



(12) 发明专利

(10) 授权公告号 CN 102822345 B

(45) 授权公告日 2016.06.08

(21) 申请号 201080050341.X

(22) 申请日 2010.09.07

(30) 优先权数据

0915523.5 2009.09.07 GB

61/240,282 2009.09.07 US

61/276,203 2009.09.08 US

(85) PCT国际申请进入国家阶段日

2012.05.07

(86) PCT国际申请的申请数据

PCT/GB2010/051493 2010.09.07

(87) PCT国际申请的公布数据

W02011/027180 EN 2011.03.10

(73) 专利权人 基因组研究有限公司

地址 英国伦敦

(72) 发明人 P·刘 W·王 J·杨

(74) 专利代理机构 永新专利商标代理有限公司

72002

代理人 左路

(51) Int. Cl.

C12N 15/85(2006.01)

C12N 5/074(2006.01)

(56) 对比文件

CN 101250502 A, 2008.08.27,

Kazutoshi Takahashi et al., .Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. 《cell》.2006,

Peili Gu et al., .Orphan Nuclear Receptor LRH-1 Is Required To Maintain Oct4 Expression at the Epiblast Stage of Embryonic Development. 《MOLECULAR AND CELLULAR BIOLOGY》.2005,

Bo Feng et al., .Molecules that Promote or Enhance Reprogramming of Somatic Cells to Induced Pluripotent Stem Cells. 《Cell Stem Cell》.2009,

审查员 潘天耀

权利要求书1页 说明书50页

序列表25页 附图36页

(54) 发明名称

细胞及其获得方法

(57) 摘要

描述了重编程的体细胞、用于重编程的方法、用于体细胞的重编程因子以及治疗因子和细胞的用途。所述的核重编程因子 [NRF] 包含下述中的一种或多种：视黄酸受体 (RAR/RXR) 家族成员的基因产物或编码视黄酸受体 (RAR/RXR) 家族成员的基因产物的多核苷酸，或其激动剂或拮抗剂；Lrh1 家族成员的基因产物或其激动剂；视黄酸或参与视黄酸的合成或代谢的基因产物，或其激动剂或拮抗剂；或参与视黄酸家族成员转运的基因产物。

1. 能够将体细胞重编程为iPS细胞的核重编程因子(NRF),其包括视黄酸受体(RAR/RXR)家族成员RAR γ 或RAR α 的基因产物,或编码该基因产物的多核苷酸,或其激动剂,以及
 - (i)Lrh1的基因产物或编码该基因产物的多核苷酸;
 - (ii)Oct4基因产物或编码其的核酸;与以下一或多种的组合:
 - C-MYC基因产物或编码其的核酸;
 - KLF4基因产物或编码其的核酸;和
 - SOX2基因产物或编码其的核酸。
2. 权利要求1的NRF,其中所述RAR γ 或RAR α 激动剂选自:CD437和AM580。
3. 权利要求1-2任一项的NRF,其包括细胞因子;和/或Esrrb。
4. 权利要求1-3任一项的NRF,其中所述NRF包括载体,所述载体含有编码所述NRF组分的多核苷酸。
5. 权利要求4的NRF,其中所述载体是质粒、活病毒载体或转座子。
6. 一种通过体细胞的核重编程制备诱导性多能干细胞的方法,所述方法包括在所述体细胞内表达权利要求1-4任一项的核重编程因子(NRF)以促进iPS细胞形成的步骤。
7. 通过权利要求6的方法获得的人诱导性多能干细胞。
8. 权利要求7的人诱导性多能干细胞,其特征在于具有以下至少一种性质:
 - 表达一种或多种多能性标志物,所述多能性标志物选自Oct4、Nanog和Rex1;
 - 不依赖FGF生长;
 - 在传代后保持正常的核型,所述核型利用光谱核型分析进行测量;
 - 当注射到小鼠中能形成畸胎瘤;
 - oct4、nanog或rex1中的一个或多个的启动子区脱甲基化;
 - 细胞可以解离成能形成次级集落的活单细胞;
 - 在M15+hLIF培养基和/或2i+LIF培养基中增殖;
 - 外源重编程因子表达不存在或不显著。
9. 权利要求7或8的人诱导性多能干细胞,其中所述细胞不包含外源重编程因子。
10. 药物组合物形式的权利要求1-4任一项的核重编程因子或权利要求9的人诱导性多能干细胞在制备用于治疗有需要的患者的药物中的用途。

细胞及其获得方法

[0001] 本发明涉及重编程体细胞、重编程方法、体细胞的重编程因子以及此类因子和细胞的用途。

[0002] 通过表达4种转录因子(Yamanaka因子)即Oct4、Sox2、c-Myc以及Klf4(1-3),可以将小鼠胚胎成纤维细胞(MEF)重编程为多能状态。之后应用同一组基因或其变体重编程小鼠(4-6)、大鼠和人(7-9)中的多种谱系的体细胞。将原代体细胞重编程为诱导性多能干细胞(iPS)是复杂且逐渐的过程,伴随着遗传和表观遗传变化(37,51)。因此,4种Yamanaka因子的转基因表达所产生的小鼠iPS克隆经常是异源的,为全部或部分重编程细胞的混合物(52)。iPS细胞自我更新和其他重要的胚胎干细胞(ES)样特性有时需要外源因子的持续表达(51,11)。因此,在无强的选择方案来鉴定完全重编程细胞的情况下,已经证明,通过简单连续传代和亚克隆难以建立种系有活性的(germline-competent)小鼠iPS细胞系。尽管利用基于Yamanaka因子的各种平台,产生了人类iPS细胞,但是可以合理地预测,这些细胞的性质可能也是异源的。因为缺少可靠的报道基因,如连接于内源多能基因的荧光标记或药物选择标记,使得难以从利用可用的重编程因子所产生的人类iPS细胞异源群体中分离完全重编程细胞。

[0003] 目前来源于胚胎的人类ES细胞,在其形态、基因表达陌生以及集落生成上不同于小鼠ES细胞(53,44)。最明显的差别是,小鼠ES细胞的多能性取决于Jak/Stat3途径通过白血病抑制因子(LIF)的激活(54)或Mek/Erk途径的抑制(40)。相比之下,人类ES细胞不响应LIF,并且仅能通过FGF和活化素维持(44)。最近利用人类ES细胞培养条件衍生了小鼠外胚层干细胞(EpiSC)(55,56),这表明人类ES细胞在许多方面与小鼠EpiSC类似,而不是与真正的多能小鼠ES细胞类似。目前缺少小鼠ES细胞的真正的人类配对物,这使得难以证明操纵ES细胞多能性的常见范例是否在其他哺乳动物种类中起作用。此外,从应用的观点看,等同于小鼠ES细胞的人类多能干细胞系的有效性,会使得将小鼠ES细胞自我更新、分化以及基因操作的知识财富直接应用于人类ES细胞成为现实。

[0004] 发明概述

[0005] 一方面,本发明提供了通过体细胞的核重编程制备诱导性多能干细胞的方法,所述方法包括使体细胞与核重编程因子[NRF]接触的步骤,所述因子包括下述中的一种或多种:

[0006] (i)视黄酸受体(RAR/RXR)家族成员的基因产物,或其激动剂或拮抗剂;

[0007] (ii)Lrh1家族成员的基因产物;或其激动剂;

[0008] (iii)视黄酸或参与视黄酸的合成或代谢的基因产物;或其激动剂或拮抗剂;

[0009] (iv)参与视黄酸家族成员转运的基因产物;

[0010] (v)编码上文(i)-(iv)中任一基因产物的多核苷酸(polynucleic acid)。

[0011] 一方面,所述方法包括使用上文的(i)和(ii)两者。

[0012] 一方面,本发明涉及将体细胞重编程为诱导多能细胞的方法,包括:

[0013] (a)使体细胞与NRF接触,NRF包含或为下述中的一种或多种:

[0014] (i)视黄酸受体(RAR/RXR)家族成员的基因产物,或其激动剂;

- [0015] (ii)Lrh1家族成员的基因产物,或其激动剂;
- [0016] (iii)视黄酸或参与视黄酸家族成员的合成或代谢的基因产物;或其激动剂或拮抗剂;
- [0017] (iv)参与视黄酸家族成员转运的基因产物;
- [0018] (v)编码上文(i)-(iv)中任一基因产物的多核苷酸;
- [0019] 以及
- [0020] (b)任选地检查或确定所述体细胞是否已经被重编程,以及
- [0021] (c)使步骤(a)的产物与RA或RAR/RXR家族成员的拮抗剂接触,或消除与NRF的接触,从而维持多能性。
- [0022] 一方面,本发明提供了NRF,其包含下述中的一种或多种:
- [0023] (i)视黄酸受体(RAR/RXR)家族成员的基因产物,或其激动剂或拮抗剂;
- [0024] (ii)Lrh1家族成员的基因产物;或其拮抗剂;
- [0025] (iii)Lrh1家族成员的基因产物;
- [0026] (iv)视黄酸或参与视黄酸合成的成或代谢的基因产物;或其激动剂或拮抗剂;
- [0027] (v)参与视黄酸家族成员转运的基因产物;
- [0028] (vi)编码上文(i)-(iv)中任一基因产物的多核苷酸;
- [0029] 一方面,NRF用于将体细胞重编程为诱导多能细胞。
- [0030] 一方面,本发明提供了包含或编码来自RAR家族成员和Lrh1家族成员的基因产物的NRF。
- [0031] 另一方面,本发明提供了包含或编码来自RAR家族成员、Lrh1家族成员、Oct家族成员以及Myc家族成员的基因产物的NRF。
- [0032] 在另一方面,本发明提供了包含或编码来自RAR家族成员、Lrh1家族成员、Oct家族成员、Klf家族成员、Myc家族成员以及Sox家族成员的基因产物的NRF。
- [0033] 在本发明的另一方面,所述NRF包含一个或多个载体,所述载体包含编码本文所述的一种或多种核重编程因子的一个核酸/多个核酸。
- [0034] 在另一方面,本发明提供了诸如人类iPS的诱导性多能干细胞,其通过本文所述方法获得或通过本文所述方法可以获得。
- [0035] 在另一方面,本发明涉及诱导性人类多能细胞,其特征为下述至少之一:
- [0036] ●表达一种或多种多能性标志物,如Oct4、Nanog、Rex1;
- [0037] ●不依赖FGF生长;
- [0038] ●在传代后保持正常核型,所述核型利用光谱核型分析进行测量;
- [0039] ●当注射到小鼠中时,能形成畸胎瘤;
- [0040] ●Oct4、Nanog、Rex1中的一个或多个的启动子区脱甲基化;
- [0041] ●细胞可以解离为能形成次生集落的活单细胞;
- [0042] ●在M15+hLIF培养基和/或2i+LIF培养基中增殖
- [0043] 仍然在另一方面,本发明提供了诱导性多能干细胞(iPS),其中所述iPS包含外源DNA序列。
- [0044] 在另一方面,本发明提供了通过分化本发明的诱导性多能干细胞而衍生的体细胞。

[0045] 在另一方面,本发明提供了这样的细胞,其基因组已经被修饰为允许调节编码视黄酸受体(RAR/RXR)家族成员的核酸或Lrh1家族成员的基因产物的表达。

[0046] 在另一方面,本发明提供了组织、器官或非人类动物,其衍生自或包含通过分化本发明的诱导性多能干细胞而衍生的体细胞。

[0047] 在另一方面,本发明提供了药物组合物,其包含与药物可接受赋形剂组合的本文所述的核重编程因子、载体、细胞或组织。

[0048] 在另一方面,本发明提供了本文所述的核重编程因子或iPS细胞或衍生自iPS细胞的体细胞或组织或器官在医药中的用途和本文所述的核重编程因子或iPS细胞或衍生自iPS细胞的体细胞或组织或器官在制备用于治疗有需要的患者的药物中的用途。

[0049] 在另一方面,本发明提供了预防或治疗有需要的患者的疾病的方法,所述方法包括将药学可接受量的本发明的核重编程因子、iPS细胞、体细胞、组织或器官之一递送给所述患者。

附图说明

[0050] 图1.表示重编程中的RA信号传导(signaling)。A.利用piggyBac(PB)转座的重编程策略的示意图。B-C.表达Rara或Rarg以及4种Yamanaka因子显著促进重编程,而通过表达Rara-DN抑制RA信号传导阻断重编程。D.重编程短暂需要RA信号传导。E-F.重编程细胞中多能性基因的表达(E)和在Nanog和Rex1基因座处的甲基化(F)。G.Rarg激动剂CD437改善iPSC质量。

[0051] 图2.Rarg(R)和Lrh-1(L)协同促进重编程。A.Tet-On重编程策略的示意图。B.6种因子(TRE-OCKS和TRE-RL)的4天表达足以完全激活内源Oct4表达用于获得不依赖Dox的iPSC。C.转染后6天的集落图像。D.共表达Rarg和Lrh-1(RL)而不是单独的Rarg或Lrh-1改善iPSC质量。

[0052] 图3.不依赖Dox的小鼠iPSC的表征。A.对iPS20-A1细胞免疫染色来检测Oct4和Nanog。B.小鼠iPSC、亲代MEFs和野生型ES细胞中Oct4、Nanog和Rex1的qRT-PCR。C.iPSC(iPS20-A1)中Oct4和Nanog的启动子几乎完全脱甲基化。D.衍生自iPSC的畸胎瘤含有所有3种胚层的细胞类型。E.iPSC对嵌合体中的种系的贡献。

[0053] 图4.在无血清、无饲养细胞条件下,将MEF重编程为基态(ground state)iPSC。A.用PB-TRE-OCKS(4F)或PB-TRE-OCKS加PB-TRE-RL(6F)重编程的AP⁺集落的数量。B.利用qRT-PCR分析iPSC的基因表达。C.用于Nanog和SSEA-1表达的iPSC的免疫染色(6F)。D.免疫染色所检测的iPSCs(6F)向代表3种胚层的细胞类型的体外分化。

[0054] 图5.不依赖Dox的独特的人类iPSC的产生和表征。A.RAREoct序列在几种哺乳动物种类中保守。B.利用Tet-On 6因子平台重编程HDFn细胞。C.典型的人类iPSC集落形态和AP染色。D.亲代HDFn、人类iPSC以及H1hESC细胞中多能性基因的qRT-PCR分析。E.用于ES细胞表面标志物和多能性因子的人类iPSC的免疫染色。F.人类iPSC的体外分化。抗体。G.由人类iPSC分化的畸胎瘤。H.通过基因表达(qRT-PCR)所测量的人类iPSC的信号传导依赖性。I.生长于不同条件下的人类iPSC的基因表达变化。

[0055] 图6.A.携带转录因子cDNA的PB转座子。B.Oct4-IRES-Puro-Egfp敲入等位基因。C.Oct4-IRES-Puro-Egfp MEF细胞和Oct4-IRES-Puro-Egfp敲入ES细胞的流式细胞术分析。

D. 在转染中增加携带Rarg的转座子的量降低OCKS重编程效率。E. 在2种嘌呤霉素浓度存活的重编程小鼠细胞的流式细胞术分析。

[0056] 图7. Rarg和Lrh-1协同促进重编程。A. 携带由编码2A(口足病病毒2A自切割肽)的DNA所连接的多重cDNA的PB转座子。B. 典型的iPSC集落。C. 转染10天后iPSC集落的碱性磷酸酶染色。D. 用于荧光素酶报道基因测定的DNA构建体图。

[0057] 图8. 小鼠iPSC多能性的表征。A. 免疫染色iPS20-A1细胞以检测SSEA1和Nanog。B. 小鼠iPSC中内源多能性基因的强健(robust)表达。C. 小鼠iPSC系中外源重编程因子表达的RT-PCR分析。

[0058] 图9. 在无饲养细胞和无血清的条件下MEF的重编程。A. 转染的MEF。B. 苏木素和曙红使畸胎瘤的石蜡切片染色。

[0059] 图10. 用6因子平台(CAG启动子版本)产生独特的人类iPSC细胞。A. 形成于M15加LIF培养基或2i/LIF培养基中的人类iPSC集落, 其与常规小鼠ES细胞集落相似。B. 免疫染色所检测的人类iPSC中内源多能性蛋白质的表达。C. RT-PCR所检测的人类iPSC中多能性基因的表达。D. 在畸胎瘤中人类iPSC向3种胚层细胞类型的分化。E. 大量传代(>20代)后, 人类iPSC的正常核型。F. 人类iPSC的Y染色体基因分型分析证实人类iPSC的HDFn来源。

[0060] 图11. 利用6因子Tet-On系统(不依赖Dox)所产生的独特的人类iPSC的表征。A. 人类iPSC中基因表达的RT-PCR分析。B. 人类iPSC的SSEA-4、Tra-1-60以及Tra-1-81表达的FACS分析。C. 人类iPSC中重编程因子表达的RT-PCR分析表明无外源重编程因子表达。D. 在长期体外培养后, 人类iPSC具有正常的核型。E. 基因捕获PB转座子插入HPRT基因座中。

[0061] 图12. Rarg显性失活(Dominant-negative)等位基因阻断ES细胞分化。A. 携带强CAG启动子/增强子和一对剪接受体的PB转座子的示意图。B. Rex1-Puro-IRES-Egfp敲入小鼠ES细胞系的图。C. ES细胞中的遗传学筛选策略, 用于鉴定能阻断视黄酸所诱导的ES细胞分化的突变体。D. 遗传学筛选中所鉴定的Rarg处的4个独立突变。E. Rarg或Rara显性形式的过表达阻断RA诱导的ES细胞分化。F. Rarg或Rara特异性拮抗剂阻断RA诱导的ES细胞分化。

[0062] 图13. RA信号传导在将小鼠MEF重编程为iPS细胞中发挥关键作用。A. 更高质量的iPS细胞与在更高浓度的嘌呤霉素存活关联。B. Rarg-FL的过表达显著增加重编程效率。C. iPS细胞培养平板表明Rarg增加重编程效率。细胞用结晶紫染色。D. Rarg-DN的过表达也改善iPS克隆的质量, 但是降低重编程效率。E. Rarg的剂量影响重编程效率。F. Rarg或Rara特异性激动剂增强重编程效率。G. Rarg特异性激动剂改善iPS细胞的质量。

[0063] 图14. Rarg和Lrh1协同作用促进重编程。A. Rarg特异性拮抗剂处理可以改善部分重编程iPS克隆的质量。B. Lrh1的超表达促进重编程。C. 如Puro抗性iPS细胞集落的非常早期表观所示, 6种因子对MEF的快速重编程。快速实现完全重编程的多能性需要Rarg和Lrh1的表达。E. Rarg和Lrh1表达的剂量对iPS细胞的质量非常关键。

[0064] 图15. 6种因子重编程的iPS细胞的高质量。A. 利用Rarg和Lrh1所产生的小鼠iPS细胞具有ES细胞多能性标志物的强健表达。B. 6因子诱导的小鼠iPS细胞可对畸胎瘤中的所有谱系有贡献。

[0065] 图16. 利用6种因子产生高质量iPS细胞。A. 在M15加hLIF培养基中所产生的人类iPS细胞集落(上幅)。人类iPS细胞集落在2i加hLIF培养条件下扩增(底幅)。B. 人类iPS细胞表达高水平的OCT4和NANOG。C. RT-PCR分析证实, 人类iPS细胞中多能性基因的高水平表

达。4-1、4-2、4-3以及4-5是人类iPS细胞。ES：人类ES细胞对照。

[0066] 表

[0067] 表1.含有推断的RAREoct元件的小鼠基因启动子的列表。

[0068] 表2.含有推断的RAREoct元件的人类基因启动子的列表。

[0069] 表3.cDNA克隆所用的引物。

[0070] 表4.Splinkerette PCR所用的引物。

[0071] 表5.RT-PCR和DMR分析所用的引物。

[0072] 表6.Applied Bioscience为小鼠和人类基因的实时RT-PCR预设计的Taqman探针。

[0073] 表7.定制设计的人类QPCR探针。

[0074] 详细描述

[0075] 在此我们报导了小鼠和人类体细胞的快速和高效的重编程。在小鼠中,将胚胎成纤维细胞(MEF)重编程。根据形态和分子生物学标准,所产生的iPS克隆是高度同质的。当用同样的6种因子重编程人类新生儿包皮真皮成纤维细胞(HDFn)时,我们鉴定了人类多能干细胞克隆。这些细胞还可以以单细胞密度在小鼠ES细胞培养条件下亚克隆,并扩增且无任何可识别的染色体异常。我们还证明,可以在这些人类iPS细胞中实现有效转座,其效率与小鼠ES细胞相当。

[0076] 在一方面,本发明涉及体细胞的核重编程因子(NRF)。合适的是,NRF能促进由体细胞形成iPS。

[0077] 在一方面,核重编程因子包含视黄酸受体RAR/RXR家族成员(如Rar α 、Rar γ 、Rar β 、RXR α 、RXR β 、RXR γ —在本文中也称为Rara、Rarg、Rarb、RXRa、RXRb、RXRg,其在一方面,具有小鼠或人类序列)的基因产物和/或视黄酸和/或Lrh1家族成员(如Lrh1、Sf1,或核受体/nR5a类固醇激素受体家族Ftz-F1亚家族的其他成员)的基因产物和/或参与视黄酸家族成员转运的基因产物。

[0078] 提到RAR家族,包括提到RXR家族,除非根据上下文具有明显不同。在一方面,核重编程因子包含视黄酸受体RAR家族成员(如Rar α 、Rar γ 、Rar β)和Lrh1家族成员,例如全长Rarg和Lrh1。在一方面,提到RAR家族成员指RAR。

[0079] 在一方面,核重编程因子包含视黄酸,例如所有反式或RA或9-顺式 RA,合适的是其浓度为 10^{-8} – 10^{-10} M,合适的是 10^{-9} M。

[0080] 本文提到家族成员和其基因产物,包括编码所述基因产物的多核苷酸,如RAR家族成员和Lrh1家族成员,包括一物种中相同基因/蛋白质家族的成员、不同物种的家族成员和其变体,如具有取代、缺失或添加的蛋白质,合适的是,所述成员或其变体在功能上是等同的,因为它们单独或与其他因子组合能促进由本文所述体细胞形成iPS,这可以合适地通过本申请所述方法评估。这类基因产物包括在氨基酸水平或核苷酸水平与本发明的NRF序列具有至少70%、优选至少80%、90%或95%同源性或相同性的序列,所述序列能单独或与其他因子组合促进由本文所述的体细胞形成iPS,这可以合适地通过本申请所述的方法评估。优选地,RAR家族成员是人类全长野生型Rar γ 或Rar α 序列。优选地,LRH1家族成员是人类全长野生型LRH1序列。

[0081] 在一方面,在本发明的任何方面提到蛋白质或基因产物,指该蛋白质或基因产物的人类全长野生型序列。

[0082] 在一方面,由体细胞形成iPS的促进通过Oct4基因表达的激活来评估,例如在ES细胞中所观察到的水平,或通过具有本文所公开的任何特征的iPS细胞的形成来评估。

[0083] 在一方面,如本文所述,基因产物是具有人类或小鼠序列的蛋白质,或是其变体。

[0084] 在NRF组分的上下文中,在本文中提到的基因产物并非限于由多核苷酸的表达所制备的蛋白质或多肽,而是包括例如可以直接合成的蛋白质或其片段。

[0085] 提到基因产物还包括可以由基因产生的其他非蛋白质物质,例如由DNA向RNA的转录而引起的多核苷酸片段。

[0086] 提到视黄酸受体(RAR/RXR)家族成员的基因产物,包括具有视黄酸受体(RAR/RXR)家族成员的基因产物的一些活性的RAR/RXR家族成员突变体的基因产物。

[0087] 在一方面,本发明涉及体细胞的核重编程因子,其包含RAR/RXR家族成员和/或Lrh1家族成员的突变体所编码的蛋白质,所述突变体如蛋白质功能降低或废除的突变体或活性增强的突变体。如上文所讨论的,具有亲本蛋白质的一些活性的突变体,可以视为该蛋白质的变体。

[0088] 在一方面,本发明涉及体细胞的核重编程因子,其包含RAR家族成员和/或Lrh1家族成员的显性失活突变体所编码的蛋白质。合适的是,显性失活突变体蛋白能阻断视黄酸(RA)、合适地高浓度如1.0 μ M的RA所诱导的ES细胞分化。在一方面,核重编程因子包含Rarg或Rara的显性失活突变体所编码的蛋白质。在一方面,所述蛋白质是C末端缺失部分的Rarg蛋白。在一方面,所述突变体是RARG的突变体,其对应于Rarg的最后内含子(内含子9)的氨基酸序列已经缺失。在一方面,所述蛋白质是Rara的显性失活形式,合适地是Rara-DN。

[0089] 野生型RARG的序列(SEQ ID No.1)

[0090]

```
ATGGCCACCAATAAGGAGAGACTCTTTGCGCCCGGTGCCCTGGGGCCTGGATCTGGTTACCCAGGAGCAGGCTTCCC
ATTCGCCTTCCCAGGTGCACTCAGAGGGTCGCCACCATTTGAGATGCTGAGCCCTAGCTTCCGGGGCCTGGGCCAGC
CTGACCTCCCCAAGGAGATGGCTTCTCTCTCGGTGGAGACACAGAGCACCAGCTCGGAGGAGATGGTACCCAGCTCT
CCCTCACCCCCACCACCTCCTCGGGTCTATAAGCCATGCTTTGTATGCAATGACAAGTCTTCTGGCTACCACTATGG
GGTCAGCTCCTGTGAAGGCTGCAAGGGCTTCTTCAGACGCAGCATTTCAGAAAAACATGGTGTATACATGTACCGTG
ACAAAACTGTATCATCAACAAGGTCACCAGAAATCGATGCCAGTACTGCAGGCTACAAAAGTGTTCGAAGTGGGC
ATGTCCAAGGAAGCTGTAAGGAACGATCGAAACAAGAAGAAAAAGGAGGTAAAAGAGGAGGGCTCGCCGACAGCTA
TGAAGTGAAGTCCACAGTTAGAGGAACTCATCACCAGGTCAGCAAAGCCCACCAGGAGACTTTTCCCTCACTCTGCC
AGCTGGGCAAGTACACCACGAACCTCAGTGCAGATCACCGGTGCAGCTGGACCTGGGGCTGTGGGACAAGTTCAGC
GAGCTGGCCACCAAATGCATCATCAAGATTGTGGAGTTTGCGAAGCGGCTGCCTGGTTTTACAGGGCTCAGCATTGC
CGACCAGATCAGCTGCTCAAGGCTGCTTGTCTGGACATCCTAATGCTGCGGATCTGTACAAGGTATACCCAGAGC
AGGACACTATGACATTCTCGGATGGGCTGACCCTGAACCGAACCCAGATGCACAATGCTGGCTTTGGGCCCCTTACA
GACCTCGTCTTTGCCTTTGCCGGGCAGCTGCTGCCCTGGAGATGGATGACACCGAGACTGGGCTACTTAGTGCTAT
CTGCCTCATCTGTGGAGACCGAATGGACCTGGAAGAGCCCGAGAAGGTGGACAAGCTGCAGGAGCCCCTGCTGGAAG
CCCTGAGGCTCTATGCCCGGCAGCGAGACCCAGCCAACCTACATGTTCCCAAGGATGCTGATGAAAATCACCGAC
CTCCGGGGCATCAGCACTAAGGGAGCAGAAAGGGCTATAACCCTGAAGATGGAGATTCCAGGCCCCGATGCCACCCCT
GATCCGAGAGATGCTGGAGAACCCGAGATGTTTGAGGACGACTCCTCGAAGCCTGGCCCCCACCCTCAAGGCTTCCA
GTGAGGACGAAGCTCCAGGGGGCCAGGGCAAAAGGGGCCAAAGTCCCCAACCTGACCAGGGGCCCTGA
```

[0091] 突变的Rarg的序列(SEQ ID No.2)

[0092]

ATGGCCACCAATAAGGAGAGACTCTTTGCGCCCGGTGCCCTGGGGCCTGGATCTGGTTACCCAGGAGCAGGCTTCCC
ATTCGCCTTCCCAGGTGCACTCAGAGGGTCGCCACCATTTGAGATGCTGAGCCCTAGCTTCCGGGGCCTGGGCCAGC
CTGACCTCCCCAAGGAGATGGCTTCTCTCTCGGTGGAGACACAGAGCACCAGCTCGGAGGAGATGGTACCCAGCTCT
C C C T C A C C C C C A C C A C C T C C T
CGGGTCTATAAGCCATGCTTTGTATGCAATGACAAGTCTTCTGGCTACCACTATGGGGTCAGCTCCTGTGAAGGCTG
CAAGGGCTTCTTCAGACGCAGCATTGAGAAAAACATGGTGTATACATGTCACCGTGACAAAACTGTATCATCAACA
AGGTACCCAGAAATCGATGCCAGTACTGCAGGCTACAAAAGTGTTCGAAGTGGGCATGTCCAAGGAAGCTGTAAGG
AACGATCGAAACAAGAAGAAAAAGGAGGTAAAAAGAGGAGGGCTCGCCCGACAGCTATGAACTGAGTCCACAGTTAGA
GGAACCTCATCACAAGGTGAGCAAGCCCACAGGAGACTTTTCCCTCACTCTGCCAGCTGGGCAAGTACACCACGA
ACTCCAGTGCAGATCACCGGGTGCAGCTGGACCTGGGGCTGTGGGACAAGTTCAGCGAGCTGGCCACCAAATGCATC
ATCAAGATTGTGGAGTTTGCGAAGCGGCTGCCTGGTTTACAGGGCTCAGCATTGCCGACCAGATCACGCTGCTCAA
GGCTGCTTGTCTGGACATCCTAATGCTGCGGATCTGTACAAGGTATACCCAGAGCAGGACACTATGACATTCTCGG
ATGGGCTGACCCTGAACCGAACCCAGATGCACAATGCTGGCTTTGGGCCCTTACAGACCTCGTCTTTGCCTTTGCC
GGGCAGCTGCTGCCCCTGGAGATGGATGACACCGAGACTGGGCTACTTAGTGCTATCTGCCTCATCTGTGGAGACCG
AATGGACCTGGAAGAGCCCCGAGAAGGTGGACAAGCTGCAGGAGCCCCTGCTGGAAGCCCTGAGGCTCTATGCCCGGC
GACGGAGACCCAGCCAACCCTACATGTTCCCAAGGATGCTGATGAAAATCACCGACCTCCGGGGCATCAGCACTAAG
GGATGATGATGA

[0093] 在一方面,本发明涉及RAR/RXR的拮抗剂和/或Lrh1家族成员的拮抗剂,以及所述拮抗剂在制备和维持IPS细胞中的用途。一种合适的Rarg拮抗剂是CD2665。一种合适的Rara拮抗剂是R0-41-5253。在另一方面,所述拮抗剂与DNA甲基化抑制剂如5-氮胞苷和5-氮杂-2'-脱氧胞苷组合使用。

[0094] 在一方面,重编程因子包含RAR和/或Lrh1家族成员的激动剂,例如Rara的激动剂,如AM580,或Rarg的激动剂,其为CD437。

[0095] 在一方面,核重编程因子包含上文所述RAR家族成员、Lrh1家族成员、视黄酸、它们的激动剂和它们的突变体和/或变体中的一种或多种的组合,例如RAR(全长)和Lrh1(全长)家族成员的组合,例如Rarg和Lrh1的组合。

[0096] 合适的是,Rarg和Lrh1的作用是协同的,因为,例如,如本文所述,对于所获得的iPS细胞的数量和/或质量,所观察到其对iPS重编程的作用大于相加。合适的是,通过监测本文所述的Oct4表达水平和/或产生具有本文所列的iPS细胞的一种或多种特征的iPS的能力来评估iPS重编程。

[0097] 核重编程因子还可以包括其他附加组分。合适的是,NRF能将体细胞重编程形成iPS细胞,并且合适地频率比单独的Yamanaka因子更高。Yamanaka因子公开于EP1970446中,其教导通过援引并入本文。

[0098] NRF可以包含一种或多种Yamanaka因子或编码它们的核酸。在一方面,核重编程因子包含诸如C-MYC的MYC蛋白,或编码它的核酸。在一方面,核重编程因子包含诸如KLF4的KLF蛋白或编码它的核酸。在一方面,核重编程因子包含诸如SOX2的SOX基因,或编码它的核酸。在一方面,核重编程因子包含Oct4或编码它的核酸。在一方面,核重编程因子包含C-

MYC、KLF-4、Oct4和SOX2或编码C-MYC、KLF-4、Oct4和SOX2的核酸。

[0099] 在本文中提到的任何基因或蛋白质,如C-MYC、Oct4、KLF-4和SOX2,包括与上文所述C-MYC、KLF-4、Oct 4和SOX2家族成员享有相同性的基因或蛋白质,合适地编码至少在某种程度上与全长野生型序列具有共同的重编程活性的蛋白质。

[0100] 包括在本发明的NRF中的蛋白质包括诸如添加、缺失或取代突变体的变体,所述变体合适地在至少某种程度上与全长野生型序列具有共同的重编程活性。

[0101] 已经鉴定了可以取代Yamanaka因子的组分和各种因子。当将本发明描述为利用一种或多种Yamanaka因子时,应当理解到,本发明还考虑了包括Yamanaka因子每种因子的替换。例如,通过孤儿核受体Nr5a2的过表达,该受体能取代oct4,已经报道了不依赖Oct4的重编程。据报道,另一孤儿核受体Esrrb能取代Klf4。因此,本发明涉及包含本发明的核重编程因子和Nr5a2或Esrrb组合的NRF或包含编码核重编程因子和Nr5a2或Esrrb的核酸的NRF。

[0102] 在一方面,核重编程因子包含下述的组合:

[0103] (i)选自下述中任一种或多种的基因产物或编码它们的核酸:上述RAR/RXR家族成员、Lrh1家族成员、它们的激动剂、它们的拮抗剂、以及它们的突变体和变体,优选RAR和Lrh1家族成员、视黄酸以及调节RA信号途径或被RA信号途径调节的因子。

[0104] (ii)能将多能性赋予给分化的细胞或将成体细胞转化成多能细胞的因子的基因产物或编码所述因子的多核苷酸,所述因子优选为下述中的一种或多种:Oct家族基因、Klf家族基因、以及Myc家族基因和Sox家族基因,或其功能等同物。

[0105] 在一方面,Klf-4的功能等同物是Esrrb。在一方面,Oct4的功能等同物是Nr5a2。

[0106] 在一方面,NRF包含RAR家族成员、Lrh1家族成员、Oct家族成员和Myc家族成员的基因产物或编码所述基因产物的一或多种多核苷酸。

[0107] 在一方面,NRF包含RAR家族成员、Lrh1家族成员、KLF-4以及SOX2的基因产物,但不包含C-MYC和Oct 4。

[0108] 优选地,在利用这类NRF将人类体细胞重编程为IPS细胞的本发明方法中,在重编程体细胞的过程中实现KLF-4和SOX2表达,但是所产生的人类IPS细胞系在染色体中不具有KLF-4和/或SOX2插入。在一方面,KLF-4和/或SOX2没有整合于待重编程的细胞中,例如编码这类因子的转座子没有整合于本发明的人类iPSC系中。

[0109] 在一方面,NRF包含或编码OCT4、CMYC、LRH1以及RARG,且在另一方面,不包含其他2种Yamanaka因子中的任何一种因子。

[0110] 在一方面,NRF包括至少Oct家族基因、Klf家族基因和Myc家族基因或它们的基因产物。合适的家族成员包括Oct 3/4、Klf4、c-Myc家族基因和Sox2的基因产物或编码它们的多核苷酸。

[0111] 在一方面,本发明的NRF包含或编码所有4种Yamanaka因子,即Oct家族基因产物、Klf家族基因产物、Myc家族基因产物以及Sox家族基因产物,或这些基因产物的能促进将体细胞重编程为iPS细胞的功能等同物。

[0112] 在一方面,核重编程因子还包含或编码下述中的一种或多种:

[0113] 细胞因子,其和Myc家族基因的基因产物一起,或可选地代替Myc家族基因的基因产物。作为更优选的实施方案,提供了上文所述的因子,其中所述细胞因子是碱性成纤维细胞生长因子(bFGF)和/或干细胞因子(SCF)。

[0114] 除了Oct家族基因、Klf家族基因、Myc家族基因以及Sox家族基因中各自的基因产物外的TERT基因的基因产物；

[0115] 除了Oct家族基因、Klf家族基因、Myc家族基因、Sox家族基因以及TERT基因中各自的基因产物外，选自下述基因的一种 或多种类型的基因的一种或多种基因产物：SV40大T抗原、HPV16E6、HPV16E7以及Bmi1；

[0116] 一种或多种选自下述的基因的一种或多种基因产物：Fbx15、Nanog、ERas、ECAT15-2、Tcl1以及 β -连环蛋白；

[0117] 一种或多种选自下述的基因的一种或多种基因产物：ECAT1、Esg1、Dnmt3L、ECAT8、Gdf3、Sox15、ECAT15-1、Fth117、Sal14、Rex1、UTF1、Stella、Stat3以及Grb2。

[0118] 优选的NRF包含或编码可合适地送递到体细胞中并且可以在所述细胞内表达的全长Oct 3/4、Klf4、c-Myc家族基因和Sox2、Lrh1和Rarg。

[0119] 在一方面，上文所述的NRF还包含或编码RXR，如RXR α 、 β 或 γ 。

[0120] 在本发明的另一方面，NRF包含RA信号途径的能促进本文所述的体细胞重编程的组分。这类组分可以是RA的下游效应物、RARG或LRH1的上游或下游的调节剂或影响RA产生的分子。本申请所公开的测定，允许鉴定所述途径的合适组分（即能将体细胞重编程从而形成本文所述iPS细胞的那些组分），并允许确定实现重编程的那些组分的合适浓度。

[0121] 本发明的NRF还包括内源基因或内源基因产物的调节剂，如RAR/RXR家族成员和Lrh1家族成员的基因或基因产物的调节剂。例如，可以增强LRH1和/或RARG的内源表达。通过插入或操作驱动表达的启动子，可以实现宿主基因表达的操作，从而提供必须的增强的重编程效果。这类启动子可以包括MSCV（逆转录病毒）的LTS、CAG以及诱导型启动子Tet-On。这类NRF可以用于本发明的所有所述方面。作为实例，包含或表达4种Yamanaka因子或其功能等同物和Lrh1的核重编程因子可以与内源Rarg表达得到增强或可以通过用细胞外因子处理细胞而得以增强的细胞联合使用。这类细胞外因子可以是例如化学制品或环境条件的变化。因此，本发明还涉及这样的细胞，其基因组已经得到修饰，从而允许内源（RAR/RXR）家族成员和/或Lrh1家族成员的表达响应于细胞外因子而得到增强或受到调节。

[0122] NRF可以与其他化合物或药物联合使用，所述其他化合物或药物可以是促进重编程或实际上改善NRF的送递效率的化合物或药物。

[0123] NRF可以包含本文所述的基因产物如蛋白质，为单独表达的或融合 蛋白形式。

[0124] 在一方面，NRF可以包含或组成为编码上文所述的重编程因子的组分的多核苷酸，所述多核苷酸可以送递到体细胞。因此，可以理解的是，重编程因子可以包含蛋白质成分（composition），或经设计允许合适的蛋白质在细胞内表达的核酸成分，或实际上多核苷酸和蛋白质的组合。

[0125] 合成的或纯化的化学品也可以是本发明NRF的一部分。例如，如果使用蛋白质的激动剂或拮抗剂，则如果合适，NRF可以包含经设计而激活或抑制蛋白质功能的化学品。

[0126] 如果使用多核苷酸，且如果多个基因产物形成NRF，则这些可以编码于相同或不同的多核苷酸片段上。例如，一方面，Rarg和Lrh1基因产物可以编于一DNA构建体上，且1、2、3或4种Yamanaka因子可以编码于单独的多核苷酸片段上。

[0127] 编码NRF的蛋白质组分的多核苷酸可以是裸DNA或与送递试剂复合的DNA或采用适合送递到体细胞中的诸如质粒或转座子或病毒载体的载体形式。

[0128] 为了完整,在这个意义上,本发明还涉及编码所述NRF的组分的多核苷酸,则本发明还涉及编码所述NRF组分的变体的多核苷酸,所述组分能单独或组合地促进由体细胞形成iPS。

[0129] 本发明还涉及包含本发明的多核苷酸和载体的细胞,但不限于体细胞,并包括例如适合产生用于转化或转染的核酸的诸如细菌细胞的细胞。

[0130] NRF的各种组分可以单独或组合地在调节系统的控制下表达。在一方面,一种或多种组分,例如4种Yamanaka因子(或其等同物)的任何组分,Rarg或Lrh1,单独或组合,可以在Tet-on®系统(Clontech)的控制下表达。因此,在一方面,控制例如Lrh1或Rarg表达的启动子,受到启动子内响应四环素的存在或不存在的元件的控制。

[0131] 本发明的重编程因子,可以与其他重编程技术如低氧联合使用。

[0132] 作为实例,本发明的NRF可以包含下述(作为蛋白质,或编码所述蛋白质的DNA或RNA,或多核苷酸和蛋白质的混合物):

[0133] ●Rarg和Lrh1

[0134] ●Rarg、Lrh1、Oct4和cMyc

[0135] ●Rarg、Lrh1、Sox2和Klf4

[0136] ●Rarg、Lrh1、Oct4、cMyc、Sox2和Klf4.

