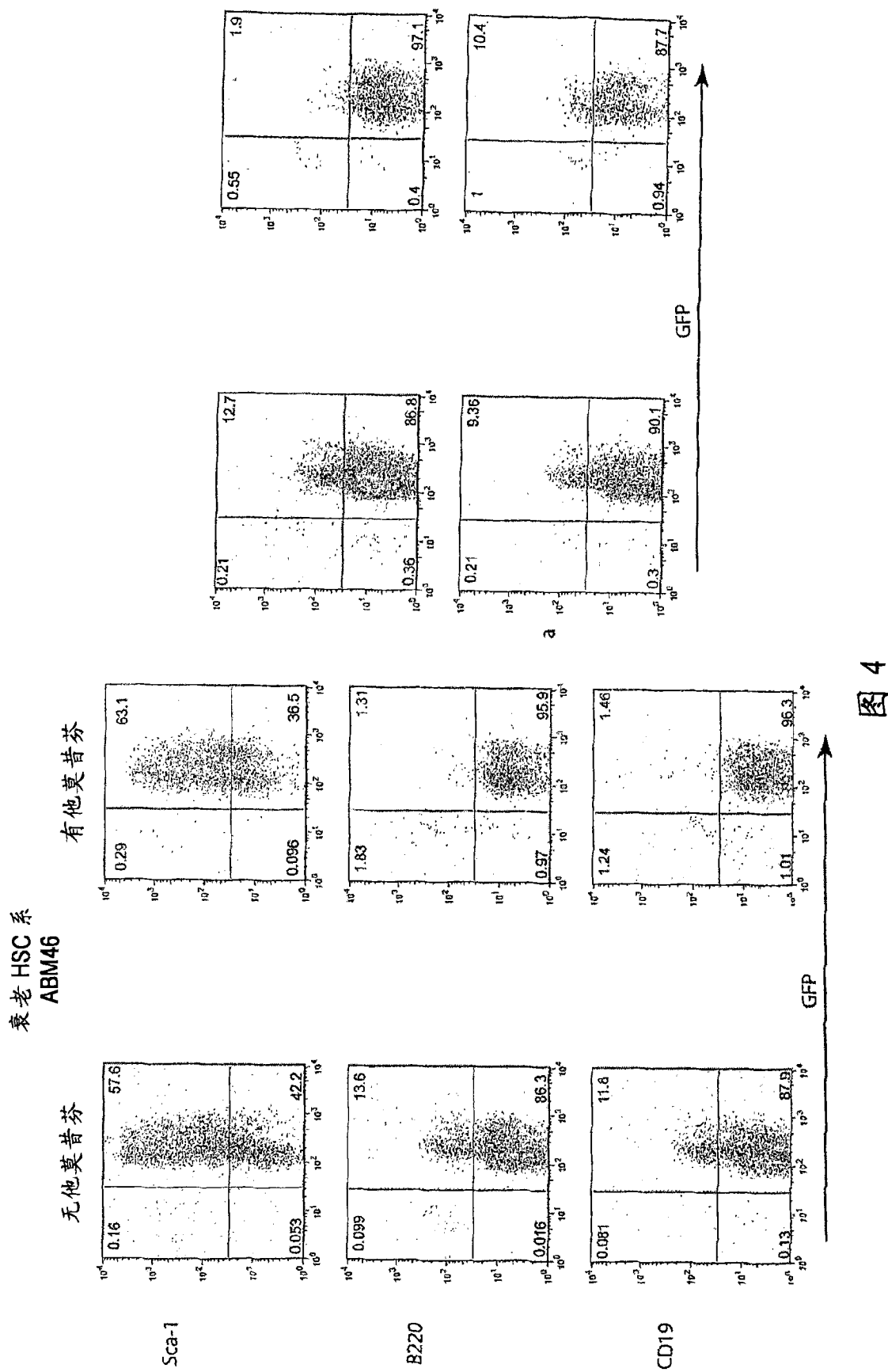


图 3



