

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷



[12] 发明专利申请公开说明书

[21] 申请号 03802055.6

C12N 15/00

A01K 67/027

C12N 15/90

C12N 9/22

C12N 15/85

A01K 67/033

C12N 15/10

[43] 公开日 2005 年 5 月 4 日

[11] 公开号 CN 1612931A

[22] 申请日 2003.1.9 [21] 申请号 03802055.6

[30] 优先权

[32] 2002. 1. 9 [33] US [31] 60/347,107

[32] 2002. 1. 9 [33] GB [31] 0200419.0

[32] 2002. 5. 16 [33] GB [31] 0211242.3

[86] 国际申请 PCT/GB2003/000065 2003.1.9

[87] 国际公布 WO2003/056912 英 2003.7.17

[85] 进入国家阶段日期 2004.7.9

[71] 申请人 米诺斯生物系统有限公司

地址 英国柴郡

[72] 发明人 R·克雷 C·萨瓦基斯

F·格罗斯维尔德

[74] 专利代理机构 中国专利代理(香港)有限公司

代理人 刘 玥 徐雁漪

权利要求书 5 页 说明书 48 页 附图 15 页

[54] 发明名称 遗传操作方法

[57] 摘要

本发明介绍了诱导转基因胚胎中转座的方法，该方法包括下列步骤：(a)产生第一成年转基因生物体，其基因组中包含一个或多个拷贝的转座子；(b)产生第二成年转基因生物体，其基因组中包含一个或多个拷贝的、编码与所述转座子同源的转座酶基因和/或能够调控编码转座酶的基因表达的序列；(c)将第一成年转基因生物体和第二转基因成年生物体相杂交，产生子代，导致包含下列结果在内的结果：使该子代的一个或多个细胞基因组中产生(i)一个或多个拷贝的转座子和(ii)编码与所述转座子相关连的转座酶基因，其中编码转座酶的基因位于一个或多个诱导性调控序列的控制之下，该调控序列能够使得转座酶表达；和(d)诱导所述胚胎中编码转座酶的所述基因的表达，导致子代的至少一部分组织或细胞中的所述转座子发生移动。利用该方法，可以在胚胎发育的预定期内方便地诱导所

述转座子的移动，单细胞中的突变基因可以在随后的细胞分裂中复制，从而产生与被转座基因基本同质的细胞群。

1. 产生转基因子代和诱导转座的方法，该方法包括下列步骤：
- 5 (a) 产生第一成年转基因生物体，其基因组中包含一个或多个拷贝的转座子；
- (b) 产生第二成年转基因生物体，其基因组中包含一个或多个拷贝的、编码与所述转座子相关联的转座酶的基因和/或能够调控编码转座酶的基因表达的元件；
- 10 (c) 将第一成年转基因生物体和第二转基因成年生物体相杂交，产生子代，使该子代的一个或多个细胞的基因组中包含 (i) 一个或多个拷贝的转座子和 (ii) 编码与所述转座子相关联的转座酶的基因，其中编码转座酶的基因位于一个或多个调控序列的控制之下，该调控序列能够使得转座酶表达；和
- 15 (d) 诱导所述子代中编码转座酶的所述基因的表达，导致子代的至少一部分组织或细胞中的所述转座子发生移动。
2. 权利要求 1 中要求保护的方法，其中子代是雌性转基因生物体。
3. 权利要求 2 中要求保护的方法，其中使转座酶表达的一个或多个调控序列是来源于 ZP3 调控序列的序列。
- 20 4. 权利要求 1 中要求保护的方法，其中子代是雄性转基因生物体。
5. 权利要求 4 中要求保护的方法，其中使转座酶表达的一个或多个调控序列是来源于 H1t 基因调控序列的序列。
6. 产生转基因胚胎和诱导转座的方法，该方法包括下列步骤：
- 25 (a) 产生第一成年转基因生物体，其基因组中包含一个或多个拷贝的转座子；
- (b) 产生第二成年转基因生物体，其基因组中包含一个或多个拷贝的、编码与所述转座子相关联的转座酶的基因和/或一个或多个能够调控编码转座酶的基因表达的序列；
- 30 (c) 将第一成年转基因生物体和第二转基因成年生物体相杂交，产生胚胎，所述胚胎的一个或多个细胞的基因组中包含 (i) 一

个或多个拷贝的转座子和 (ii) 编码与所述转座子相关联的转座酶的基因，其中编码转座酶的基因位于一个或多个调控序列的控制之下，该调控序列能够使转座酶表达；和

(d) 诱导所述胚胎中编码转座酶的所述基因的表达，导致胚胎的至少一部分组织或细胞中的所述转座子发生移动。

7. 根据权利要求 1 至 6 中任意一项的方法，其中所述的第二转基因生物体的基因组中包含编码关联性转座酶的基因。

8. 根据权利要求 1 至 6 中任意一项的方法，其中所述的第一转基因生物体的基因组中包含一个或多个拷贝的转座子及编码关联性转座酶的基因，第二转基因生物体包含诱导性转座酶表达所必需的一个或多个拷贝的调控序列。

9. 根据权利要求 8 的方法，其中所述的转座子和编码关联性转座酶的所述基因以单一构建体的形式提供到第一生物体的基因组中，转座子移动时编码转座酶的基因遭到破坏。

10. 根据前面权利要求中任意一项的方法，其中的转座子和/或编码转座酶的基因位于染色质开放元件中，使得转座子和/或编码转座酶的基因和/或调控编码转座酶的基因表达的基因位于开放染色质结构中。

11. 根据权利要求 10 的方法，其中染色质开放元件为普遍作用的染色质开放元件 (UCOE)、基因座控制区 (LCR)、CpG 岛或隔离子。

12. 根据前面权利要求中任意一项的方法，其中的调控序列是发育控制序列，可在发育的预定阶段活化。

13. 根据权利要求 1 至 12 中任意一项的方法，其中的调控序列是诱导性调控序列，选自四环素诱导性表达系统、雌激素诱导性表达系统、蜕皮素诱导性表达系统和 lac 操纵子阻遏系统。