[0137] 为了避免怀疑,NRF可以包含单组分蛋白质,或包含编码单个蛋白质的核酸,或编码本文所述的单个激动剂或拮抗剂。

[0138] 诸如Rarg的某些蛋白质的表达水平,影响重编程效率。因此,本发明还涉及确定合适的NRF重编程水平的方法,所述方法包括改变NRF或其组分的浓度,并直接或例如通过监测Oct4的表达并选择用于重编程的NRF的合适浓度来监测iPS细胞的产生。本发明还包括用于本文所公开的重编程的方法,其中用上文所评估的适当量的NRF来实现重编程。

[0139] 本文所公开的各种元件的合适浓度包括例如Ro-41-5253RARa拮抗剂1 μ M、CD2665RARg拮抗剂1 μ M、AM580RARa激动剂10nM、CD437RARg激动剂100nM、约1 $\times 10^{-9}$ M的所有的反式RA和9-顺式RA。

[0140] 本发明的NRF可以包含载体,诸如质粒、活病毒载体或转座子,载体具有编码本文所公开的NRF组分。特别是,载体可以是包含编码本文所述核重编程因子组分的核酸的载体。

[0141] 本发明的载体通常适合将编码NRF的组分的核酸递送到体细胞中,并且还可以含有允许或促进所述核酸所编码的蛋白质表达的必需序列。

[0142] 载体可以表达来自染色体外基因座的NRF的组分,或可以被设计为整合到染色体中,并从染色体内表达。

[0143] 诸如PiggyBac的转座子可以适合将核酸递送到体细胞中的用途,并且还可以适合将外源DNA递送到本发明的iPS细胞中的用途。

[0144] 本发明还涉及通过体细胞的核重编程制备诱导性多能干细胞的方法,所述方法包括使本文所述的核重编程因子与体细胞接触的步骤。

[0145] 所述方法可以任选地包括例如通过选择或筛选具有本文所述特性的细胞而选择或筛选iPS细胞的另一步骤。

[0146] 使NRF与体细胞接触可以以多种方式发生。例如,可以将核重编程因子添加于体细

胞的培养物或将编码核重编程因子的组分的载体引入体细胞。载体例如可以是质粒或转座子或病毒载体。

[0147] 如果NRF包含蛋白质,则合适的是,将NRF暴露于培养的体细胞,例如持续3、4、5、6、7、8、9或10天。在一方面,不超过8、9、10、11、12、13、14或15天。

[0148] 体细胞可以是任何合适的细胞,且待编程的体细胞的类型并无特别限制。例如,可以使用成熟的体细胞或祖细胞以及胚胎期的体细胞。当诱导性多能干细胞用于疾病的治疗性治疗时,使用分离自待治疗的患者的体细胞是合适的。例如,可以使用与疾病有关的体细胞、参与疾病的治疗性治疗的体细胞等。体细胞可以是任何合适的哺乳动物细胞,如,作为实例,人类、小鼠、大鼠、猪、绵羊或牛。

[0149] 在一方面,一旦体细胞被重编程,则与所述体细胞接触的外源重编程因子(也称为外源因子或外源NRFs)的表达被关闭。这可以通过使用诸如Tet-On的诱导型启动子实现,所述启动子通过从培养基中移除Dox而使得待关闭的外源重编程因子表达。合适的是,在使体细胞与NRF接触约4天后关闭外源重编程因子的表达,在该阶段已经将体细胞重编程并足以能转化为iPS细胞而无需来自外源重编程因子的其他输入。

[0150] 合适的是,在用NRF处理细胞后的10天内,如在3、4、5、6、7、8或9天内,并且合适的是,在处理2天后,鉴定并选择iPS细胞,例如,这可以通过本文所述的Oct 4表达来评估,。

[0151] 根据本发明的方法,用于选择出现于培养基中的诱导性多能干细胞的方法并无特别限制,且合适的是,可以使用任何熟知的方式,例如利用药物抗性作为指数,可以利用药物抗性基因等作为标志物基因来分离诱导性多能干细胞。含有未分化状态和多能性的ES细胞的各种培养基以及不能维持这类特性的各种培养基在本领域中是已知的,且利用合适培养基的组合可以有效分离诱导性多能干细胞。本领域技术人员利用广泛应用于ES细胞的证实方式可以容易地证实分离的诱导性多能干细胞的分化和增殖能力。

[0152] 在另一方面,本发明涉及维持多能性的合适量的RA激动剂或拮抗剂(优选RA拮抗剂)在任何细胞系(如人类或小鼠的任何细胞系)的ES细胞或iPS细胞培养基中的用途。本发明还涉及包含RA激动剂或拮抗剂的细胞培养基。

[0153] 一方面,本发明提供了通过短暂加强RA信号传导将细胞重编程为诱导性多能细胞的方法。在一方面,加强RA信号传导可以包括使体细胞与下述中的一种或多种接触:

[0154] i 视黄酸受体(RAR/RXR)家族成员的基因产物或,该基因产物的激动剂;

[0155] ii Lrh1家族成员的基因产物;或该基因产物的激动剂;

[0156] iii 视黄酸或参与视黄酸家族成员的合成或代谢的基因产物;或该基因产物的激动剂;

[0157] iv 参与视黄酸家族成员转运的基因产物;

[0158] v 编码(i)-(iv)中任一的多核苷酸。

[0159] 合适的是,将RA信号传导加强3、4、5、6、7或8天。在一方面,不超过8天。

[0160] 本领域技术人员可以容易地确定基因产物的合适剂量,从而优化iPS产生。诸如上文所列家族成员的诸如蛋白质的基因产物或其片段或变体可以直接添加于细胞,或可以例如通过插入表达期望的基因产物的表达载体在所操作的细胞内表达,或可以通过启动子插入或操作而操作宿主基因表达。

[0161] 因此,本发明还涉及体细胞,其中一种或多种下述基因产物的表达水平已经通过

宿主细胞基因组的转染或基因操作(例如通过靶向插入事件)而得到修饰:

[0162] i 视黄酸受体(RAR/RXR)家族成员的基因产物,或该基因产物的激动剂;

[0163] ii Lrh1家族成员的基因产物,或该基因产物的激动剂;

[0164] iii 参与视黄酸家族成员的合成或代谢的基因产物;或该基因产物的激动剂;

[0165] iv 参与视黄酸家族成员转运的基因产物。

[0166] 本发明还涉及通过本文所述方法获得的或可获得的诱导性多能干细胞。iPS可以来自任何物种,但是合适的是,获得自作为初始材料的小鼠或人体细胞。

[0167] 体细胞可以是任何合适的细胞,例如成纤维细胞。

[0168] 特别是本发明涉及通过本文所述方法获得的或可获得的诱导性人类多能干细胞。

[0169] 本发明合适地具有下述特性中的一种或多种:

[0170] iPS细胞合适地是未分化的多能细胞。多能性可以利用下述方式评估,例如检测诸如Oct4、Nanog、Rex1的一种或多种多能性标志物的表达,或使Oct4、Nanog、Rex1中的一种或多种的启动子区脱甲基化。

[0171] 在本申请的语境中,Oct 4的表达表示与报道ES细胞系(其中IRES-Puro-Egfp盒靶向Oct4基因座)中的表达水平相当的表达,如下文实施例所述和图2a和图7所示。这些报道ES细胞对2.0 μ g/ml的嘌呤霉素是抗性的。

[0172] 合适的是,iPS细胞可以解离成能在适当的培养基中进行细胞分裂的活单细胞。

[0173] iPS细胞可以合适地生长于M15+hLIF培养基(对于600ml M15,500或504ml(82%) GIBCO™敲除™ D-MEM(Invitrogen,目录号:10829018)、90ml(15%)胎牛血清(所测试的ES细胞)、6ml(1%)青霉素-链霉素-谷氨酰胺(100X)、液体(Invitrogen,目录号:10378-016)、4.3ul β -巯基乙醇(Sigma,目录号:M 7522)LIF(1X10⁶单位)中。

[0174] iPS细胞可以合适地生长于2i+LIF培养基中。(2i+LIF培养基是2i补充了1u/mL人重组LIF的2i培养基。2i培养基是补充了1uM PD0325901和3 μ MCHIR99021的N2B27。N2B27培养基是DMEM/F12和Neurobasal培养基的1:1的混合物,DMEM/F12补充了修饰的N2(胰岛素25 μ g/mL、apo-转铁蛋白100 μ g/mL、黄体酮6ng/mL、腐胺16 μ g/mL、亚硒酸钠30nM、牛血清白蛋白组分V 50 μ g/mL),Neurobasal培养基(Neurobasal medium)补充了B27。DMEM/F12、Neurobasal培养基以及B27均来自Gibco)。

[0175] 本发明的iPS细胞合适地能在合适的细胞培养条件下分化成体细胞。

[0176] 合适的是,本发明的iPS细胞可以用于在胚胎注射后产生嵌合体,从而证明种系传递。

[0177] 当注射到小鼠中时,iPS细胞合适地能形成畸胎瘤。

[0178] 在一方面,iPS细胞生长合适地不依赖人类ES细胞通常所需的FGF。

[0179] 可以将原代iPS细胞合适地解离成单细胞,随后所述单细胞在诸如M15+LIF的合适培养基中能形成次级集落和稳定的细胞系。

[0180] 本发明的iPS细胞合适地在形态上不同于在相同的条件下仅利用4种Yamanaka因子oct4、cmc、sox2和Klf4制备的那些细胞。

[0181] 在一方面,iPS在预失活的状态下包含2个X染色体。小鼠雌性基态ES细胞中的2个X染色体是预失活的,而小鼠EpiSC已经经历了X失活。用6F)Yamanaka因子LRH1和RARG)所产生的雌性iPSC具有2个活性X染色体。

[0182] iPS细胞的基因组完整性可以利用光谱核型分析证实。即使在体外培养后,细胞仍然保持正常的核型,这一事实表示遗传稳定性/基因组完整性。合适的是,在传代后(例如在至少2、3、4、5、6、7、8、9或10、15或20代后),iPS细胞具有正常的核型。

[0183] 因此,在一方面,本发明涉及特征为下述至少之一的诱导性人类多能细胞:

[0184] ●表达一种或多种多能性标志物,如Oct4、Nanog、Rex1等;

[0185] ●不依赖FGF而生长;

[0186] ●在传代后保持正常的核型,所述核型利用光谱核型分析进行测量;

[0187] ●当注射到小鼠中时能形成畸胎瘤;

[0188] ●Oct4、Nanog、Rex1中的一种或多种的启动子区脱甲基化;

[0189] ●细胞可以解离成能形成次级集落的活单细胞;

[0190] ●在M15+hLIF培养基和/或2i+LIF培养基中增值;

[0191] ●外源重编程因子(例如,LRH1、Rarg、Oct4、Sox2、Klf4以及c-Myc中的一种或多种或其等同物)表达不存在或不显著;

[0192] 在一方面,本发明的iPS细胞不表达外源重编程因子。如说明书到处所讨论的,这可以利用诸如Tet-On的诱导型启动子来实现,Tet-On允许通过从培养基中除去Dox而关闭外源重编程因子的表达。这还可以通过以下方式实现:防止外源重编程因子整合到iPS细胞的染色体中或从基因组中除去任何整合的外源重编程因子,从而使得iPS细胞不含任何外源重编程因子。在一些情况下,用诸如Tet-On的启动子关闭外源重编程因子的表达,可以允许通过RT-PCR检测但在生物学上无关紧要的外源因子的一些表达。

[0193] 在一方面,本发明的人类iPS细胞在FGF培养基中生长不好并且不保持多能性。

[0194] 在一方面,iPS产生可以通过构建含有本文所述的oct4-IRES-puro-egfp敲入的体细胞来评估。含有所述构建体的细胞被视为哺乳动物(如小鼠或人类)iPS细胞,其合适地对浓度为2ug/ml或更高的嘌呤霉素具有抗性,且优选地还是GFP阳性的,这可以合适地通过流式细胞术检测。

[0195] 在另一方面,本发明涉及通过分化本文所述的诱导性多能干细胞而衍生的体细胞,以及涉及包含通过分化本文所述诱导性多能干细胞而衍生的一个或多个体细胞的组织、器官或非人类动物。利用本领域技术人员已知的各种因子、化合物或因子和化合物的组合,可以在体外实现分化。

[0196] 可通过分化本发明的iPS细胞而获得的细胞(例如心肌细胞、胰岛素产生细胞、神经细胞等)潜在地特别有用,这是因为它们可以用于多种疾病的干细胞移植疗法,所述疾病例如心功能不全、胰岛素依赖性糖尿病、帕金森氏病、癌症以及脊髓损伤,由此可以避免移植后有关使用人类胚胎和排斥的伦理问题。

[0197] 由本文所公开的iPS体外/离体产生的细胞、组织以及器官,在评价诸如药物的化合物的功能或毒性方面是有价值的。因此,本发明还涉及iPS细胞和由此类iPS细胞衍生的组织、器官以及非人类动物在分析化合物或医药治疗的效果中的用途。

[0198] 人类iPS细胞可以用于与感染性物质有关的高通量筛选读取。一方面,遗传学修饰的或遗传学突变的(包括纯合突变体)(主要是丧失功能突变)人类iPS细胞或由iPS细胞分化的体细胞可以被病原体感染,从而鉴定抑制或增强感染的因子。

[0199] 本发明还提供了利用诱导通过上述方法所获得的诱导性多能干细胞分化而获得

的各种细胞,评估化合物、药物、毒物等的生物功能或毒性的方法。

[0200] 本发明还涉及药物组合物,其包含本文所述的核重编程因子或细胞或组织以及药物可接受赋形剂。合适的赋形剂尤其包括,药物可接受缓冲液、载体和水。

[0201] 在一方面,本发明的组织或器官或动物并非衍生于人类iPS细胞。在可选的方法,所述组织或器官衍生于人类iPS细胞。在一方面,本发明不延伸至人类胚胎或成体,但可以延伸至人类器官或组织或在胚胎状态确定前衍生自单个iPS细胞的人类细胞群。

[0202] 在一方面,iPS细胞是单个人类细胞。

[0203] 本发明的iPS细胞的基因组例如通过转座子进行合适的遗传操作。我们已经证明,可以用高效转座到染色体中的转座子转染人类iPS细胞。因此本发明涉及对iPS细胞进行遗传修饰的方法,所述方法包括将外源多核苷酸序列递送到宿主染色体中,合适的是,递送转座子或基因捕获构建体或基因靶向构建体。然而,可以将任何合适的构建体递送到本发明的iPS细胞中。

[0204] 在一方面,本发明的iPS细胞易于进行同源重组和/或靶向,从而合适地允许校正遗传突变,并易于潜在地在患者细胞中引入新的有利遗传变化。

[0205] 本发明还涉及iPS细胞,其基因组包含外源DNA序列。

[0206] 本发明还涉及医药治疗方法以及本发明在医药中的用途。

[0207] 本文公开的iPS细胞或由其产生的体细胞可以在治疗中用作药物。此外,由人类体细胞产生患者特异性人类iPS系的能力使得校正iPS细胞中患者特异性遗传异常并产生用于基因疗法的无突变细胞成为可能。由本文所述iPS产生的组织或器官也可以用于医药中。NRF也可以用作药物,其中它们被递送且在体内直接发挥作用,以及它们在产生iPS中的用途。

[0208] 因此,本发明涉及合适地如本文所述的核重编程因子或iPS细胞或由其衍生的细胞、组织或器官在医药中的用途。

[0209] 本发明还涉及治疗有需要的患者的方法,所述方法包括合适地如本文所述,递送药物可接受量的重编程因子或iPS细胞或由其衍生的细胞,或递送由本文所述的iPS细胞衍生的组织或器官。

[0210] 本发明还涉及合适地如本文所述的核重编程因子或iPS细胞或由其衍生的细胞在制备用于治疗有需要的患者的药物中的用途。

[0211] 如果体细胞分离自身体并且重编程为iPS细胞,则合适的是,所述细胞来自意图得到治疗的患者。

[0212] 本发明还涉及促进分化的方法,其利用本发明的细胞或NRF,例如,产生起始iPS细胞或细胞群,由所述iPS细胞或细胞群产生分化的细胞。一方面,RAR家族成员如Rarg、Rara的激动剂或拮抗剂或LRH1的激动剂或拮抗剂可以用于促进患者的细胞分化。

[0213] 本发明的可选方法涉及促进细胞的分化而不是重编程。如本文所公开的,全长Rarg和Rara的过表达增强RA所诱导的分化。本发明因此还涉及改善细胞的分化能力和/或生长能力的方法,所述方法包括使体细胞与诸如Rarg或Rara的全长RAR并联合RA以适合引发分化的浓度和持续时间进行接触。

[0214] 任何合适的靶细胞都可以用于本发明中。用于通过本发明的NRF治疗的合适的细胞包括肝细胞、胃细胞、皮肤细胞以及单核细胞。在一方面,用于治疗的靶细胞表达LRH1/

RARG,或具有低水平的本文所述的COUP-TFI/II和其他负效应物,或两者。

[0215] 本发明还涉及例如COUP-TFI、COUP-TFII、GCNF和G9a以及为多能基因的调节剂并与RA受体和LRH1家族成员相互作用的其他因子的miRNA和siRNA,所述miRNA和siRNA可以改善iPS细胞质量和产生效率。本发明还涉及利用野生型或突变形式的COUP-TFI、COUP-TFII、GCNF和G9a以及多能基因的其他调节剂改善iPS细胞质量和产生效率。

[0216] 本发明一方面涉及能增加iPS细胞的数量和/或质量的某些蛋白质以及蛋白质的组合,例如RAR α 、Rarg和Lrh1。在本文中提到的蛋白质或多肽应当认为包括所述蛋白质或多肽的能引发相同的或基本上相似的生物活性的功能等同物。相似的活性可以是,刺激iPS细胞产生和/或iPS质量的能力,这可以合适地利用本文所公开的方法评估。作为非限制性实例,已知全长RAR α 增强iPS产生,该蛋白质的功能等同物也增强iPS产生,合适地以利用RAR α 所观察到的水平的至少20、30、40、50%或更多增强iPS产生。功能等同物可以包括来自同一物种或不同物种的蛋白质家族的其他成员,或诸如取代、添加或缺失突变体的蛋白质变体。

[0217] 同样,在本发明涉及例如编码本文所公开的蛋白质或多肽或其功能等同物的多核苷酸方面,本发明涉及编码所述序列的任何多核苷酸,例如天然存在的序列或所述天然存在的序列的简并等同物。本发明包括这样的多核苷酸,其一条链与编码与本发明的蛋白质或多肽在功能上等价的蛋白质或多肽的多核苷酸合适地在严格性杂交条件下杂交。严格性条件可以包括例如将约45°C的6XNaCl/柠檬酸钠(SSC)应用于杂交步骤,随后50°C下用2XSSC洗涤,或例如,可选地于42°C下5XSSC、20mM NaPO₄pH6.8、50%甲酰胺下杂交并于42°C、0.2XSSC中洗涤。本领域技术人员理解,基于待杂交的序列的长度和GC核苷酸碱基含量,需要依据经验改变这些条件,并且存在用于确定此类改变的公式(参阅例如Sambrook et al, "Molecular Cloning: A Laboratory Manual", Second Edition, pages 9.47-9.51, Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press(1989))。

[0218] 应当理解到,以说明而非限制本发明的方式显示了本文所述的具体实施方案。在不偏离本发明范围的情况下,本发明的主要特征可以用于各种实施方案。本领域技术人员利用仅仅常规研究会意识到或能查明本文所述具体规程的许多等同规程。这类等同规程视为位于本发明的范围内,并为权利要求涵盖。说明书所提到的所有出版物和专利申请表示本发明所述领域技术人员的技术水平。将所有出版物和专利申请通过援引并入本文,其并入的程度如同特别和单独地表示将每一单独的出版物或专利申请通过援引并入。当与权利要求和/或说明书中的术语“包含”共同使用时,词语“a(一)”或“an(一)”的使用可以表示“一”,但是还与“一或多”、“至少一”和“一或多于一”的涵义一致。在权利要求中使用术语“或”用来表示“和/或”,除非清楚地指出仅仅指可选项或可选项相互排斥,尽管所公开的内容支持仅指可选项和“和/或”的定义。在本申请中,术语“约”用来表示,值包括装置、用来确定值的方法的误差的固有变动,或存在于研究主题中的变动。

[0219] 如本说明书和权利要求书所用,措词“包含(comprising)”(和包含的任何形式,如“包含(comprise)”和“包含(comprises)”、“具有(having)”(和具有的任何形式,如“具有(have)”和“具有(has)”、“包括(including)”(和包括的任何形式,如“包括(includes)”和“包括(include)”或“含有(containing)”(和含有的任何形式,如“含有(contains)”和“含有(contain)”)是包括在内的或开放式的,并且不排除其他未提及的要素或方法步骤。在一方面,这类开放式的术语还包含其范围内的限制或封闭式定义,例如“基本上由……组成

(consisting essentially of)”或“由……组成(consisting of)”。

[0220] 本文所用术语“或其组合(or combinations thereof)”指该术语之前的所列项目的排列和组合。例如，“A、B、C或其组合意图包括下述至少之一：A、B、C、AB、AC、BC或ABC，如果顺序在具体环境中重要，还有BA、CA、CB、CBA、BCA、ACB、BAC或CAB。以该实例继续，清楚地包括含有一个或多个项目或术语的重复，如BB、AAA、MB、BBC、AAABCCCC、CBBAAA、CABABB等。技术人员应当理解，通常不以任何组合限制项目或术语的数目，除非根据上下文显而易见。

[0221] 根据本发明公开的内容，在无需过度实验的情况下，可以制备并实施本文所公开和请求保护的所有组合物和/或方法。尽管已经根据优选的实施方案描述了本发明的组合物和方法，但是对本领域技术人员显而易见的是，在不偏离本发明的概念、实质和范围的情况下，可以改变本文所述的组合物和/或方法和步骤或步骤顺序。所有此类替代和修饰对本领域技术人员是显而易见的，并视为位于所附权利要求所限定的本发明的实质、范围和概念内。

[0222] 通过援引将本文提到的所有文献并入，至可准许的最充分程度。

[0223] 除非根据本申请的上下文显而易见，否则本公开内容的任何要素清楚地考虑了与本公开内容的任何其他要素的组合。

[0224] 参考下述实施例，进一步描述本发明，但并非限制本发明。

[0225] 实施例1

[0226] 表达视黄酸受体 γ (RARG)和肝受体同系物1(LRH-1)以及4种Yamanaka因子将小鼠和人类体细胞重编程为基态iPSC。

[0227] 摘要

[0228] 干细胞具有自我更新的能力和分化成不同细胞类型的潜力。可以通过表达4种转录因子将体细胞重编程为诱导性多能干细胞(iPSCs)，但是机制仍然不明。在此我们报导了，调节视黄酸(RA)信号传导强烈促进重编程。此外，Rarg(视黄酸受体 γ)和Lrh-1(肝受体同系物1)与4种因子的共表达，导致将小鼠胚胎成纤维细胞(MEFs)在化学限定2i培养基中直接快速重编程为基态或天然iPSC。RA信号传导与RARG和LRH-1在重编程中的关键功能在进化上是保守的，这是因为因子的这种组合将人类成纤维细胞重编程为iPSC，其在生长特性、基因表达、信号传导依赖性以及对遗传修饰的感受性方面与天然小鼠ES细胞相似。

[0229] 可以通过表达4种Yamanaka因子即Oct4、Sox2、c-Myc和Klf4，将小鼠胚胎成纤维细胞(MEFs)重编程为iPSC(1-3)。这些因子也能重编程小鼠(4-6)和人类(7-9)中的各种体细胞谱系的细胞。已经描述了最初重编程方案的许多改善，其包括替换Oct4(10)、使用化学化合物(11-14)、调节诸如Wnt信号传导的信号转导途径(15)、扰乱诸如p53的细胞周期调节剂(16,17)以及增强小鼠iPSC的种系能力(18)。重要的是，通过在四倍体互补测定中产生足月成体小鼠，小鼠iPSC已经通过最严格的多能性测试(19-21)。尽管在iPSC领域有了巨大进展，但是仍然存在技术挑战，并且有关重编程机制知之甚少。例如，在小鼠中，重编程MEF花费2-3周，且仅少量重编程细胞是真正的iPSC(22,23)。在人类中，由原代细胞重编程的iPSC与常规人类ES细胞相似，其被认为具有更多的小鼠Epi干细胞(EpiStem cell)特征(24)。

[0230] 结果

[0231] 重编程需要RA信号传导

[0232] 在脊椎动物发育过程中,视黄酸信号传导具有复杂和多效性功能(25)。延长对高浓度RA的暴露诱导小鼠胚胎干(ES)细胞和胚胎性癌(EC)细胞的分化。然而,生化研究表明,在存在低浓度的RA的情况下,RA受体(RARs)能通过结合RAREoct的RAR:RXR异源二聚体正调节Oct4表达(28,29),RAREoct是位于Oct4基因座启动子中的复合的RA响应元件(26-28)。此外,高水平的异源二聚体有效地从结合RAREoct竞争掉诸如COUP-TF的阻遏物(28)。

[0233] 为了研究RA信号传导在重编程中的作用,我们将Oct4、Sox2、Klf4、cMyc、Rara以及Rarg的cDNA克隆到PB-MSCV载体中,其中cDNA的表达受MSCV LTR控制(图6A)。尽管外源c-Myc并非重编程必须的,但是省略它会降低重编程效率并且延迟该过程(23,30)。将个体PB-MSCV-cDNA与PB转座酶质粒共转染到MEF中(图1A)。PB转座促进PB转座子有效整合到基因组中,从而确保cDNA的稳定表达(31)。

[0234] ES细胞的多能性需要适当水平的Oct4(Pou5f1)表达(32),且已知Oct4的激活是重编程中的关键事件(3)。为了监测重编程中内源Oct4基因组的激活并允许公正地评估iPSC质量,我们制备Oct4报道小鼠系,其中将IRES-Puro-GFP盒靶向Oct4基因座的3'UTR(图6B)。靶向的ES细胞对2.0 μ g/ml嘌呤霉素(Puro)选择具有抗性。这些细胞在荧光显微镜下并非可见的绿色,但是可以通过流式细胞术区分(图6C)。将Puro-GFP盒插入Oct4基因座,看起来不影响ES细胞多能性,这是因为种系有活性的嵌合体来源于这些ES细胞。因此将来自Oct4报道小鼠的MEF用于所有MEF重编程实验中,因此,可通过iPSC中的Puro抗性或GFP表达,方便地监测重编程和iPSC质量。

[0235] 由ES细胞样细胞组成的集落在转染后3三周开始出现,且截至到第30天,Rara和Rarg的表达分别使AP⁺ES细胞样集落的数量增加一和二数量级(图1B-C)。相比之下,在大多数实验中,表达Rara显性失活(DN)形式(33)完全阻断重编程(图1B),并且少数生长中的集落不能扩增成稳定的系。因此,这些结果表明RA信号传导在将MEF重编程为iPSC中的关键作用。

[0236] 接下来,我们利用RA的对Rarg具有特异性的合成激动剂CD437(34)检查了RA信号传导在重编程中的特异性和短暂需要。以推荐的浓度向培养基添加CD437达4或8天大幅增加将MEF重编程为iPSC(图1D)。然而,转染后用CD437处理MEF超过8天(图1D),或转染增加量的PB-MSCV-Rarg质粒DNA,降低了重编程效率(图6D),这反映了RA信号传导在重编程中的时间依赖方式和剂量依赖性方式。

[0237] 内源Oct4的表达,在4种Yamanaka因子重编程的iPSC中是不稳定的(11)。我们还发现,大多数iPSC样集落(PB转染后第30天),无论是由4种因子单独重编程还是由4种因子加Rarg重编程的,仅存活于1.0 μ g/ml Puro选择,且GFP表达不可检测(图6E),这表明在这些细胞中Oct4表达水平低。进一步的分析证实了它们的部分重编程状态:低水平的内源多能性基因表达(图1E),在Nanog和Rex1启动子的差异甲基化的区域(DMR)中,保留了大量DNA甲基化(图1F)。另一方面,获得了少量这样的集落:其对2.0 μ g/ml Puro具有抗性,且在多能性基因表达、Nanog和Rex1启动子中的脱甲基化以及GFP表达方面与ES细胞相似(图1E-F和图6E)。

[0238] 令人惊讶的是,通过CD437(1-4天)处理所获得的大多数集落表达GFP(图1G)并且存活于2.0 μ g/ml Puro选择。因此,调节RA信号传导不仅极大增强重编程效率,而且还改善稳定的Oct4表达所定义的iPSC质量。在随后的实验中,我们使用对2.0 μ g/ml Puro的抗性

GFP表达作为由Oct4-Puro-GFP MEF重编程的小鼠原代iPSC多能性质量的替代标准。。

[0239] 通过共表达Rarg和Lrh-1进行快速重编程

[0240] SF-1(Nr5a1)和LRH-1(Nr5a2)是Nr5a类固醇激素家族的2个成员。通过形成结合于RAREoct的异源二聚体,Sf-1和Rarg协同促进和维持Oct4在EC细胞中的表达(29)。然而,Lrh-1而非Sf-1在ES细胞中表达,并且,在胚胎发育的移植后阶段是必须的(35)。我们因此构建了两种PB载体:PB-CAG-OCKS和PB-CAG-RL,其中4种Yamanaka因子(OCKS)、Rarg和Lrh-1(RL)分别通过T2A连接(图7A)。这些因子的表达受CAG启动子的控制。将这两种PB载体共转染或单独转染到Oct4报道MEF中,用于重编程。共转染OCKS和RL 4天后,立即出现2.0μg/ml Puro^r显微 ES样集落。这些Puro^r集落继续生长成大ES样集落,并且通常从第7天到第10挑取它们用于扩增和进一步表征(图7B)。在转染后第10天,与仅表达OCKS的对照相比,存在高达20多倍的由6种因子重编程的AP⁺集落(图7C)。单独的RL转染在10天后不产生任何集落。

[0241] 以前利用基于病毒的强力霉素诱导型载体进行的研究表明,外源因子诱导的重编程是具有明确的中间步骤的逐渐过程,其中不同的细胞群可能变成iPSC(36,37)。在细胞进入自我维持多能性状态之前,经常需要表达Yamanaka因子至少10-16天的时间(36,37)。为了确定在重编程中短暂需要6中外源因子,我们将CAG启动子转换为四环素应答元件(TRE),以便这些因子的表达可以被强力霉素(Dox)诱导(图7A)。将两种PB载体即PB-TRE-OCKS和PB-TRE-RL,以及表达反向四环素反式激活剂的PB-CAG-rtTA(图7A)转染到Oct4报道MEF中。将Dox添加到培养基中,并随后在几个时间点除去(图2A)。4天的Dox处理足以获得存活于获得2.0μg/ml嘌呤霉素的重编程细胞,其在缺少Dox的情况下继续生长成iPSC集落(图2B-C)。相比之下,如果仅表达Yamanaka因子,则获得2.0μg/ml嘌呤霉素细胞需要至少12天(PB-TRE-OCKS)(图2B)。因此,共表达Rarg/Lrh-1使重编程显著加速。重要的是,根据它们对2.0μg/ml Puro的抗性和GFP表达,用这6种因子所产生的大多数iPSC的质量很高(图2D)。相比之下,Rarg或Lrh-1单独和OCKS表达不显著影响重编程动力学和iPSC质量(图2D)。

[0242] 除了Oct4报道系以外,我们还制备了Rex1-GFP小鼠系,其中将GFP-IRES-Puro盒插入Rex1基因座中,使用这些Rex1-GFP MEF并利用Dox诱导型因子进行重编程实验。Rex-1仅在ES细胞中表达,而不在EpiSC中表达(Brons et al.,2007;Tesar et al.,2007),因此代表完全重编程细胞的理想标志物。利用4种Yamanaka因子、Lrh1和Rarg进行4天Dox处理后,不依赖Dox的GFP阳性集落再次出现。

[0243] 多能性基因如Oct4的完全激活是重编程以及体细胞克隆中的关键事件(38)。因此,在报道基因测定中,我们利用连接于Oct4启动子的荧光素酶表达盒研究了Rarg和Lrh-1的表达是否的确能快速激活Oct4表达(图7D)。将携带各种cDNA的PB-CAG转座子与报道质粒共转染到MEF中,在2天后收获MEF。OCKS或单独的Rarg或Lrh-1的表达对荧光素酶报道基因表达不具有实质影响(图2E)。然而,Rarg和Lrh-1在ES细胞中的共表达,将报道基因表达增加了4-5倍(图2E),这与以前的生化发现相似,即Rarg和Sf-1协调激活Oct4在EC细胞中的表达(29)。因此,Rarg和Lrh-1的表达在6因子重编程系统中主要负责快速激活Oct4基因座,而OCKS可能在相对晚的阶段参与重编程。

[0244] 进一步表征来自MEF的用6种因子重编程的不依赖Dox的iPSC。这些iPSC表达通过免疫染色和qRT-PCR分析所检测的适当水平的多能性基因,并且在Oct4和Nanog的启动子区域具有DNA脱甲基化(图3A-C和图8A-B)。此外,当注射到F1杂种小鼠中时,这些细胞在畸胎

瘤中能分化成代表所有3个胚层的体细胞类型(图3D)。最后,一旦注射到胚泡中,这些iPSC高效地有助于嵌合体中的种系(图3E)。重要的是,大多数iPSC系不表达可检测的外源因子(图8C),并且嵌合体不发展肿瘤(12个月,n=7)。因此,对2.0 μ g/ml Puro的抗性和GFP表达,是评价来自Oct4报道MEF的原代iPSC的合适标准。因为大多数利用我们的6因子系统重编程的iPSC集落得到完全重编程,因此对于常规重编程而言不需要报道MEF。我们能容易地由野生型C57B6J MEF产生iPSC。此外,携带所有6种因子的cDNA的单独的PB转座子(OCKSRL)也有效地重编程MEF。