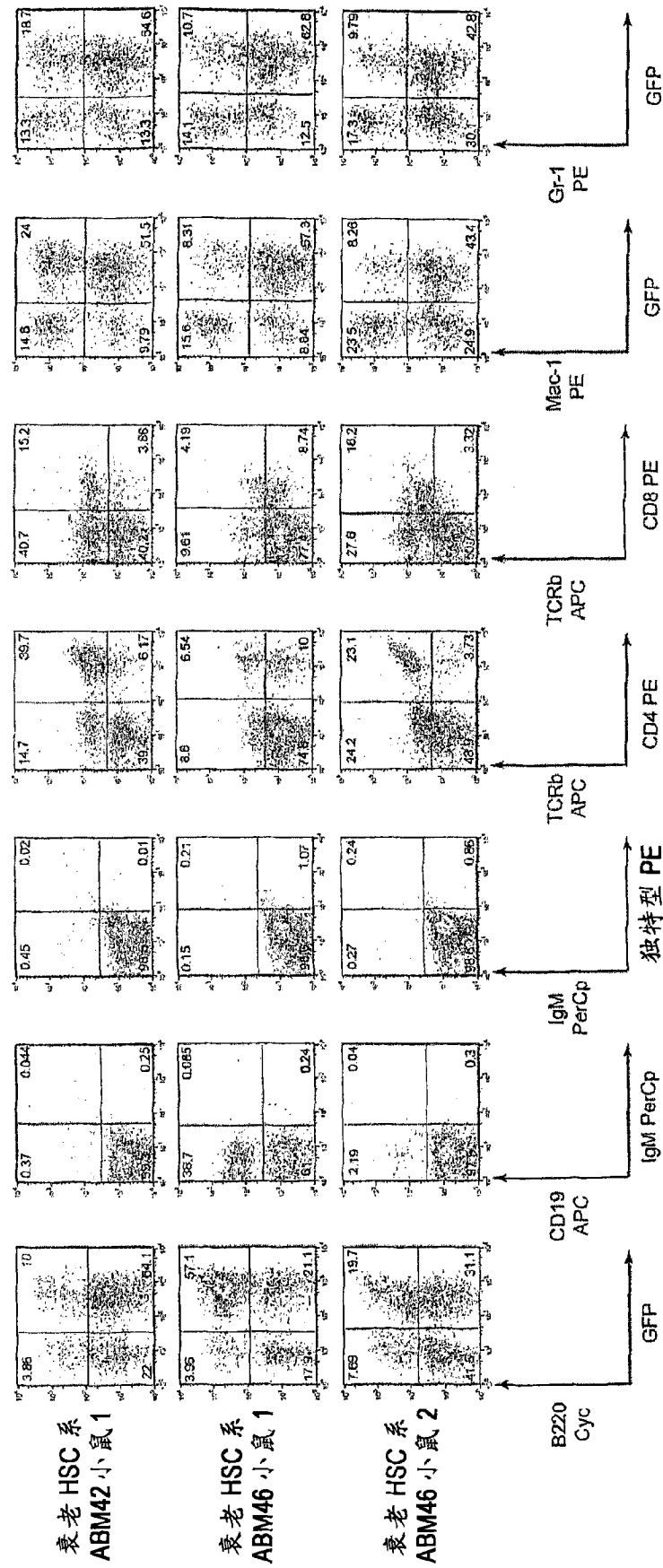
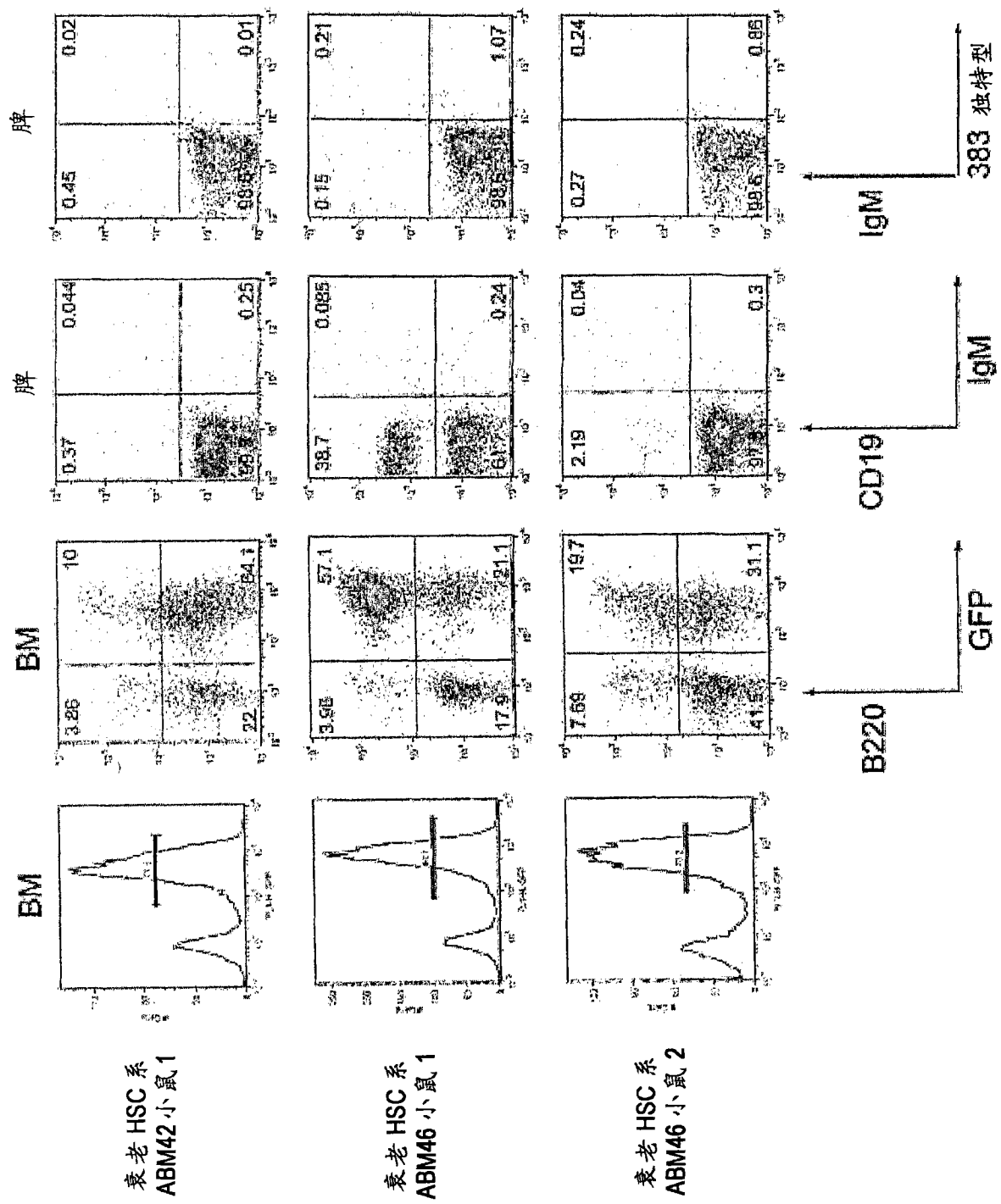