14. 根据前面权利要求中任意一项的方法，其中所述转座子移动之后或移动的结果是，因基因切除而使编码转座酶的基因的表达消失。

15. 根据前面权利要求中任意一项的方法，其中转座子和/或编码转座酶的基因插入或靠近表达的基因。

16. 根据前面权利要求中任意一项的方法，其中转座子是 2 型转座子。
17. 根据权利要求 16 的方法，其中转座子是 Minos、mariner、Hermes、piggyBac 或 Sleeping Beauty。
- 5 18. 根据权利要求 17 的方法，其中转座子是 Minos。
19. 根据前面权利要求中任意一项的方法，其中对转座子加以修饰使之包括异源核酸序列，旁侧为与转座子同源的反向末端重复。
20. 根据权利要求 19 的方法，其中转座子包含编码选择标记的核酸序列。
- 10 21. 根据权利要求 20 的方法，其中选择标记是荧光或发光多肽。
22. 根据前面权利要求中任意一项的方法，其中对编码转座酶的基因的核苷酸序列加以修饰，优化了密码子利用率。
23. 产生转基因生物体的方法，其中的转基因生物体具有与由转座子移动所修饰的基因同质的多个细胞或组织，该方法包括产生转
- 15 基因胚胎，根据前面权利要求中任意一项的方法诱导胚胎中的转座。
24. 检测和表征转基因生物体中基因突变的方法，该方法包括下列步骤：
- (a) 根据权利要求 1 至 22 中任意一项的方法产生转基因胚胎，诱导胚胎中的转座；
- 20 (b) 鉴定转基因胚胎或由其发育形成的后代中表现出变异表型特征的多个细胞的存在；
- (c) 检测一个或多个所述细胞的基因组中一个或多个转座子转座事件的位置；和
- (d) 将转座事件的位置和观察到的变异表型相关联，转座事件
- 25 的位置为与所观察到的变异表型相关的一个或多个遗传性基因座的位置的指标。
25. 分离和转基因动物表型特征相关的基因的方法，该方法包括下列步骤：
- (a) 根据权利要求 1 至 22 中任意一项的方法产生转基因胚胎，
- 30 诱导胚胎中的转座；
- (b) 鉴定转基因胚胎或由其发育形成的后代中表现出所述表型

特征的多个细胞的存在;

(c) 检测一个或多个所述细胞的基因组中一个或多个转座子转座事件的位置; 和

(d) 克隆包含插入物的遗传性基因座。

5 26. 分离转基因动物中增强子的方法, 该方法包括下列步骤:

(a) 根据权利要求 1 至 22 中任意一项的方法产生转基因胚胎, 诱导胚胎中的转座, 其中转座子包含位于最小启动子控制下的报道基因, 使得该基因在基础水平上表达;

10 (b) 评价转基因胚胎或由其形成的后代的一个或多个细胞或组织中的报道基因的表达水平;

(c) 鉴定和克隆一个或多个所述细胞或组织中的遗传性基因座, 与基础表达水平相比, 所述基因座中报道基因的调节增加或降低; 和

(d) 表征所述细胞或组织中克隆的遗传性基因座。

15 27. 分离转基因动物中外显子的方法, 该方法包括下列步骤:

(a) 根据权利要求 1 至 22 中任意一项的方法产生转基因胚胎, 诱导胚胎中的转座, 其中转座子包含缺失翻译起始序列但包括剪接受体序列的报道基因;

20 (b) 鉴定胚胎或由其形成的后代中表达报道基因的一个或多个细胞或组织; 和

(c) 从所述细胞或组织中克隆包含表达的报道基因的遗传性基因座。

28. 鉴定转基因动物中对刺激物产生应答的基因的方法, 该方法包括:

25 (a) 根据权利要求 1 至 22 中任意一项的方法产生转基因胚胎, 诱导胚胎中的转座, 其中转座子包含位于最小启动子控制下的报道基因, 使得该基因在基础水平上表达;

(b) 在无刺激物的条件下, 评价转基因胚胎或由其形成的后代的一个或多个细胞或组织中的报道基因的表达水平;

30 (c) 提供刺激物;

(d) 鉴定和克隆一个或多个所述细胞或组织中的遗传性基因

座，与基础表达水平相比，在刺激物刺激后，所述基因座中报道基因的调节增加或降低。

29. 根据权利要求 28 的方法，进一步包括额外的步骤：

(e) 表征所述细胞或组织中克隆的遗传性基因座。

- 5 30. 产生转基因生物体文库的方法，所述转基因生物体具有在发育的预定阶段诱导的细胞群中的基因突变，该方法包括根据权利要求 1 至 22 中任意一项的方法产生多个转基因胚胎，诱导胚胎中的转座，其中步骤(d)在胚胎发育的预定阶段进行。

31. 调节生物体中基因表达的方法，该方法包括：

- 10 (a) 根据权利要求 30 产生转基因生物体文库；和
 (b) 从所述文库中选择一个或多个转基因生物体，其中目的基因的表达受到一个或多个转座子转座事件的调节。

32. 根据权利要求 31 的方法产生的转基因生物体文库。

遗传操作方法

5 发明领域

本发明设计使用转座子(transposons)在生物体中传递遗传信息的方法。具体地说,它涉及在发育的预定期诱导细胞遗传改变的方法。

背景技术

10 高通量(high through-put) DNA 测序技术及复杂的数据捕捉和计算机分析的发展使人们完成了包括果蝇和人类的全基因组测序。已经鉴定到了没有相关生物学功能的、新的“预测性”基因序列。要阐明哪些基因是人类疾病治疗和诊断的治疗性靶点,首先需要了解其功能信息。

15 目前,鉴定个体基因功能及其与疾病状态的功能关系是生物技术和医药工业的当务之急。鉴定疾病相关基因将促进新药或药物发现靶点的发展,提供疾病的诊断性或预示性标志物,指导医生开处方。其中后者在具有复杂遗传因素的疾病中将尤为有用。当患者间遗传的差异可以测定时,可以建立个人用药程序,鉴定特定患者对药物作用的
20 应答。

有许多鉴定基因功能的方法正在应用,多数依赖于健康和疾病状态中基因结构和基因表达图谱的对比分析。这类途径花费大、耗时,结果经常是主观性的,缺少将基因表达变化和功能性疾病相关的体内事件相联系的有力证据。基因功能的证实需要在动物模型中进行研究,这些动物模型系统将原因(即基因序列中突变,缺失或插入)直接
25 与整体动物中的可检测效应(即行为、发育、代谢等)相联系。

小鼠和其他哺乳动物中的基因功能研究目前仅限于:

A) 在来源于胚胎干细胞(ES 细胞)的“敲除”转基因小鼠中对个体基因进行辛苦的突变分析,其中胚胎干细胞通常包含通过病毒
30 感染导入的一个或多个标记基因。

B) 通过烷基化试剂进行体内随机突变,随后进行全基因组序

列分析来鉴定多个突变。

敲除方法是有效的，该方法根据与已知功能基因的序列同源性来推测其功能，但是该方法耗时、耗力。

- 5 烷基化方法完全依赖于鉴定突变位点及对前面所测定行为特点和代谢数据的变化进行分类的全基因组测序。随后须将鉴定到的表型和目标小鼠基因组中的可能上百种烷基化事件中的一个事件相关联。该方法也非常耗时，需要产生并维持大的小鼠文库，并且局限于近交的小鼠系（比较性评论参见 Abuin et al. (2002) TIB 20, 36-42）。

获得突变的另一个方法是将外源 DNA 导入基因组中。

- 10 转座子是能够从物种基因组中的一个位置跳跃或转座到另一个位置上的天然遗传元件。转座子的移动(mobilisation)依赖于转座酶的表达，转座酶能够与转座子 DNA 旁侧的序列相结合，导致 DNA 从基因组中的一个位点上切除，并重新插入到基因组中的其他位置上。向基因序列中的插入会导致基因功能的改变，这种改变反过来可导致整个生物体表型的可检测性改变。