[0245] 在2i培养基中将MEF直接重编程为基态iPSC

[0246] 重编程MEF通常在含有血清或血清替代品的标准小鼠ES细胞培养基中进行。这些培养基的复杂属性不允许准确确定单个化合物或生长因子在重编程中的功能。在化学限定培养基中直接重编程体细胞,还会促进治疗性iPSC产生。因此,我们试图在化学限定培养基N2B27-LIF中进行重编程,N2B27-LIF能维持小鼠ES细胞的多能性(39)。用单独的PB-TRE-OCKS或与PB-TRE-RL一起转染MEF,并将其在明胶化平板中在具有Dox的N2B27-LIF培养基内培养高达17天。早在表达6种因子后的第5天,便出现ES细胞样集落,在第14-17天挑选所述集落,用于在添加两种抑制剂即PD(ERK1/2抑制剂PD0325901)和CH(GSK3抑制剂CHIR99021)的N2B27-LIF(2i/LIF培养基)中扩增。在N2B27培养基中添加2i可选择并维持基态小鼠和大鼠ES细胞(40,41)。在每次转染中通过表达OCKS和RL,我们获得平均50个iPSC集落,其中大多数集落(83%)能在2i/LIF中扩增成系(图4A)。相比之下,AP⁺集落直到转染PB-TRE-OCKS后至少12天才出现,并且仅有少量AP⁺集落(28%)能扩增成稳定的系。

[0247] 接下来,我们试图在2i/LIF培养基中将MEF直接重编程为iPSC。通过表达OCKS加RL而不是单独的OCKS,基态iPSC在培养时显著地快速出现(图9A),这表明,一些MEF,在表达6种因子后不久,就已处于重编程阶段,它们在2i/LIF培养基中,被进一步促向基态多能性(11,36,37)。接下来,我们检查了在2i/LIF中重编程的iPSC的质量。这些iPSC表达适当水平的多能性基因(图4B-C),并且在体外(图4D)和在畸胎瘤(图9B)中分化成所有3个胚层的体细胞类型。因此,通过表达Rarg和Lrh-1调节视黄酸信号传导,将MEF快速重编程为基态或天然iPSC。

[0248] OCKS和RL能产生具有独特特性的人类iPSC

[0249] OCKS加上RL通过有效诱导内源Oct4表达将MEF迅速重编程为高质量的iPSC的能力促使我们探讨这些因子是否也在重编程人类体细胞中发挥相似的作用。计算机分析确定,Oct4基因座处的RAREoct在几种哺乳动物基因组中是高度保守的(图5A)。实验证据也支持RAREoct在调节Oct4在人类细胞中表达的可能作用(42)。有意思的是,在小鼠基因组和人类基因组中,在约30-40个基因座中鉴定了RAREoct样元件(表S1和S2),但是Oct4基因座是唯一一个在两个基因组中发现含有RAREoct位点的基因座,这表明了该元件在Oct4基因组中的功能意义。

[0250] 因此,我们探讨了产生与小鼠ES细胞相似人类iPSC的可能性。首先通过利用CAG启动子连续表达6种因子,我们产生了人类iPSC。通过共表达PB转座酶,将携带OCKS和RL cDNA的PB-CAG转座子转座到人类新生儿包皮真皮成纤维细胞(HDFn)中。早在转染后10天,便在小鼠ES细胞培养基中形成集落(M15-LIF),有一些集落在形态上与小鼠ES细胞集落相似(例如,紧凑性提高的集落、高核质比以及显著的核仁)(图10A)。在12-14天,挑选这些集落,并

随后利用标准小鼠ES细胞方案进行培养(图10A)。一旦用胰蛋白酶解离成单细胞悬浮液并接种于ST0饲养细胞上,便有效形成次级集落,再接种效率高达70%。我们能由许多原代集落建立稳定的系,其能在M15-LIF培养基中在饲养细胞上维持至少50代。在仅利用OCKS的对照中,原代集落在形态上不同于用OCKS加上RL所产生的那些集落(图10A)。与以前的报道一致(7),我们从来未能由这些OCKS集落在M15-LIF培养基中建立稳定的系。除了它们能在常规小鼠ES细胞培养基中增值以外,这些人类iPSC还表达关键的多能性基因(图10B-C),并且还能维持在2i/LIF培养基中,这证实了它们与基态小鼠ES细胞的相似性(40)(图10A)。为了确定这些人类iPSC的分化潜力,我们将其皮下注射到免疫低下的NSG小鼠(43)中。尽管利用CAG启动子连续表达外源因子,但是这些细胞形成具有代表所有3个胚层的细胞类型的畸胎瘤(图10D)。此外,即使在长期的体外培养后,这些iPSC保持了正常的核型(图10E)。利用Y染色体多态性标志物,我们还证实了iPSC的HDFn来源(图10F)。

[0251] 为了获得不依赖外源因子表达的人类iPSC,我们使用了Tet-On系统,其中将6种因子的人类cDNA克隆到PB-TRE质粒中。将这些PB-TRE-cDNA质粒与BP转座酶和rtTA质粒共转染到HDFn细胞中。添加Dox,诱导外源因子表达(图5B)。转染后10-15天,出现含有ES细胞样细胞的集落。将这些集落进行胰蛋白酶处理,并再次接种到饲养平板中的无Dox的M15-LIF培养基中。随后将不依赖Dox的次级集落在饲养细胞上在添加两种抑制剂PD和CH的KSR-LIF培养基(KSR-2i-LIF)中扩增成系(40,41)。根据一实验,我们建立了维持在KSR-2i-LIF培养基中的5个不依赖Dox的iPSC系(图5C)。与小鼠ES细胞相似,这些人类iPSC易于扩增,并且以1至4的拆分比,每隔3-4天进行传代。为了确定这些不依赖Dox的iPSC的多能性,我们检查了关键多能性基因的表达。它们以可与人类ES细胞相当的水平表达内源Oct4、Sox2和Nanog(图5D和图11A),并且对于人类ES细胞表面标志物为阳性染色(图5E和图11B)。重要的是,即使在RT-PCR分析中,外源因子的表达也是不可检测的(图11C)。最后,可以将不依赖Dox的人类iPSC容易地在体外和体内分化为3个胚层的细胞(图5F-G)。

[0252] 与在小鼠ES细胞中维持多能性依赖Lif/Jak/Stat途径相反,将衍生于胚胎的常规人类ES细胞培养于含有FGF2(碱性FGF)的培养基中,单独的LIF不足以维持多能性(44,45)。不依赖Dox的人类iPSC并不需要培养基中的FGF。实际上,它们生长的很好,并且即使在存在FGF受体抑制剂SU54021的情况下,也在KSR-2i-LIF培养基中保持多能性基因表达(图5H)。常规的ES细胞对JAK抑制剂(JAKi)具有抗性,但是在存在BMP4的情况下分化。相比之下,我们的人类iPSC在存在BMP4的情况下增殖良好,但是如果将阻断Stat3磷酸化的JAK添加到KSR-2i-LIF培养基中,它们丧失OCT4和NANOG的表达,并且快速分化(图5H)。最后,与人类ES细胞相比,我们的人类iPSC在KSR-2i-LIF培养基中表达较低水平的谱系特异性基因,如PAX6、GATA6和SOX17(图5H)。

[0253] 为了确定不依赖Dox的人类iPSC是否具有变成FGF依赖性的潜力,我们将培养基变成KSR-FGF2。与生长于KSR-2i-LIF培养基相比,在传代3次后,这些细胞仍然生长良好,并且不表现出明显的分化。它们在形态上与人类ES细胞相似,并且表达适当水平的多能性基因,如OCT4和NANOG(图5I)。然而,一旦将培养基变回KSR-2i-LIF,这些iPSC便变成分化的并且表达高水平的SOX17和SOX1,伴随着OCT4和NANOG表达的丧失(图5I),这与人类ES细胞的行为相似(44)。因此这些数据表明,生长于KSR-2i-LIF培养基中的人类iPSC具有转变成与常规人类ES细胞相似的细胞的潜力,但是反之不是这样。因此,基于它们的生长特性、信号传

导依赖性以及基因表达,利用我们的6因子平台所产生的不依赖Dox的人类iPSC,比常规人类ES细胞更不成熟,并且与基态小鼠ES细胞相似(40,41)。因此,我们将这些人类iPSC称为桑格人类iPSC(Sanger Human iPSC)或SH-iPSC。

[0254] SH-iPSC与小鼠ES细胞相似的事实表明,它们可能可用于遗传研究。在利用这些细胞进行大量体外操作之前,我们利用光谱核型分析检查了SH-iPSC的基因组完整性。即使在大量的体外培养(20代)后,这些细胞仍然保持正常的核型(图11D),这表明,它们在遗传上是稳定的。因此我们进行PB转座,从而通过将PB-PGK-Neo转座子和Pbase质粒共转染到SH-iPSC中将外源DNA引入这些细胞。由一次电穿孔,我们获得约500,000个G418^r集落,其占存活于电穿孔的细胞的约10%,转座效率与利用小鼠ES细胞相当(31)。

[0255] 在SH-iPSC中的有效PB转座,为进行基因组范围的诱变筛选提供了机会。为了开始探讨这种可能性,我们将PB-SA- β geo基因捕获转座子和PB转座酶质粒转染到SH-iPSC中。SA- β geo盒整合到表达基因座中,能使 β geo表达和捕获事件评分为G418抗性。由一次电穿孔,我们回收了22,000个G418^r集落,其代表0.3%的电穿孔幸存细胞。因为我们的人类iPSC系衍生于包皮成纤维细胞并且因此是XY,因此我们利用丧失HPRT活性的细胞中的6TG抗性,研究了X染色体上HPRT基因座处的突变效率。在22,000个G418抗性集落中,我们获得了1个6-TG集落。该克隆的分子分析表明,PB插入内含子2中并且干扰HPRT转录和剪接(图11E)。因此,SH-iPSC适合基因操作和筛选。

[0256] 可以将我们6因子重编程系统的独特能力归因于RA信号传代和相关因子LRH-1在多能干细胞中的作用。与视黄酸的充分表征的促进分化活性相比,RA信号传导可以通过RAREoct正调节Oct4表达。高水平的RAR:RXR异源二聚体特异性结合RAREoct,其克服了COUP-TF的抑制(28)。通过RA信号传导调节Oct4表达的可选途径是通过RAREoct内的SF-1结合位点。RARG/LRH-1异源二聚体结合RAREoct,并且通过有效竞争掉阻遏物协同激活Oct4表达(29)。这些类固醇激素受体二聚体随后将共激活剂如CBP/p300、P/CAF和SRC1/TIF2(25,46)募集到Oct4基因座,从而促进进一步的染色质改造。因此,体细胞重编程中最关键的障碍之一,即内源Oct4基因座的激活(47),在外源因子表达之后即易于克服。有意思的是,注意到已经证明,表达高水平LRH-1和低水平COUP-TF的细胞,如小鼠肝细胞(48),具有较高的重编程效率(5)。

[0257] 我们的数据清楚地表明,激活Oct4基因座是早期重编程事件并且是关键的事件。已知通过表达4种Yamanaka因子激活Oct4基因座是不稳定的,这在常规ES细胞培养基中导致许多部分重编程的集落(11)。利用Oct4报道MEF进行重编程,允许直接观察Oct4基因座激活。当仅表达4种因子时,少量细胞在流式细胞术中是GFP⁺,这表明这些细胞中Oct4激活。然而,在接下来的8天中,GFP⁺细胞数量并未实质改变,这与4种因子介导的长期重编程过程一致。相比之下,在表达6种因子的MEF中,从转染后第3天起,GFP⁺以指数方式增加。而且,少量MEF在表达6种因子后(24小时)几乎立即获得高水平的Oct4表达。然而,这种快速的Oct4激活并不能维持,并且在Yamanka因子未持续表达更多天的情况下,其可能失败,这是因为仅在Dox诱导实验中在外源因子表达3-4天后才获得自我维持的iPSC集落(图2)。

[0258] Lrh1不足的ES细胞未表现出严重的多能性缺陷,因此其不是多能性线路(pluripotency circuit)的主要必须组分(35)。然而,其在ES细胞分化中对于Oct4表达的维持是必须的(35)。最近的证据还表明,Lrh1在典型的Wnt信号传导的下游发挥作用,并且

除了Oct4外还调节其他多能性因子,如Nanog和Tbx3(64)。Lrh1的这些关键功能,可能在其能在重编程中替代Oct4中发挥重要作用(10)。然而,我们的数据表明,这种替代可能是环境依赖性的,这是因为单独的Lrh1以及4种Yamanaka因子的表达,在iPSC质量或动力学方面,并不显著改善重编程。仅Rarg和Lrh1之间的协同相互作用深刻影响快速产生基态小鼠iPSC的重编程,并且能产生与基态小鼠ES细胞密切相似的人类iPSC。有意思的是,尽管Rarg在包括MEF在内的许多细胞类型中表达,但是其最高表达见于ES细胞中,MEF或人类成纤维细胞中的内源Rarg看起来并不足以与外源诱导的Lrh1合作来促进快速且有效的重编程。

[0259] 最近的报告表明,通过重编程的次级成纤维细胞(iPSC-衍生的)获得的人类iPSC可以保持在添加有毛喉素的2i/LIF培养基中(49)。这些细胞看起来与基态小鼠ES细胞相似(40),因此提供了基态或天然多能性存在于人类中的独立性证据。尽管他们的人类iPSC可以变成不依赖Dox的,但是它们生长缓慢,并且不能维持超过15-20代。另一最近的研究也描述了仍然依赖于外源因子表达的小鼠ES样人类iPSC(50),其在某种程度上看起来与利用PB-CAG-cDNA获得的PB-CAG人类iPSC相似(图10)。与这些报告相比,SH-iPSC直接由人类原代成纤维细胞重编程,不依赖表达外源重编程因子,并且可以在含有2i/LIF的培养基中维持超过50代。此外,它们不依赖FGF,但是依赖LIF-JAK-STAT途径。因此SH-iPSC与预想的基态人类多能干细胞相似,并且对于人类基因组的功能研究应该是有用的。

[0260] 材料和方法

[0261] 质粒载体的构建

[0262] 为了制备PB-TRE、PB-MSCV和PB-CAG载体,由pTight (Clontech)扩增Tet应答元件(TRE),由pMSCV-Neo(Clontech)扩增MSCV LTR,并由pBluescript-CAG载体扩增CAGG启动子(未发表的数据),并将其克隆到PB-bpA载体中(未发表的数据)。

[0263] 由最初的逆转录病毒载体(Addgene)扩增(引物位于补充表3中)4种小鼠和人类Yamanaka因子的cDNA,并分别将其克隆到PB-TRE、PB-MSCV和PB-CAG转座子载体中。由IMAGE克隆(Geneservice)扩增小鼠和人类Rarg、Lrh1以及Sf1,并将其克隆到转座子载体中。

[0264] 小鼠和人类ES和iPSC培养

[0265] 通常将小鼠ES细胞和iPSC培养于M15培养基中:敲除DMEM、15%胎牛血清(FBS, Hyclone)、1X谷氨酰胺-青霉素-链霉素(GPS, Invitrogen)、1X非必需氨基酸(NEAA, Invitrogen)、0.1mM β -巯基乙醇(3-ME, Sigma)以及 10^6 U/ml LIF(Millipore)。

[0266] 将通过PB-CAG载体重编程的人类iPSC培养于具有 10^6 U/ml人类重组LIF的M15培养基中。将通过PB-TRE载体重编程的人类iPSC维持DMEM/F12中,其具有1XGPS(Invitrogen)、20%敲除血清替代品(Invitrogen)、1X NEAA、0.1mM β -巯基乙醇(β -ME, Sigma)以及两种抑制剂CHIR99021(5 μ M)和PD0325901(1 μ M)。

[0267] 将人类ES细胞系BG01V/hOG(来自Invitrogen)和H1(来自WiCell)培养于hESC培养基中:具有Glutamax (Invitrogen)的DMEM/F 12、20%敲除血清替代品(Invitrogen)、1X NEAA、0.1mM β -巯基乙醇(β -ME, Sigma)以及4.0ng/ml FGF2(Invitrogen)。

[0268] 用于重编程的MEF细胞和HDFn细胞的制备

[0269] 由12.5d.p.c.Oct4-IRES-Puro-Egfp胚胎制备MEF。为了减低胚胎与胚胎之间的差异,将来自具有相同基因型的几个胚胎的MEF混合在一起,用于在M10培养基中扩增。在对MEF进行计数、等分以及冻干之前,使其传代一次。在电穿孔之前,将 1×10^6 个MEF接种到一个

明胶化的15-cm组织培养板中。当MEF汇合70-80%时,将它们用胰蛋白酶处理并收集用于电穿孔。M10:敲除DMEM、10%胎牛血清(FBS, Hyclone)、1XGPS、1X非必需氨基酸(NEAA, Invitrogen)。

[0270] HDFn细胞购自Invitrogen,并维持在补充了低血清生长补充物(Invitrogen)的Media 106中。在计数、等分以及冷冻之前,使原代HDFn培养物传代一次。在电穿孔之前,将 5×10^5 个HDFn细胞接种于3个T75组织培养瓶中。当HDFn细胞融合70-80%时,将它们进行胰蛋白酶处理并收集用于电穿孔。

[0271] 转染和细胞培养

[0272] 利用Amaza机器(Lonza)根据制造商的方案(程序A-023)进行MEF转染。在电穿孔后,将MEF接种于添加LIF的M15中的ST0饲养细胞上。对于Tet-On实验,转染后24小时,添加含有强力霉素($1.0 \mu\text{g/ml}$)的M15,并每隔一天替换。在第7-第10天,将iPSC集落挑选到96孔板中,并根据标准小鼠ES细胞培养条件扩增细胞。

[0273] 利用Amaza机器根据制造商的方案(程序U-020)实现HDFn细胞的转染。电穿孔后,将HDFn细胞接种于添加了hLIF的M15中的ST0饲养细胞上。对于Tet-On实验,转染后24小时,添加含有强力霉素($2.0 \mu\text{g/ml}$)的M15,并每隔一天替换。通常在第10天挑选通过PB-CAG载体重编程的人类iPSC集落,并在接种成24孔模式前用胰蛋白酶解离成单细胞悬浮液。通常在第30天(再接种后10天)挑选通过PB-TRE载体重编程的人类iPSC集落,并在接种于24孔板之前,用胰蛋白酶解离成单细胞悬浮液。由次级集落建立稳定的系,并根据标准小鼠ES细胞培养条件来维持。

[0274] 使用白细胞碱性磷酸酶试剂盒(Sigma),通过碱性磷酸酶染色,观察小鼠和人类iPSC集落。

[0275] 亚硫酸氢盐基因组测序

[0276] 利用EpiTect亚硫酸氢盐试剂盒(Qiagen),根据制造商的推荐,进行亚硫酸氢盐处理。PCR引物列于补充表5中。将扩增产物克隆到pGEM-T-easy(Promega)中。用针对各自启动子的M13正向引物和M13反向引物,对随机选择的克隆进行测序。

[0277] RT-PCR

[0278] 利用Rneasy微型试剂盒(Qiagen),分离RNA。随后定量样品,并用gDNA WipeOut处理。利用QuantiTect反向转录试剂盒(Qiagen),制备cDNA的第一链。对于每个RT-PCR反应,我们使用了50-100ng cDNA和补充表5所列的引物5。标准的PCR条件为: 94°C 30s、 60°C 30s、 68°C 30s;X 30个循环。对于实时PCR,我们使用了Taqman基因表达测定。Taqman探针也购自Applied Biosciences(补充表Table 6)。在9700HT快速实时PCR相同(Applied Biosciences)中,进行所有定量PCR。利用 ΔCt 方法,相对于小鼠Gadph,测定小鼠基因表达。利用 ΔCt 方法,相对于人类GADPH基因,测定人类基因表达。

[0279] 免疫染色和流式细胞术

[0280] 将生长于12孔饲养板中的iPSC用PBS洗涤,于4%PFA/PBS中在室温下固定10min,用0.3%Triton X-100在PBS中渗透10min,在5%驴血清中封闭1h,并添加一抗,于 4°C 过夜。Nanog(1:150, Abcam)、SSEA-3、SSEA-4、TRA-1-60、TRA-1-81(1:10, 由Peter W. Andrews博士惠赠)。将细胞在PBS中洗涤,并添加二抗(Alexa488IgG或IgM, 1:1000; Alexa594IgG, 1:1000),并在室温下孵育1小时。

[0281] 将生长于96孔饲养板中的小鼠iPSC进行胰蛋白酶处理,并重悬浮于M15中。将iPSC 1,000rpm离心3分钟,通过将板在组织上倒置,除去培养基。将iPSC重悬浮于PBS中,并通过Cytomics FC-500(Beckman Coulter)进行分析。

[0282] 将生长于6孔板中的人类iPSC进行胰蛋白酶处理,并重悬浮于M15中。将iPSC用PBS洗涤一次、离心,并与SSEA-4-FITC、TRA-1-60-PE或TRA-1-81-FITC抗体(BD Bioscience)一起孵育1小时。随后洗涤iPSC,并将其重悬浮于PBS中,并通过Cytomics FC-500(Beckman Coulter)进行分析。

[0283] 畸胎瘤形成

[0284] 将小鼠iPSC悬浮于M10中,并将 1×10^6 个细胞皮下注射到F1(129S5/C57B6)杂种小鼠的两侧背侧面。切开畸胎瘤,在磷酸盐缓冲的福尔马林中固定过夜,并在切片之前包埋于石蜡中。将人类iPSCs(1×10^6)皮下注射到NSG小鼠(NOD.Cg-Prkdcscid Il2rgtm1Wjl/SzJ.The Jackson Laboratory)的两侧背侧面。注射后8周,收获畸胎瘤,用于固定和切片。用苏木精和曙红使切片染色。根据UK's 1986Animals Scientific Procedure Act(动物科学步骤法案)和当地的伦理道德委员会规章(local institute ethics committee regulations)进行所有动物实验。

[0285] Splinkerette PCR

[0286] 按以前所述,进行Splinkerette PCR,来确定PB基因组整合位点。用M13正向引物和反向引物,在两个末端,对TA克隆的PCR产物进行测序。利用BLAST确定PB插入位点。利用补充表4所述的引物,进行除去外源因子的iPS系的基因组DNA PCR。

[0287] 统计分析

[0288] 将数据显示为平均值和标准偏差。用Excel 2008(Microsoft)或Prism(Graphpad)进行所有的统计分析。

[0289] 实施例2

[0290] 在这里,我们报导,表达两种新的关键转录因子Rarg和Lrh1,以及4种Yamanaka因子,在少至3天内快速激活内源Oct4基因座,并且允许快速且有效地重编程小鼠胚胎成纤维细胞(MEF)。所产生的iPS克隆,基于形态和分析分析,具有高质量,并且能在体内包括种系中分化成各种细胞类型。当应用相同的策略重编程人类新生儿包皮真皮成纤维细胞(HDFn)时,我们轻易获得了依赖LIF但是不依赖FGF的人类iPS细胞。此外,人类iPS细胞增殖充分,并使其以单细胞密度在常规小鼠ES细胞培养条件下亚克隆,并在无任何可辨别的染色体异常的情况下进行大量扩增。因此,我们的新的iPS细胞平台为在两个哺乳动物物种中实现全部的多能性提供了新的高效方法。对于探讨人类基因组功能,所产生的人类iPS细胞具有极好的资源。

[0291] Rarg显性失活等位基因阻断ES细胞分化

[0292] 为了鉴定在基因突变时能够阻断ES细胞分化的基因,我们利用携带强CAG启动子/增强子和一对剪接受体的piggyBac DNA转座子(PB),在小鼠ES细胞中进行基因筛选(Fig 12a)。将该转座子插入到小鼠基因组中,可以上调插入位点附近的基因或者干扰它们,这取决于插入等位基因的结构。为了有效排除培养物中的分化的ES细胞,我们将Puro-IRES-Egfp盒靶向Rex1基因座(Fig 12b)。因此,未分化的ES细胞在流式细胞术中对嘌呤霉素(Puro)选择会具有抗性并且是GFP⁺,而由于在分化后快速丧失Rex1表达,分化的细胞会被

Puro选择杀死并变成GFP⁻。

[0293] 用1.0 μ M所有反式视黄酸(RA)将突变ES细胞的文库分化4天,并随后用嘌呤霉素进行选择(图12c)。由独立的突变ES细胞库,鉴定了几个Puro^rES细胞克隆。Splinkrette PCR分析在编码Rarg、Pparg以及Gclc的基因中鉴定了PB整合位点。对于Pparg,PB插入使全部的蛋白质编码序列过表达。相比之下,Rarg基因座具有4个独立的PB插入,所有的插入都发现于内含子9(最后一个内含子)中(图12d)。这些PB插入截短了Rarg,其失去外显子10所编码的氨基酸。过表达这种截短的Rarg证实了其阻断RA所诱导的ES细胞分化的能力(图12e)。这种截短的Rarg可能发挥野生型Rarg的显性失活形式(Rarg-DN)的功能。全长形式的Rarg和Rara (Rarg-FL和Rara-FL)的过表达增强了RA所诱导的分化,然而所报道的Rara的显性失活形式(Rara-DN)(20),即使在高浓度RA的情况下,也能阻断ES细胞分化。Rarg和Rara特异性拮抗剂CD2665和Ro-41-5253也能阻断ES细胞分化(图12f)。因此这些数据证明了Rarg和Rara在ES细胞分化中的关键作用。

[0294] 通过Rarg的信号传导增强了重编程效率

[0295] 我们接下来测试在我们的筛选中所鉴定的基因是否在iPS细胞产生中发挥作用。为了直接评价iPS细胞的质量,我们将IRES-Puro-Egfp盒靶向ES细胞中Oct4基因座的3' UTR,用于制备MEF细胞(图6a)(57)。这些细胞在流式细胞术中是绿色的,并且在2 μ g/mL Puro中增殖良好(图6b,且数据未显示)。

[0296] 我们将4种Yamanaka因子单独克隆到PB转座子中,其中每一cDNA的表达受MSCV的LTR(PB-MSCV-cDNAs,图6c)的控制。随后将这些4种转座子和PB转座酶(PBase)质粒转移到Oct4-IRES-Puro-Egfp MEF中。在3周内,出现了与常规ES细胞的集落相似的iPS细胞集落。大多数iPS集落存活于1.0或1.5 μ g/mL Puro,但是很少数能生长于2.0 μ g/mL Puro中。此外,大多数iPS集落是GFP阴性的,这表明,这些细胞中的内源Oct4基因座不表达至ES细胞中的水平。然而,生长于2.0 μ g/mL中的这些少数iPS集落的细胞,大部分是GFP阳性的,其荧光强度与亲代Oct4-IRES-Puro-Egfp敲入ES细胞中的荧光强度相当(图13A)。RT-PCR和DMR测序证实,生长于2.0 μ g/mL Puro中的这些iPS克隆,在多能性基因表达和这些基因的启动子处的DNA甲基化方面,比1.0或1.5 μ g/mL Puro所选择的克隆,与常规ES细胞更相似(数据未显示)。这些结果表证明了许多以前的研究并与之一致,即4种Yamanaka因子所诱导的大多数iPS细胞仅是部分重编程的,并且内源Oct4基因座并未完全激活。因此我们使用2.0 μ g/mL或更高浓度的Puro,联合流式细胞术,来在随后实验中定量评估iPS细胞质量。

[0297] 接下来我们将全长形式的Rarg、Rara、Pparg以及Gclc,以及显性失活的Rarg和Rara克隆到PB-MSCV转座子中,以便研究它们对Yamanaka因子所介导的重编程的影响。Pparg和Gclc的过表达看起来不影响Yamanaka因子介导的MEF的重编程(数据未显示)。递送Rarg-DN以及4种Yamanaka因子,产生了比对照稍微多的iPS集落(图13B)。然而,显著的是,Rarg全长cDNA的表达使iPS集落的数量增加约100倍(图13B和13C)。表达Rara-FL也显著增强重编程效率,尽管程度比Rarg-FL低。相比之下,Rara-DN完全阻断了Yamanaka因子所诱导的重编程(图13B)。这些数据清楚地表明重编程过程中的视黄酸信号传导。

[0298] 与单独的4种Yamanaka因子所诱导的低质量iPS相似,添加Rarg-FL或Rara-FL过表达不明显改善iPS细胞质量,这是因为大多数集落是GFP阴性的,并且仅少数集落存活于2.0 μ g/mL Puro选择。另一方面,来自过表达Rarg-DN的iPS克隆,在整个重编程过程中即使无任

何选择的情况下,具有与用2.0 μ g/mL Puro所选择的iPS克隆相当的相似的GFP阳性细胞百分比(图13D)。

[0299] 表达Rarg-DN改善iPS细胞质量,而Rara-DN过表达完全阻断iPS过程,这一事实表明,RA信号传导水平是重要的。因此我们测试了Rarg的表达水平是否影响重编程效率。重编程效率的显著降低与PB-MSCV-Rarg-FL DNA量的增加相关(图13E)。而且利用较强的CAG启动子表达Rarg-FL将重编程效率降低至对照的水平,对照仅表达Yamanaka因子(数据未显示)。这些结果证实,Rarg过表达对重编程的影响是剂量敏感性的。

[0300] 为了进一步研究Rarg在iPS细胞中的作用,我们在MEF重编程中,联合4种Yamanaka因子,测试了合成的Rarg特异性激动剂CD437。添加100nM的CD437高达8天显著增加iPS细胞集落的数量。此外,iPS细胞的质量,如FACS所测量的,获得显著改善(图13E)。然而,利用CD437处理细胞超过8天(第1-12天)大量降低iPS集落的数量。当给予8天时Rara特异性激动剂AM580也可以改善重编程效率,但是对iPS克隆的质量影响很小(图13F)。这些结果表明,Rarg信号传导在初始阶段或早期阶段促进重编程,但是过度的Rarg信号传导在后期阶段对iPS细胞具有负面影响。尽管表达Rarg-FL或Rara-FL或利用它们的激动剂改善重编程效率,但是它们看起来不缩短重编程时间,这是因为iPS集落的形成仍然需要约3周的时间。

[0301] Rarg和Lrh1协同促进重编程

[0302] 为了探讨Rarg信号传导的下调是否是Rarg-DN iPS克隆更好质量的主要原因,在具有或不具有5-Aza的情况下,我们用Rarg特异性拮抗剂CD2665处理独立地部分重编程的克隆。我们注意到,10 μ M CD2665显著改善部分重编程的克隆的质量,但是当与5-Aza一起使用时,其作用更明显(图14A)。然而,在处理4天后,该克隆仍然主要是GFP阴性的。这促使我们认为,Rarg信号传导的下调可能仅是Rarg-DN iPS细胞具有好得多的质量的部分原因。

[0303] 在小鼠胚胎瘤细胞中(EC),Rarg与孤儿受体Sf1(Nr5a1)形成复合体,所述孤儿受体特异性结合Oct4启动子中的关键应答元件RAREoct,并协同上调Oct4表达(59)。因为表达Rarg极大地促进iPS细胞产生,因此我们决定检查Sf1是否也发挥相似的作用。因为ESf1通常不在小鼠ES细胞或早期小鼠胚胎中表达,因此我们还在实验中包括了其亚家族成员Lrh1(Nr5a2),该成员在ES细胞和早期胚胎中都表达(62)。

[0304] 因此,我们将Sf1和Lrh1cDNA分别克隆到PB-MSCV转座子中,并与Yamanaka因子一起共转染到MEF中。如图14B所示,表达Lrh1显著改善了重编程效率,但是其程度不如Rarg改善的程度。表达Sf1也增加iPS集落的数量,然而许多集落看起来是分化的(数据未显示)。共表达Lrh1和Rarg不进一步增加iPS集落的数量。另人惊讶的是,许多iPS在早至转染后10天而不是通常3周出现,其形态与野生型小鼠ES细胞几乎不能区分(数据未显示)。

[0305] 为了促进在将来切除PB转座子所携带的外源因子,我们制备了PB-CAG-Oct4-2A-cMyc-2A-Klf4-2A-Sox(PB-CAG-OCKS)转座子,其中4种Yamanaka因子用T2A肽序列连接,并由单个CAG启动子表达。PB-CAG-OCKS所诱导的一些iPS集落,在转染后4-5天内,在显微镜下可见。然而,这些集落中的细胞后来在形态上变成异质的(图7B),并且显然仍然不是完全重编程的,因为在第10天挑选的大部分集落(90%)对2 μ g/mL是敏感的,并且含有低百分比的GFP阳性细胞(图2D)。这与以前的出版物一致,即如果仅使用Yamanaka因子,则iPS变成完全重编程的需要许多传代。

[0306] 接下来,我们制备了PB-CAG-Rarg-2A-Lrh1(PB-CAG-RL)转座子,来由单个启动子

共表达Rarg和Lrh1。将PB-CAG-OCKS和PB-CAG-RL共转染到MEF中4天后,在转染后的4-5天内,在显微镜下可见很小的ES样集落。这些集落看起来在形态上是同质的,并且在生长方面与常规小鼠ES细胞非常相似(图7B)。此外,与仅转染了PB-CAG-OCKS的对照相比,有大致多于7倍的iPS细胞集落。有意思的是,在第4-6天之间,我们经常观察到彼此并排形成的相同大小的2个、4个甚至8个很小的集落,在第8天时,其合并成1个或2个大集落,这反映了可能同时重编程来自一个亲代MEF细胞的子细胞。有趣的是,几乎所有这些集落对2 μ g/mL Puro选择都具有抗性,并且由高百分比的GFP阳性细胞组成(图2D),这表明由内源Oct4基因座的强Oct4表达。

[0307] 为了研究能以多快激活内源Oct4基因座,我们在转染后的几个时间点,对MEF进行Puro选择。早至用PB-CAG-OCKS和PB-CAG-RL转染后3天,选择出Puro抗性iPS细胞集落(图14C),由此证实在此条件下Oct4基因座的快速激活。在仅用PB-CAG-OCKS转染的对照MEF中,则仅在Puro选择在第8天或以后开始时出现Puro抗性集落。即使对于这些Puro抗性集落而言,我们经常在原代集落以及亚克隆过程中观察到大部分细胞被杀死,这表明这些部分重编程的克隆的异质性。

[0308] 为了确定建立完全重编程iPS克隆是否需要外源因子的持续表达,我们制备了PB-TRE-Oct4-2A-cMyc-2A-Klf4-2A-Sox2(PB-TRE-OCKS)和PB-TRE-Rarg-2A-Lrh1(PB-TRE-RL)构建体,其中通过强力霉素(DOX)诱导了重编程因子的表达。将两种PB转座子构建体,以及PB-CAG-rtTA载体和Pbase质粒,共转染到MEF中。转染后6、8或10天,立即应用DOX诱导,我们在第10天应用2 μ g/mL Puro选择。在用DOX诱导仅6天的MEF中,获得Puro抗性集落(图14D),这表明这6种因子表达6天就足以将内源Oct4基因座激活到高水平。相比之下,在仅用PB-TRE-OCKS和PB-CAG-rtTA转染的MEF中,在相同的DOX诱导方案下,回收了非Puro抗性集落。因此,仅4种Yamanaka因子表达高达10天,不足以完全激活内源Oct4基因座。为了进一步证明内源Oct4基因座的快速激活,我们重复了6因子Tet/On实验,但是仅在转染后4天用DOX处理MEF。DOX除去后,立即应用Puro选择来选择具有内源Oct4激活的细胞。出现了几个Puro抗性集落,当扩增时,其表现出优异的ES细胞形态。这些结果清楚地表明,表达6种转录因子导致内源Oct4基因座的快速激活,并且导致高质量iPS细胞的产生。

[0309] 为了研究iPS细胞需要多少Rarg和Lrh1,我们将PB-CAG-OCKS以及携带由MSCV或CAG启动子驱动的Rarg-FL、Rarg-DN、Lrh1或Rarg-2A-Lrh1的PB转座子转染到MEF中。iPS质量直接与所用的启动子相关是显然的,这是因为PB-MSCV-RL不显著改善iPS质量(Fig 14E)。单独的Rarg或Lrh1,不管使用哪个启动子,都对iPS质量无任何影响。然而,当由MSCV启动子驱动时,Rarg-DN能显著改善iPS质量。

[0310] 接下来我们制备了PB-CAG-Oct4-2A-cMyc-2A-Rarg-2A-Lrh1-Klf4-2A-Sox2(PB-TRE-OCRLKS),来由单个启动子表达所有6种转录因子。在相似的时间窗内,这种构建体从MEF产生相似数量的高质量iPS细胞。因此我们将MC1-tk盒克隆到PB-CAG-OCRLKS中,从而形成PB-CAG-OCRLKS-MC1-tk转座子。与利用PB-CAG-OCRLKS相似,许多iPS细胞集落在转染到通常在第8天挑选的MEF中4-5天后是可见的。扩增的iPS细胞对更昔洛韦(Ganc)敏感,这表明MC1-tk盒是功能性的。PB转座子通过Pbase切自整合位点,并且在原处不留下足迹。因此我们通过用pL623转染iPS细胞暂时表达Pbase,并且使所述细胞经受Ganc选择。通常,单次转染通常获得数百个Ganc抗性集落。这些Ganc抗性细胞的表征证实消除了这些细胞的PB转