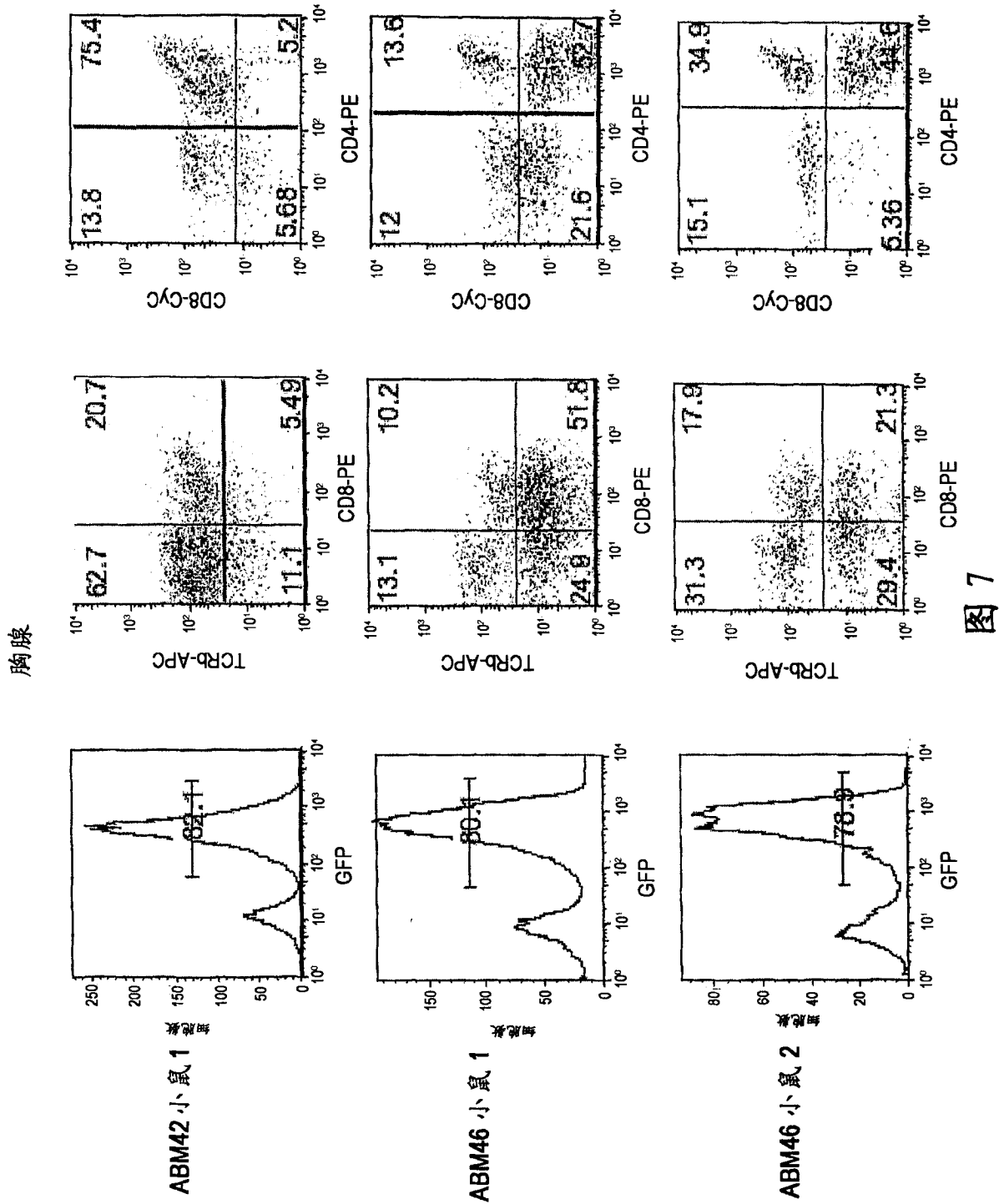