- 20 在三个“经典”的模型动物即蝇、虫和小鼠中，已经在黑腹果蝇和线虫中建立了基于转座子的有效插入方法。转化了果蝇遗传，果蝇中 *P* 元件介导的转基因和插入诱变的导入[Spradling & Rubin (1982) Science 218, 341-347]转变了果蝇的遗传特征，为其他真核生物中类似方法学的建立提供了范例。但是，*P* 元件的宿主范围非常有限，因此过去十年里，采用了其他的元件作为各种复杂的真核生物中基因转移和/或诱变的载体，这些真核生物包括线虫、植物、哺乳动物、斑马鱼之类的鱼和鸟。

- 25 从 *D. hydei* 中分离出来的 Minos 为 2 型转座子，是 Tc1 元件家族的成员。Minos 已经用于黑腹果蝇、*C. capitata* 和史氏按蚊的胚系(germ line)转化 (Loukeris, T.G. et al (1995) Proc Natl Acad Sci USA, 92, 9485-9; Loukeris, T.G. et al (1995) Science, 270, 2002-5; Catteruccia F. et al. (2000) Nature 405:959-962)，通过瞬时移动分析，发现它在黑腹果蝇、埃及伊蚊、史氏按蚊和家蚕的胚胎及黑腹果蝇、埃及伊蚊、冈比亚按蚊和草地夜蛾的细胞系中具有活性(Catteruccia, F. et al (2000) Proc Natl Acad Sci USA, 97, 2157-2162; Klinakis et al (2000) EMBO

Reports 1:416-421; Shimizu et al. Insect Mol Biol 2000 Jun; 9(3):277-81)。

Gelbart WM, Blackman RK, Prog Nucleic Acid Res Mol Biol (1989): 36:37-46 中对黑腹果蝇的 *hobo* 元件作了介绍。

- 5 Ivics et al (1997) Cell 91, 501-510 和 Horie et al (2001), Proc. Natl. Acad. Sci. USA, 98(16):9191-9196 中介绍了 Salmonid 型转座子, 如 *Sleeping Beauty* (SB) 转座子, 一个从鱼重新构建的 Tc1/mariner 样转座元件。

- 10 *Mairiner* 是最初从 *Drosophila mauritiana* 分离的转座子, 但是后来在多种无脊椎动物和脊椎动物中也发现了该元件。国际专利申请 WO99/09817 中描述了如何使用 *maieiner* 转化生物体。

Hermes 来源于普通家蝇。美国专利 5,614,398 中介绍了其在产生转基因昆虫中的应用, 该专利整体形式在此引用作为参考。

- 15 *PiggyBac* 是来源于杆状病毒宿主 *Trichoplusia ni* 的转座子。Handler et al., (1998) PNAS (USA) 95:7520-5 和美国专利 6,218/185 中介绍了其在 *Medfly* 胚系转化中的应用。

欧洲专利申请 0955364 (Savakis et al., 其内容在此引入作为参考) 介绍了 *Minos* 转化细胞、植物和动物的应用。另外还介绍了包含一个或多个 *Minos* 插入的转基因小鼠的产生。

- 20 国际专利申请 WO99/07871 介绍了 *C. elegans* 的 Tc1 转座子转化 *C. elegans* 和人细胞系的应用。

黑腹果蝇中的果蝇 P 元件在增强子捕获和基因标记中的应用也有介绍, 参见 Wilson et al., (1989) Genes Dev. 3:1301; Spradling et al., (1999) Genetics 153:135。

- 25 在优先领域里介绍的技术中, 认为使用用于诱导转座子跳跃 (或转座) 的关联转座酶是必要的。其中描述的转基因动物具有以顺式或反式形式提供的转座酶, 例如通过与转座酶基因共转化来提供。

- 30 转座元件介导的标准转化方法是通过将两种质粒的混合物注射到胚盘形成前的胚胎中, 两种质粒中的其中一种表达转座酶 (辅助因子), 但是不能转座; 另外一种携带有旁侧为元件 (供体) 反向末端重复的目的基因。通过检测显性标记基因的表达来发现受注射动物已

转化的子代。

PCT/EP01/03341 (WO 01/71019) 介绍了如何使用转座元件来产生转基因动物。根据该方法, 通过将转基因生物体相杂交来提供转座酶的功能; 其中一个生物体提供转座子功能, 另一个生物体提供转座酶功能, 以便产生所需细胞或组织中包含转座子和转座酶的生物体。使用组织特异性染色质开放结构域以组织特异性的方式指导转座酶活性, 在体细胞组织中产生多个独立的转座事件 (参见 Zagoraiou et al (2001) P.N.A.S. 98 11474-11478)。

转座作用可以加以“标记”, 使得能够快速检测到复合体基因组中的位置变化, 并通过测序分析测定旁侧基因。这可以在原因 (即特定基因或调控元件中的插入事件) 和效应 (即可检测表型的改变) 之间建立直接的联系。但是, 通过转座作用诱导基因修饰的传统方法具有不利因素: 发生转座的组织是含独特转座作用的个体细胞的嵌合体。这样的结果是难以对转座事件所导致的表型结果进行分析, 因为每个转座事件都是独特的。因而, 可以看出控制转座事件、从而在大量细胞中提供相同基因修饰的方法将对本技术领域具有很大的贡献。

发明概述

根据本发明的第一方面, 提供了产生转基因子代和诱导转座的方法, 该方法包括以下步骤:

- (a) 产生第一成年转基因生物体, 其基因组中包含一个或多个拷贝的转座子;
- (b) 产生第二成年转基因生物体, 其基因组中包含一个或多个拷贝的、编码与所述转座子相关联(cognate)的转座酶基因和/或能够调控编码转座酶的基因表达的元件;
- (c) 将第一成年转基因生物体和第二转基因成年生物体相杂交, 产生子代, 使该子代的一个或多个细胞基因组中包含 (i) 一个或多个拷贝的转座子和 (ii) 编码与所述转座子相关联的转座酶的基因, 其中编码转座酶的基因位于一个或多个诱导性调控序列的控制之下, 该调控序列能够使得转座酶表达; 和
- (d) 诱导所述子代中编码转座酶的所述基因的表达, 导致子代的

至少一部分组织或细胞中的所述转座子发生移动。

换句话说，本发明提供了通过转座子移动产生转基因子代的方法，该方法包括以下步骤：

- (a) 产生子代，使该子代的一个或多个细胞基因组中包含 (i) 一个或多个拷贝的转座子和 (ii) 编码与所述转座子相关联的转座酶的基因，其中编码转座酶的基因位于一个或多个诱导性调控序列的控制之下，该调控序列能够使得转座酶表达；和

(b) 诱导所述子代中编码转座酶的所述基因的表达，导致子代的至少一部分组织或细胞中的所述转座子发生移动。

- 适当地说，第一成年转基因生物体可以包含稳定整合有“静止”转座子的转基因系。可以使用标准的 ES 细胞技术产生这种转基因系。可以通过与第二成年转基因生物体杂交来诱导静止转座子的转座。相应地，本发明提供了一种能够快速产生上千种突变子代如小鼠突变体的方法。