座子,因此,这些细胞不含外源转录因子。

[0311] 6种因子所诱导的iPS细胞是多能性的

[0312] 为了进一步表征最初通过表达6种因子所产生的但是不含外源因子的iPS细胞,我们进行RT-PCR来检测这些细胞中多能基因的表达式。iPS细胞以与亲代Oct4-IRES-Puro-Egfp ES细胞表达水平相似的水平强烈表达小鼠ES细胞多能性标志物(图15A)。我们还检测了关键多能性基因Oct4、Nanog以及Rex1的启动子中的DNA甲基化模式。如图3C所示,启动子在这些细胞中完全脱甲基化。

[0313] 将6因子诱导的iPS细胞皮下注射到129/C57B6杂种F1小鼠中,来检测它们的体内分化潜力。所有小鼠在2周的时间内发展了畸胎瘤,并且在一个月后处死小鼠,用于肿瘤取样。在显微镜下,在畸胎瘤中发现了代表所有3个胚层的细胞类型(图15B)。

[0314] 此外,利用胚胎注射,很容易地获得了高质量嵌合体。当将嵌合小鼠与白化病C57B6野生型小鼠回交时,鉴定了携带刺豚鼠毛色的种系传递幼鼠,这表明注射的iPS细胞的有效种系贡献(图3E)。

[0315] 利用6种因子产生高质量人类iPS细胞

[0316] 通过有效诱导内源Oct4表达快速重编程小鼠成纤维细胞,促使我们探讨RARG和LRH1是否也能在人类中改善重编程效率和质量。通过比较人类、小鼠和大鼠基因组序列,我们确定,RAR-SF1结合位点即RAREoct在所有3个物种中是高度保守的(图5A),这表明RARG-LRH1可能在人类细胞中调节Oct4激活。我们构建了携带6种人类因子cDNA的PB-CAG转座子。与使用小鼠cDNA相比,这些PB转座子能由MEF快速产生高质量小鼠iPS细胞(数据未显示)。

[0317] 接下来我们将含有6种人类cDNA的PB-CAG转座子与PBase质粒共转染到人类新生儿包皮真皮成纤维细胞(HDFn)中。然后将这些人类成纤维细胞培养于补充了bFGF的常规人类ES培养基中或补充了人类白血病抑制因子(hLIF)的常规小鼠ES细胞培养基(M15)中。我们在补充了bFGF的人类ES培养基中鉴定了人类ES细胞样集落。令人惊讶的是,小鼠ES样集落出现于添加了hLIF的M15培养基(图16A)中。因此,我们利用标准小鼠ES方案处理了这些集落。通过胰蛋白酶处理,将人类iPS集落解离成单细胞悬浮液,并再次接种到STO饲养细胞上。能形成稳定系的次级集落易于在M15+hLIF中从一半在第10天挑选的原代集落建立。在仅使用了4种Yamanaka因子的对照组中,原代集落在形态上不同于用6种因子所产生的集落(图16A),我们从未能由在第10天挑选的集落建立稳定的系。

[0318] ERK激酶和GSK3信号传导级联(2i)的抑制剂支持天然多能干细胞的体外生长(40),但是不引发如同EpiSC一样的多能干细胞(57)。已经成功地用2i培养基来从困难小鼠株系(40)和从大鼠胚胎(14,58)衍生ES细胞。因此我们测试了我们的人类iPS细胞是否能存活于2i+hLIF培养基并且在2i+hLIF培养基中增殖。如图16A所示,人类iPS细胞在2i+hLIF培养基中增殖良好,并且能由STO饲养细胞上的单细胞形成ES样集落。

[0319] 为了测定人类iPS细胞的亚克隆效率,利用胰蛋白酶解离来自一个克隆的细胞并计数。将100或1000个细胞接种到M15+hLIF、20%KSR(敲除血清替代品)+hLIF或2i+hLIF培养条件。在这些条件下高达60%的细胞能再次形成集落(克隆效率)。尽管在2i+hLIF中所形成的集落通常更小,但是并未在这3种不同培养条件下的集落数量方面发现明显差异(图16A)。我们挑选了次级集落,并在M15+hLIF或2i+hLIF中扩增它们。扩增的克隆生长正常并且在两种培养条件下保持未分化。当解离并稀释成单细胞密度时,它们能形成三代集落,这

表明它们大量体外操作的强健性。我们的人类iPS细胞在常规小鼠ES细胞培养基中健壮生长,这一事实表明,人类iPS细胞被与小鼠ES细胞中相同的核心分子调节机制调节。即使在大量培养后,iPS细胞的基因组仍然保持完整,这可通过FISH分析确定(图10E)。

[0320] 我们还通过检查关键ES细胞特异性基因的表达,进一步评价了人类iPS细胞。如图16B和16C所示,这些基因以高水平表达,并且DNA甲基化分析证实,关键的人类多能性基因的启动子未甲基化。

[0321] 人类iPS细胞的行为与小鼠ES细胞相似,并且因此对于人类基因组的基因操作极具潜力。为了探讨这种可能性,我们首先测试了将外源DNA引入这些细胞中的PB转座。我们将PB-SB-PGK-Neo转座子与Pbase质粒一起共转染到人类iPS细胞中(31)。由一次电穿孔,我们可以获得约520,000个G418抗性集落,其由5%的幸存细胞组成。这种转座效率与在小鼠ES细胞中观察到的效率相当(31)。

[0322] 我们随后将PB-SB-SA- β geo捕获转座子和Pbase质粒转染到人类iPS细胞中。SA- β geo盒允许在人类iPS细胞中进行基因捕获,因此能在基因组的插入位点中断基因表达。利用一次电穿孔,我们收获了9,000个G418抗性集落,其代表存活于电穿孔的0.09%的细胞。我们的人类iPS细胞系衍生于包皮成纤维细胞,因此都是雄性的。因为HPRT活性的丧失能通过6TG抗性评分,我们因此研究了HPRT基因座处PB-SB-SA- β geo转座的突变效率。在9,000个G418抗性集落中,回收了1个6-TG集落。

[0323] 讨论

[0324] 有效激活内源Oct4是体细胞重编程(3)和成功的核转移(38)的最重要的已知条件之一。在胚胎发育过程中,Oct4基因座逐步被几层抑制沉默,其包括转录阻遏物的结合、组蛋白甲基化和脱甲基化以及DNA重新甲基化(47)。因此,将体细胞重编程为多能状态需要除去这些Oct4表达抑制层中的每一个。已知2种Yamanaka因子即Klf4和cMyc除了其他功能外还参与染色质改造(59)。以通过小分子抑制剂来进行组蛋白去乙酰化、组蛋白甲基化或DNA甲基化为目标的尝试,也在一定程度上改善了重编程效率(51,35,13)。然而,并未研究如何促使转录阻抑物复合体如COUP-TF和GCNF与Oct4启动子分离并同时与转录激活剂结合于Oct4启动子。因此,利用目前的平台的iPS细胞中Oct4表达并不能总是得到适当的恢复(60)。

[0325] 生物化学证据已经表明,视黄酸信号传导通过Oct4启动子中的RA应答元件(RARE)即RAREoct正调节Oct4表达(26)。RAR:RXR异源二聚体特异性结合RAREoct,并且在存在低浓度RA的情况下,部分负责激活Oct4转录(28)。在分化过程中,高浓度RA具有高得多的RAREoct亲和性,其所诱导的孤儿受体ARP-1/COUP-TFII和EAR-3/COUP-TFI,从RAREoct取代RAR:RXR,并通过募集共阻遏物主动沉默Oct4启动子(28)。RAR:RXR异源二聚体的过表达能克服COUP-TF对Oct4的抑制(28)。

[0326] RAREoct含有与RAR识别位点重叠的SF1结合位点,这为RAR激活Oct4提供了另一机制。SF1和RAR可以形成特异性结合RAREoct的复合体并协同激活Oct4(29)。有意思的是,RAR-SF1复合体的形成和其对Oct4的激活都不需要RA(29)。此外,已经证明,SF1与COUP-TFI在几个其他基因中竞争相同的顺式作用元件(61,62)。因此,RAR和SF1/LRH1复合体可能通过在结合RAREoct中拮抗COUP-TF而发挥Oct4表达的正调节剂的作用。

[0327] 综合这些出版的生物化学数据和我们的观察结果,我们提出了Rarg和Lrh1在体细

胞重编程中的工作模式。当单独表达Rarg时,因为Lrh1和Sf1通常不在MEF中表达,因此该机制主要募集RXR来结合RAREoct并激活内源Oct4转录。然而,RAR:RXR异源二聚体与RAREoct相对低的结合亲和性使得其难以与诸如COUP-TF和GCNF的阻遏物进行竞争。因此,Oct4要么并未完全被激活,要么激活并非是稳定的。这解释了Rarg过表达增强重编程效率,但是并未改善iPS细胞质量。共表达Rarg和Lrh1允许Rarg-Lrh1复合体解除阻遏物结合RAREoct并有效激活Oct4转录。随后已知与RAR相互作用的共激活剂,如CBP/p300、P/CAF以及SRC1/TIF2(36)被募集到Oct4基因座,来促进染色质进一步重塑。因此,Rarg-Lrh1复合体稳定地占据Oct4启动子中的关键元件不仅上调内源Oct4的转录水平,还促进Oct4启动子染色质结构的重塑。因此,体细胞重编程中最关键的障碍即内源Oct4激活(47),在重编程的早期阶段得到克服,这显著地加速了重编程过程并改善iPS细胞质量。

[0328] 小鼠ES细胞系由预移植胚泡的内细胞群(ICM)建立,并能对成体小鼠组织的所有谱系做贡献。另一方面,小鼠EpiSC分离自移植后胚胎,即使 当将它们注射到免疫缺陷小鼠中时能产生畸胎瘤,它们也不能定殖胚泡的ICM。EpiSC可能代表不同的多能细胞类型,其存在于小鼠胚胎发生的不同阶段、位于不同的微环境中,并且因此响应来自ES细胞的不同外来信号(63)。因为小鼠EpiSC是利用人ES细胞培养基衍生,现有的人类ES和iPS细胞系可能代表干细胞层级更低级的另一引发的多能状态(63)。引人注目的是,本文所报道的人类iPS细胞系在小鼠ES培养条件下生长强健并且行为与小鼠ES细胞相似,这表示小鼠和人类中天然多能干细胞的共同调节。

[0329] 综述,我们在此报道了在激活Oct4基因座中Rarg和Lrh1的协同相互作用导致小鼠和人类体细胞的快速和有效重编程。所产生的小鼠和人类iPS细胞质量卓越,同时人类iPS细胞与常规小鼠ES细胞高度相似,但是不同于现有人类ES细胞和iPS细胞。这些人类iPS细胞生长强健,扩增这些细胞和对其进行基因操作,会使得它们对于探讨基因座功能和产生患者特异性、无突变且无外源因子的人类iPS细胞而言有吸引力。

[0330] 材料和方法

[0331] 质粒载体构建

[0332] 为了制备PB-TRE、PB-MSCV和PB-CAG载体,由pTight(Clontech)扩增Tet应答元件(TRE),由pMSCV-Neo(Clontech)扩增MSCV LTR,并由pBluescript-CAG载体扩增CAGG启动子(未发表的数据)并将其克隆到PB-bpA载体中(未发表的数据)。

[0333] 由最初的逆转录病毒载体(Addgene)扩增4种小鼠和人类Yamanaka因子,并将其分别克隆到PB-TRE、PB-MSCV以及PB-CAG转座子载体中。由IMAGE克隆(Geneservice)扩增小鼠和人类Rarg、Lrh1以及Sf1,并将其克隆到转座子载体中。

[0334] MEF制备

[0335] 将12.5d.p.c.Oct4-IRES-Puro-Egfp胚胎断头、去内脏、用0.25%胰蛋白酶解离并接种于M10(敲除DMEM,10%胎牛血清(FBS)、青霉素链霉素和以及glutamax和非必须氨基酸(NEAA))中。HEF1购自Invitrogen并维持在补充了低血清补充物的Media 106(Invitrogen)中。

[0336] 转染和细胞培养

[0337] 电穿孔后,将MEF接种于STO饲养板中的添加了小鼠LIF的M15(敲除DMEM、15%FBS、青霉素-链霉素、glutamax、 β -巯基乙醇(β ME)、NEAA)中。转染后24h添加含有强力霉素(1 μ g/

mL)的M15,并每日更换。在第10日以96孔模式挑选iPS集落,并根据标准小鼠ES细胞培养条件进行维持。

[0338] 以相似的方式进行HDFn转染,除了将成纤维细胞接种于添加了人类的M15中。转染后24h添加含有强力霉素(2 μ g/mL)的M15,并在第10天以24孔模式挑选集落。扩增产生次级集落的克隆,并根据标准小鼠ES细胞培养条件进行维持。

[0339] Splinkerette PCR

[0340] 如所述,进行确定PB基因组整合位点的Splinkerette PCR。用M13正向和反向引物对TA克隆的PCR产物进行双向测序。利用BLAST,确定PB插入基因座。利用所述引物,对除去因子的PBiPS系的基因组进行PCR。

[0341] RT-PCR

[0342] RNA利用Reasy微型试剂盒(Qiagen)来分离、定量,并用gDNA Wipe Out处理,用QuantiTect反向转录试剂盒(Qiagen)制备cDNA。对于每一RT-PCR反应,我们利用50-100ng所列的cDNA和引物。标准的PCR条件为:94 $^{\circ}$ C 30s、58-62 $^{\circ}$ C 30s、72 $^{\circ}$ C 15-30s;X 30-35循环。

[0343] 免疫染色和流式细胞术

[0344] 使细胞生长于12饲养板。用PBS洗涤细胞、于25 $^{\circ}$ C在4%PFA/PBS中固定10min、用0.3%Triton X-100在PBS中于25 $^{\circ}$ C渗透10min并在5%驴血清中封闭1h,并于4 $^{\circ}$ C添加一抗过夜。在PBS中洗涤样品,并于25 $^{\circ}$ C添加二抗1h(cy3IgG,1:1,000;Alexa488IgG or IgM,1:400;Alexa594IgG,1:200)。

[0345] 体外分化测定

[0346] 解离人类PB-iPS系,并用于通过在AggreWell 400平板(StemCell Technologies)中的无hLIF的M15中聚集产生拟胚体。生长14天后,将拟胚体接种于Matrigel包被的盖玻片或4室载玻片上。在再培养10天后,进行免疫组织化学。

[0347] 畸胎瘤形成

[0348] 将细胞系悬浮于含有10%FBS的DMEM中,并将100 μ l(1×10^6 个细胞)皮下注射到用异氟烷麻醉的裸鼠(CByJ.Cg-Foxn1nu/J)两侧背侧面。注射6周后,将畸胎瘤解剖、在10%的磷酸盐缓冲的福尔马林中固定过夜并包埋在石蜡中。切片用苏木精和曙红染色。

[0349] 图例

[0350] 图1.重编程中的RA信号传导。A.通过piggyBac(PB)转座进行重编程的策略的示意图。用PB-MSCV-cDNA转座子和转座酶质粒转染Oct4报道MEFt,并将其直接接种于ST0饲养细胞上。通常在3周内出现ES样集落。M10和M15分别是MEF和ES细胞的培养基。B-C.表达Rara或Rarg以及4种Yamanaka因子显著促进重编程,同时通过表达Rara-DN抑制RA信号传导阻断重编程。iPSC通过碱性磷酸酶染色观察集落。重编程对照为PB-PGK-Neo加Yamanaka因子。**: $P < 0.005$;***: $P < 0.0005$ 。FL:全长cDNA;DN:显性失活形式。D.在重编程中短暂需要RA信号传导。在重编程的各种阶段添加Rarg的激动剂(CD437)。第1-4天:转染后从第1天到第4天添加激动剂。DMSO为CD437的溶剂,将其添加于对照平板中。E-F.多能性基因的表达(E)和重编程细胞中Nanog和Rex1基因座处的DNA甲基化(F)。仅对2.0 μ g/ml Puro选择具有抗性的细胞是完全重编程的iPSC,而仅对1.0 μ g/ml Puro具有抗性的细胞是部分重编程的。ES:Oct4-IRES-Puro-Egfp敲入ES细胞;MEF:Oct4-IRES-Puro-Egfp报道MEF细胞。G.通过Rarg激动剂CD437

改善iPSC质量。利用流式细胞术,通过Oct4基因座表达GFP分析iPSC(第1-4天的CD437)的质量。

[0351] 在图1(D)的柱状图中,第一长方形表示第1-4天,第二个长方形表示第1-8天,第三个长方形表示第1-12天,第4个长方形表示第4-8天且第5个长方形表示第4-12天。

[0352] 图2. Rarg(R)和Lrh-1(L)协同促进重编程。A. Tet-On重编程策略的示意图。用PB-TRE-cDNA转座子、PB-CAG-rtTA以及PB转座酶质粒转染MEF,立即将其接种于STO饲养细胞上。转染后24小时,添加Dox(1.0 μ g/ml)。转染后4、6、8以及10天,从培养基中除去Dox,同时立即应用嘌呤霉素选择(2.0 μ g/ml)来选择Oct4基因座完全激活的细胞。嘌呤霉素在24小时内杀死MEF和ES细胞。B. 6种因子(TRE-OCKS和TRE-RL)表达4天足以完全激活内源Oct4表达,用于获得不依赖Dox的iPSC。相比之下,4种Yamanaka因子(TRE-OCKS)直到表达至少12天,才出现Puro^riPSC集落。C. 转染后6天的集落的图像。比例尺:10.0 μ m。D. 共表达Rarg和Lrh-1(RL)而不是单独表达Rarg或Lrh-1改善了iPSC质量。在转染后第10天,挑选单个iPSC集落,并使其生长于96孔板中,用于在流式细胞术中表达GFP。***:p<0.0005。E. 在荧光素酶报道基因测定中Rarg和Lrh-1的协同相互作用激活Oct4启动子。将连接于Oct4启动子的荧光素酶报道基因转染到MEF中。48小时后测量荧光素酶活性。针对在转染中使用PB-Neo质粒的阴性对照,使荧光素酶表达标准化。误差条(bar)为平均值 $\pm$ SD。

[0353] 图3. 不依赖Dox的小鼠iPSC的表征。A-B. 多能性基因在小鼠iPSC中的强健表达。A. 检测Oct4和Nanog的iPS20-A1细胞的免疫染色。B. 小鼠iPSC、亲代MEF和野生型ES细胞中的Oct4、Nanog和Rex1的qRT-PCR。表达相对于Gapdh,并且针对亲代Oct4报道小鼠ES细胞中的基因表达而标准化。Dox诱导6种因子(OCKS+RL)4天(A1)、6天(B1)或8天(C1)产生所示的iPSC系。MEF:Oct4-IRES-Puro-Egfp报道MEF细胞;ES细胞:Oct4-IRES-Puro-Egfp敲入ES细胞。C. iPSC(iPS20-A1)中Oct4和Nanog的启动子激活完全脱甲基化。D. 衍生自iPSC的畸胎瘤含有所有3个胚层的细胞类型。图版表示软骨细胞(上部左侧)、角质细胞(上部右侧)、肠样上皮细胞(底部左侧)以及神经细胞(底部右侧)。所有的照片均在相同的放大率下拍摄(x200)。E. iPSC对嵌合体中种系的贡献。使嵌合体与野生型野生型C57BL6雌性(白化病者)杂交。种系传递幼鼠为刺鼠。

[0354] 在图3(B)中的每一柱状图中,第一个长方形表示mESC,第二个长方形表示MiPS20-A1,第三个长方形表示MiPS20-B2,以及第四个长方形表示MiPS20-C1。

[0355] Fig 4. 在无血清、无饲养细胞的条件下,将MEF重编程为基态iPSC。A. 用PB-TRE-OCKS(4F)或PB-TRE-OCKS加PB-TRE-RL(6F)重编程的AP⁺集落的数量。将转染后的MEF接种于明胶化的6孔板(2 $\times$ 10⁵个/孔)的M15/LIF/Dox中,达1天,转到N2B27/LIF/Dox中高达17天。B. 利用qRT-PCR进行的基因表达分析。表达相对于Gapdh,并且针对野生型ES细胞而标准化。误差条表示平均值 $\pm$ SD。C. 用于Nanog和SSEA-1表达的iPSC(6F)的免疫染色。D. iPSC(6F)向代表3个胚层的细胞类型的体外分化,这可以通过免疫染色来检测。抗体:TuJ、神经元特异性III型 β -微管蛋白;SMA,平滑肌 α -肌动蛋白;AFP, α -胎蛋白。

[0356] 在图4(B)中的每一柱状图中,第一个长方形表示ES,第二个长方形表示4FiPS,第三个长方形表示6FiPS,且第四个长方形表示MEF。

[0357] 图5. 不依赖Dox的独特的人类iPSC的产生和表征。A. RAREoct序列在几种哺乳动物中的保守性。B. 利用Tet-On6因子平台重编程HDFn细胞。将PB-TRE-cDNA转座子、PB-CAG-

rtTA和转座酶质粒转染到HDF中,将HDF接种于STO饲养细胞上。转染后24小时,添加Dox(2.0 $\mu\text{g}/\text{ml}$)。Dox诱导通常持续15-20天。在转染后高达30天,挑选集落用于扩增。或者,一旦除去Dox,立即将集落用胰蛋白酶进行处理,并再次接种至新的饲养板,用于次级集落的形成。C. 典型的人类iPSC集落形态和AP染色。比例尺:10.0 μm 。D. 亲代HDFn、人类iPSC以及H1hESC细胞中的多能性基因的qRT-PCR分析。HiPS15-5和-10是由一次转染产生的不依赖Dox的iPSC系。E. 针对ES细胞表面标志物和多能性因子的人类iPSC的免疫染色。F. 人类iPSC的体外分化。抗体:Tuj、神经元特异性III型 β -微管蛋白;SMA,平滑肌 α -肌动蛋白;AFP, α -胎蛋白。G. 由人类iPSC分化的畸胎瘤。左侧图版:软骨(中胚层);中部图版:神经组织(外胚层);有纤毛的腺上皮(内胚层)。所有照片均在相同的放大率下拍摄(x200)。H. 通过基因表达(qRT-PCR)所测量的人类iPSC的信号传导依赖性。在存在FGF受体抑制剂(FGFRi)但是没有添加JAK抑制剂(JAKi)的情况下,多能性基因的强健表达。用生长于KSR-FGF2(bFGF)或KSR-2i-LIF培养基(2i/LIF)中的H1hESC细胞作为对照。I. 生长于不同条件下的人类iPSC的基因表达变化。如果再次培养于2i/LIF培养基中,则适应FGF的iPSC快速分化。这些细胞丧失OCT4和NANOG的表达,伴随SOX17和SOX1的表达。相对于GADPH进行基因表达的定量分析。用生长于FGF2中的H1hESC细胞的表达作为对照。误差条为平均值 $\pm$ SD。

[0358] 在图5(D)中的前3个柱状图中,第一个长方形表示H1hESC,第二个长方形表示HiPS15-5,且第三个长方形表示HiPS15-10。在图5(D)中第4个柱状图中,第一个长方形表示HDFn,第二个长方形表示H1hESC,第三个长方形表示HiPS15-5,且第四个长方形表示HiPS15-10。

[0359] 在图5(H)中的每一柱状图中,如果存在,则第一个长方形表示H1hESC细胞中的FGF2,第二个长方形表示HiPS15-10中的2i/LIF,第三个长方形表示HiPS15-10细胞中的2i/LIF至FGF2,第四个长方形表示HiPS15-10细胞中的FGF2至2i/LIF。

[0360] 在图5(I)中的每一柱状图中,如果存在,则第一个长方形表示H1hESC细胞中的FGF2,第二个长方形表示H1hESC细胞中的2i/LIF,第三个长方形表示HiPS15-10中的2i/LIF,第四个长方形表示HiPS15-10细胞中的2i/LIF/BMP4,第五个长方形表示HiPS15-10细胞中的2i/LIF/JAKi,以及第六个长方形表示HiPS15-10细胞中的2i/LIF/FGFRi。

[0361] 图6.A.携带转录因子cDNA的PB转座子。PB:piggyBac转座子;TR:PB的末端重复;MSCV:小鼠干细胞病毒的LTR;CAG:CAG启动子;TRE:四环素应答元件;pA:聚腺苷酸化信号传导序列。B.The Oct4-IRES-Puro-Egfp敲入等位基因。将IRES-Puro-Egfp盒靶向Oct4基因座的3'UTR来监测并选择该基因座的激活。C.Oct4-IRES-Puro-Egfp MEF细胞Oct4-IRES-Puro-Egfp敲入ES细胞的流式细胞术分析。在分析中仅靶向的ES细胞是绿色的。D.在转染中携带Rarg的转座子DNA量的增加降低OCKS的重编程效率。对照:无任何PB-MSCV-Rarg的单独的OCKS。E.存活于两种嘌呤霉素浓度的重编程小鼠细胞的流式细胞术分析。大多数1.0 $\mu\text{g}/\text{ml}$ Puro^r集落不表达GFP。另一方面,存活于2.0 $\mu\text{g}/\text{ml}$ 嘌呤霉素选择的重编程细胞都是GFP⁺。利用生长于96孔板的细胞直接进行流式细胞术分析。

[0362] 图7.Rarg和Lrh-1协调促进重编程。A.PB转座子携带多个被编码2A的DNA链接的cDNA,2A即口蹄疫病毒2A自切割肽。CAG:CAG启动子;pA:聚腺苷酸化信号传导序列;rtTA:反向四环素应答转录激活剂(rtTA);TRE:四环素应答元件。B.利用单独的PB-CAG-OCKS(上部的2幅)或与PB-CAG-RL一起(底部的2幅)重编程的典型iPSC集落。转染后10天采集图像。C.

转染后10天的iPSC集落的碱性磷酸酶染色。6种因子(CAG-OCKS和CAG-RL)比仅利用4种Yamanaka因子(CAG-OCKS)多产生高达20多倍的集落(右幅)。D. 用于荧光素酶报道基因的DNA构建体的图。460-bp的Oct4启动子DNA片段位于RAREoct位点的侧翼,将其克隆于luc2编码序列之前。

[0363] 图8.小鼠iPSC多能性的表征。A.检测SSEA1和Nanog的iPS20-A1细胞的免疫染色。B.内源多能性基因在小鼠iPSC中的强健表达。用 α -肌动蛋白作为PCR对照,利用RT-PCR检测表达。iPS20-A1、B1以及C1是DOX诱导6种因子(OCKS+RL)4天(A1)、6天(B1)或8天(C1)所产生的3种iPS细胞系。MEF:Oct4-IRES-Puro-Egfp敲入MEF;ES:Oct4-IRES-Puro-Egfp敲入ES细胞。C.外源重编程因子在小鼠iPSC系中的表达的RT-PCR分析。对照:在诱导外源因子表达的DOX存在的情况下生长的小鼠iPSC;ES,亲代Oct4-IRES-Puro-Egfp敲入ES细胞。Oct4-cMyc-外源的、cMyc-Klf4-外源的、Klf4-Sox2-外源的,是检测PB-TRE-OCKS构建体中cDNA中连接片段的3个引物对。Rarg-Lrh-1-外源的是检测PB-TRE-RL构建体的侧翼序列的引物对。在大多数iPSC系中,外源因子的表达是不可检测的。

[0364] 图9.MEF在无饲养细胞无血清的条件下的重编程。A.MEFh用PB-TRE-OCKS(4-因子)或PB-TRE-OCKS加PB-CAG-RL(6-因子)转染,并接种于N2B27加具有DOX的LIF的培养基中,达所示天数。一旦挑选集落就将它们培养于2i/LIF培养基中。通过表达4因子同时用DOX处理8天,没有获得集落。在另一实验中,用2i/LIF培养基直接培养新鲜转染的MEF。添加Dox高达17天,以便通过表达OCKS出现任何集落。然而,仅当使用6种外源因子时获得集落。B.苏木色素和曙红染色的畸胎瘤石蜡切片含有所有3个胚层的衍生物。上皮分化(伴随角蛋白per1的形成)和神经分化(伴随花饰状物和神经管的形成)显示于上排图中。内胚层衍生物(箭头所示)以沿杯状细胞、肠上皮样细胞以及潘氏细胞排列的颗粒状结构表示。内胚层衍生物由软骨、免疫脂肪组织和肌肉组成。所有照片均在相同的放大率下拍摄(200x)。

[0365] 图10.用6因子平台(CAG启动子版本)产生独特的人类iPSC细胞。A.在M15加LIF培养基或在2i/LIF培养基中形成的人类iPSC集落,其与常规小鼠ES细胞集落相似。上部左幅:用4种Yamanaka因子重编程的典型集落;上部右幅:利用6种因子重编程的iPSC集落。在转染后10天采集图像。底部左幅:将原代iPSC集落之后的次级集落解离并再接种于饲养细胞上;底部右幅:以单细胞密度亚克隆后的典型人类iPSC集落。B.免疫染色所检测的内源多能性蛋白质在人类iPSC中的表达。C.RT-PCR所检测的多能性基因在人类iPSC中的表达。HiPS1-3和4、HiPS6-4和16是4种独立的iPSC系。HDF:人类真皮成纤维细胞;hESC:人类ES细胞。D.在畸胎瘤中人类iPSC向3个胚层的细胞类型的分化。上部左幅:具有偶然的花饰品状物(箭头)的神经组织;上部右幅:纤维肌性纤维(箭头);底部左幅:宽松的间质(箭头);底部右幅:有纤毛的颗粒状上皮(箭头)。所有照片均以相同的放大率拍摄(x200)。E.大量传代(>20代)后人类iPSC的正常核型。上部图:DAPI;底部图:MFISH。F.人类iPSC的Y染色体基因分型证实了人类iPSC的HDFn来源。用16种人类Y染色体标志物进行基因分型。

[0366] 图11.利用6因子Tet-On系统(不依赖Dox)所产生的独特的人类iPSC的表征。A.用人类GADPH作为对照,人类iPSC中基因表达的RT-PCR分析。HDF:人类真皮成纤维细胞;hESC:H1人类ES细胞。B.人类iPSC的SSEA-4、Tra-1-60以及Tra-1-81表达的FACS分析。左幅:HDF对照,其不表达SSEA-4、TRA-1-60或TRA-1-81。C.人类iPSC中外源因子表达的RT-PCR分析。对照:生长于诱导外源因子表达的DOX中的iPSC;hESC:H1人类ES细胞。PCR扩增后,2种系不具

有明显的外源OCT4、SOX2和MYC、KLF4表达。D.在大量的体外培养(>20代)后,人类iPSC具有正常的核型。上部图:DAPI;底部图:MFISH。E.将基因捕获PB转座子插入HPRT基因座,破坏了人类iPSC中的HPRT活性。在HPRT基因座的内含子2中发现了基因捕获PB转座子整合。底部图:PB转座子和HPRT基因座DNA连接的序列。

[0367] 图12.Rarg显性失活等位基因阻断ES细胞分化

[0368] A.携带强CAG启动子/增强子和一对剪接受体的PB转座子的示意图。

[0369] B.Rex1-Puro-IRES-Egfp敲入小鼠ES细胞系的图。

[0370] C.小鼠ES细胞基因筛选的策略,来鉴定能够阻断视黄酸所诱导的ES细胞分化的突变体。

[0371] D.基因筛选中所鉴定的Rarg基因座处的4个独立的突变。

[0372] E.RA诱导Rarg或Rara显性形式的过表达所阻断的ES细胞分化。

[0373] F.RA诱导Rarg或Rara特异性拮抗剂所阻断的ES细胞分化。

[0374] 在图12(E)中每一柱状图中,如果存在,则第一个长方块表示PB-MSCV-Rara-DN,第二个长方块表示PB-MSCV-Rarg-DN,第三个长方块表示PB-PGK-Neo,第四个长方块表示PB-MSCV-Rara-FL,以及第五个长方块表示PB-MSCV-Rarg-FL。

[0375] 在图12(F)中的每一柱状图中,如果存在,则第一个长方块表示单独的RA,第二个长方块表示RA+Ro-41-5253以及第三个长方块表示RA+CD2665。

[0376] 图13.RA信号传导在将小鼠MEF重编程为iPS细胞中发挥关键作用。

[0377] A.iPS细胞的较高质量与存活于更高浓度的嘌呤霉素相关。

[0378] B.过表达Rarg-FL显著增加重编程效率。

[0379] C.iPS细胞培养板显示Rarg增强重编程效率。细胞用结晶紫染色。

[0380] D.Rarg-DN的过表达也改善iPS克隆的质量但是降低重编程效率。

[0381] E.Rarg的剂量影响重编程效率。

[0382] F.Rarg或Rara特异性激动剂增强重编程效率。

[0383] G.Rarg特异性激动剂改善iPS细胞的质量。

[0384] 在图13(F)的柱状图中,第一个长方块表示第1-4天,第二个长方块表示第1-8天,第三个长方块表示第1-12天,第四个长方块表示第4-8天,第五个长方块表示第4-12天,以及第六个长方块表示第8-12天。

[0385] 图14.Rarg和Lrh1协同作用促进重编程

[0386] A.Rarg特异性拮抗剂处理能改善部分重编程的iPS克隆的质量。

[0387] B.Lrh1的过表达促进重编程。

[0388] C.Puro抗性iPS细胞集落非常早的出现显示通过6种因子的MEF快速重编程。

[0389] D.快速实现完全的重编程多能性需要Rarg和Lrh1的表达。

[0390] E.Rarg和Lrh1表达的剂量对于iPS细胞很关键。

[0391] 在图14(A)的柱状图中,在图13(F)的柱状图中,第一个长方块表示第1天,第二个长方块表示第2天,第三个长方块表示第3天,以及第四个长方块表示第4天。

[0392] 图15.6种因子重编程的iPS细胞的高质量

[0393] A.利用Rarg和Lrh1所产生的小鼠iPS细胞具有ES细胞多能性标志物的强健表达。

[0394] B.6因子所诱导的小鼠iPS细胞能对畸胎瘤中的所有谱系有贡献。

[0395] 图16.利用6种因子产生高质量人类iPS细胞

[0396] A.在M15加hLIF培养基中所产生的人类iPS细胞集落(上部图)。在2i加hLIF培养条件下扩增的人类iPS细胞集落(底部图)。

[0397] B.人类iPS细胞表达高水平的OCT4和NANOG。

[0398] C.RT-PCR分析证实了人类iPS细胞种多能性基因的高水平表达。4-1、4-2、4-3以及4-5是人类iPS细胞。ES:人类ES细胞对照。

[0399] 参考文献

[0400] 1.K.Takahashi,S.Yamanaka,Cell126,663(Aug 25,2006).

[0401] 2.K.Okita,T.Ichisaka,S.Yamanaka,Nature448,313(Jul 19,2007).

[0402] 3.M.Wernig et al.,Nature 448,318(Jul 19,2007).

[0403] 4.J.Hanna et al.,Cell133,250(Apr 18,2008).

[0404] 5.T.Aoi et al.,Science 321,699(Aug 1,2008).

[0405] 6.M.Wernig et al.,Nat Biotechnol26,916(Aug,2008).

[0406] 7.K.Takahashi et al.,Cell131,861(Nov 30,2007).

[0407] 8.I.H.Park et al.,Nature 451,141(Jan 10,2008).

[0408] 9.J.Yu et al.,Science 318,1917(Dec 21,2007).

[0409] 10.J.C.Heng et al.,Cell stem cell6,167(Feb 5).

[0410] 11.J.Silva et al.,PLoS biology 6,e253(Oct 21,2008).

[0411] 12.D.Huangfu et al.,Nat Biotechnol26,795(Jul,2008).

[0412] 13.Y.Shi et al.,Cell Stem Cell2,525(Jun 5,2008).

[0413] 14.J.K.Ichida et al.,Cell stem cell5,491(Nov 6,2009).

[0414] 15.A.Marson et al.,Cell stem cell3,132(Aug 7,2008).

[0415] 16.H.Hong et al.,Nature 460,1132(Aug 27,2009).

[0416] 17.J.Hanna et al.,Nature462,595(Dec 3,2009).

[0417] 18.J.Han et al.,Nature 463,1096(Feb 25,2010).

[0418] 19.M.J.Boland et al.,Nature 461,91(Sep 3,2009).

[0419] 20.L.Kang,J.Wang,Y.Zhang,Z.Kou,S.Gao,Cell stem cell5,135(Aug 7,2009).

[0420] 21.M.Stadtfield et al.,Nature 465,175(May 13).

[0421] 22.A.Meissner,M.Wernig,R.Jaenisch,Nat Biotechnol25,1177(Oct,2007).

[0422] 23.M.Nakagawa et al.,Nature biotechnology26,101(Jan,2008).

[0423] 24.J.Yu,J.A.Thomson,Genes & development 22,1987(Aug 1,2008).

[0424] 25.K.Niederreither,P.Dolle,Nature reviews 9,541(Jul,2008).

[0425] 26.E.Pikarsky,H.Sharir,E.Ben-Shushan,Y.Bergman,Mol Cell Biol14,1026(Feb,1994).

[0426] 27.I.Sylvester,H.R.Scholer,Nucleic Acids Res 22,901(Mar 25,1994).

[0427] 28.E.Ben-Shushan,H.Sharir,E.Pikarsky,Y.Bergman,Mol Cell Biol15,1034(Feb,1995).

[0428] 29.E.Barnea,Y.Bergman,J Biol Chem275,6608(Mar 3,2000).