图 5





体外产生的 LT-HSC 细胞系: MYC-ER+Bcl-2 (pMIG-MYC.ER+pMIT-Bcl-2)

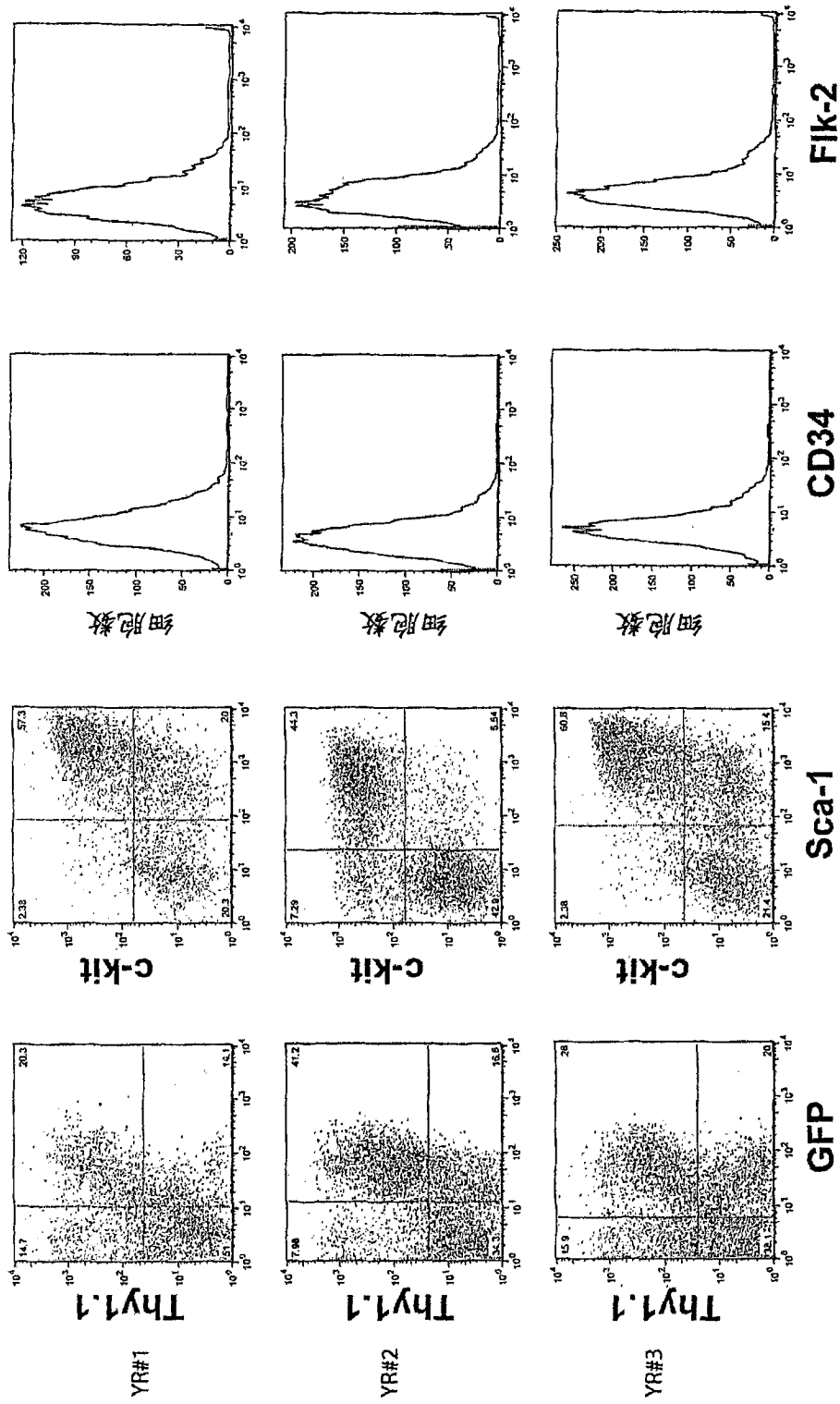


图 8

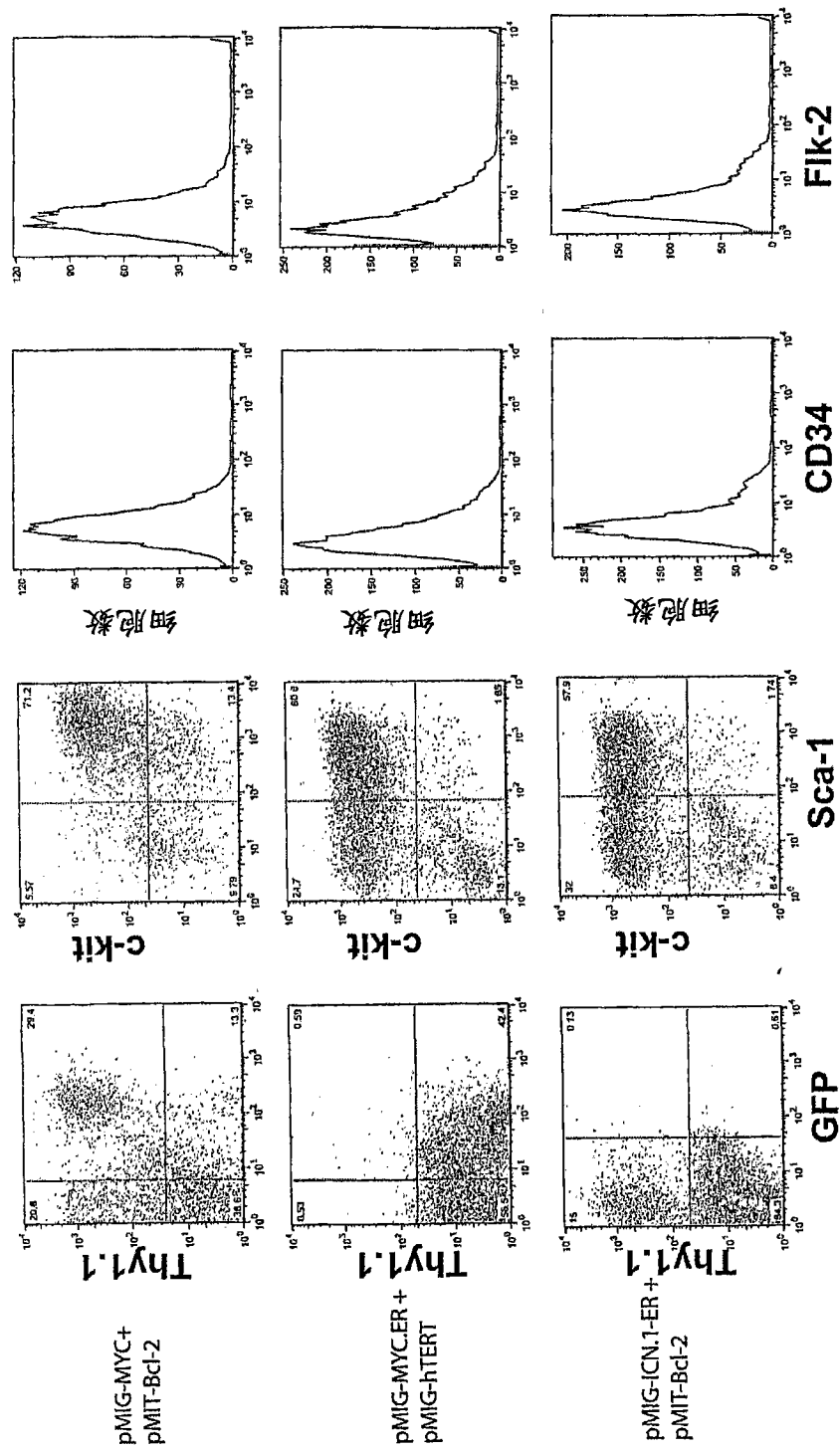


图 9