- “子代”指第一转基因生物体和第二转基因生物体之间的繁殖产物。

在一个优选的实施方案中，允许转座酶表达的一个或多个调控序列是能够使转座酶在胚系发育过程中特异性表达的序列。相应地，子代的胚细胞发生了转座事件。

- 图 13 所示为利用雌性动物提供转座子和转座酶的体内转座作用的示意图。转座发生在卵母细胞中。将雌性动物和雄性动物杂交后获得突变体。

- 相应地，在一个实施方案中，子代是上述第一转基因生物体和第二转基因生物体之间杂交产生的雌性转基因生物体。在该实施方案中，允许转座酶表达的一个或多个调控序列是能够使转座酶在雌性子代卵子发生过程中特异性表达的序列。因而，卵子发生过程中诱导了转座酶的表达。该事件反过来在卵母细胞中引起胚系转座事件的发生，从而产生具有插入序列的卵母细胞。

- 在该实施方案中，允许转座酶表达的一个或多个调控序列来源于卵母细胞发育过程中表达基因的调控序列。适当的调控序列包括那些控制卵母细胞基因，如 Zp3、Zp1、Zp2、Gdf9、Bmp15、Figla 和 Mater

表达的序列（例如,可参见 Rajkovic and Matzuk, *Molecular and Cellular Endocrinology*, 187(2002), 5-9）。其他的适当调控序列可以来源于 Oct-4 的调控序列。

图 14 所示为利用雌性动物提供转座子、雄性动物提供转座酶的体内转座作用示意图。转座发生在精子中。

相应地,在另一个实施方案中,子代是上述第一转基因生物体和第二转基因生物体之间杂交产生的雄性转基因生物体。在该实施方案中,允许转座酶表达的一个或多个调控序列是能够使转座酶在雄性子代精子发生过程中特异性表达的序列。因而,精子发生过程诱导了转座酶的表达。该事件反过来引起在精母细胞中胚系转座事件的发生,从而产生具有插入序列的精母细胞。

在该实施方案中,允许转座酶表达的一个或多个调控序列来源于精母细胞发生过程中表达基因的调控序列。适当的调控序列包括那些控制精母细胞特异性 mRNA 如 H1t 基因转录物表达的序列（Bartell et al. *Biol of Reproduction* 2000; Aug; 63(2);409-16）。

适当地,使胚系中发生了转座事件的子代相交配,产生能够鉴定其转座事件的后代。具有胚系转座作用的子代可以和正常配对物(mate)相交配,或和自身具有胚系转座作用的配对物相交配。

在本发明一个进一步的实施方案中,“子代”是胚胎。相应地,在该实施方案中,提供了产生转基因胚胎和诱导转座的方法,该方法包括下列步骤:

(a) 产生第一成年转基因生物体,其基因组中包含一个或多个拷贝的转座子;

(b) 产生第二成年转基因生物体,其基因组中包含一个或多个拷贝的、编码与所述转座子相关联的转座酶的基因和/或能够调控编码转座酶的基因表达的元件;

(c) 将第一成年转基因生物体和第二转基因成年生物体相杂交,产生子代,使该子代的一个或多个细胞基因组中包含 (i) 一个或多个拷贝的转座子和 (ii) 编码与所述转座子相关联的转座酶的基因,其中编码转座酶的基因位于一个或多个调控序列的控制之下,该调控序列能够使得转座酶表达; 和

(d) 诱导所述胚胎中编码转座酶的所述基因的表达，导致胚胎的至少一部分组织或细胞中的所述转座子发生移动。

换句话说，本发明提供了通过转座子移动来产生转基因胚胎的方法，该方法包括以下步骤：

- 5 (a) 产生一种胚胎，使该胚胎一个或多个细胞的基因组中包含 (i) 一个或多个拷贝的转座子和 (ii) 编码与所述转座子相关联的转座酶的基因，其中编码转座酶的基因位于一个或多个调控序列的控制之下，该调控序列能够使得转座酶表达；和

- 10 (b) 诱导所述胚胎中编码转座酶的所述基因的表达，导致胚胎的至少一部分组织或细胞中的所述转座子发生移动。

这里所述的“胚胎”可以理解为从单一的受精卵或合子发育至出生或脊椎动物或无脊椎动物的孵出或植物的萌发时的结构。因而，在本发明的范围内，“胚胎”应该还包括哺乳动物胎儿。

- 15 图 15 所示为使用雌性动物提供转座子、雌性动物提供转座酶的体内转座作用示意图。转座在卵细胞或早期胚胎中发生。

- 可以在胚胎发育的任何时间诱导胚胎细胞或组织中转座子的移动，例如在胚胎发育阶段的预定期诱导转座。通过在发育的这种阶段诱导转座，单细胞中的突变基因可以在随后的细胞分裂中复制，产生转座基因基本同质(homogeneous)的一个或多个细胞群。因而，转座的基因可以存在于特定组织或组织群的某些或全部细胞中。本发明因此能够产生转基因胚胎和生物体，这种转基因胚胎和生物体包含一个或多个对于个体突变为同质的克隆细胞群。

- 25 因此，本发明的第二方面提供了产生转基因生物体的方法，其中转基因生物体中的多个细胞或组织对于通过转座移动加以修饰的基因来说是纯合的，该方法包括产生一个转基因胚胎，并根据本发明的第一方面中的方法诱导转基因胚胎中的转座。

- 30 染色质结构域对于转录复合体来说是可接近的，并且在胚胎发育以及成年生命过程和异常生长情况如肿瘤的不同时间、不同细胞组织类型中，转座事件很可能是不同的，因此本发明的方法通过在发育的不同时间调控转座作用还增加了基因组范围转座的可能性。

可以使用染色质开放结构域，例如作用无处不在的染色质开放元

件(UCOEs)(PCT/GB99/02357(WO 0005393))、基因座控制区(LCRs)
(Fraser, P. & Grosveld, F. (1998). Curr. Opin. Cell Biol. 10, 361-365)、
CpG 岛(CpG islands)或隔离子(insulators)来控制转座子和/或编码转座
酶的基因的表达, 进一步增加在特定基因组区域中发生转座的可能
5 性。

在本发明的优选实施方案中, 将转座子和/或编码转座酶的基因
和/或调控编码转座酶的基因表达的序列掺入染色质开放结构域中,
增加胚胎靶组织中发生转座的可能性, 从而产生组织中转座事件同质的
细胞群。例如, 当需要在特定的组织中诱导转座时, 将转座子和/
10 或编码转座酶的基因和/或调控编码转座酶的基因表达的序列掺入基
因座控制区中, 实现对组织中转基因表达的组织特异性控制。当不存在
特异性 LCR 而需要诱导组织中的转座、和/或需要诱导特定组织中
早期转座时, 将转座子和/或编码转座酶的基因和/或调控编码转座酶
的基因表达的序列掺入 UCOE 中。在本发明的具体优选实施方案中,
15 将转座子和编码转座酶的基因都掺入到了染色质开放结构域中。这样
有力地增强了胚胎发育过程中整个基因组中特定基因座上的转座效
率, 从而产生特定转座事件同质的一个或多个细胞群。