[0429] 30.M.Wernig,A.Meissner,J.P.Cassady,R.Jaenisch,Cell stem cell2,10(Jan

- 10,2008).
- [0430] 31.W.Wang et al.,Proc Natl Acad Sci USA105,9290(Jul 8,2008).
- [0431] 32.H.Niwa,J.Miyazaki,A.G.Smith,Nat Genet24,372(Apr,2000).
- [0432] 33.S.Tsai et al.,Genes & development 6,2258(Dec,1992).
- [0433] 34.M.Gianni,S.Zanotta,M.Terao,S.Garattini,E.Garattini,Biochemical and biophysical research communications196,252(Oct 15,1993).
- [0434] 35.P.Gu et al.,Mol Cell Biol25,3492(May,2005).
- [0435] 36.M.Stadtfield,N.Maherali,D.T.Breault,K.Hochedlinger,Cell stem cell2, 230(Mar6,2008).
- [0436] 37.T.Brambrink et al.,Cell stem cell2,151(Feb 7,2008).
- [0437] 38.M.Boiani,S.Eckardt,H.R.Scholer,K.J.McLaughlin,Genes & development 16,1209(May 15,2002).
- [0438] 39.Q.L.Ying,J.Nichols,I.Chambers,A.Smith,Cell 115,281(Oct 31,2003).
- [0439] 40.Q.L.Ying et al.,Nature 453,519(May 22,2008).
- [0440] 41.M.Buehr et al.,Cell135,1287(Dec 26,2008).
- [0441] 42.H.M.Yang et al.,J Cell Biochem 101,1198(Aug 1,2007).
- [0442] 43.F.Ishikawa et al.,Blood 106,1565(Sep 1,2005).
- [0443] 44.J.A.Thomson et al.,Science(New York,N.Y 282,1145(Nov 6,1998).
- [0444] 45.M.Amit et al.,Dev Biol 227,271(Nov 15,2000).
- [0445] 46.T.Perlmann,R.M.Evans,Cell190,391(Aug 8,1997).
- [0446] 47.K.Hochedlinger,K.Plath,Development 136,509(Feb,2009).
- [0447] 48.<http://biogps.gnf.org/>.
- [0448] 49.J.Hanna et al.,Proceedings of the National Academy of Sciences of the United States of America,(May 4,2010).
- [0449] 50.C.Buecker et al.,Cell stem cell6,535(June 4,2010,2010).
- [0450] 51.T.S.Mikkelsen et al.,Nature 454,49(Jul 3,2008).
- [0451] 52.A.Meissner,M.Wernig,R.Jaenisch,Nat Biotechnol25,1177(Oct,2007).
- [0452] 53.M.J.Evans,M.H.Kaufman,Nature 292,154(Jul 9,1981).
- [0453] 54.A.G.Smith et al.,Nature 336,688(Dec 15,1988).
- [0454] 55.I.G.Brons et al.,Nature 448,191(Jul 12,2007).
- [0455] 56.P.J.Tesar et al.,Nature 448,196(Jul 12,2007).
- [0456] 57.G.Guo et al.,Development 136,1063(Apr,2009).
- [0457] 58.P.Li et al.,Cell135,1299(Dec 26,2008).
- [0458] 59.S.Yamanaka,Cell Stem Cell1,39(Jun 7,2007).
- [0459] 60.N.Feldman et al.,Nat Cell Biol8,188(Feb,2006).
- [0460] 61.R.N.Yu,M.Ito,J.L.Jameson,Mol Endocrinol12,1010(Jul,1998).
- [0461] 62.K.Zeitoun,K.Takayama,M.D.Michael,S.E.Bulun,Mol Endocrinol13,239 (Feb,1999).
- [0462] 63.J.Nichols,A.Smith,Cell Stem Cell4,487(Jun 5,2009).

[0463] 64.Wagner RT,Xu X,Yi F,Merrill BJ,Cooney AJ.2010.Canonical Wnt/beta-Catenin Regulation of Liver Receptor Homolog-1(Lrh-1)Mediates Pluripotency Gene Expression.Stem Cells.

[0464] 表1

[0465]

序列	模型名称	位置	链
GXP_1805437 [GXP_1805437] (1 - 601)	RXRF_SF1F_02	328 - 355	(+)
Slco2a1, GXL_555, GeneID: 24059, 小家鼠(Mus musculus) chr. 9			
溶质载体有机阴离子转运蛋白家族, 成员 2a1			
GXP_2043001 [GXP_2043001] (1 - 601)	RXRF_SF1F_02	51 - 78	(+)
Gpr156, GXL_13708, GeneID: 239845, 小家鼠 chr. 16			
G protein-coupled receptor 156			
GXP_48251 [GXP_48251] (1 - 601)	RXRF_SF1F_02	273 - 246	(-)
Pgm1, GXL_39900, GeneID: 66681, 小家鼠 chr. 5			
葡萄糖磷酸变位酶 1			

[0466]

GXP_54608 [GXP_54608] (1 - 896)	RXRF_SF1F_02	863 - 890	(+)
Dna2, GXL_45025, GeneID: 327762, 小家鼠 chr. 10			
DNA 复制解旋酶 2 同系物(酵母)			
GXP_70835 [GXP_70835] (1 - 689)	RXRF_SF1F_02	94 - 67	(-)
Cox6a2, GXL_58865, GeneID: 12862, 小家鼠 chr. 7			
细胞色素 c 氧化酶, 亚基 VI a, 多肽 2			
GXP_76405 [GXP_76405] (1 - 1461)	RXRF_SF1F_02	1042 - 1015	(-)
Pfn4, GXL_63527, GeneID: 382562, 小家鼠 chr. 12			
抑制蛋白家族, 成员 4			
GXP_76829 [GXP_76829] (1 - 1168)	RXRF_SF1F_02	38 - 11	(-)
Slc43a1, GXL_63856, GeneID: 72401, 小家鼠 chr. 2			
溶质载体家族 43, 成员 1			
GXP_84648 [GXP_84648] (1 - 674)	RXRF_SF1F_02	173 - 146	(-)
Cd209f, GXL_70439, GeneID: 69142, 小家鼠 chr. 8			
CD209f antigen			
GXP_92569 [GXP_92569] (1 - 781)	RXRF_SF1F_02	124 - 97	(-)
Trpv2, GXL_77017, GeneID: 22368, 小家鼠 chr. 11			
瞬时受体电位阳离子通道, 亚家族 V, 成员 2			
GXP_1248142 [GXP_1248142] (1 - 876)	RXRF_SF1F_02	759 - 732	(-)
Hrh1, GXL_88265, GeneID: 15465, 小家鼠 chr. 6			
组胺受体 H1			
GXP_105499 [GXP_105499] (1 - 601)	RXRF_SF1F_02	520 - 547	(+)
Ube4b, GXL_88402, GeneID: 63958, 小家鼠 chr. 4			
泛素化因子 E4B, UFD2 同系物(酿酒酵母(S. cerevisiae))			
GXP_407485 [GXP_407485] (1 - 601)	RXRF_SF1F_02	94 - 121	(+)
Chst3, GXL_96707, GeneID: 53374, 小家鼠 chr. 10			
碳水化合物(软骨素 6/角质素) 磺基转移酶 3			
GXP_158497 [GXP_158497] (1 - 1378)	RXRF_SF1F_02	46 - 19	(-)
Mesp1, GXL_132109, GeneID: 17292, 小家鼠 chr. 7			
中胚层后方同源物 1			
GXP_1803377 [GXP_1803377] (1 - 603)	RXRF_SF1F_02	240 - 213	(-)
Anpep, GXL_132245, GeneID: 16790, 小家鼠 chr. 7			
丙氨酰(膜)氨肽酶			
GXP_1799422 [GXP_1799422] (1 - 603)	RXRF_SF1F_02	325 - 298	(-)
Nup210l, GXL_133945, GeneID: 77595, 小家鼠 chr. 3			

[0467]

核孔蛋白 210-样			
GXP_2045579 [GXP_2045579] (1 - 601)			
Muc1, GXL_134083, GeneID: 17829, 小鼠 chr. 3	RXRF_SF1F_02	87 - 114	(+)
粘蛋白 1, 跨膜			
GXP_1806208 [GXP_1806208] (1 - 601)			
Rsl1, GXL_145407, GeneID: 20147, 小鼠 chr. X	RXRF_SF1F_02	300 - 330	(+)
视网膜裂(X-连锁, 不成熟的) 1 (人类)			
GXP_2047904 [GXP_2047904] (1 - 601)			
Pdc8a, GXL_153622, GeneID: 18584, 小鼠 chr. 7	RXRF_SF1F_02	126 - 153	(+)
磷酸二酯酶 8A			
GXP_816489 [GXP_816489] (1 - 601)			
Arfgef2, GXL_169927, GeneID: 99371, 小鼠 chr. 2	RXRF_SF1F_02	166 - 139	(-)
ADP-核糖基化因子鸟嘌呤核苷酸交换因子 2 (蓝藻素 A-抑制的)			
GXP_1460689 [GXP_1460689] (1 - 601)			
Gtpbp5, GXL_169951, GeneID: 52856, 小鼠 chr. 2	RXRF_SF1F_02	582 - 555	(-)
GTP 结合蛋白 5			
GXP_433031 [GXP_433031] (1 - 601)			
Ccdc21, GXL_181575, GeneID: 70012, 小鼠 chr. 4	RXRF_SF1F_02	15 - 42	(+)
含有卷曲螺旋结构域的 21			
GXP_446388 [GXP_446388] (1 - 601)			
Prtg, GXL_183874, GeneID: 235472, 小鼠 chr. 9	RXRF_SF1F_02	127 - 154	(+)
Protogenin 同系物 (红原鸡(Gallus gallus))			
GXP_1801525 [GXP_1801525] (1 - 601)			
Trpv4, GXL_210485, GeneID: 63873, 小鼠 chr. 5	RXRF_SF1F_02	137 - 164	(+)
瞬时受体电位阳离子通道, 亚家族 V, 成员 4			
GXP_1798493 [GXP_1798493] (1 - 668)			
Gpr176, GXL_240571, GeneID: 381413, 小鼠 chr. 2	RXRF_SF1F_02	205 - 232	(+)
G 蛋白偶联受体 176			
GXP_2046947 [GXP_2046947] (1 - 604)			
Tbl2, GXL_240884, GeneID: 27368, 小鼠 chr. 5	RXRF_SF1F_02	210 - 183	(-)
转导蛋白(β)-样 2			
GXP_2044366 [GXP_2044366] (1 - 614)			
Sik25, GXL_242978, GeneID: 59041, 小鼠 chr. 1	RXRF_SF1F_02	298 - 271	(-)
丝氨酸/苏氨酸激酶 25 (酵母)			
GXP_1795374 [GXP_1795374] (1 - 1306)	RXRF_SF1F_02	1046 -	(+)

[0468]

H2-DMa, GXL_251594, GeneID: 14998, 小家鼠 chr. 17		1073	
组织相容性 2, II 类, 基因座 DMa			
GXP_299576 [GXP_299576] (1 - 977)			
Pou5f1, GXL_251685, GeneID: 18999, 小家鼠 chr. 17	RXRF_SF1F_02	760 - 787	(+)
POU 结构域, 5 类, 转录因子 1			
GXP_1793956 [GXP_1793956] (1 - 624)			
C7, GXL_258854, GeneID: 109828, 小家鼠 chr. 15	RXRF_SF1F_02	382 - 355	(-)
补充组分 7			
GXP_2044577 [GXP_2044577] (1 - 606)			
Pbx1, GXL_271528, GeneID: 18514, 小家鼠 chr. 1	RXRF_SF1F_02	152 - 125	(-)
前 B-细胞白血病转录因子 1			
GXP_411169 [GXP_411169] (1 - 601)			
Tanc2, GXL_273504, GeneID: 77097, 小家鼠 chr. 11	RXRF_SF1F_02	135 - 162	(+)
三角形四肽重复, 锚蛋白重复以及含有卷曲螺旋结构域的 2			
GXP_2043287 [GXP_2043287] (1 - 601)			
H2-DMb1, GXL_278291, GeneID: 14999, 小家鼠 chr. 17	RXRF_SF1F_02	342 - 369	(+)
组织相容性 2, II 类, 基因座 Mb1			
GXP_816740 [GXP_816740] (1 - 619)			
Impa1, GXL_281247, GeneID: 55980, 小家鼠 chr. 3	RXRF_SF1F_02	195 - 222	(+)
肌醇(肌)-1(或 4)-单磷酸酶 1			
GXP_2047788 [GXP_2047788] (1 - 601)			
Hsd17b14, GXL_286930, GeneID: 66065, 小家鼠 chr. 7	RXRF_SF1F_02	549 - 576	(+)
羟基类固醇(17-β)脱氢酶 14			
GXP_2042863 [GXP_2042863] (1 - 601)			
Snx29, GXL_461294, GeneID: 74478, 小家鼠 chr. 16	RXRF_SF1F_02	213 - 186	(-)
分拣微管连接蛋白 29			
GXP_412534 [GXP_412534] (1 - 601)			
100040879, GXL_659082, GeneID: 100040879, 小家鼠 chr. 12	RXRF_SF1F_02	597 - 570	(-)
推测的基因, 100040879			
GXP_431619 [GXP_431619] (1 - 601)			
Tmod1, GXL_661320, GeneID: 21916, 小家鼠 chr. 4	RXRF_SF1F_02	259 - 232	(-)
原肌球蛋白调节蛋白 1			
GXP_2046709 [GXP_2046709] (1 - 610)	RXRF_SF1F_02	42 - 69	(+)
Slc4a4, GXL_661656, GeneID: 54403, 小家鼠 chr. 5			

[0469]

溶质载体家族 4 (anion exchanger), 成员 4

在 38 条序列中总计发现 38 个匹配。检索的序列: 65829 (43681672 bp)。

[0470]

表2

[0471]

序列	模型名称	位置	链
GXP_38642 [GXP_38642] (1 - 601) PRPF31, GXL_32057, GeneID: 26121, 智人(Homo sapiens) chr. 19 PRP31 前体 mRNA 加工因子 31 同系物(酿酒酵母)	RXRF_SF1F_02	32 - 62	(+)
GXP_1255917 [GXP_1255917] (1 - 629) PITPNM1, GXL_41522, GeneID: 9600, 智人 chr. 11 磷脂酰肌醇转移蛋白, 膜相关 1	RXRF_SF1F_02	196 - 169	(-)
GXP_52656 [GXP_52656] (1 - 916) POU5F1, GXL_43374, GeneID: 5460, 智人 chr. 6 POU 5 类同源框 1	RXRF_SF1F_02	695 - 722	(+)
GXP_1814152 [GXP_1814152] (1 - 1138) TMEM102, GXL_45908, GeneID: 284114, 智人 chr. 17 跨膜蛋白 102	RXRF_SF1F_02	208 - 181	(-)
GXP_1815496 [GXP_1815496] (1 - 613) TNFRSF11A, GXL_47222, GeneID: 8792, 智人 chr. 18 肿瘤坏死因子受体亚家族, 成员 11a, NFkB 激活剂	RXRF_SF1F_02	31 - 58	(+)
GXP_1260208 [GXP_1260208] (1 - 602) CDH5, GXL_89968, GeneID: 1003, 智人 chr. 16 钙粘素 5, 2 型 (血管内皮)	RXRF_SF1F_02	184 - 155	(-)
GXP_116394 [GXP_116394] (1 - 685) NSUN5, GXL_97642, GeneID: 55695, 智人 chr. 7 NOL1/NOP2/Sun 结构域家族, 成员 5	RXRF_SF1F_02	41 - 68	(+)
GXP_1818489 [GXP_1818489] (1 - 602) BAT2D1, GXL_120915, GeneID: 23215, 智人 chr. 1 含有 1 的 BAT2 结构域	RXRF_SF1F_02	534 - 561	(+)
GXP_160425 [GXP_160425] (1 - 956) DPEP3, GXL_133499, GeneID: 64180, 智人 chr. 16 二肽酶 3	RXRF_SF1F_02	187 - 214	(+)

[0472]

GXP_160480 [GXP_160480] (1 - 886) PYDC1, GXL_133531, GeneID: 260434, 智人 chr. 16 含有 1 的 PYD (脓素结构域)	RXRF_SF1F_02	6-33	(+)
GXP_1810489 [GXP_1810489] (1 - 602) CCT2, GXL_141723, GeneID: 10576, 智人 chr. 12 含有 TCPI 的伴侣蛋白, 亚基 2 (β)	RXRF_SF1F_02	228 - 255	(+)
GXP_1273930 [GXP_1273930] (1 - 1366) ATXN7L1, GXL_182315, GeneID: 222255, 智人 chr. 7 ataxin 7-样 1	RXRF_SF1F_02	364 - 391	(+)
GXP_221757 [GXP_221757] (1 - 601) C9orf86, GXL_184953, GeneID: 55684, 智人 chr. 9 染色体 9 可读框 86	RXRF_SF1F_02	419 - 392	(-)
GXP_1259198 [GXP_1259198] (1 - 733) SNX1, GXL_230316, GeneID: 6642, 智人 chr. 15 分拣微管连接蛋白 1	RXRF_SF1F_02	61 - 88	(+)
GXP_1479250 [GXP_1479250] (1 - 635) GALNTL1, GXL_247708, GeneID: 57452, 智人 chr. 14 UDP-N-乙酰基-α-D-半乳糖胺: 多肽 N-乙酰基氨基半乳糖基转移酶-样 1	RXRF_SF1F_02	84 - 111	(+)
GXP_2051860 [GXP_2051860] (1 - 601) C14orf166, GXL_247796, GeneID: 51637, 智人 chr. 14 染色体 14 可读框 166	RXRF_SF1F_02	67 - 40	(-)
GXP_297369 [GXP_297369] (1 - 639) MB, GXL_249810, GeneID: 4151, 智人 chr. 22 肌球蛋白	RXRF_SF1F_02	60 - 33	(-)
GXP_2054048 [GXP_2054048] (1 - 601) MAG, GXL_262148, GeneID: 4099, 智人 chr. 19 髓磷脂相关糖蛋白	RXRF_SF1F_02	17 - 44	(+)
GXP_210437 [GXP_210437] (1 - 602) RNF213, GXL_316752, GeneID: 57674, 智人 chr. 17 环指蛋白 213	RXRF_SF1F_02	207 - 234	(+)
GXP_488164 [GXP_488164] (1 - 655) YDJC, GXL_316849, GeneID: 150223, 智人 chr. 22 YdjC 同系物(细菌)	RXRF_SF1F_02	228 - 201	(-)

[0473]

GXP_1499697 [GXP_1499697] (1 - 605) CCDC125, GXL_383423, GeneID: 202243, 智人 chr. 5 含有卷曲螺旋结构域的 125	RXRF_SF1F_02	302 - 275	(-)
GXP_1497474 [GXP_1497474] (1 - 663) RPL22L1, GXL_694179, GeneID: 200916, 智人 chr. 3 核糖体蛋白 L22-样 1	RXRF_SF1F_02	50 - 23	(-)
GXP_14398 [GXP_14398] (1 - 753) ALG13, GXL_695854, GeneID: 79868, 智人 chr. X 天冬酰胺-连接的糖基化 13 同系物(酿酒酵母)	RXRF_SF1F_02	164 - 137	(-)
GXP_91298 [GXP_91298] (1 - 707) CORO7, GXL_744275, GeneID: 79585, 智人 chr. 16 冠蛋白 7	RXRF_SF1F_02	626 - 599	(-)
GXP_122795 [GXP_122795] (1 - 901) CRHR1, GXL_744353, GeneID: 1394, 智人 chr. 17 促肾上腺皮质激素释放激素受体 1	RXRF_SF1F_02	307 - 280	(-)
GXP_1821114 [GXP_1821114] (1 - 602) ANKRD39, GXL_744677, GeneID: 51239, 智人 chr. 2 锚定蛋白重复结构域 39	RXRF_SF1F_02	80 - 107	(+)
GXP_1821115 [GXP_1821115] (1 - 601) ANKRD39, GXL_744677, GeneID: 51239, 智人 chr. 2 锚定蛋白重复结构域 39	RXRF_SF1F_02	249 - 276	(+)
在 27 条序列中总计发现 27 个匹配。检索的序列: 78566 (52385643 bp).			

[0474] 表3

[0475]

引物名称	引物序列	目的
HM5pAa	CGAAGAGTAACCGTTGCTAGGAGAGACCGTGGCTGAATGAGACTG GTGTGACACTAGTGG	Splinkerette 接头
HM5pBb-Sac3A	gagcCAGTAGTGTGACACCGAGTCTGTAATTTTTTTTCAAAAAA	Splinkerette 接头
HM5p1	CGAAGAGTAACCGTTGCTAGGAGAGACG	第1轮PCR引物
HM5p2	GTGGCTGAATGAGACTGGTGTGAGC	第2轮PCR引物
HM5p3	GGTGTGACACTAGTGG	测序引物
PB-L-Sp1	CAGTACACCTTACCGCATTGACAAAGCAGGC	第1轮PCR引物, 和HM5p1一起使用
PB-L-Sp2	SAGAGAGCAATATTTCAAGAAATGCATGCGT	第2轮PCR引物, 和HM5p2一起使用
PB-L-Sp3	GCGCCTTACTGAGCTTAACCTTTAA	测序引物
PB-R-Sp1	GCTCGATATACAGAGCCGATAAAGACATGC	第1轮PCR引物, 和HM5p1一起使用
PB-R-Sp2	ACGCATGATTATCTTTACGGTACGTCACAA	第2轮PCR引物, 和HM5p2一起使用
PB-R-Sp3	CGTCACAAATGATTATCTTCTAGG	测序引物

[0476] 表4

[0477]

引物名称	引物序列	目的
mOct4-F-BglII	GAAGATCTCTCCACCTTCCCCATGGCTGGACACC	mOct4 ORF 克隆
mOct4-R-EcoRI	CGGAATTCTTGATCAACAGCATCACTGAGCTTC	
mSox2-F-BglII	GAAGATCTCTTGTATAACATGATGGAGACGG	mSox2 ORF 克隆
mSox2-R-EcoRI	CGGAATTCTTTCACATGTGCGACAGGGGCAGT	
mKlf4-F-BglII	GAAGATCTCACCATGGCTGTCAGCGACGCTCTGCTC	mKlf4 ORF 克隆
mKlf4-R-EcoRI	CGGAATTCACATCCACTACGTGGGATTTAAAA	
mcMyc-F-BglII	GAAGATCTCACCATGCCCCTCAACGTGAACTTCACC	mcMyc ORF 克隆
mcMyc-R-EcoRI	CGGAATTCTTATGCACCAGAGTTTCGAAGCTGTTCG	
mRarg-FL/DN-F-EcoRI	GGGAATTCCCGCAGCTACCATGGCCACC	mRarg- FL/DN 克 隆
mRarg-DN-R-XbaI	GCTCTAGATCATCATCATCCCTTAGTGCTGATGC	mRarg-DN 克隆
mRarg-FL-R-XbaI	GCTCTAGAGCCCAACCCACAACGGGG	mRarg-FL 克隆
mRara-FL/DN-F-EcoRI	GGGAATTGCTGCTTGGCATGGCCAGCA	mRara- FL/DN 克 隆
mRara-DN-R-XbaI	GCTCTAGATCATCATCATGGGATCTCCATCTTCAA	mRara-DN 克隆
mRara-FL-R-XbaI	GCTCTAGAGTGTGAGGTGGTCATGGGGAT	mRara-FL 克隆
mSf1-F-EcoRI	GGGAATTCCCGCGGGCATGGACTATTCGTACG	mSf1 ORF 克隆
mSf1-R-XbaI	GCTCTAGAGGCACCCAGGCTCAAGTCTGCTTG	
mLrh1-F-EcoRI	GGGAATTCTTTCGCTAAGAATGTCTGCTAGTTTGG	mLrh1 ORF 克隆
mLrh1-R-XbaI	GCTCTAGAGGGGACTTAGGCTCTTTTGGCATGC	
hOCT4-F-EcoRI	GGGAATTCCACCATGGCGGGACACCTGGCTTCAG	hOCT4 ORF 克隆
hOCT4-R-XbaI	GCTCTAGAACCTCAGTTTGAATGCATGGGAGAGC	
hSOX2-F-EcoRI	GGGAATTCCACCATGTACAACATGATGGAGACGGAGCTG	hSOX2 ORF 克隆
hSOX2-R-XbaI	GCTCTAGATCACATGTGTGAGAGGGGCAGTGTGC	
hKLF4-F-EcoRI	GGGAATTCCACCATGGCTGTCAGTGACGCGCTGCTCCC	hKLF4 ORF 克隆
hKLF4-R-XbaI	GCTCTAGATTAAAAATGTCTCTTCATGTGTAAGGCGAG	

[0478]

hCMYC-F-EcoRI	GGGAATTCCACCATGCCCCCTCAACGTTAGCTTCACCAA	hCMYC ORF 克隆
hCMYC-R-XbaI	GCTCTAGATCACGCACAAGAGTTCCGTAGCTGTTCAAG	
hLRH1-F-EcoRI	GGGAATTCATGTCTTCTAATTCAGATACTGGGG	hLRH1 ORF 克隆
hLRH1-R-XbaI	GCTCTAGATTATGCTCTTTTGGCATGCAAC	
hRARG-F-EcoRI	GGGAATTCATGGCCACCAATAAGGAGCG	hRARG ORF 克隆
hRARG-R-XbaI	GCTCTAGATCAGGCTGGGGACTTCAGGCC	

[0479] 表5

[0480]

引物名称	引物序列	目的
mOct4-RT-En-F	TCTTTCCACCAGGCCCGGCTC	用于内源 mOct3/4 的 RT-PCR
mOct4-RT-En-R	GCGGGCGGACATGGGGAGATCC	
mSox2-RT-En-F	TTGCCTTAACAAGACCACGAAA	用于内源 mSox2 的 RT-PCR
mSox2-RT-En-R	TAGAGCTAGACTCCGGGCGATGA	
mKlf4-RT-En-F	GCGAACTCACACAGGCGAGAAACC	用于内源 mKlf4 的 RT-PCR
mKlf4-RT-En-R	TCGCTTCCTCTTCTCCGACACA	
mNanog-RT-En-F	CAGGTGTTTGAGGGTAGCTC	用于 mNanog 的 RT-PCR
mNanog-RT-En-R	CGGTTCATCATGGTACAGTC	
mRex1-RT-F	ACGAGTGGCAGTTTCTTCTTGGGA	用于 mRex1 的 RT-PCR
mRex1-RT-R	TATGACTCACTTCCAGGGGGCACT	
mEsg1-RT-F	GAAGTCTGGTTCCTTGGCAGGATG	用于 mEsg1 的 RT-PCR
mEsg1-RT-R	ACTCGATACTGGCCTAGC	
mFgf4-RT-F	CGTGGTGAGCATCTTCGGAGTGG	用于 mFgf4 的 RT-PCR
mFgf4-RT-R	CCTTCTTGGTCCGCCCCGTTCTTA	
mDax1-RT-F	TGCTGCGGTCCAGGCCATCAAGAG	用于 mDax1 的 RT-PCR
mDax1-RT-R	GGGCACTGTTCACTTCAGCGGATC	
mGdf3-RT-F	GTTCCAACCTGTGCCTCGCGTCTT	用于 mGdf3 的 RT-PCR
mGdf3-RT-R	AGCGAGGCATGGAGAGAGCGGAGCAG	
mUtf1-RT-F	GGATGTCCCGGTGACTACGTCTG	用于 mUtf1 的 RT-PCR

[0481]

mUlf1-RT-R	GGCGGATCTGGTTATCGAAGGGT	
mOct4-DMR-F2	TGGGTTGAAATATTGGGTTTATTT	mOct4 启动子的 DMR 引物
mOct4-DMR-R2	CTAAAACCAAATATCCAACCATA	
mNanog-DMR-F	GATTTTGTAGGTGGGATTAATTGTGAATTT	mNanog 启动子的 DMR 引物
mNanog-DMR-R	ACCAAAAAACCCACACTCATATCAATATA	
mRex1-DMR-F2	TATATTAATGTTGGAAAAAGTTTAGGTAAT	mRex1 启动子的 DMR 引物
mRex1-DMR-R2	AACTCCTTAAACCCCTCCCTTTTTAAATAA	
hOCT4-RT-endo-F	GACAGGGGGAGGGGAGGAGCTAGG	用于内源 hOCT3/4 的 RT-PCR
hOCT4-RT-endo-R	CTTCCCTCCAACCAGTTGCCCCAAAC	
hSOX2-RT-endo-F	GGGAAATGGGAGGGGTGCAAAAGAGG	用于内源 hSOX2 的 RT-PCR
hSOX2-RT-endo-R	TTGCGTGAGTGTGGATGGGATTGGTG	
hKLF4-RT-endo-F	ACGATCGTGGCCCCGAAAAGGACC	用于内源 hKLF4 的 RT-PCR
hKLF4-RT-endo-R	GCGTCCTGGGAAGGGAGATCCGGAGC	
hCMYC-RT-endo-F	GCGTCCTGGGAAGGGAGATCCGGAGC	用于内源 hCMYC 的 RT-PCR
hCMYC-RT-endo-R	TTGAGGGGCATCGTCGCGGGAGGCTG	
hNANOG-RT-F	CAGCCCCGATTCTTCCACCAGTCCC	用于 hNANOG 的 RT-PCR
hNANOG-RT-R	CGGAAGATTCCCAGTCGGGTTCACC	
hREX1-RT-F	CAGATCCTAAACAGCTCGCAGAAT	用于 hREX1 的 RT-PCR
hREX1-RT-R	GCGTACGCAAATTAAAGTCCAGA	
hTERT-RT-F	CCTGCTCAAGCTGACTCGACACCGTG	用于 hTERT 的 RT-PCR
hTERT-RT-R	GGAAAAGCTGGCCCTGGGGTGGAGC	
hDNMT3B-RT-F	TGCTGCTCACAGGGCCCGATACTTC	用于 hDNMT3 的 RT-PCR
hDNMT3B-RT-R	TCCTTTCGAGCTCAGTGCACCACAAAAC	

[0482] 表5的参考文献:

[0483] Takahashi,K.&Yamanaka,S.Induction of pluripotent stem cells from小鼠 embryonic and adult fibroblast cultures by defined factors.Cell126,663-76 (2006).

[0484] Takahashi,K.et al.Induction of pluripotent stem cells from adult

human fibroblasts by defined factors. Cell 131,861-72(2007).

[0485] 表6

[0486]

测定 ID	基因标识符	基因名称
Mm00658129_gH	Pou5f1,mCG19893	POU 结构域, 5 类, 转录因子 1, Gene mCG19893 Celera Annotation
Mm00516104_m1	Klf4,mCG20055	Kruppel-样因子 4 (肠), Gene mCG20055 Celera Annotation
Mm02384862_g1	Nanog,Nanogpd,mCG132219,mCG56641	Nanog 同源框, 与 Nanog 同源框相似, Gene mCG132219 Celera Annotation, Gene mCG56641 Celera Annotation
Mm03053975_g1	Zfp42,mCG10211	锌指蛋白 42, Gene mCG10211 Celera Annotation
4352341E	β -actin (Actb)	小鼠 ACTB (肌动蛋白, β) 内源对照 (VIC® / MGB Probe, Primer Limited)
Hs01053049_s1	SOX2,hCG2021649	SRY (性别决定区 Y)-盒 2, Gene hCG2021649 Celera Annotation
Hs02387400_g1	NANOG,hCG1745097,hCG1730824	Nanog 同源框, Gene hCG1745097 Celera Annotation, Gene hCG1730824 Celera Annotation
Hs00999632_g1	POU5F1,hCG25999	POU 5 类同源框 1, Gene hCG25999 Celera Annotation
Hs00358836_m1	KLF4,hCG1738095	Kruppel 样因子 4 (肠), Gene hCG1738095 Celera Annotation
Hs01931905_g1	DPPA3	发育多能性相关 3
Hs00415443_m1	NODAL,hCG23654	节点同系物 (mouse), Gene hCG23654 Celera Annotation
Hs00232018_m1	GATA6,hCG37191	GATA 结合蛋白 6, Gene hCG37191 Celera Annotation
Hs00751752_s1	SOX17,hCG40422	SRY (性别决定区 Y)-盒 17, Gene hCG40422 Celera Annotation
Hs01057642_s1	SOX1,hCG1640458	SRY (性别决定区 Y)-盒 1, Gene hCG1640458 Celera Annotation
Hs01088114_m1	PAX6,hCG26499	成对的盒 6, Gene hCG26499 Celera Annotation
4326317E	GAPDH	TaqMan® 内源对照, GAPDH

[0487] 表7

[0488]

名称	序列
----	----

hCMYC-taq-endo-F	CCTGAGCAATCACCTATGAACTTG
hCMYC-taq-endo-R	TTATGCCCAAAGTCCAATTTGA
hCMYC-taq-endo-探针	CAAATGCAACCTCACAACCTTGGCTG
hKLF4-taq-endo-F	TTTCACACTGTCTTCCCGATGA
hKLF4-taq-endo-R	TCCTGATTATCCACTCACAAGATGA
hKLF4-taq-endo-探针	CCAGCCAGAAAGCACTACAATCATGGTCAA
hOCT4-taq-endo-F	CACTGTACTCCTCGGTCCCTTTC
hOCT4-taq-endo-R	CAACCAGTTGCCCCAAACTC
hOCT4-taq-endo-探针	CTTCCCCCTGTCTCCGTCACCACTC
hSOX2-taq-endo-F	ACAGCAAATGACAGCTGCAAA
hSOX2-taq-endo-R	AAGTCCAGGATCTCTCTCATAAAAGTTT
hSOX2-taq-endo-探针	ATCCACACTCACGCAAAAACCGCG
hLRH1-taq-endo-F	AAAGCTGAACTGAAACAATTCTCAAG
hLRH1-taq-endo-R	GCTCGGGCCTTCAAAGGA
hLRH1-taq-endo-探针	TGCATCAGCTGTACCTACAATAGCCCCTCC
hRARG-taq-endo-F	GAGCCTGGGTTTGGACTCTAAA
hRARG-taq-endo-R	CCTCTTGCCCTGGGAAGTCT
hRARG-taq-endo-探针	CTCAGCACTGCCCCATGGGTCC

序列表

<110> 基因组研究有限公司

<120> 细胞及其获得方法

<130> 16199.18

<150> GB0915523.5

<151> 2009-09-07

<150> US61/276,203

<151> 2009-09-08

<150> US61/240,282

<151> 2009-09-07

<160> 111

<170> PatentIn version 3.5

<210> 1

<211> 1377

<212> DNA

<213> Homo sapiens

<400> 1

[0001] atggccacca ataaggagag actcttttggc cccgggtgcc tggggcctgg atcttggttac 60

ccaggagcag gcttcccatl cgccttccca ggtgcactca gagggtcgcc accatttgag 120

atgtgagacc ctagtctccg gggcctgggc cagcctgacc tccccaggga gatggcttct 180

ctctcggttg agacacagag caccagctcg gaggagatgg taaccagctc tccctcacc 240

ccaccacctc ctcggtteta taagccatgc ttgtatgca atgacaagtc ttctggctac 300

cactatgggg tcagctcctg tgaaggctgc aagggttct tcagacgcag caticagaaa 360

aacttggtgt atacatgtca ccgtgacaaa aactgtatca tcaacaaggt caccagaaat 420

cgatgccagt actgcaggct acaaaagtgt ttggaagtgg gcattgtcaa ggaagctgta 480

aggaacgate gaaacaagaa gaaaaaggag gtaaaaggagg agggctcgcc cgacagctat 540

gaactgagtc cacagttaga ggaactcact accaaggtea gcaaaagcca ccaggagact 600

tttccctcac tctgccagct gggcaagtac accacgaact ccagtgcaga tcaccgggtg 660

cagctggacc tggggctgtg ggacaagtgc agcgagctgg ccaccaaatg catcatcaag 720

attgtggagt ttgcgaagcg gctgcctggg ttacagggc tcagcattgc cgaccagatc 780

acgtctctca aggtctctg tctggacatc ctaatctgc ggatctgtac aaggatatcc 840

ccagagcagg acactatgac attctcgcat gggtgaccc tgaaccgaac ccagatgcac 900

aatgttggtt ttgggcacct tacagacctc gtctttgctt ttgcggggca gctgctgccc 960

ctggagatgg atgacaccga gactgggcta cttagtctca tctgctcat ctgtggagac 1020

	cgaatggacc tgggaagagcc cgagaaggtg gacaagetge aggagcccct gctggaagcc	1080
	ctgaggetct atgcccggcg acggagaccc agccaaccct acatgttccc aaggatgetg	1140
	atgaaaatca ccgacctccg gggcateage actaaggag cagaaagggc tataaccctg	1200
	aagaatggaga ttccaggecc gatgccaccc ctgatccgag agatgettga gaaccggag	1260
	atgtttgagg acgactctc gaagectgge ccccaeccca aggettccag tgaggacgaa	1320
	gctccagggg gccagggcaa aaggggcaa agtccccaac ctgaccaggg gccctga	1377
	<210> 2	
	<211> 1188	
	<212> DNA	
	<213> Homo sapiens	
	<400> 2	
	atggccacca ataaggagag actctttgcg cccggtgccc tggggcctgg atctggttac	60
	ccaggagcag gcttcccatt cgccttccca ggtgcactca gagggtcgcc accatttgag	120
	atgetgagcc ctagcttccg gggcctgggc cagcctgacc tccccaggga gatggttct	180
	ctctcggtgg agacacagag caccagctcg gaggagatgg taccagctc tccctaccc	240
	ccaccacctc ctgggtcta taagccaigc ttigtatgca atgacaagtc ttctggtac	300
[0002]	cactatgggg tcagctctg tgaaggetge aaggcttct tcagacgcag cattcagaaa	360
	aacatggtgt atacatgtca ccgtgacaaa aactgtatca tcaacaaggt caccagaaat	420
	cgatgccagt acigcaggt acaaaagtgt ttccaagtgg gcatgtccaa ggaagctgta	480
	aggaacgata gaaacaagaa gaaaaaggag gtaaaaggagg agggctcgcc cgacagctat	540
	gaactgagtc cacagttaga ggaactcacc accaaggtea gcaaagccca ccaggagact	600
	tttccctcac tctgccagct gggaagtae accacgaact ccagtgcaga tcaccgggtg	660
	cagctggacc tggggctgtg ggacaagtgc agcgagctgg ccaccaaatg catcatcaag	720
	attgtggagt ttggaagcg gctgccigtgt ttacagggc tcagcatfge cgaccagatc	780
	acgtctctca aggetgett tctggacatc ctaatgetge ggatctgtac aaggtatacc	840
	ccagagcagg acactatgac attctcgat gggtgaccc tgaaccgaac ccagatgcac	900
	aatgtggct ttggcccct tacagacctc gtctttgctt ttgcgggca getgtgccc	960
	ctggagatgg atgacacga gactgggcta cttagtgeta tctgctcat ctgtggagac	1020
	cgaatggacc tgggaagagcc cgagaaggtg gacaagetge aggagcccct gctggaagcc	1080
	ctgaggetct atgcccggcg acggagaccc agccaaccct acatgttccc aaggatgetg	1140
	atgaaaatca ccgacctccg gggeateage actaagggat gatgatga	1188