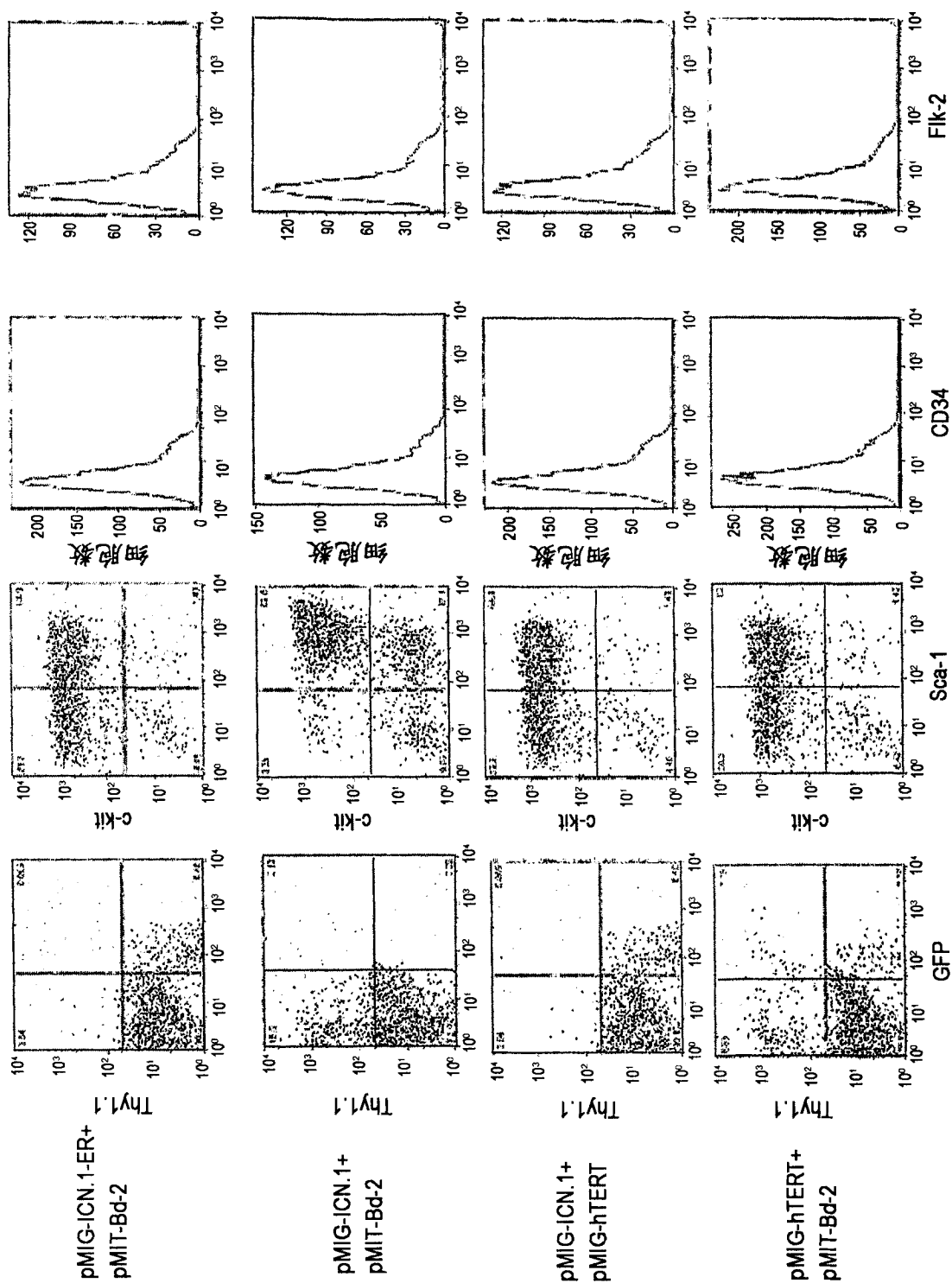


图 10

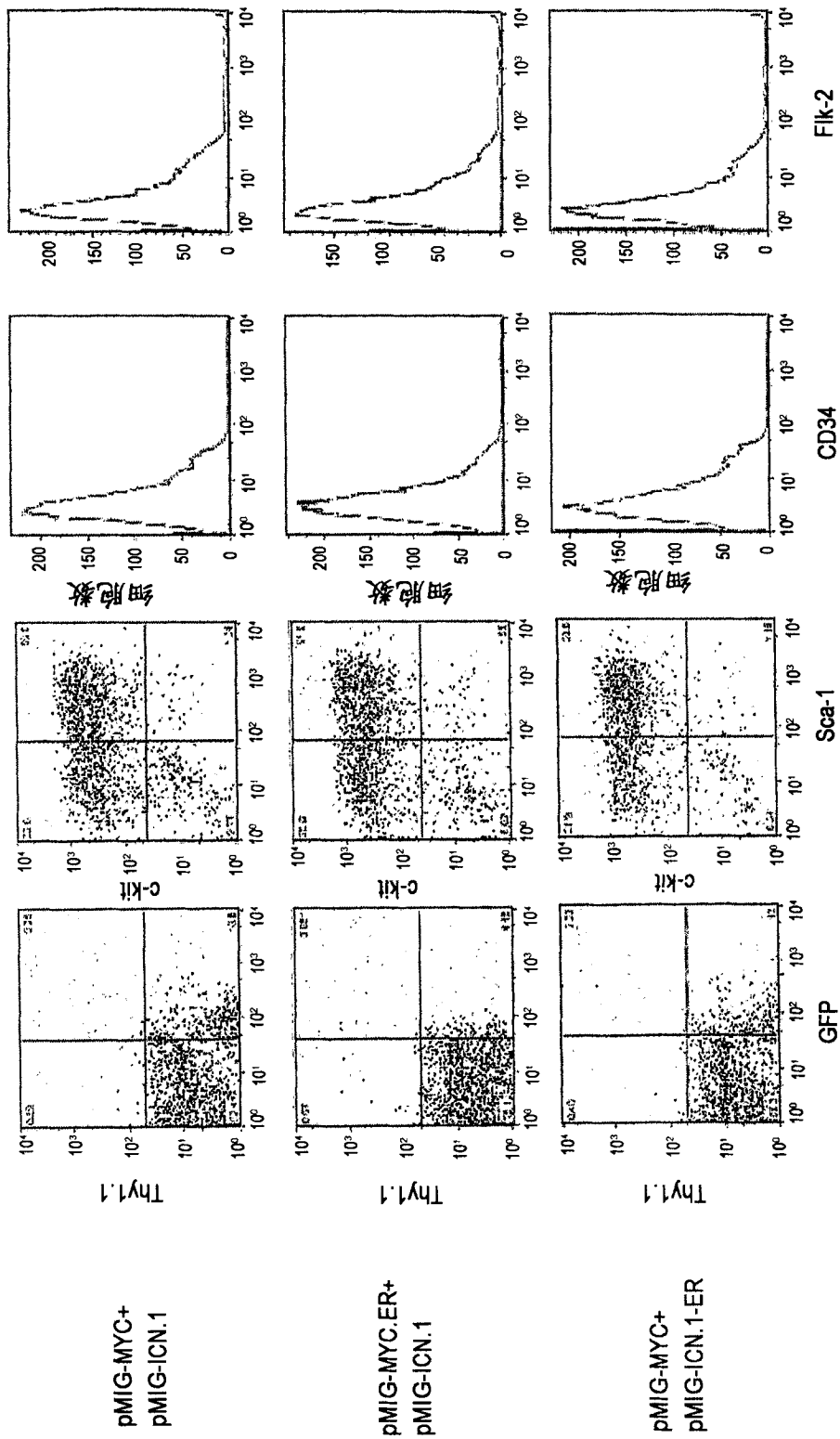