因此, 本发明在第三方面中进一步提供了产生转基因胚胎和诱导
转座的方法, 该方法包括以下步骤:

20 (a) 产生第一成年转基因生物体, 其基因组中包含一个或多个
拷贝的转座子;

(b) 产生第二成年转基因生物体;

(c) 将第一成年转基因生物体和第二转基因成年生物体相杂
交, 产生胚胎, 使该胚胎的一个或多个细胞基因组中包含 (i) 一个
25 或多个拷贝的转座子和 (ii) 编码和所述转座子相关联的转座酶的基
因, 其中编码转座酶的基因位于一个或多个诱导性调控序列的控制之
下, 该调控序列能够使得转座酶表达; 和

(d) 诱导所述胚胎中编码转座酶的所述基因的表达, 导致胚胎
的至少一部分组织或细胞中的所述转座子发生移动。

30 在本发明的优选实施方案中, 通过将第一生物体和第二生物体相
杂交, 产生胚胎。其中第一生物体包含一个或多个拷贝的转座子, 第

二生物体是一个转基因生物体，该生物体基因组中包含一个或多个拷贝的、编码转座酶的可调控基因。在一个可替代的实施方案中，通过将第一生物体和第二生物体相杂交，产生胚胎。其中第一生物体是转基因生物体，包含一个或多个拷贝的转座子和编码关联性转座酶的基因；第二生物体包含一个或多个拷贝的、为转座酶表达所必需的调控元件。

在一个优选的实施方案中，转座子和编码转座酶的基因可以以单一构建体的形式提供，使转座子移动时能够破坏编码转座酶的基因，从而限制转座子的进一步移动。可以将一个转座子的反向重复序列置于内含子中来达到这种目的。其中内含子以一定的方向中断转座酶基因，使转座酶基因在转座子移动时遭到破坏。该载体使我们能够采用单一的杂交步骤来产生包含调控因子、转座酶基因和转座子的转基因生物体。另外，转座作用导致转座酶的完全失活，使新的插入物即使在诱导物存在的情况下也能保持稳定。图1示意性地说明了这种载体在本发明中的应用。

还可使用 Cre/lox 功能的掺入（其细节在 Sauer, Methods of Enzymology; 1993, Vol.225, 890-900 中综述）和不同转座子/转座酶组合来消除主要的转座酶功能。但是，在本发明的进一步实施方案中，转座时不破坏转座酶的基因，例如可施用诱导物使转座子能够进一步移动。

在本发明的方法中，可以使用熟练技术人员已知的任意系统诱导转座。通过应用内源性物质或通过实施内源性信号，如发育调控信号来诱导转座酶基因表达，从而诱导转座。

编码受控制的转座酶基因的一个或多个调控序列可以是诱导性调控序列。例如，适当的诱导系统包括基于 tet 的系统、lac 操纵子-阻遏子系统、基于蜕皮素的系统和基于雌激素，下文中提供了细节。外源诱导物能够以任意方便的方式提供，如通过向母性动物或胚胎中注射，或者以食品或水添加剂的形式提供给母体动物。可以在胚胎发育过程中的一个或多个时间上诱导转座。因此，可以在发育的一个或多个阶段中仅施用一次诱导物，或重复使用。

在本发明的可替代实施方案中，胚胎发育中的特定阶段生成的基因调控信号可以诱导编码转座酶基因的表达。可以将编码转座酶的基因置于基因调控序列，如发育调控序列或启动子的控制之下，来实现这种控

制。举例来说, 这些发育调控序列或启动子包括对特定基因调控信号产生应答的发育调控特异启动子, 如瞬时表达的发育调节蛋白。当编码转座酶的基因位于这种控制之下时, 编码转座酶的基因表达仅在基因调控信号产生时发生, 例如发生在生成瞬时表达的发育调控蛋白时; 或者, 5 可替代的是, 发生在这种信号蛋白导入胚胎时, 例如通过注射或喂养母体动物导入胚胎时。

使用本发明的方法, 可以控制编码转座酶基因的表达时间, 从而控制诱导胚胎中转座的时间。

10 在需要严密控制转座酶基因表达的持续时间和效率、进一步限制转座时间的实施方案中, 编码转座酶的基因可以和转座子存在于同一个构建体中, 在转座子移动时编码转座酶的基因遭到破坏, 阻止进一步转座酶的生成, 从而限制进一步的转座。

可以通过选择诱导时间在胚胎发育的预定阶段诱导转座子的移动。例如, 可以在发育的非常早期, 如合子阶段、四细胞胚、64 细胞胚胎等阶段, 或发育的更晚阶段诱导转座。 15

在一个实施方案中, 需要将转座酶置于一个或多个调控序列的控制之下来诱导转座, 其中的调控序列将驱动早期受精卵如二细胞阶段或四细胞阶段受精卵中转座酶基因的表达。控制转座酶在该阶段表达的适当调控序列可以来源于该阶段中表达活化的基因调控序列。这些 20 基因包括 Zp3、Oct-4 (Kirchhof et al. Biol Reprod 2000, Dec; 63(6):1698-705) 和母体效应基因, 如 Zp1、Zp2、Gdf9、Bmp15、Figla 和 Mater。其他的适当调控序列包括 hsp70.1 (Bevilacqua et al. Development 2000; Apr; 127(7):1541-51)。

根据发育阶段的不同, 发生了转座的细胞可以进行分裂, 经过进一步的几轮细胞分裂, 产生初始转座事件同质的细胞群。在多次诱导转座的情况下, 细胞群对于两个或多个转座事件中每个转座事件来说是同质的。因此, 根据发育阶段的不同, 插入事件可以存在于一个或多个组织、完整组织或组织群的细胞群中。插入事件的精确性质将决定它是否影响某些或全部胚胎组织或成年组织中的功能基因表达。因此, 可以在胚胎发育过程中监测成年细胞和组织中修饰基因的基因表达模式。 30

另外，在来源于干细胞的快速生长的成年细胞和组织中，典型的是在组织再生过程中或细胞和组织培养过程中，产生初始转座事件同质的类似细胞群。这种细胞和组织包括，但不局限于消化道衬里、肝脏和血液的细胞，这些细胞能够快速更新和/或从成年动物的干细胞再生。类似的，可以在肿瘤中产生转座事件纯合的细胞群。本发明的方法可以加以变通，在这些成年细胞中产生初始转座事件纯合的细胞群。

因此，在本发明的进一步实施方案中，可以对本发明中第一或第三方面的步骤(d)加以变通，诱导新生儿、青年或成年生物体中编码转座酶的基因的表达，使转座子在生物体的组织或细胞的至少一部分中转座子的移动，来代替诱导胚胎中转座酶的表达，或者优选的是，作为诱导胚胎中转座酶表达的附加手段。因此，本发明的方法能够产生在新生儿、青年或成年生物体发育的预定阶段中诱导的转座事件同质的细胞群。在这种实施方案中，编码转座酶的基因优先位于基因座控制区的控制之下，以实现转基因的组织特异性控制。例如，在需要诱导肝脏细胞中转座的情况下，转座子和/或编码转座酶的基因可以位于肝脏细胞中表达相关的基因座控制区的控制之下。