<210> 3
 <211> 200
 <212> DNA
 <213> Mus musculus

<400> 3
 tcccccccc acagctctge tcttccaccc acccaggggg cggggccaga ggtcaagget 60
 agagggtggg attggggagg gagaggigaa accgtcccta ggtgagccgt ctttccacca 120
 ggcceccgge tgggggtgce caecttcccc atggetggac acctggette agacttgece 180
 ttctcacccc caccaggtgg 200

<210> 4
 <211> 191
 <212> DNA
 <213> Rattus norvegicus

<400> 4
 taccceact gctctactcc tccacccacc cagggggcgg ggcagaggt caaggetaga 60
 ggggtgggatt ggggaggag agaggigcat aggtgagtcg tcttccacc aggcceccgg 120
 ctgggggagc ccacttccc catggetgga cactggcctt cagacttege cttctacccc 180
 ccactgggtg g 191

[0003]

<210> 5
 <211> 200
 <212> DNA
 <213> Sus barbatus

<400> 5
 acccccaccg acccttcccc caeccatcca gggggcgggg ccagaggta aggctagtgg 60
 gtgggattgg ggaggagag aggtgtcgag cagtcctt ggagagccct ggttttactg 120
 ggcceccgge ttggggcgee ttccttcccc atggcgggae acctggette cgaactgece 180
 ttctegcccc cgcggggcgg 200

<210> 6
 <211> 200
 <212> DNA
 <213> Bos taurus

<400> 6
 acccccacgg tcccttcccc ccacccatcc agggggcggg gccagaggtc aaggctagtg 60
 ggtgggattg gggagggaga gaggtgttga gcagtctcta ggagatccct cgttttcta 120
 ggcceccgge tgggggtgce ttccttcccc atggcgggae acctggette tgacttgece 180
 ttctegcccc cgcggggcgg 200

[0004]	<210>	7	
	<211>	200	
	<212>	DNA	
	<213>	Gorilla gorilla	
	<400>	7	
		tcctccacac tcctctctct ccacccatcc agggggcggg gccagaggte aaggetagtg	60
		ggtgggactg gggagggaga gaggggttga gtagtccctt cgaagccct catttaccac	120
		ggcccccgge ttggggcgcc ttcttcccc atggcgggac acctggcttc ggatttcgce	180
		ttctcgcccc ctccaggtgg	200
	<210>	8	
[0004]	<211>	98	
	<212>	DNA	
	<213>	Pan troglodytes	
	<400>	8	
		tcctccacac tcctctctct ccacccatcc agggggcggg gccagaggte aaggetagtg	60
		ggtgggactg gggagggaga gaggggttga gtagtccc	98
	<210>	9	
	<211>	200	
	<212>	DNA	
	<213>	Homo sapiens	
	<400>	9	
		tcctccacac tcctctctct ccacccatcc agggggcggg gccagaggte aaggetagtg	60
		ggtgggactg gggagggaga gaggggttga gtagtccctt cgaagccct catttaccac	120
		ggcccccgge ttggggcgcc ttcttcccc atggcgggac acctggcttc agatttcgce	180
		ttctcgcccc ctccaggtgg	200
[0004]	<210>	10	
	<211>	61	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	Splinkerette adaptor IMSpAa	
	<400>	10	
		cgaagagtaa ccgttgctag gagagaccgt ggcctgaatga gactggtgtc gacactagtg	60
		g	61
	<210>	11	
[0004]	<211>	48	
	<212>	DNA	
	<213>	Artificial Sequence	

	<220>		
	<223>	Splinkerette adaptor HMSpBb-Sau3A1	
	<400>	11	
		gataccacta gtgtcgacac cagtctctaa tttttttttt caaaaaaa	48
	<210>	12	
	<211>	28	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	1st round PCR primer HMSp1	
	<400>	12	
		cgaagagtaa ccgttgetag gagagacc	28
	<210>	13	
	<211>	25	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	2nd round PCR primer - HMSp2	
	<400>	13	
[0005]		gtggetgaat gagactggtg tcgac	25
	<210>	14	
	<211>	25	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	Sequencing primers HMSp3	
	<400>	14	
		gcgctttact cgacctaaac tttaa	25
	<210>	15	
	<211>	30	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	1st round PCR primers, used together with HMSp1 - PB-L-Sp1	
	<400>	15	
		cctcgatata cagaccgata aaacacatgc	30
	<210>	16	
	<211>	30	
	<212>	DNA	
	<213>	Artificial Sequence	

	<220>		
	<223>	2nd round PCR primers, used together with HMs2 - PB-L-Sp2	
	<400>	16	
		gagagagcaa tattteaaga atgcatgcgt	30
	<210>	17	
	<211>	25	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	Sequencing Primer - PB-L-Sp3	
	<400>	17	
		gcgctttact cgacetaaac tttaa	25
	<210>	18	
	<211>	30	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	1st round PCR primers, used together with HMs1 - PB-R-Sp1	
	<400>	18	
[0006]		ectcgatata cagaccgata aaacacatgc	30
	<210>	19	
	<211>	30	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	2nd round PCR primers, used together with HMs2 - PB-R-Sp3	
	<400>	19	
		acgcatgatt atctttaacg tacgtcaca	30
	<210>	20	
	<211>	26	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	Sequencing Primer	
	<400>	20	
		cgtcacaata tgattatctt tctagg	26
	<210>	21	
	<211>	34	
	<212>	DNA	
	<213>	Artificial Sequence	

<220>		
<223>	mOct4-F-BglIII Primer	
<400>	21	
	gaagatctct ccaccttccc catggetgga cacc	34
<210>	22	
<211>	33	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	mOct4-R-EcoRI Primer	
<400>	22	
	cggaattctt gatcaacagc atcactgagc ttc	33
<210>	23	
<211>	31	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	mSox2-F-BglIII Primer	
<400>	23	
	gaagatctct tgtataacat gatggagagc g	31
[0007]		
<210>	24	
<211>	32	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	mSox2-R-EcoRI Primer	
<400>	24	
	cggaattctt tcacatgtgc gacaggggca gt	32
<210>	25	
<211>	36	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	mKlf4-F-BglIII Primer	
<400>	25	
	gaagatctca ccattgetgt cagcgacget ctgctc	36
<210>	26	
<211>	32	
<212>	DNA	
<213>	Artificial Sequence	

<220>		
<223>	mKlf4-R-EcoRI Primer	
<400>	26	
	cggaattcac atccactacg tgggatttaa aa	32
<210>	27	
<211>	36	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	mcMyc-F-BglII Primer	
<400>	27	
	gaagatctca ccattgcccct caacgtgaac ttcaac	36
<210>	28	
<211>	36	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	mcMyc-R-EcoRI Primer	
<400>	28	
	cggaattctt atgcaccaga gtttcgaagc tgttcg	36
[0008]		
<210>	29	
<211>	28	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	mRarg-FL/DN-F-EcoRI	
<400>	29	
	gggaattccc gcagctacca tggccacc	28
<210>	30	
<211>	34	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	mRarg-DN-R-XbaI	
<400>	30	
	gccttagatc atcatcatcc cttagtgtg atgc	34
<210>	31	
<211>	27	
<212>	DNA	
<213>	Artificial Sequence	

	<220>		
	<223>	mRarg-FL-R-XbaI Primer	
	<400>	31	
		gctctagagc ccaacccac aacgggg	27
	<210>	32	
	<211>	28	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	mRara-FL/DN-F-EcoRI Primer	
	<400>	32	
		gggaattgc tgettggcat ggccagca	28
	<210>	33	
	<211>	35	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	mRara-DN-R-XbaI Primer	
	<400>	33	
[0009]		gctctagatc atcatcatgg gatctccatc ttcaa	35
	<210>	34	
	<211>	30	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	mRara-FL-R-XbaI Primer	
	<400>	34	
		gctctagagt gtegaggtgg tcatggggat	30
	<210>	35	
	<211>	32	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	mSfl-F-EcoRI Primer	
	<400>	35	
		gggaattccc ggggcatgg actattcgta cg	32
	<210>	36	
	<211>	32	
	<212>	DNA	
	<213>	Artificial Sequence	

<220>		
<223>	mSfl-R-XbaI Primer	
<400>	36	
	getctagagg cacceaggct caagtctgct tg	32
<210>	37	
<211>	35	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	mLrh1-F-EcoRI Primer	
<400>	37	
	gggaattctt tgetaagaa tgtctgetag tttgg	35
<210>	38	
<211>	33	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	mLrh1-R-XbaI Primer	
<400>	38	
	getctagagg ggacttaggc tcttttggca tgc	33
[0010]		
<210>	39	
<211>	34	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	hOCT4-F-EcoRI Primer	
<400>	39	
	gggaattcca ccatggcggg acacctgget tcag	34
<210>	40	
<211>	34	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	hOCT4-R-XbaI Primer	
<400>	40	
	getctagaac etcagtttga atgcatggga gage	34
<210>	41	
<211>	39	
<212>	DNA	
<213>	Artificial Sequence	

	<220>		
	<223>	hSOX2-F-EcoRI Primer	
	<400>	41	
		gggaattcca ccatgtacaa catgatggag acggagctg	39
	<210>	42	
	<211>	34	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	hSOX2-R-XbaI Primer	
	<400>	42	
		gctctagatc acatgtgtga gaggggcagt gtgc	34
	<210>	43	
	<211>	38	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	hKLF4-F-EcoRI Primer	
	<400>	43	
[0011]		gggaattcca ccatggctgt cagtgaacgc ctgtctcc	38
	<210>	44	
	<211>	38	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	hKLF4-R-XbaI Primer	
	<400>	44	
		gctctagatt aaaaatgtct ctctcatgtgt aaggcgag	38
	<210>	45	
	<211>	38	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	hCMYC-F-EcoRI Primer	
	<400>	45	
		gggaattcca ccatgcccc caacgttgcg ttacacaa	38
	<210>	46	
	<211>	38	
	<212>	DNA	
	<213>	Artificial Sequence	

	<220>		
	<223>	hCMYC-R-XbaI Primer	
	<400>	46	
		gctctagatc acgcacaaga gtcccgtagc tgttcaag	38
	<210>	47	
	<211>	33	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	hLRHI-F-EcoRI Primer	
	<400>	47	
		gggaattcat gtcttctaatt tcagatactg ggg	33
	<210>	48	
	<211>	30	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	hLRHI-R-XbaI Primer	
	<400>	48	
[0012]		gctctagatt atgctctttt ggcattgcaac	30
	<210>	49	
	<211>	28	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	hRARG-F-EcoRI Primer	
	<400>	49	
		gggaattcat ggcaccaat aaggagcg	28
	<210>	50	
	<211>	29	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	hRARG-R-XbaI Primer	
	<400>	50	
		gctctagatc aggctgggga ctccagcc	29
	<210>	51	
	<211>	23	
	<212>	DNA	
	<213>	Artificial Sequence	

	<220>		
	<223>	mOct4-RT-En-F Primer	
	<400>	51	
		tctttccacc aggccccgg etc	23
	<210>	52	
	<211>	23	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	mOct4-RT-En-R Primer	
	<400>	52	
		tgcgggcgga catggggaga tcc	23
	<210>	53	
	<211>	23	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	mSox2-RT-En-F Primer	
	<400>	53	
		ttgecttaaa caagaccacg aaa	23
[0013]	<210>	54	
	<211>	23	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	mSox2-RT-En-R Primer	
	<400>	54	
		tagagctaga ctccgggcga tga	23
	<210>	55	
	<211>	24	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	mKlf4-RT-En-F Primer	
	<400>	55	
		gcgaactcac acaggcgaga aacc	24
	<210>	56	
	<211>	23	
	<212>	DNA	
	<213>	Artificial Sequence	

[0014]	<220>		
	<223>	mKlf4-RT-En-R Primer	
	<400>	56	
		tcgcttcctc ttectecgac aca	23
	<210>	57	
	<211>	20	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	mNanog-RT-En-F Primer	
	<400>	57	
		cagggtgttg agggtagctc	20
	<210>	58	
	<211>	20	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	mNanog-RT-En-R Primer	
	<400>	58	
		cggttcatca tggtagctc	20
	<210>	59	
	<211>	24	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	mRex1-RT-F Primer	
	<400>	59	
		acgagtggca gttttcttctt ggga	24
	<210>	60	
	<211>	24	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	mRex1-RT-R Primer	
	<400>	60	
		tatgactcac ttccagggg cact	24
	<210>	61	
	<211>	24	
	<212>	DNA	
	<213>	Artificial Sequence	

	<220>		
	<223>	mEsg1-RT-F Primer	
	<400>	61	
		gaagtetggt tccttggcag gatg	24
	<210>	62	
	<211>	20	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	mEsg1-RT-R Primer	
	<400>	62	
		actegataca ctggcctaga	20
	<210>	63	
	<211>	23	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	mFgf4-RT-F Primer	
	<400>	63	
		cgtggtgage atctteggag tgg	23
[0015]	<210>	64	
	<211>	23	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	mFgf4-RT-R Primer	
	<400>	64	
		ccttcttggg ccgccegttc tta	23
	<210>	65	
	<211>	24	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	mDax1-RT-F Primer	
	<400>	65	
		tgetgcgggc caggecatca agag	24
	<210>	66	
	<211>	24	
	<212>	DNA	
	<213>	Artificial Sequence	

	<220>		
	<223>	mDax1-RT-R Primer	
	<400>	66	
		gggeactgtt cagttcagcg gac	24
	<210>	67	
	<211>	24	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	mGdf3-RT-F primer	
	<400>	67	
		gttccaacct gtgcctcgcg tctt	24
	<210>	68	
	<211>	26	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	mGdf3-RT-R Primer	
	<400>	68	
[0016]		agcgaggcat ggagagagcg gacgag	26
	<210>	69	
	<211>	23	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	mUtf1-RT-F Primer	
	<400>	69	
		ggatgtcccg gtgactaagt ctg	23
	<210>	70	
	<211>	23	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	mUtf1-RT-R Primer	
	<400>	70	
		ggcggatctg gttatcgaag ggt	23
	<210>	71	
	<211>	24	
	<212>	DNA	
	<213>	Artificial Sequence	

	<220>		
	<223>	mOct4-DMR-F2 Primer	
	<400>	71	
		tgggttgaaa tattgggttt attt	24
	<210>	72	
	<211>	23	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	mOct4-DMR-R2 Primer	
	<400>	72	
		ctaaaaccaa atatecaacc ata	23
	<210>	73	
	<211>	30	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	mNanog-DMR-F Primer	
	<400>	73	
[0017]		gattttgtag gtgggattaa ttgtgaattt	30
	<210>	74	
	<211>	30	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	mNanog-DMR-R Primer	
	<400>	74	
		accaaaaaaa cccacactca tatcaatata	30
	<210>	75	
	<211>	30	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	mNanog-DMR-R Primer	
	<400>	75	
		accaaaaaaa cccacactca tatcaatata	30
	<210>	76	
	<211>	30	
	<212>	DNA	
	<213>	Artificial Sequence	

	<220>		
	<223>	mRex1-DMR-F2 Primer	
	<400>	76	
		tatattaatg ttggaaaaag tttaggtaat	30
	<210>	77	
	<211>	30	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	mRex1-DMR-R2 Primer	
	<400>	77	
		aactecttaa accctccct ttttaataa	30
	<210>	78	
	<211>	24	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	hOCT4-RT-endo-F Primer	
	<400>	78	
[0018]		gacaggggga ggggaggaga tagg	24
	<210>	79	
	<211>	26	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	hOCT4-RT-endo-R Primer	
	<400>	79	
		cttcctcca accagttgcc ccaaac	26
	<210>	80	
	<211>	26	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	hSOX2-RT-endo-F Primer	
	<400>	80	
		gggaaatggg aggggtgca aagagg	26
	<210>	81	
	<211>	26	
	<212>	DNA	
	<213>	Artificial Sequence	

	<220>		
	<223>	hSOX2-RT-endo-R Primer	
	<400>	81	
		ttgcgtgagt gtggatggga ttggtg	26
	<210>	82	
	<211>	25	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	hKLF4-RT-endo-F Primer	
	<400>	82	
		acgatacgtgg ccccggaata ggacc	25
	<210>	83	
	<211>	26	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	hKLF4-RT-endo-R Primer	
	<400>	83	
		gcgtcctggg aaggagatc cggagc	26
[0019]	<210>	84	
	<211>	26	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	hCMYC-RT-endo-F Primer	
	<400>	84	
		gcgtcctggg aaggagatc cggagc	26
	<210>	85	
	<211>	26	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	hCMYC-RT-endo-R Primer	
	<400>	85	
		ttgaggggca tcttcgcggg aggetg	26
	<210>	86	
	<211>	25	
	<212>	DNA	
	<213>	Artificial Sequence	

	<220>		
	<223>	hNANOG-RT-F Primer	
	<400>	86	
		cagcccccgat tcttcaccca gtecc	25
	<210>	87	
	<211>	25	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	hNANOG-RT-R Primer	
	<400>	87	
		cggaagattc ccagtcgggt tcaac	25
	<210>	88	
	<211>	24	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	hREX1-RT-F Primer	
	<400>	88	
		cagatcctaa acagctcgca gaat	24
[0020]	<210>	89	
	<211>	23	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	hREX1-RT-R Primer	
	<400>	89	
		gcgtacgcaa attaaagtec aga	23
	<210>	90	
	<211>	26	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	hTERT-RT-F Primer	
	<400>	90	
		cctgtctcag ctgactcgac accgtg	26
	<210>	91	
	<211>	25	
	<212>	DNA	
	<213>	Artificial Sequence	

	<220>		
	<223>	hTERT-RT-R Primer	
	<400>	91	
		ggaaaagctg gccctggggt ggagc	25
	<210>	92	
	<211>	25	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	hDNMT3B-RT-F Primer	
	<400>	92	
		tgctgctcac agggcctgat acttc	25
	<210>	93	
	<211>	28	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	hDNMT3B RT R Primer	
	<400>	93	
[0021]		tcctttcgag ctcagtgcac cacaaaac	28
	<210>	94	
	<211>	24	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	hCMYC-taq-endo-F QPCR probe	
	<400>	94	
		cctgagcaat caactatgaa cttg	24
	<210>	95	
	<211>	22	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	hCMYC-taq-endo-R QPCR Probe	
	<400>	95	
		ttatgcccaa agtccaattt ga	22
	<210>	96	
	<211>	26	
	<212>	DNA	
	<213>	Artificial Sequence	

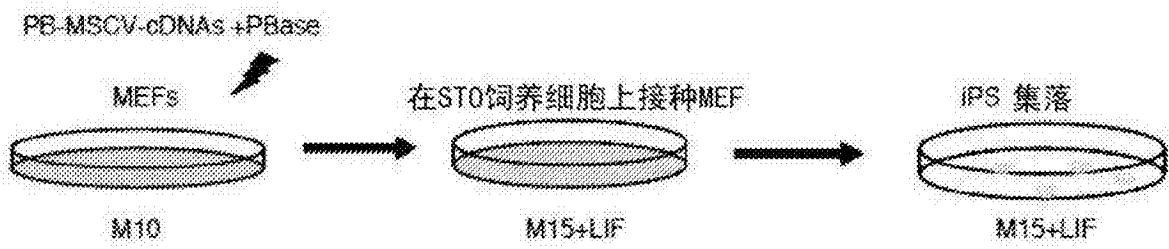
	<220>		
	<223>	hCMYC-taq-endo-probe	
	<400>	96	
		caaatgcaac ctcacaacct tggetg	26
	<210>	97	
	<211>	22	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	hKLF4-taq-endo-F probe	
	<400>	97	
		tttcacactg tcttcccgat ga	22
	<210>	98	
	<211>	25	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	hKLF4-taq-endo-R probe	
	<400>	98	
[0022]		tcttgattat ccactcaca gatga	25
	<210>	99	
	<211>	30	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	hKLF4-taq-endo-probe	
	<400>	99	
		ccagccagaa agcactacaa tcatggtaa	30
	<210>	100	
	<211>	23	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	hOCT4-taq-endo-F Probe	
	<400>	100	
		cactgtactc cteggtcct ttc	23
	<210>	101	
	<211>	20	
	<212>	DNA	
	<213>	Artificial Sequence	

	<220>		
	<223>	hOCT4-taq-endo-R Probe	
	<400>	101	
		caaccagttg ccccaaaactc	20
	<210>	102	
	<211>	26	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	hOCT4-taq-endo-probe	
	<400>	102	
		ctttccccct gtctccgtea ccactc	26
	<210>	103	
	<211>	21	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	hSOX2-taq-endo-F Probe	
	<400>	103	
[0023]		acagcaaatg acagctgcaa a	21
	<210>	104	
	<211>	28	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	hSOX2-taq-endo-R Probe	
	<400>	104	
		aagtcagga tctctctcat aaaagttt	28
	<210>	105	
	<211>	24	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	hSOX2-taq-endo-probe	
	<400>	105	
		atccacactc acgcaaaaac cgeg	24
	<210>	106	
	<211>	26	
	<212>	DNA	
	<213>	Artificial Sequence	

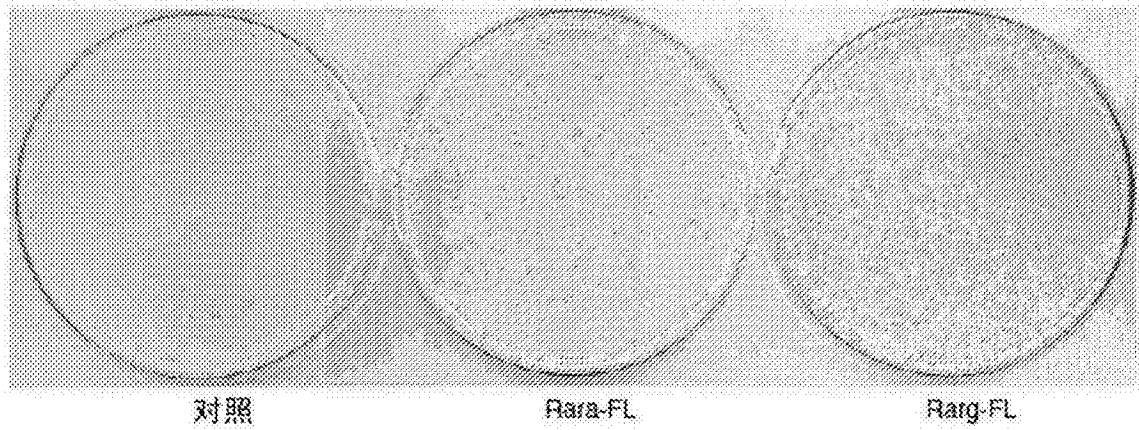
	<220>		
	<223>	hLRHI-taq-endo-F probe	
	<400>	106	
		aaagctgaac tgaacaatt ctaag	26
	<210>	107	
	<211>	18	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	hLRHI-taq-endo-R Probe	
	<400>	107	
		gctegggect tcaaagga	18
	<210>	108	
	<211>	30	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	hLRHI-taq-endo-probe	
	<400>	108	
[0024]		tgcacagct gtacctaca tagccctcc	30
	<210>	109	
	<211>	22	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	hRARG-taq-endo-F Probe	
	<400>	109	
		gagcctgggt ttggacteta aa	22
	<210>	110	
	<211>	20	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	hRARG-taq-endo-R probe	
	<400>	110	
		cctcttgecc tgggaagtct	20
	<210>	111	
	<211>	22	
	<212>	DNA	
	<213>	Artificial Sequence	

	<220>	
	<223> hRARG-taq-endo-probe	
[0025]	<400> 111	
	ctcagcactg ccccatgggt cc	22

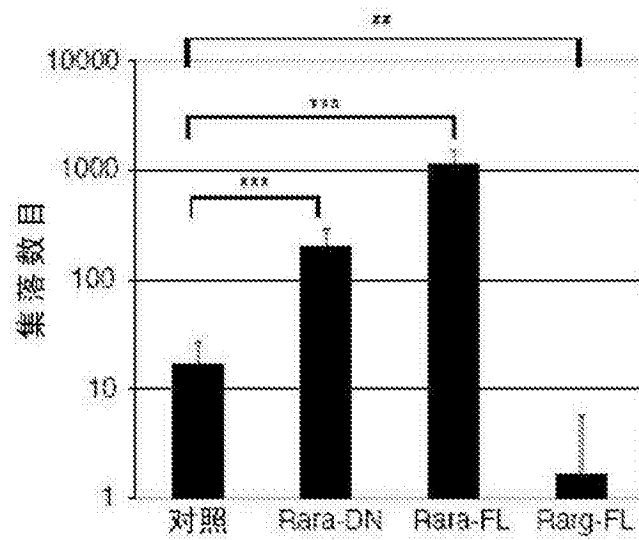
A



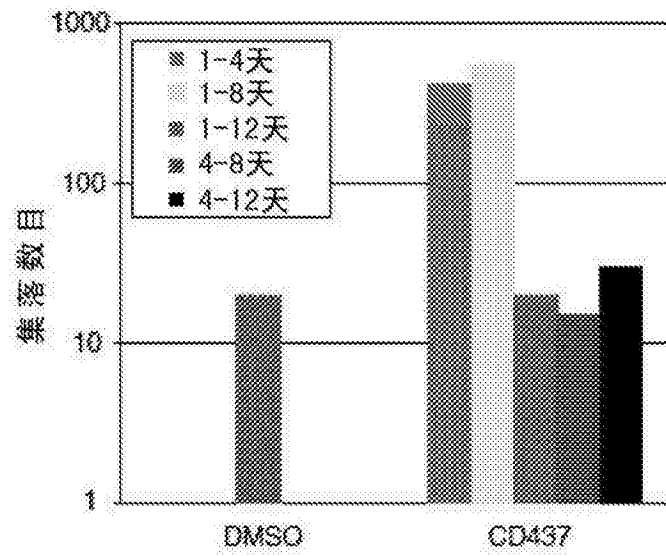
B



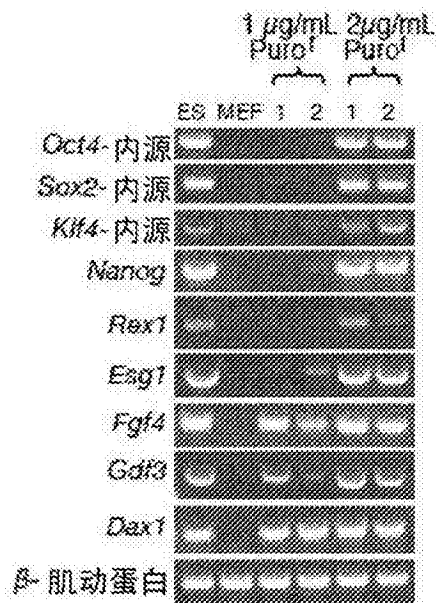
C



D



E



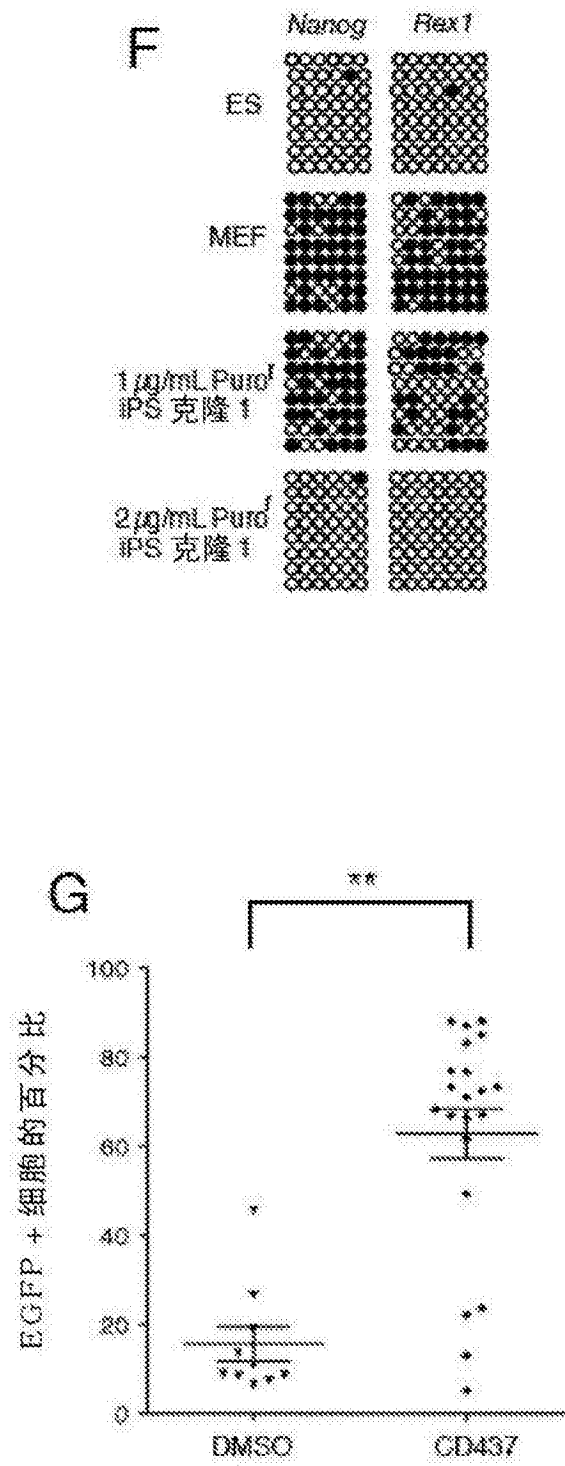
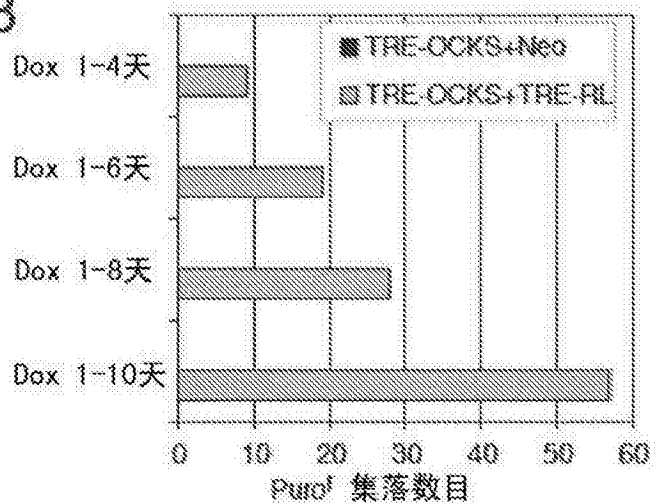


图1

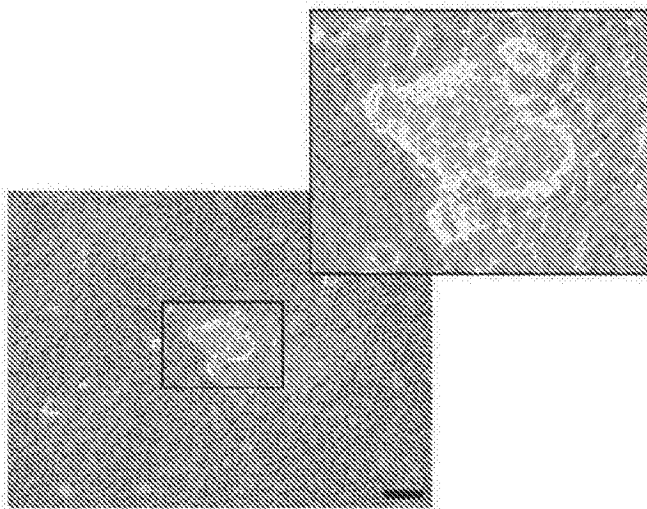
A



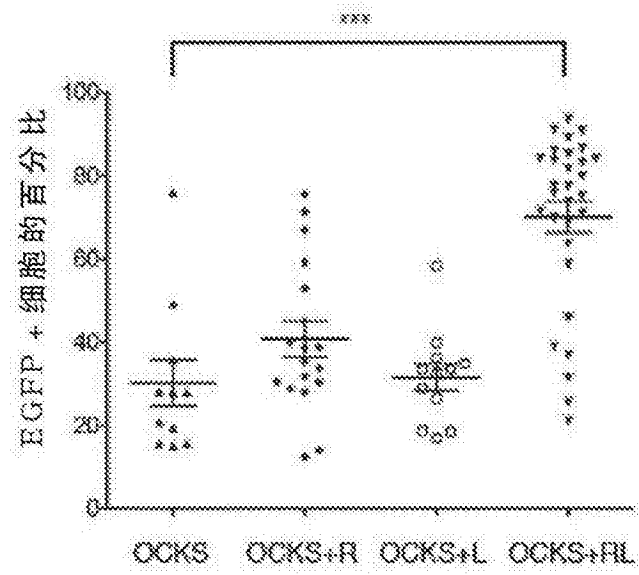
B



C



D



E

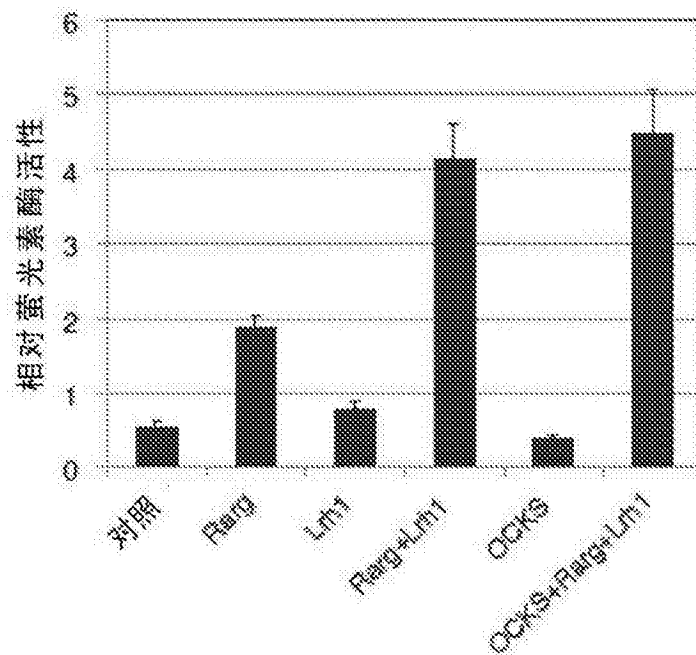
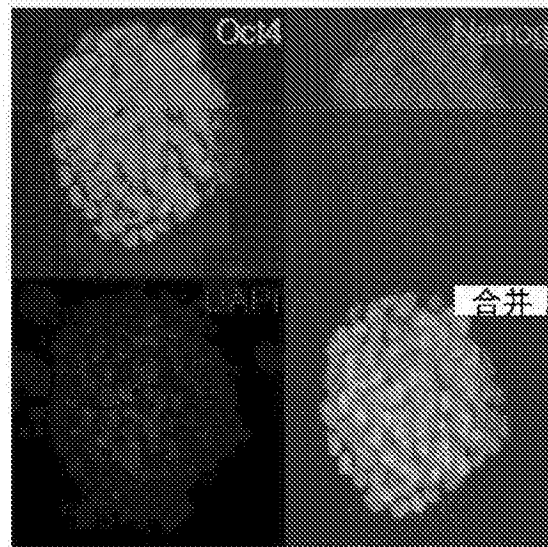
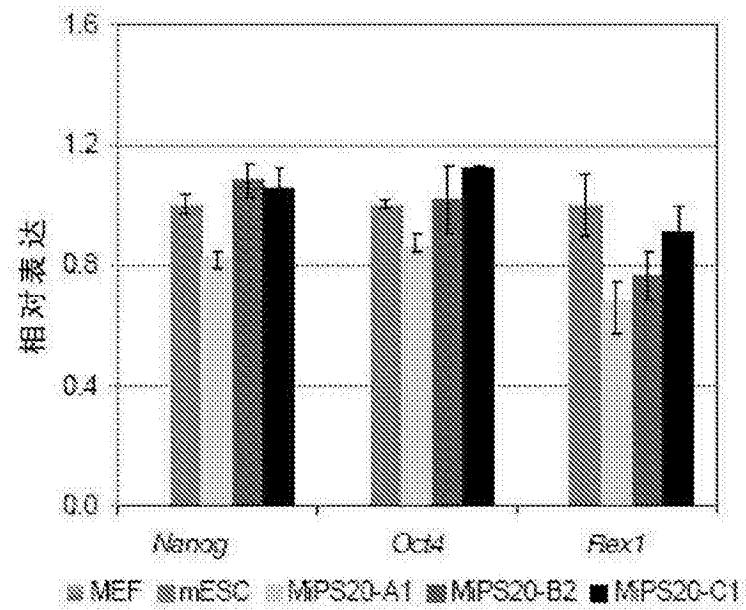