图 11

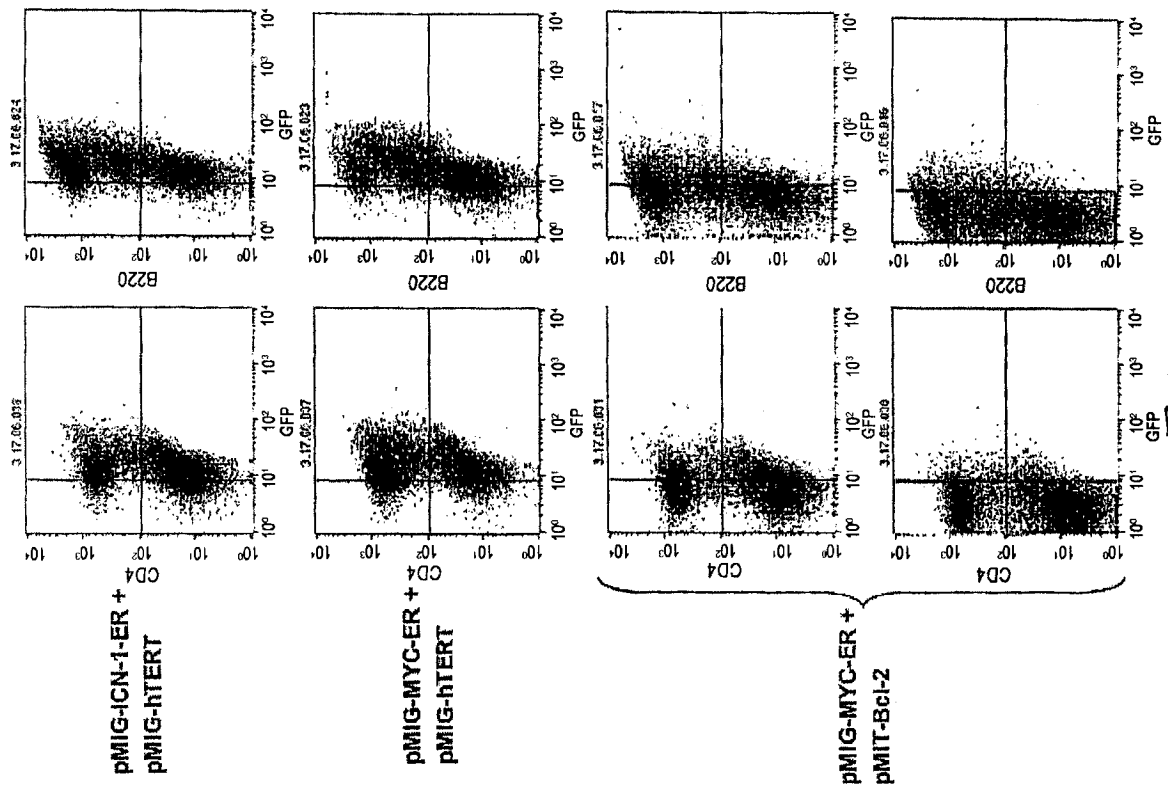


图 12