在合子的早期发育阶段诱导转座事件的情况下，有可能建立发生转座作用的ES细胞系。这些ES细胞系可以测序、分析并保存待用。

根据本发明转基因生物体中进行转座插入所产生的遗传突变，可引起生物体中新的表型变异。使用本发明的方法，可以产生转基因胚胎，其中一个或多个细胞簇或组织对于不同的转座事件是同质的，各转座事件可具有或不具有表型效应。因此，与产生的没有发生插入事件的相应细胞或组织表型相比，使用本发明的方法发育而来的胚胎或成年动物包含一个或多个细胞簇或细胞群，每个细胞群表现出一个表型变异。

当然，转座事件对转基因胚胎的影响在某种程度上取决于转座发生的发育阶段。例如在单合子阶段诱导转座时，胚胎从中发育的所有细胞对于插入事件来说都是同质的。因此，如果转座事件，例如插入事件产生导致各细胞死亡的表型变化，就不会发育形成胚胎。当在发育的较晚期诱导转座事件时，各插入事件存在于源自发生了转座事件

的细胞分裂的细胞簇或组织中。从而使每个细胞都表现出相同的表型变异。例如，如果一个细胞中诱导了转座事件，一个特定组织或特定器官的某些或所有细胞都来源于该细胞，插入事件的表型结果则局限于细胞、特定组织或特定器官。在一个细胞中诱导了转座事件、一个特定组织或特定器官中仅某些细胞都源于该细胞的情况下，转座事件所导致的表型变化将比组织中所有细胞都发生转座事件所导致的表型变化小。当转座事件使细胞致死时，细胞将不能存活。如果存在于组成特定组织或器官的所有细胞中，该组织或细胞将不具有功能，或不发育，并且胚胎不能存活。换句话说，转座事件可具有非致死的表型结果。例如，转座事件具有调节受影响细胞中酶功能的作用，产生与未受影响细胞相比代谢方面的相对变化。这可以导致产生一种器官，如肝脏，其中具有差异表型的组织部分与具有正常表型或第二差异表型的相同组织部分相邻。因此，生物体中差异表型的分布依赖于诱导转座的胚胎发育阶段。另外，在本发明的一些实施方案中，诱导“第二轮”转座可有益于检测元件的切除所致的表型转换，或者，更重要的是检测相互作用基因中新的插入所导致的表型修饰。

转基因生物体的细胞、组织或器官的表型变化可以追溯到那些细胞、组织或器官中的转座事件。

相应地，本发明的第四方面提供了检测和鉴定转基因生物体中遗传突变的方法，该方法包括以下步骤：

(a) 根据本发明的第一或第三方面中的方法产生转基因胚胎并诱导其发生转座，或者根据本发明的第二方面中的方法产生转基因生物体；

(b) 鉴定转基因胚胎或从中发育形成的后代中表现为不同表型的多个细胞的存在；

(c) 检测一个或多个所述细胞的基因组中一个或多个转座子转座事件发生的位置；将转座事件的位置和观察到的不同表型相关联，转座事件的位置是与所观察不同表型相关的一个或多个遗传性基因座的位置。

“转座事件”是转座子移动所致的基因组序列改变，包括插入事件、切除事件或染色体断裂。

通过探测转座子的核酸序列，筛选转座子的存在来检测插入事件。也可以通过切除后剩下的“标记”序列来鉴定切除。

本发明的第五方面提供了分离与转基因动物的多个细胞中表型特征相关的基因的方法，该方法包括下列步骤：

5 (a) 根据本发明的第一或第三方面中的方法产生转基因胚胎，诱导胚胎中的转座作用，或根据本发明的第二方面中的方法产生转基因生物体；

 (b) 鉴定转基因胚胎或从中发育形成的后代中表现为所述表型特征的多个细胞的存在；

10 (c) 检测一个或多个所述细胞的基因组中一个或多个转座子转座事件发生的位置；

 (d) 克隆包含插入的遗传性基因座。

 可以对转座子插入进行定位来精确地鉴定发生了修饰的基因座。对旁侧区进行测序就能够鉴定数据库中的基因座而可能不必对基因座进行
15 测序。

 在本发明的优选实施方案中，转座子可以是天然的转座子。优选2型转座子，如 Minos。最有优势的是，它是 Minos。可选择的转座子包括 *mariner*、*Hermes*、*piggyBac*、*hobo* 和 *salmonid* 型转座子，如 *Sleeping Beauty*。

20 本发明中还可以使用修饰的转座子，其引入包含一个或多个异源编码序列和/或表达控制序列。这种编码序列可以包括可选择及/或不选择的标记基因，它们可以促进基因组中转座子的鉴定和整合有转座子的基因座的克隆。适当的标记包括荧光和/或发光多肽，如 GFP 及其衍生物、荧光素酶、 β -半乳糖苷酶或氯霉素乙酰转移酶 (CAT)。

25 例如，通过插入包含标记基因的转座子，可将这种标记用在体内增强子或沉默子捕获和外显子捕获中，其中的标记基因的表达水平受邻近增强子或外显子的调节。EP 0955364 中介绍了用于外显子和增强子捕获的构建体。使用本发明的方法，转座事件同质的多个细胞或组织可以表现出调节标记基因的表达，从而能够有效地捕获增强子和
30 /或沉默子和/或外显子。另外，在仅有特定类型组织的一部分细胞对于转座事件为纯合情况的实施方案中，通过和不表现出这种调节的相

同转基因动物中相同类型细胞或组织相比较，鉴定对标记基因的调节。

相应地，本发明在第六方面中提供了分离转基因动物中增强子或沉默子的方法，该方法包括下列步骤：

- 5 (a) 根据本发明的第一或第三方面中的方法产生转基因胚胎并诱导其转座，或者根据本发明的第二方面中的方法产生转基因生物体，其中转座子包含处于最小启动子控制下的报道基因，使该基因具有基础水平的表达；

- 10 (b) 评价转基因胚胎或从中发育形成的后代的一个或多个细胞或组织中的报道基因表达水平；

 (c) 鉴定和克隆一个或多个所述细胞中与基础表达水平相比报道基因的调节增加或降低的遗传性基因座；和

 (d) 标准所述细胞或组织中克隆的遗传性基因座。

- 15 在第七方面中，提供了分离转基因动物中外显子的方法，该方法包括下列步骤：

 (a) 根据本发明的第一或第三方面中的方法产生转基因胚胎并诱导其发生转座，或者根据本发明的第二方面中的方法产生转基因生物体，其中转座子包含缺失翻译起始序列但包括剪接受体序列的报道基因；