图2

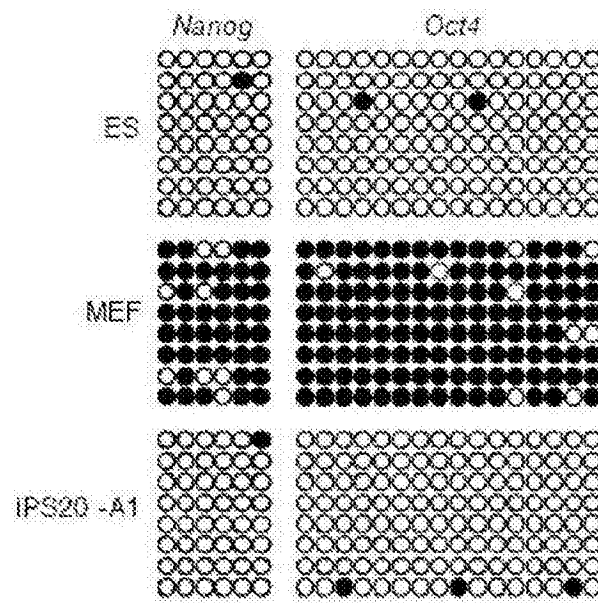
A



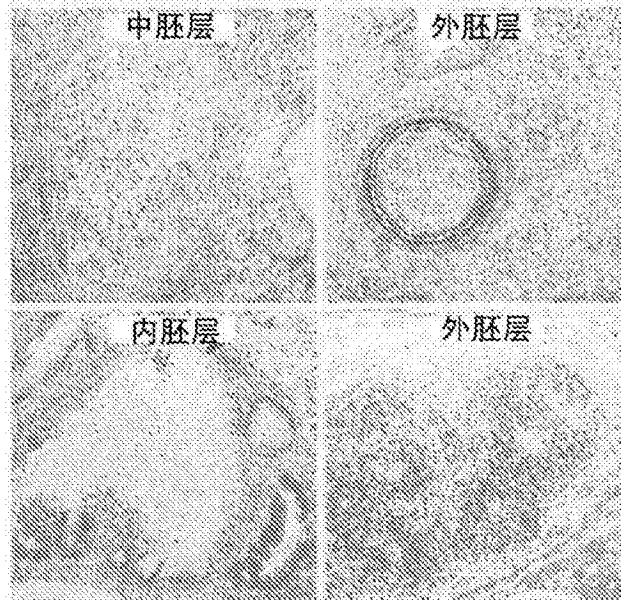
B



C



D

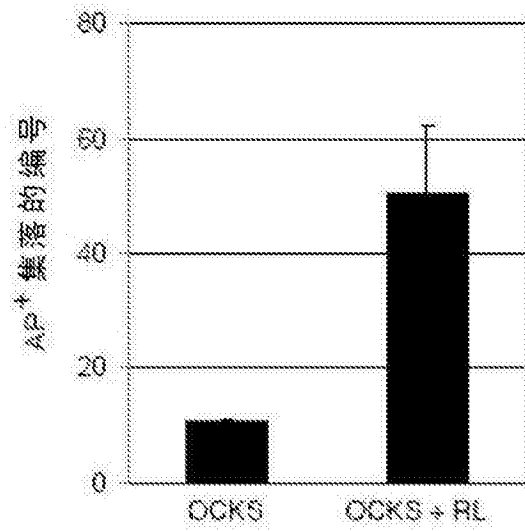


E

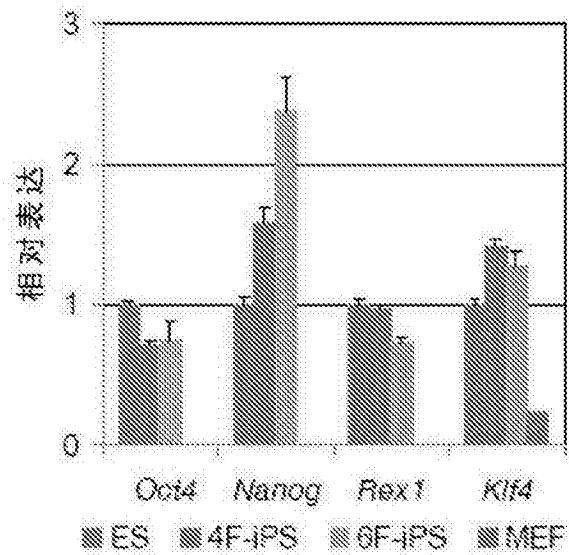


图3

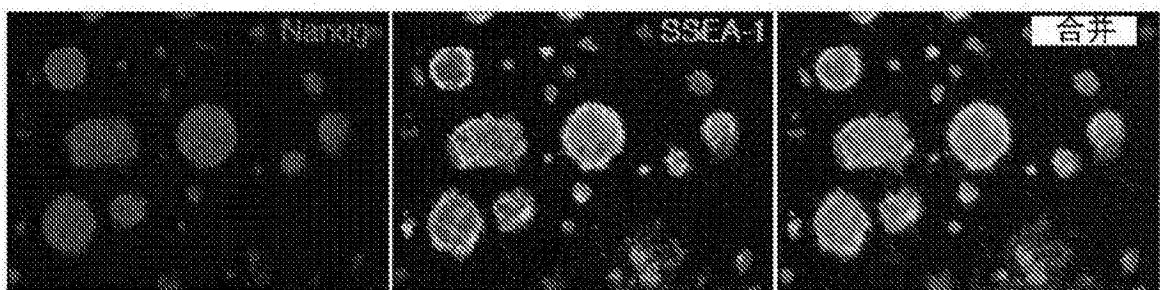
A



B



C



D

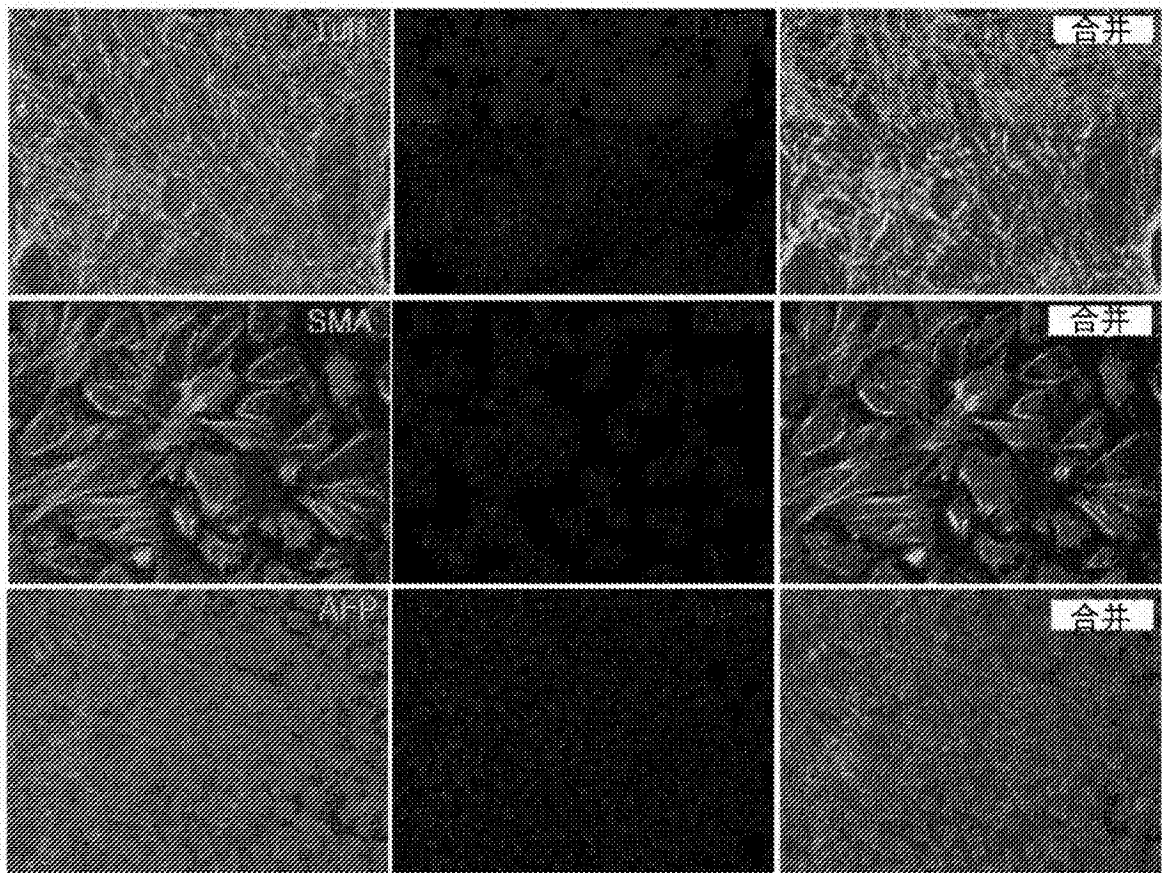
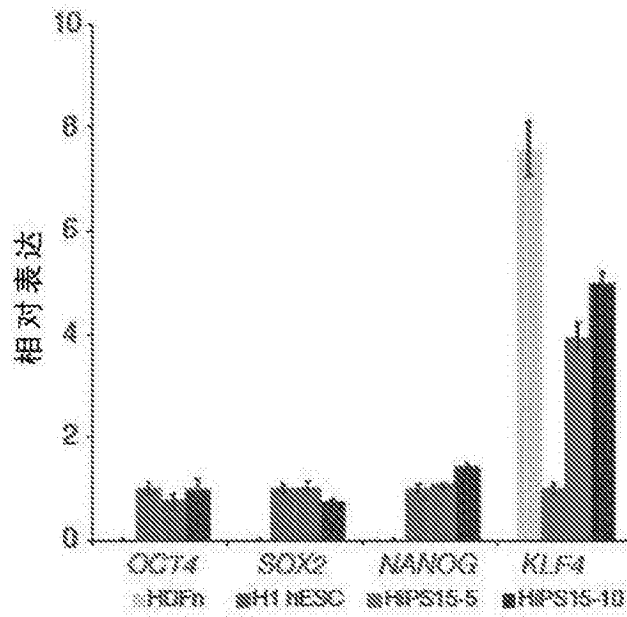
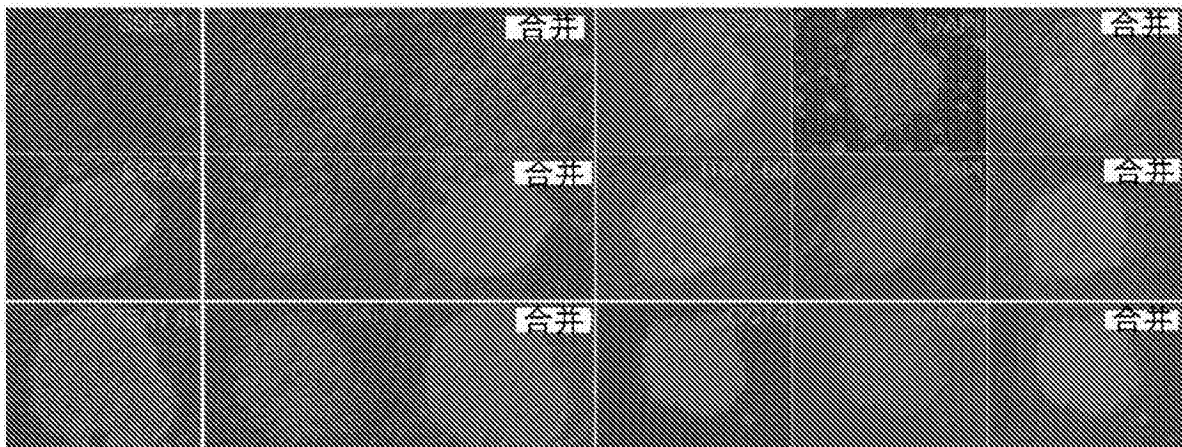


图4

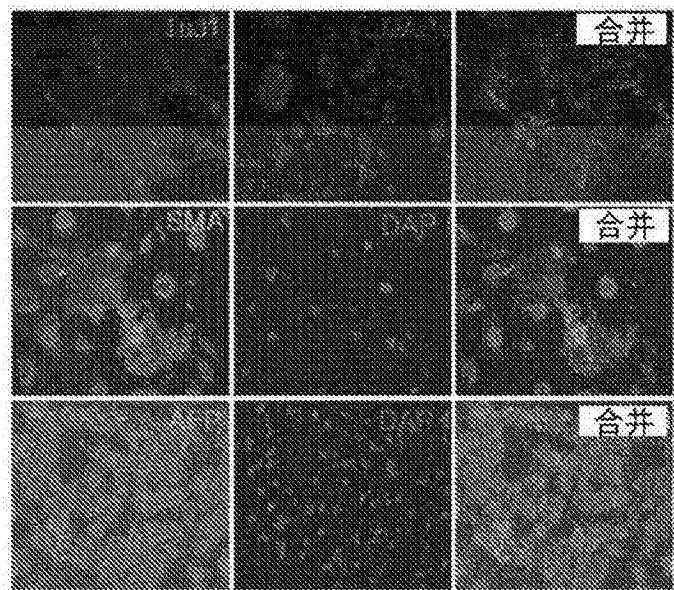
D



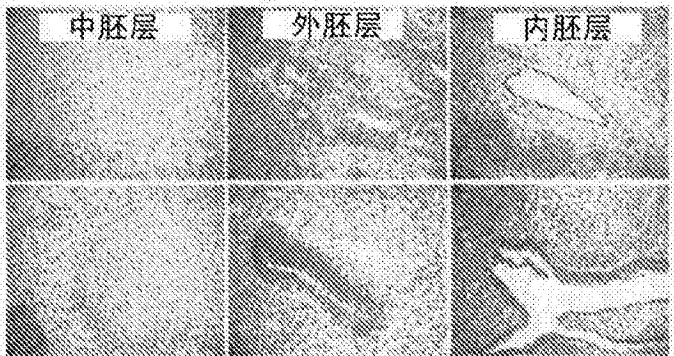
E



F



G



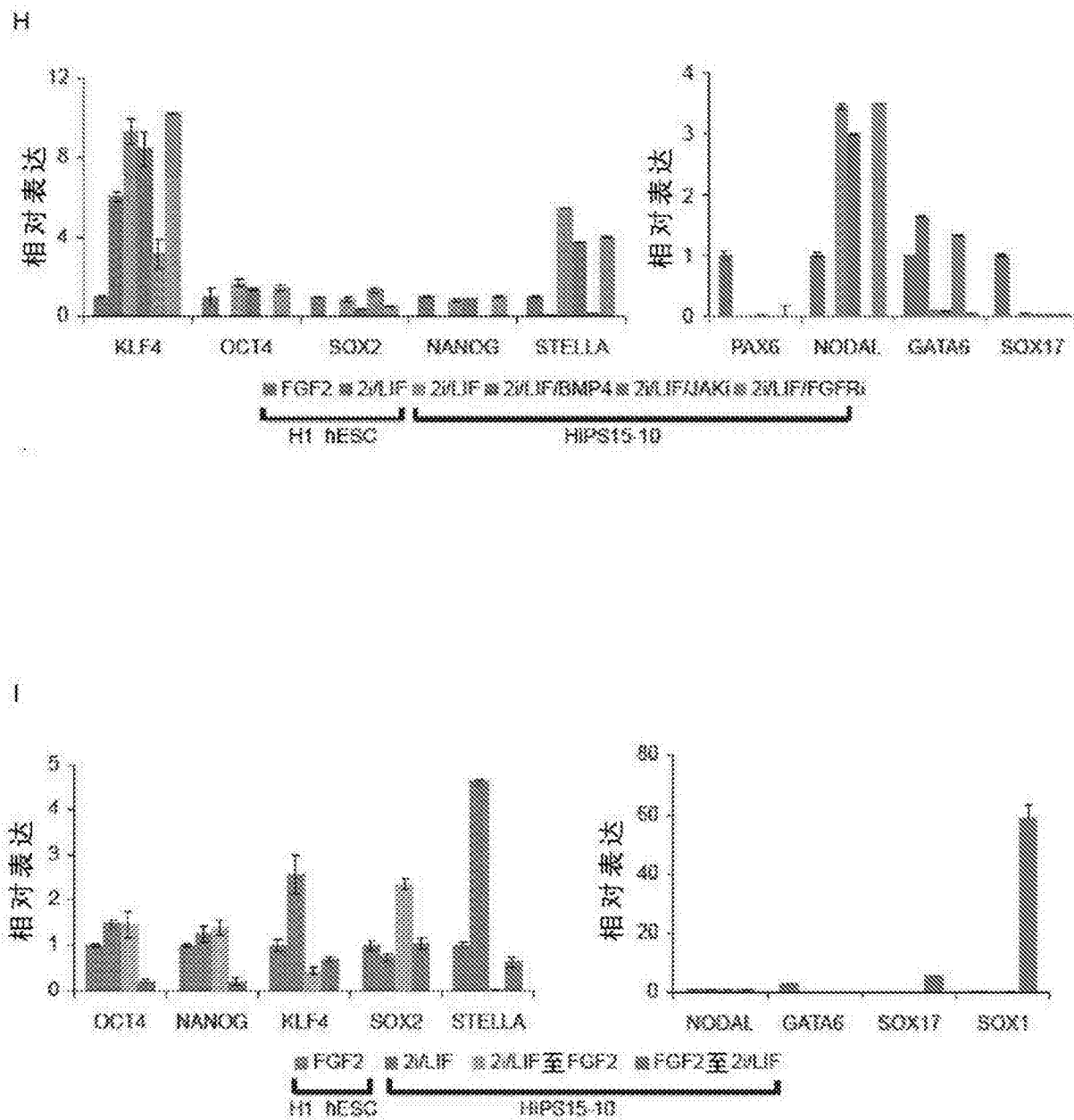
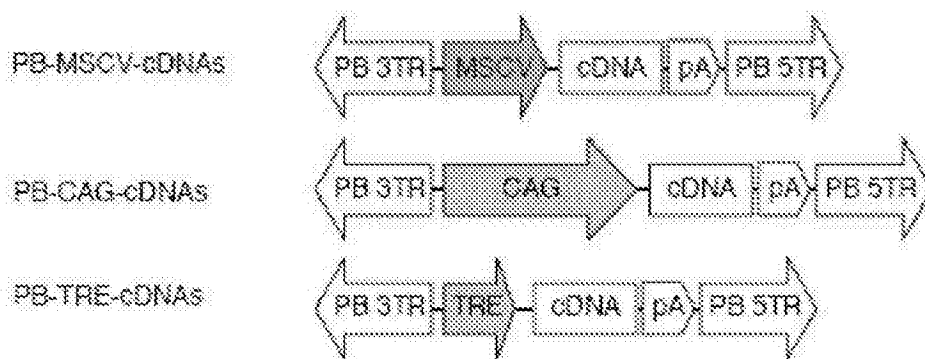


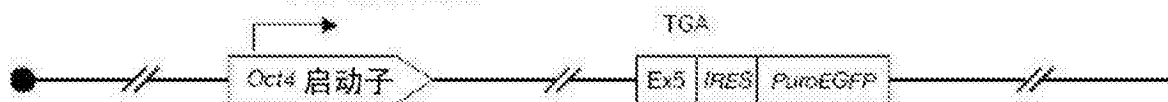
图5

A

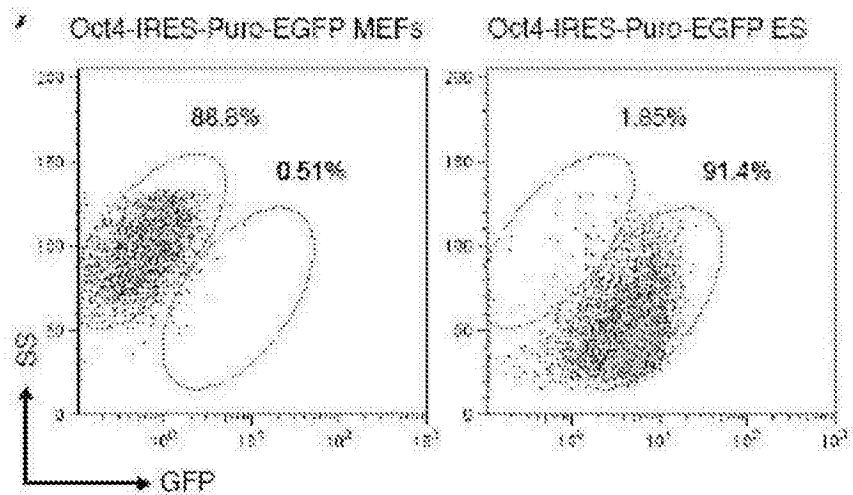


B

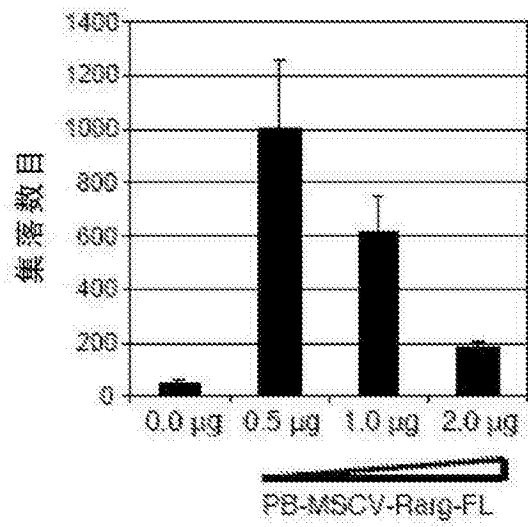
The Oct4-IRES-PuroEGFP 敲入等位基因



C



D



E

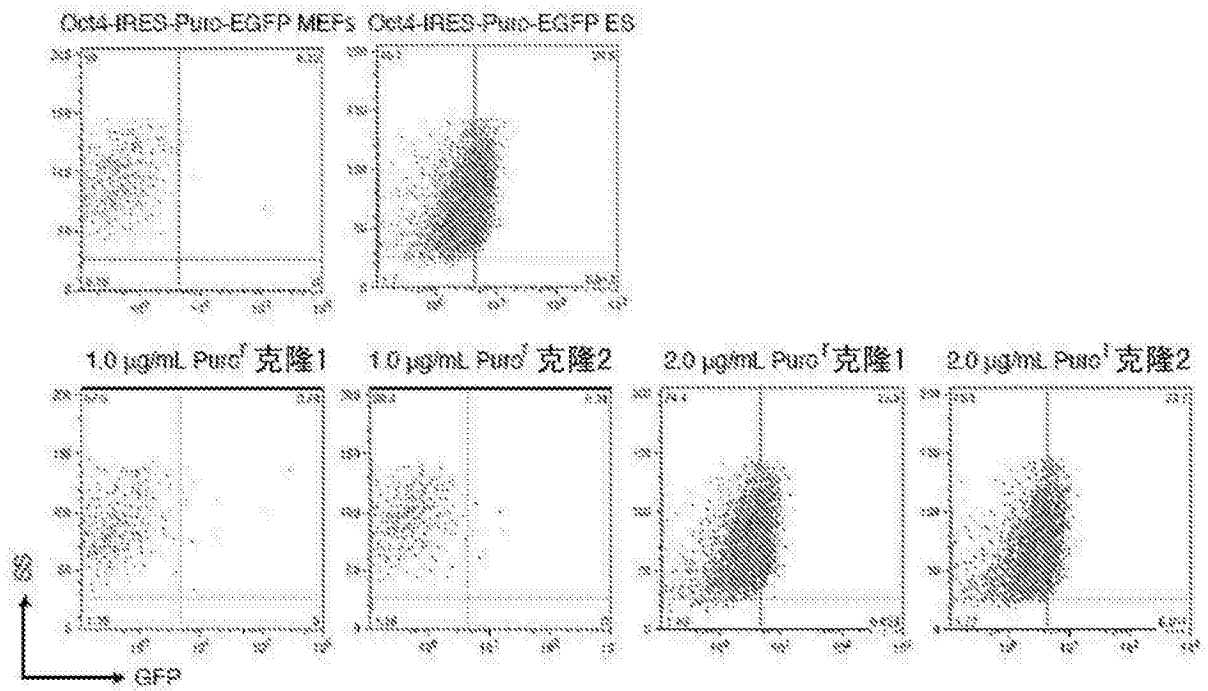
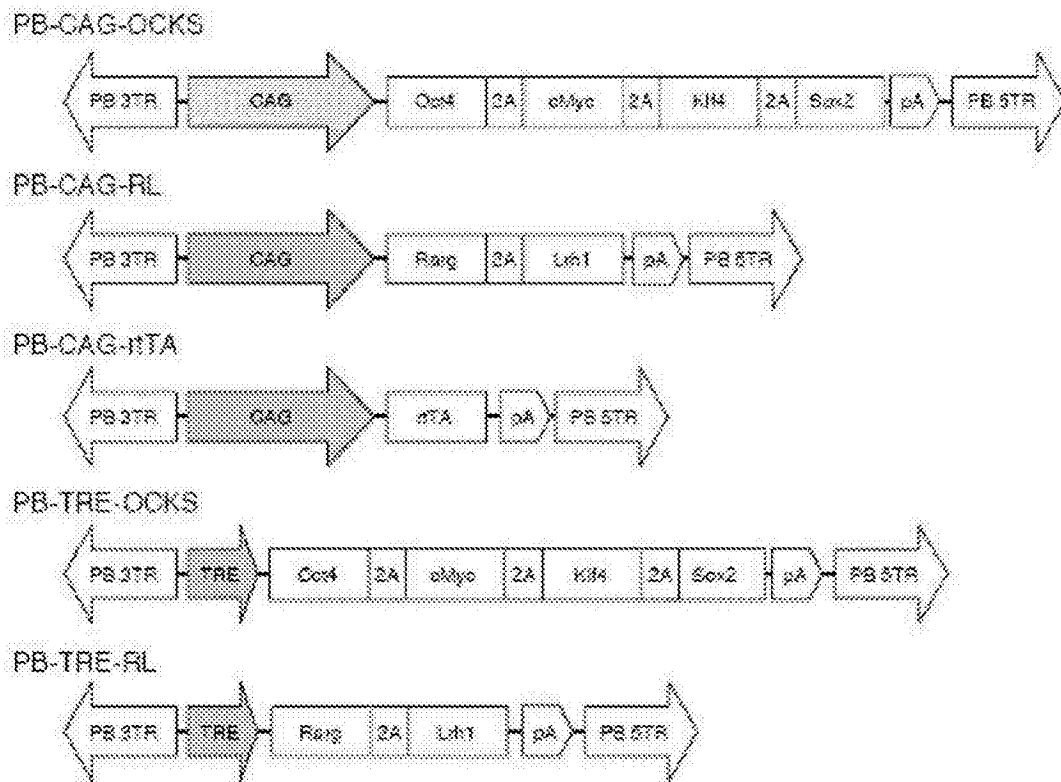


图6

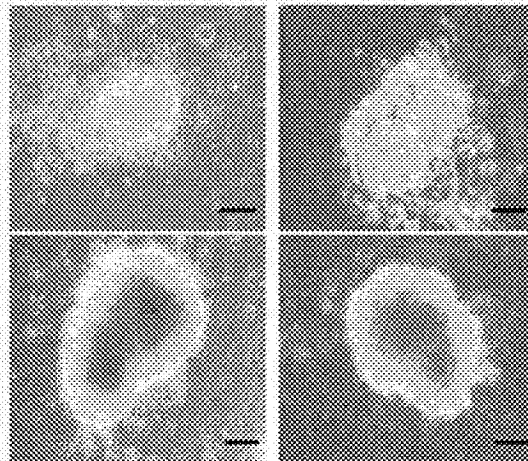
A



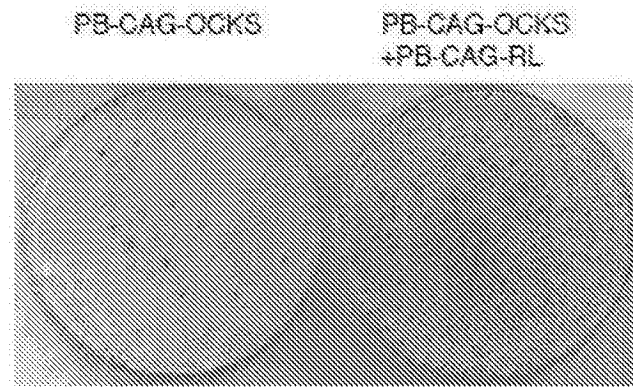
B

阶段

PB-CAG-OCKS

PB-CAG-OCKS
+ PB-CAG-RL

C



D

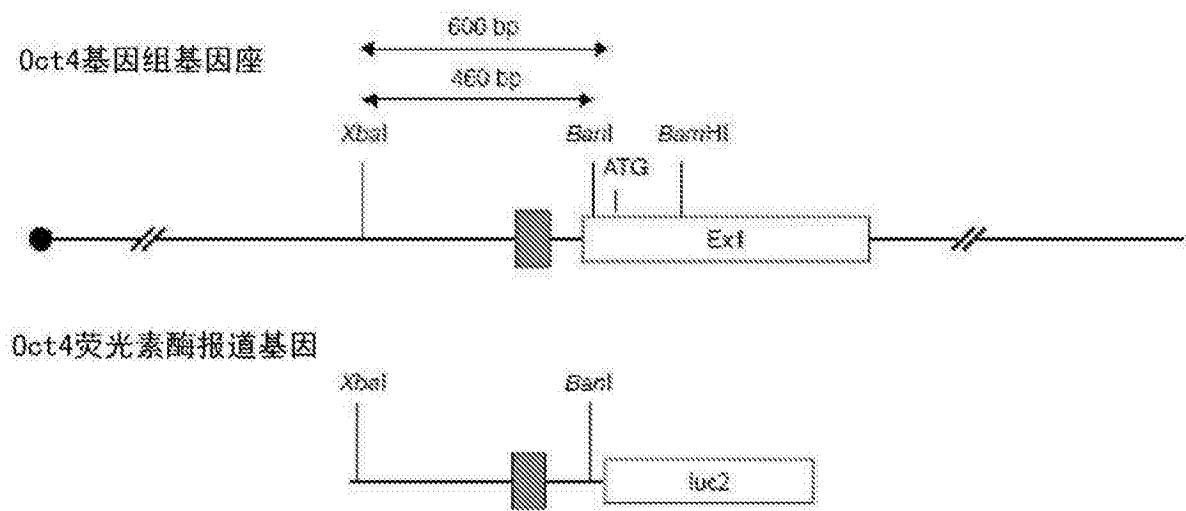
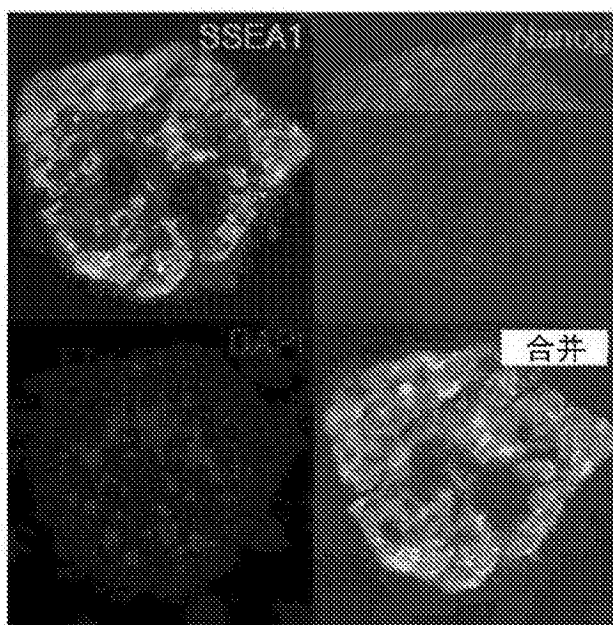
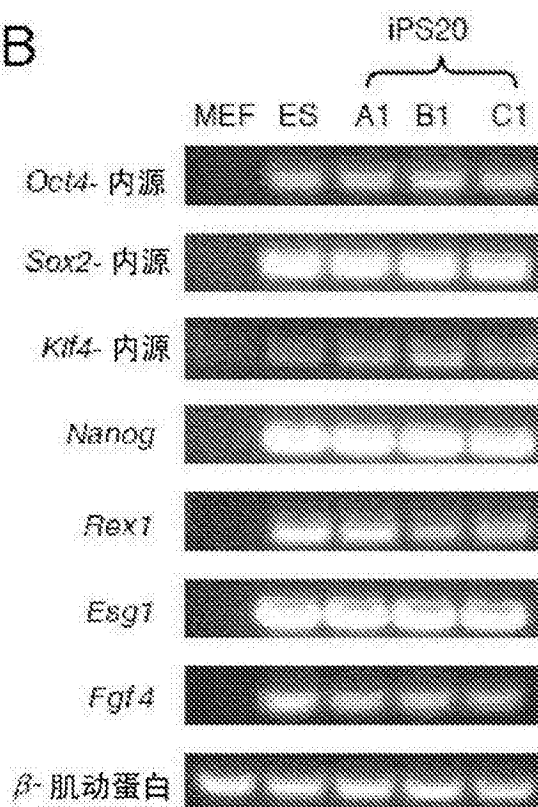


图7

A



B



C

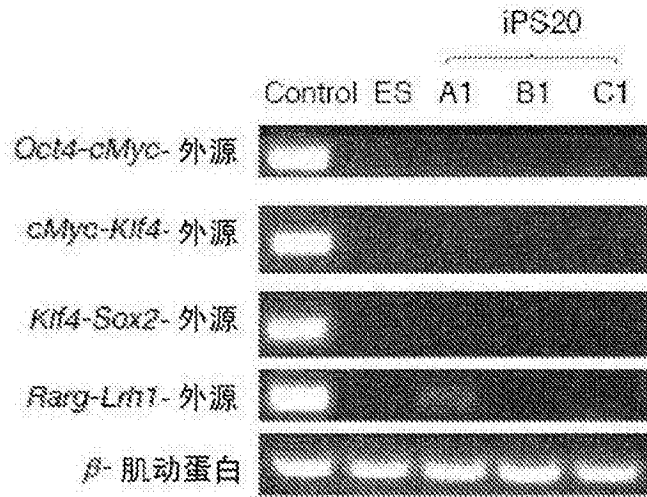


图8

A

在无血清培养基中培养17天后 AP^{+} 集落的数量

培养基	PB-TRE-OCKS	PB-TRE-OCKS+PB-TRE-FL
LIF/Dox(4d) to 2i/LIF	0:0:0	0:0:0
LIF/Dox(8d) to 2i/LIF	0:0:0	5:2:1
LIF/Dox(12d) to 2i/LIF	1:0:0	4:5:0
2i/LIF/Dox(17d)	0:0:0	4:0:2