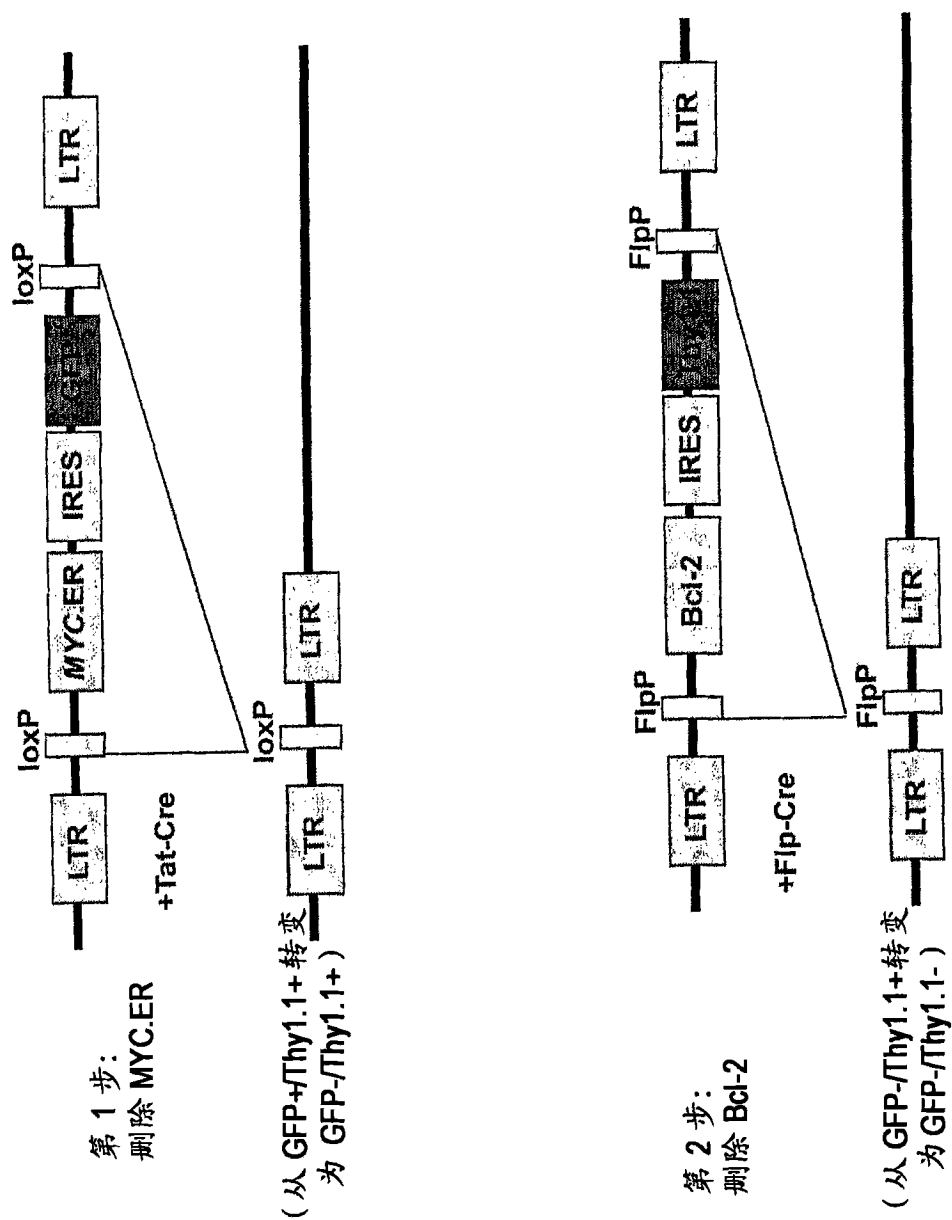


图 13

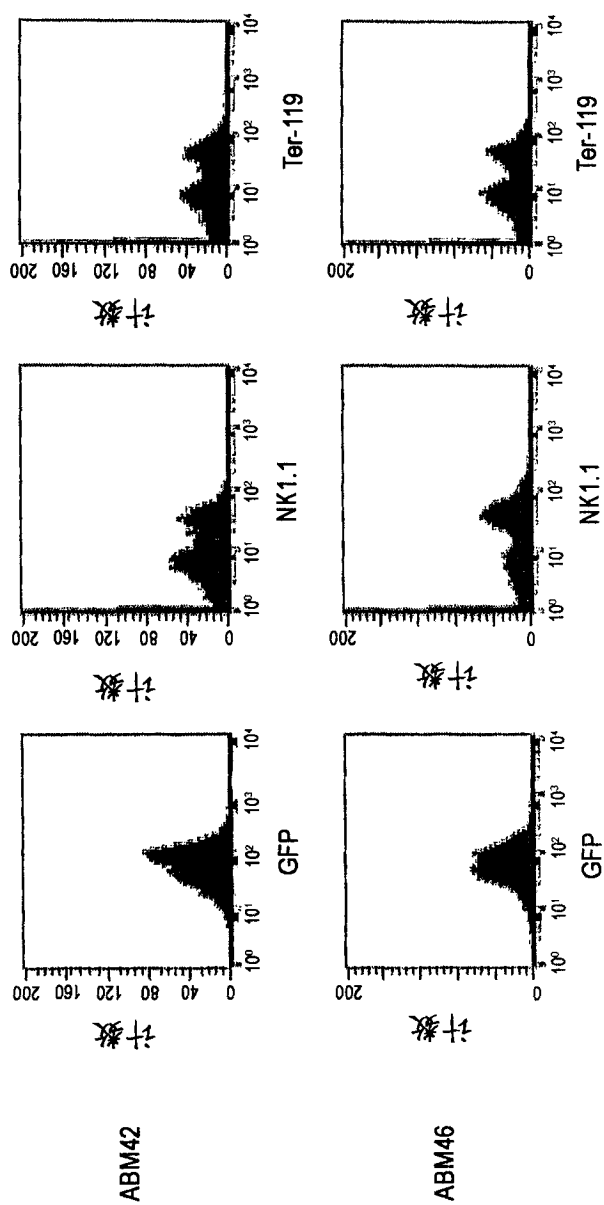


图 14