- 20 (b) 鉴定转基因胚胎或从中发育所形成后代中表达报道基因的一个或多个细胞或组织；和

 (c) 从所述细胞或组织中克隆包含表达报道基因的遗传性基因座。

- 25 图 2、3 和 4 示意性地显示了能够用于产生本发明中用到的胚胎的基因捕获构建体。

- 图 2 显示了位于诱导性 TetO 启动子控制下的转座酶构建体和包含位于最小启动子(minimal promoter)控制之下、编码自体荧光蛋白(AFP)的标记基因的转座子。在包含两个构建体的转基因胚胎中，转座酶的表达可以为诱导性 TetO 启动子的活化所诱导，从而实现转座子构建体的转座。通过检测标记基因的表达，鉴定向基因组中增强子位点处或增强子位点附近的插入。
- 30

图 3 说明了位于诱导性 TetO 启动子控制下的转座酶构建体和包含缺失翻译其始序列、包括剪接受体序列的 AFP 荧光报道基因的转座子。在包含两个构建体的转基因胚胎中，转座酶的表达可以为诱导性 TetO 启动子的活化所诱导，从而实现转座子构建体的转座。通过检测标记基因的表达，鉴定向内含子中以适当取向的插入。

在一个优选实施方案中，转座子可用于上调基因的表达。例如，可以对转座子加以修饰，使之包括一个增强子或其他转录活化元件。这种转座子在基因附近的移动或插入上调基因或基因座的表达。该实施方案在分离癌基因发明尤其具有优势，克隆肿瘤中的癌基因可通过转座子的定位加以鉴定。

图 4 显示用于产生本发明该方面所用胚胎的基因活化系统。图 4 说明了位于诱导性 TetO 启动子控制下的转座酶构建体和同样位于诱导性 TetO 启动子控制下的 AFP 荧光报道基因的转座子。在包含两种构建体的转基因胚胎中，转座酶的表达可以为诱导性 TetO 启动子的活化所诱导，从而实现转座子构建体的转座。另外，如果包含 AFP 的转座构建体插入到了异位基因的上游，则该基因可以活化，TetO 启动子诱导时可以观察到基因活化的表型效应。

在通过转座作用诱导基因修饰的传统方法中产生了细胞嵌合体，其中各细胞具有独特转座作用诱导的基因修饰，难以对缘于转座事件的表型进行研究。类似的，天然和人工刺激物对转座细胞的影响也难以进行研究，对研究结果也难以进行解释。但是，相比之下，本发明的方法的产生转基因胚胎或生物体中一个或多个细胞簇或组织对于单一转座事件是同质的，因此通过将同质细胞簇中的报道基因表达和周围细胞或组织中的报道基因表达相比，易于观察药用或天然刺激物的作用。因而，本发明的方法可以用于研究发生了转座事件的细胞对刺激物(stimuli)的应答，刺激物为天然刺激物，如生理性刺激物，例如激素、细胞因子和生长因子；或人造刺激物，如药物发现方法、毒理学研究等中的药物。事实上，本发明的方法能够实时观察细胞簇或组织对刺激物的应答。

相应地，本发明的第八方面中提供了鉴定转基因动物中对刺激物的基因应答的方法，该方法包括：

(a) 根据本发明的第一或第三方面中的方法产生转基因胚胎，诱导胚胎中的转座作用，或者根据本发明的第二方面中的方法产生转基因生物体，其中转座子包含位于最小启动子控制下的报道基因，该基因在基础水平上表达；

5 (b) 在不存在刺激物的条件下，评价转基因胚胎或从中发育形成的后代的一个或多个细胞或组织中的报道基因表达水平；

(c) 提供刺激物；

(d) 鉴定和克隆一个或多个所述细胞中的遗传性基因座，与基础表达水平相比，受刺激物的刺激后报道基因的调节增加或降低。

10 该方法可进一步包括额外的步骤：

(e) 表征所述细胞或组织中克隆的基因座。

因此本发明的该方法可以用在分子介入新靶点的鉴定中，这些新靶点包括人类、植物或动物疾病治疗的靶点、杀虫剂、除草剂、抗真菌剂和抗细菌剂开发的靶点。

15 一个进一步应用是发现负责发病（例如小鼠疾病模型中）的基因。如果一个基因（如激酶或受体）的活化参与疾病的发病过程，则 50% 基因失活就有可能缓解或逆转疾病的一个或多个表型。因此，在疾病背景下，插入灭活一个或多个拷贝这种基因的转座子的细胞群将检测为健康的细胞簇。

20 转座子可以插入到基因中。优选将转座子插入到转录的基因中，使所述转座子定位于开放染色质中。转座子旁侧可以伴有染色质开放结构域元件，如提供组织特异性表达的基因座控制元件（Fraser, P. & Grosveld, F. (1998). Curr. Opin. Cell Biol. 10, 361-365）或提供非组织特异性表达的普遍作用 (ubiquitously-acting) 的染色质开放元件 -
25 (UCOEs)（例如参见 WO 0005393）。可以用在本发明方法中的其他染色质开放元件包括 CpG 富含岛，正常情况下它与管家基因或组织特异性基因或隔离子相关。

另外，转座子自身可以在转座子末端之间包含染色质开放结构域。这导致整合有转座子的染色质活化，促进诱导性转座酶以细胞或
30 组织特异的方式靠近染色质。

类似地，转座酶构建体可以包含，或旁侧带有染色质开放结构域

元件。

胚胎发育和成年生命过程中调控转座的能力增加了发育的不同时间、不同组织中存在的多个染色质结构域中发生转座事件的可能性。

本发明的方法可以有利地用于产生基因修饰生物体文库，其中每个基因修饰生物体中一个或多个细胞群或组织对于发育预定阶段转座移动所导致的基因修饰是同质的。因此，在本发明的一个进一步方面中，提供了一种方法生成转基因生物体文库，其中每个基因修饰生物体中一个或多个细胞群或组织对于转座移动所导致的基因修饰是同质的。该方法包括根据本发明的第一或第二或第三方面的方法用转座子移动来修饰细胞，其中步骤(d)在胚胎发育的预定阶段进行。

通过这种方法生成的转基因生物体文库是本发明的一个进一步方面。

附图简述

图 1 示意性地说明自身灭活自主转座子构建体在本发明的一个实施方案中的应用。第一转基因生物体 (A) 和第二转基因生物体 (B) 相杂交，其中 A 基因组中包含一个调控因子构建体，B 基因组中包含具有目的基因的转座子。转座子的一个反向重复序列位于打断转座酶基因的内含子中。诱导杂交子代 (C) 发生转座子移动时，转座酶基因遭到破坏，从而稳定目的基因的转座作用，使之即使在诱导物存在的条件下也不再发生进一步的转座事件。

图 2 示意性地说明用于本发明增强子捕获方法中的编码转座酶的构建体和转座子，编码转座酶的构建体中转座酶基因位于诱导性 TetO 启动子的控制之下，转座子所包含的 AFP 荧光标记基因位于最小启动子的控制之下。

图 3 说明用于本发明外显子捕获方法中的转座酶构建体和转座子，其中转座酶构建体中转座酶基因位于诱导性 TetO 启动子的控制之下，转座子所包含的 AFP 荧光标记基因缺失翻译起始序列，但包括剪接受体序列。