B

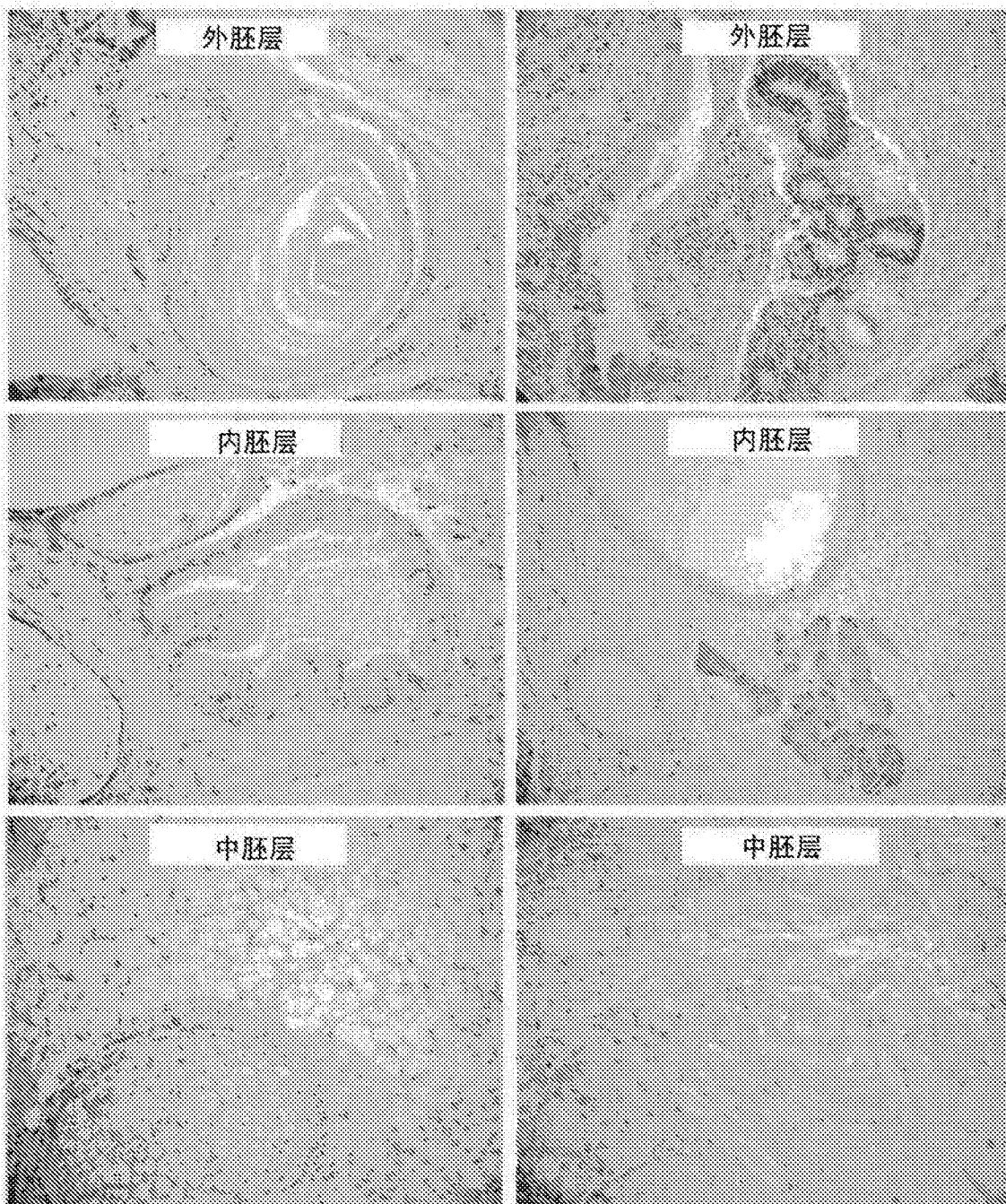
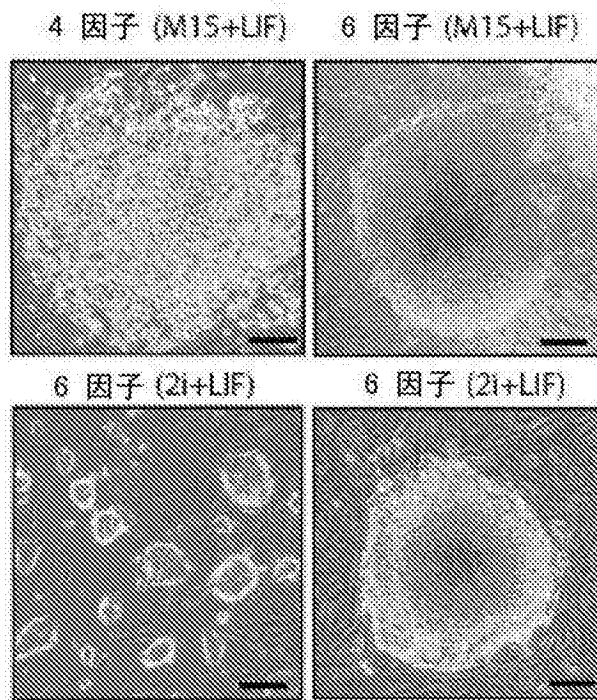
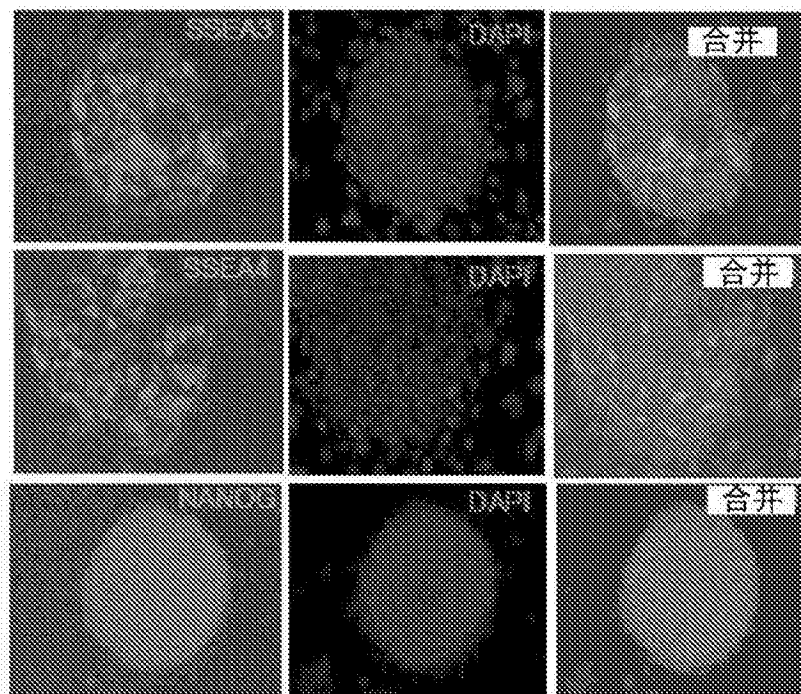


图9

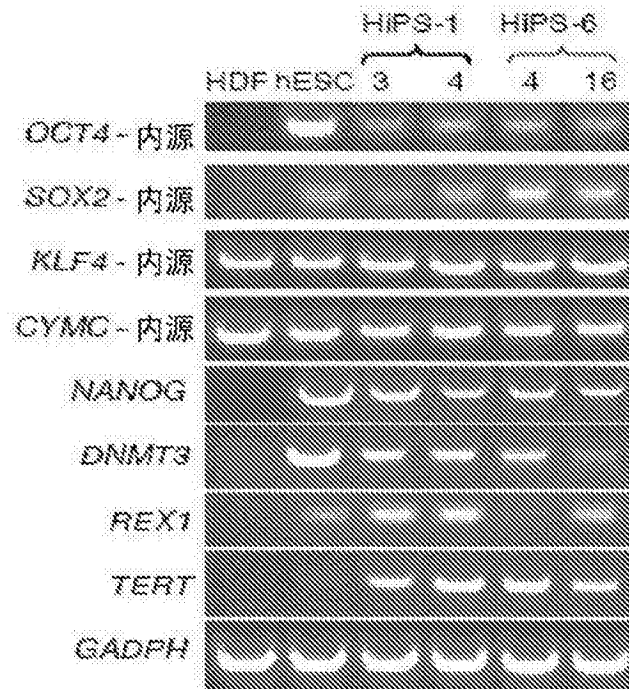
A



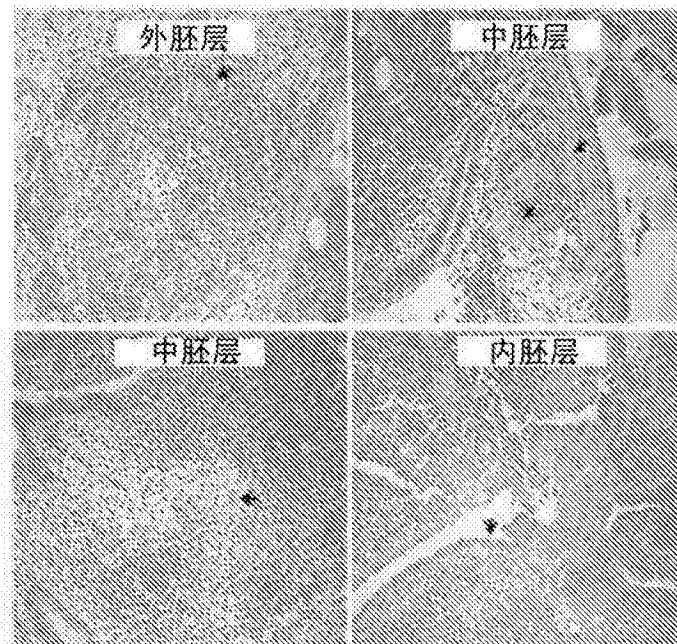
B



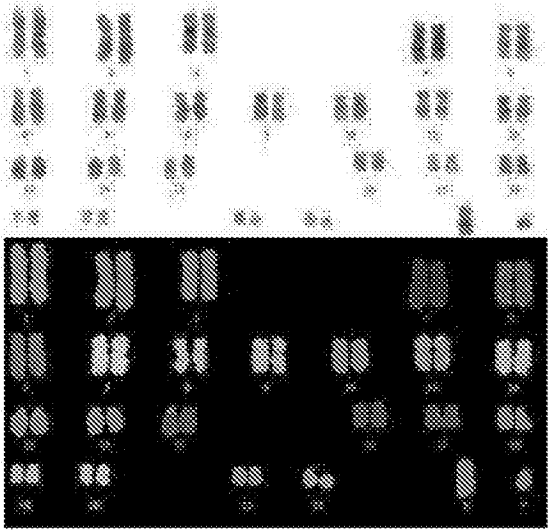
C



D



E



F

染色体	核型															
	X,Y	X,Y	X,Y	X,Y	X,Y	X,Y	X,Y	X,Y	X,Y	X,Y	X,Y	X,Y	X,Y	X,Y	X,Y	X,Y
正常	XY	XY	XY	XY	XY	XY	XY	XY	XY	XY	XY	XY	XY	XY	XY	XY
45,X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
45,X,-X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
45,X,-Y	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
45,X,-X,-Y	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

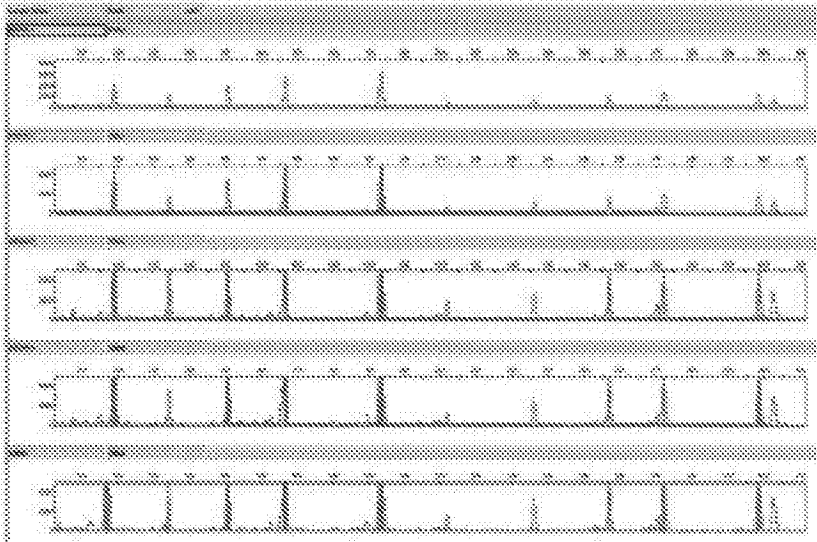
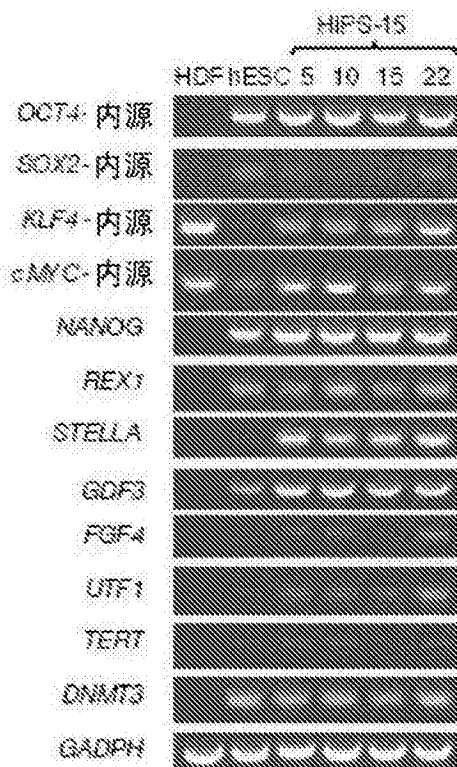
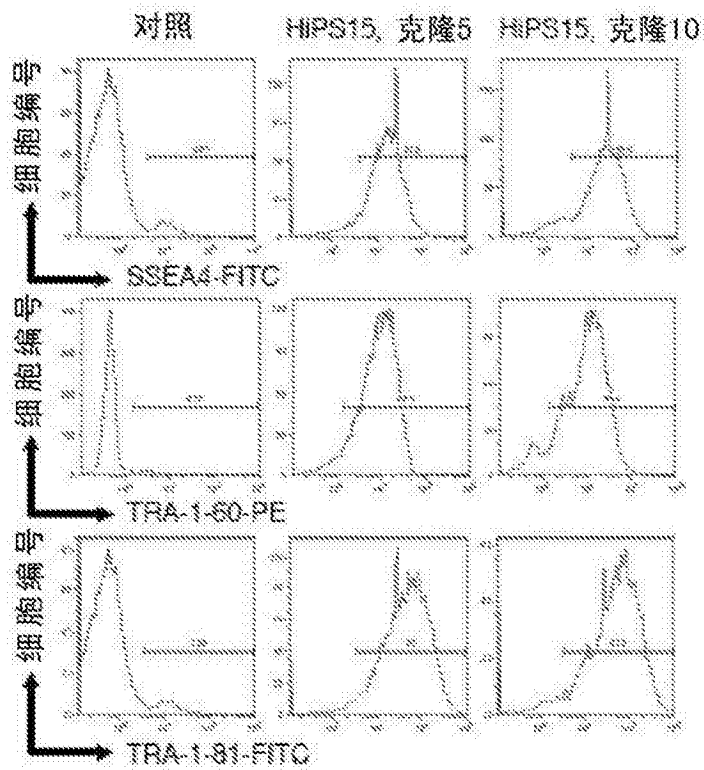


图10

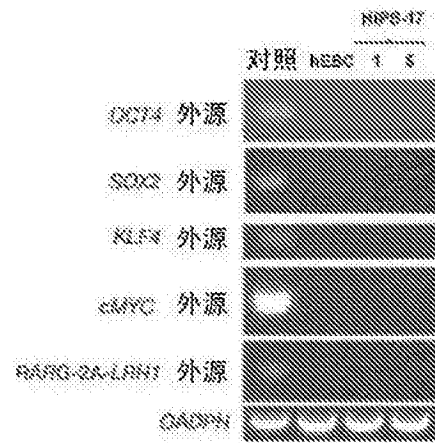
A



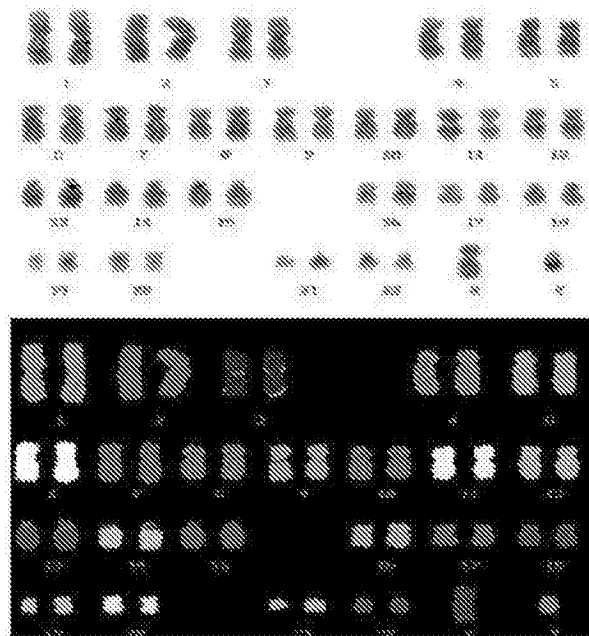
B



C



D



E

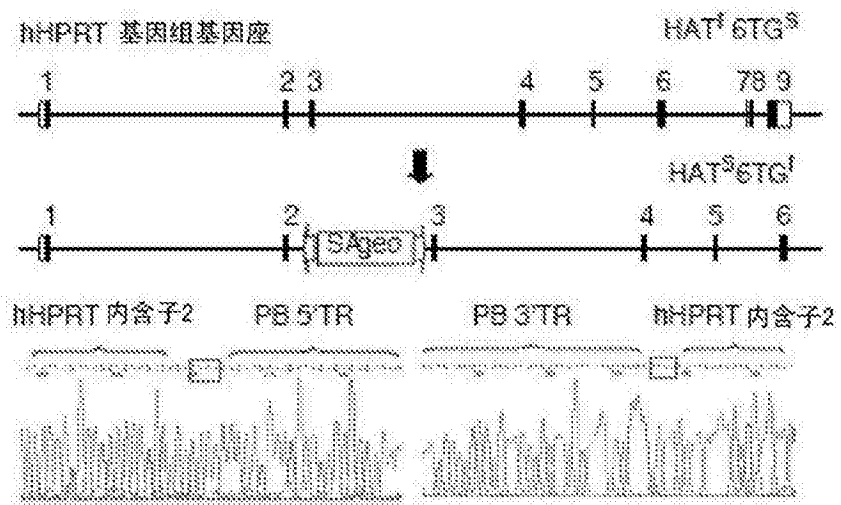
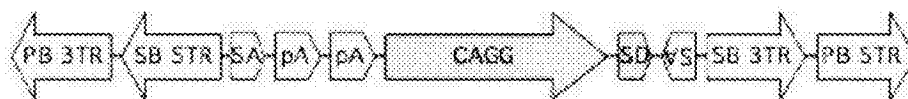
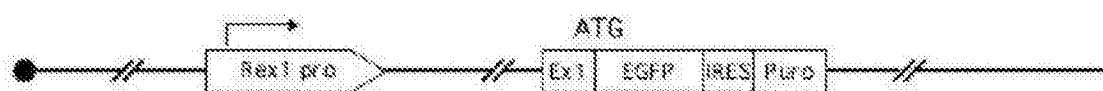


图11

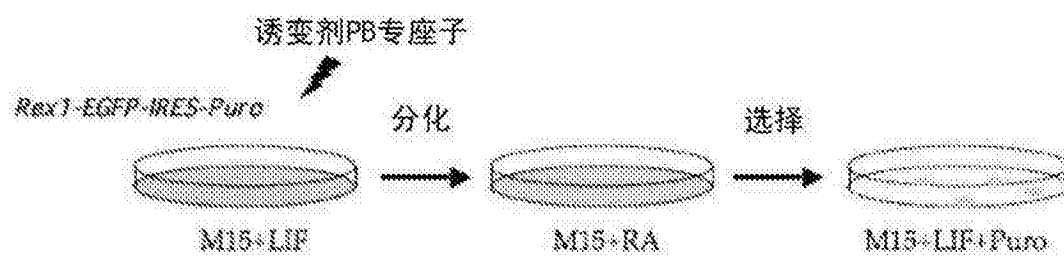
A

pA8-CAGGS

B

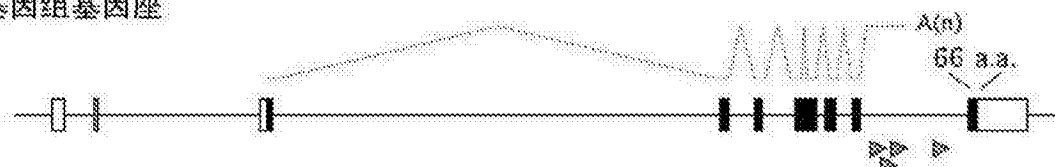
Rex1-EGFP-IRES-Puro 敲入等位基因

C

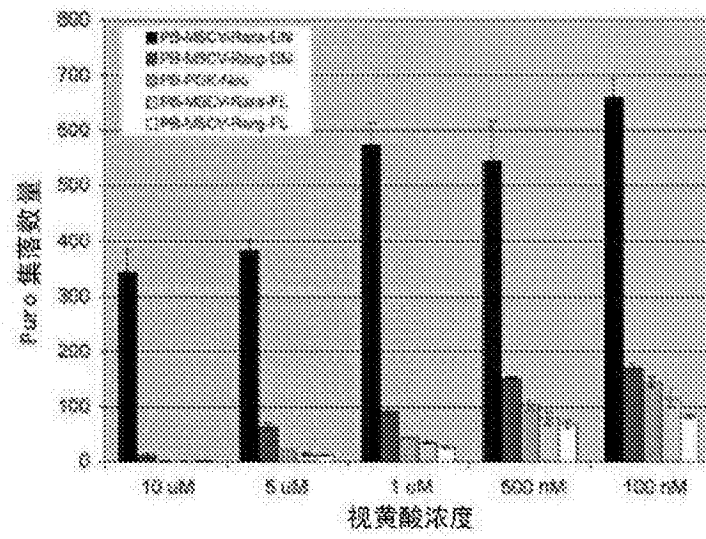


D

基因组基因座



E



F

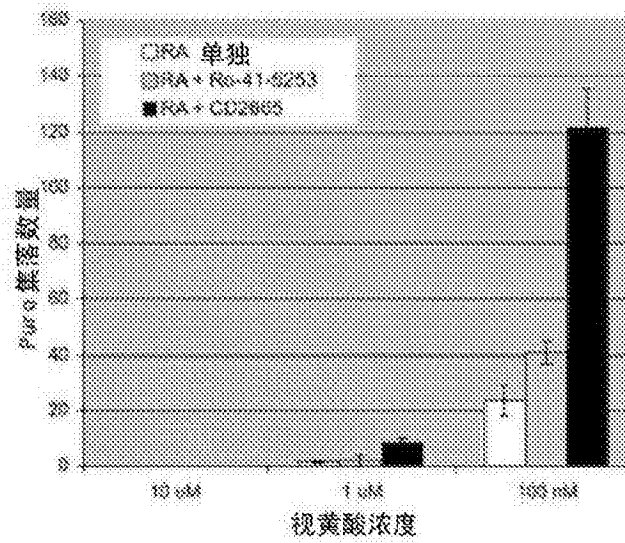
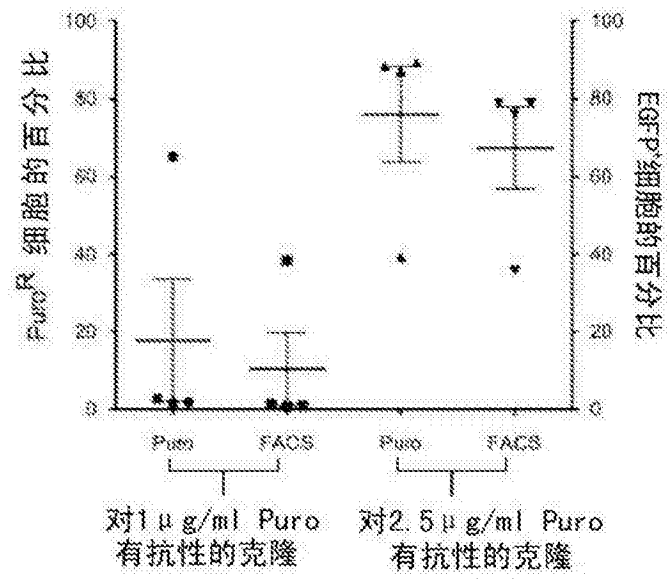
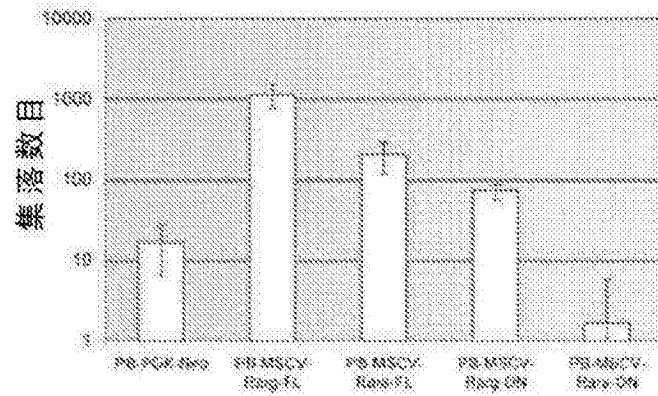


图12

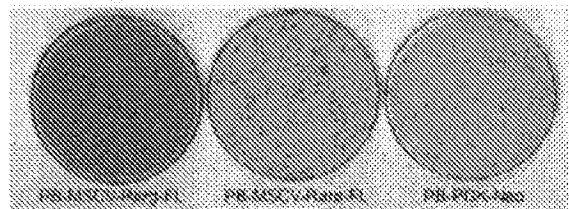
A



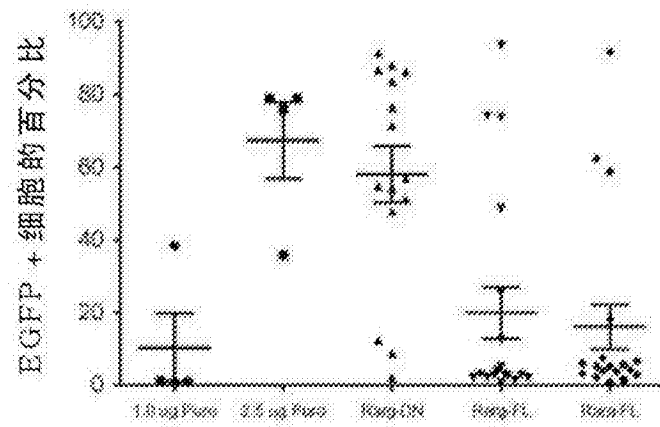
B



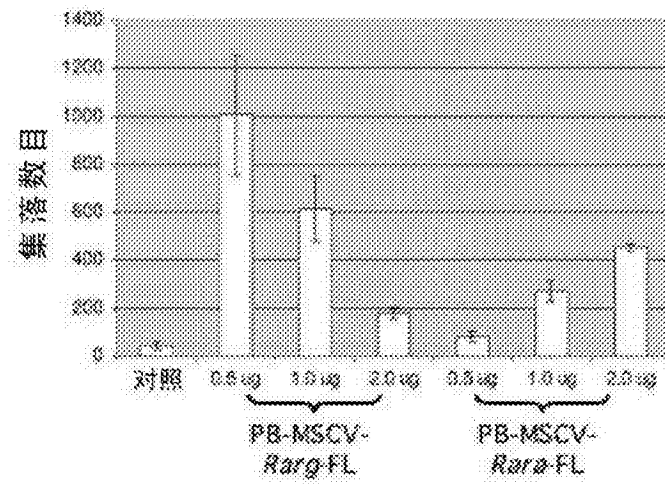
C



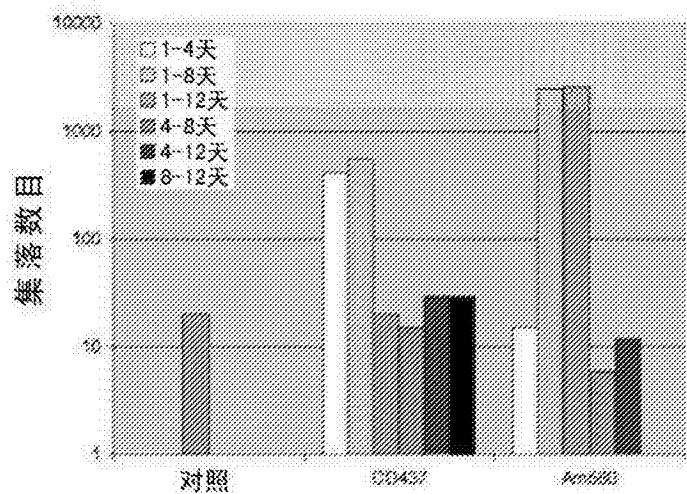
D



E



F



G

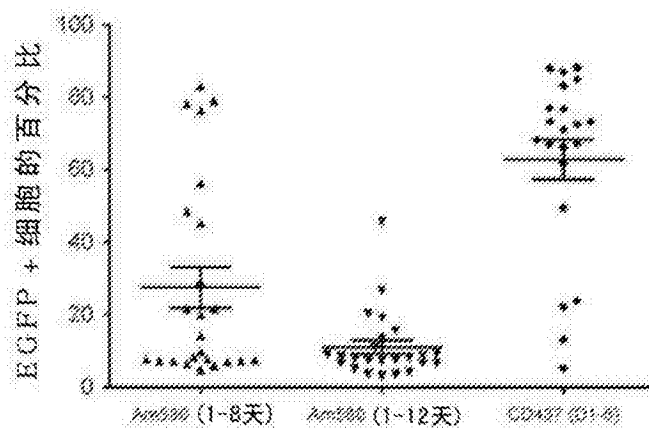
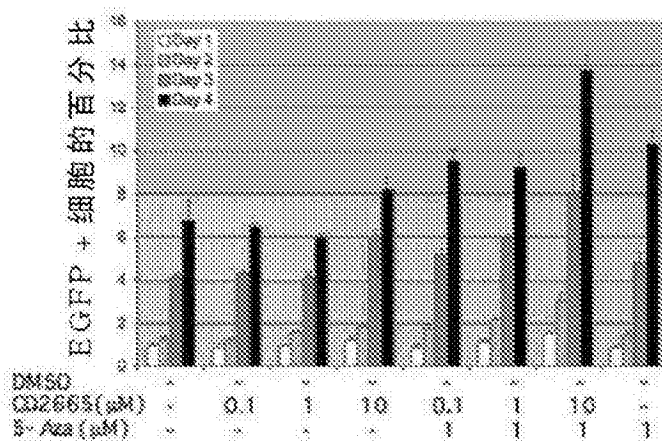
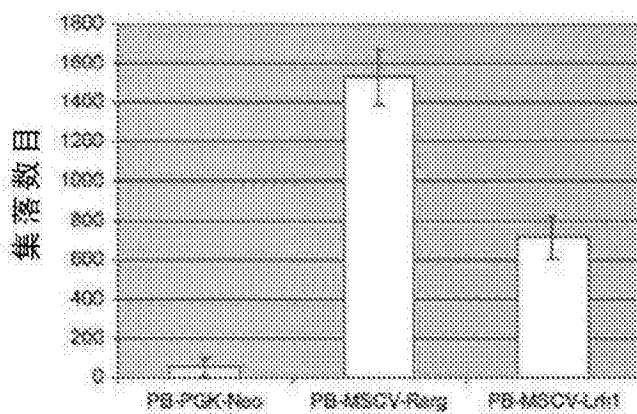


图13

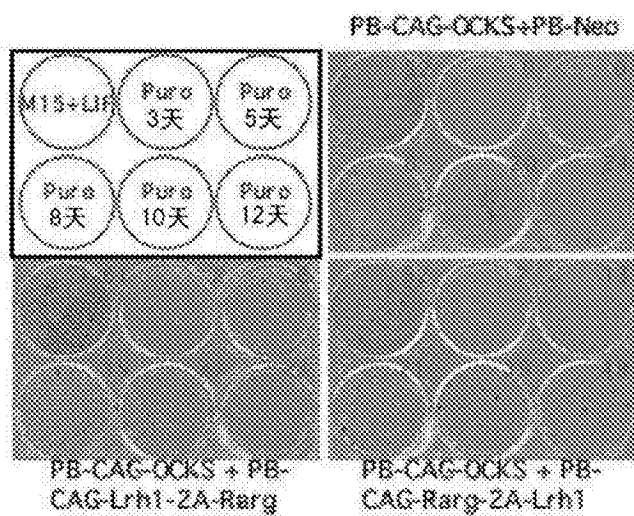
A



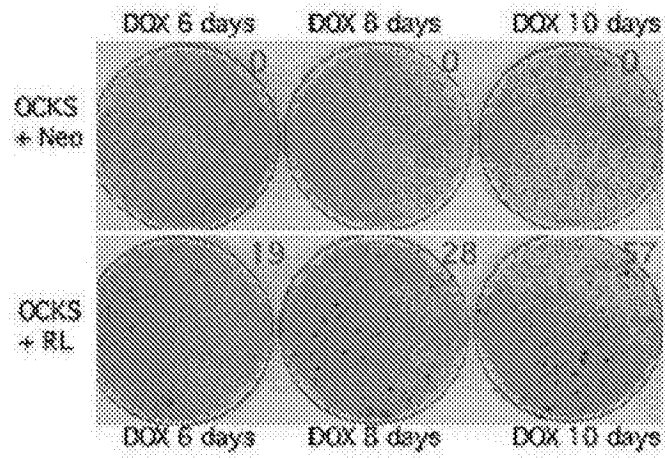
B



C



D



E

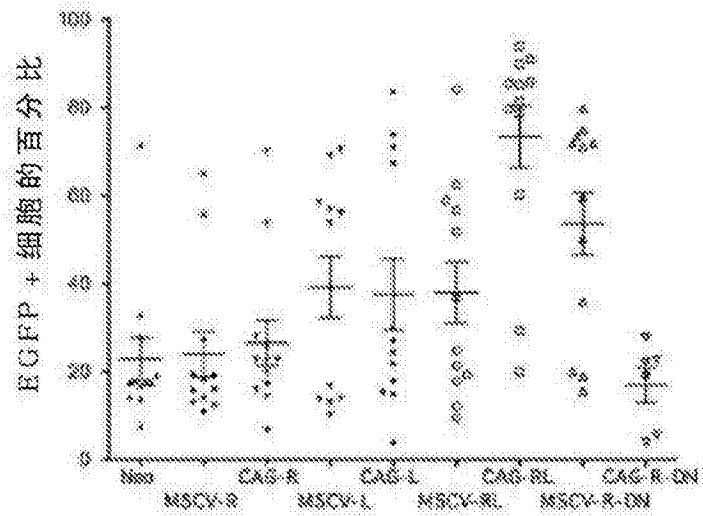
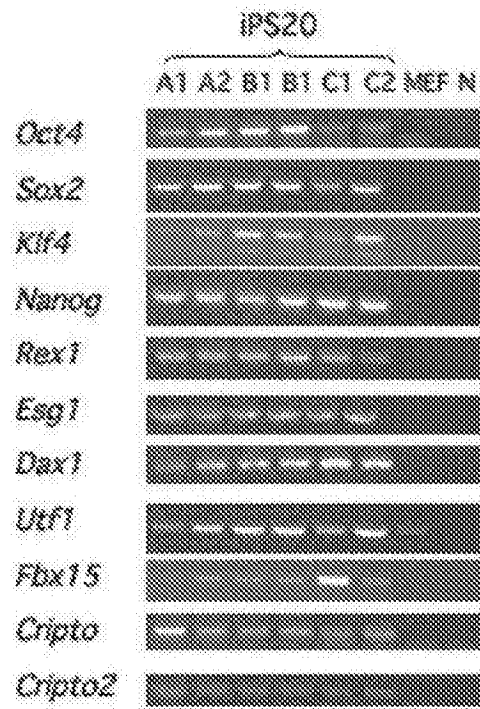


图14

A



B

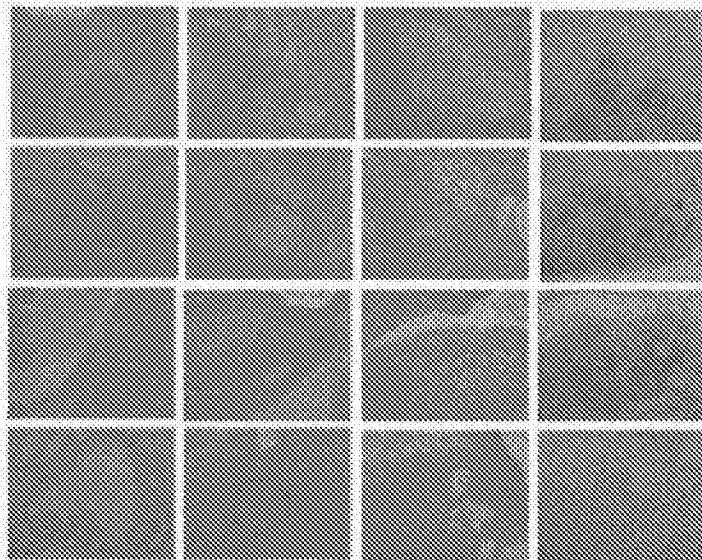
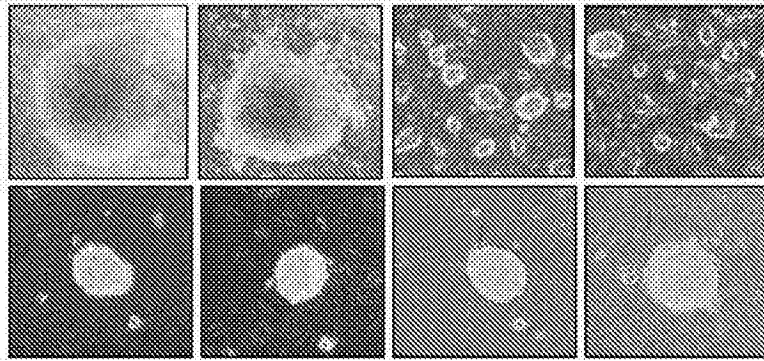
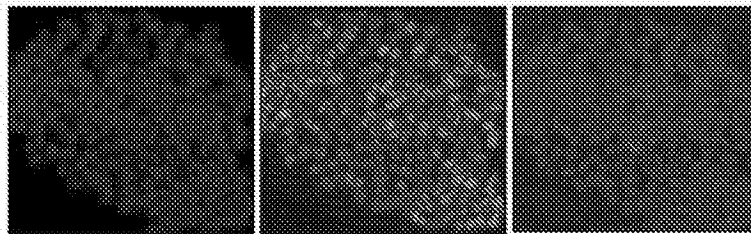


图15

A



B



C

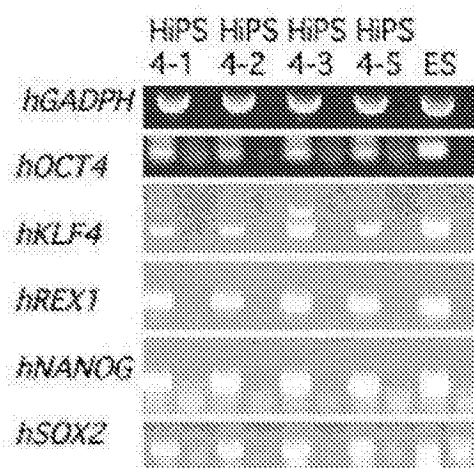


图16