图 4 示意性地说明了用于基因捕获以鉴定异位基因的转座酶构建体和转座子。

图 5 显示 Minos 来源的载体 pMiCMVGFP。Minos 反向末端重复以粗的黑箭头表示。这些箭头外面的白色方框表示 *D. hydei* 基因组中原始 Minos 元件的旁侧序列。箭头指向用于检测 Minos 切除的引物位置。小箭头指出了 GFP 和转座酶基因的转录方向。黑杠表示用作

5 探针的片段。

图 6 说明了用于本发明实施例 4 所述的一个实施方案中 Minos 转座酶表达盒的结构。ZP3 基因的 6.5 kb 5' 旁侧区和启动子与 Minos 转座酶 cDNA (ILMi) 和人 β 珠蛋白的第二内含子以及聚腺苷酸化位点相连接。图中标出了相关的限制性酶位点。

10 图 7 显示用 RT-PCR 分析检测 Minos 转座酶的转录。图中标出了进行分析的不同组织。小鼠次黄嘌呤磷酸核糖转移酶 (HPRT) 用作内部对照。pUC18 DNA MspI 消化物用作标记物 (M)，阴性对照中使用 H_2O 。

图 8 说明发生转座事件的不同后代的 Southern 印迹分析。所有的

15 DNA 样品都用 *Bgl* II 消化，用 ^{32}P 标记的 GFP 探针进行探测。泳道 1 是对照（染色体 14 上带有多拷贝的 GFP 转座子、在卵母细胞发育中表达 Minos 转座酶的雌性小鼠）。泳道 2、5、6、7、8、10 和 11 对应于双转基因雌性动物和野生型雄性动物的子代。它们均代表不同的转座事件。泳道 2 和 5：具有 2 个转座事件的小鼠。泳道 3 和 4 对应于泳道 2 中具有 2 个转座事件的小鼠后代，表现出了插入物的分离（泳道 3 为 2 号染色体上，泳道 4 为 14 号染色体靠近着丝粒部位）。泳道 9：泳道 8 中所示小鼠产生的后代，表现出了分离。

20

图 9 显示小家鼠基因组中亲本和四个不同 Minos 插入物的序列。新插入的转座子旁侧的染色体序列以大写字母表示，转座子序列以小写字母表示，靶位点重复以红色表示。标出了插入物的染色体定位和 Celera 数据库中的支架号。

25

图 10 显示 Minos 转座事件的 FISH 分析。如实施例 4 中材料和方法所述用 4',6-二氨基-2-苯基吲哚对染色体进行染色，用 GFP 探针进行探测。A 组：具有两个转座事件的小鼠 8218-01（在 2 号染色体上和 14 号染色体的着丝粒部位；参见图 8 泳道 2）；红色（箭头所指）：Tyramide 扩增后的染色（参见实施例 4 中材料与方法）。B 组：具

30

有两个转座事件的小鼠 8218-02（一个在 18 号染色体上，一个在 14 号染色体的着丝粒附近，靠近转座子串联体的起始位置；参见图 8 泳道 5）。绿色（箭头所指）：FITC 染色。黄色箭头所指为转座事件。

图 11 显示用于在精子或卵中特异性表达或普遍性表达转座酶的敲入（A）和常规构建体（B）。在图 11 中，加了 β 珠蛋白 3'第二外显子（红色）、介入序列（绿色）和第三外显子（红色）。

图 12 显示转基因小鼠和/或克隆入逆转录病毒质粒中的捕获构建体(trap construct)。

图 13 是说明使用雌性动物提供转座子和转座酶的体内转座作用示意图。转座作用发生在卵母细胞中。

图 14 是说明使用雌性动物提供转座子，雄性动物提供转座酶的体内转座作用示意图。转座发生在精子中。

图 15 是说明使用雄性动物提供转座子，雌性动物提供转座酶的体内转座作用。转座发生在受精卵或胚胎中。

发明详述

虽然这里提到的技术一般都是本技术领域中的已知，但具体可参考 Sambrook et al., 分子克隆：实验指南（1989）和 Ausubel et al., 分子生物学中的简短方案（1999）4th 版，John Wiley & Sons 公司。

转座子

本发明的方法中可以使用任意转座子。优选 2 型转座子，更优选的是选自 *Minos*、*mariner*、*Hermes*、*piggyBac* 和 *Sleeping Beauty* 转座子。有利的是，转座子是 *Minos*。每个转座子及其天然的相关转座酶一起使用是有利的，但也可以考虑使用修饰和/或改进的转座酶。*Minos* 转座子及其相关转座酶在美国专利 5,840,865 和欧洲专利申请 EP 0955364 中作了详细介绍。

转座子优选包含编码异源多肽的核酸序列。在转座子整合时，该序列将和转座子一起整合到细胞的基因组中。另外，当后来切除、再次移动时它和转座子一起切除。在一个优选的实施方案中，异源多肽为选择性标记物。这样就能够鉴定整合有转座子的细胞，确切地绘制

整合位点。

标记基因

优选的标记基因包括编码荧光多肽的基因。例如，本发明中可以使用刺丝胞动物的绿色荧光蛋白（“GFP”），作为生物发光中的能量转移受体发挥作用。这里所用的绿色荧光蛋白是发出绿色荧光的蛋白质，蓝色荧光蛋白是发出蓝色荧光的蛋白质。已经从西北太平洋水母 *Aequorea victoria* 中、从海洋三色堇 *Renilla reniformis* 中、从 *Phialidium gregarium* 中分离到的 GFP（Ward et al., 1982, Photochem. Photobiol., 35:803-808; Levine et al., 1982, Comp. Biochem. Physiol., 72B:77-85）。最近还从珊瑚虫中分离到了荧光蛋白（登记号 AF168419、AF168420、AF168421、AF168422、AF168423 和 AF168424）。

已经通过对 *Aequorea victoria* 来源的天然发生 GFP 进行氨基酸序列修饰，获得了多种具有有用激发和发射光谱的水母相关 GFP（Prasher et al., 1992, Gene, 111:229-233; Heim et al., 1994, Proc. Natl. Acad. Sci. U.S.A., 91:12501-12504; PCT/US95/14692）。举例来说，水母相关的荧光蛋白包括野生型 *Aequorea Victoria* GFP，其核苷酸序列和推测的氨基酸序列的 Genbank 登记号为 L29345、M62654、M62653，还包括其他人工改造版本的水母相关绿色荧光蛋白，上文中列出了其中一些这样的绿色荧光蛋白。这些绿色荧光蛋白中，P4、P4-3、W7 和 W2 的荧光波长明显比野生型的短。

可以使用的其他标记基因包括选择标记基因，如编码新霉素、嘌呤霉素或潮霉素的基因，或者是对应的选择基因，如胞嘧啶脱氨酶基因或硝酸还原酶基因。

本领域的熟练技术人员都了解可以使用的大量标记基因。可以使用任意的适当标记基因，对于本发明来说，没有必要进行特意的选择。

插入和切除事件的鉴定

可以通过测序分析鉴定转座子和转座子被切除的位点。例如，Minos 典型地整合在 TA 碱基对部位，切除时产生两个靶 TA 序列，

位于转座子末端四个核苷酸的旁侧。可以通过测序、PCR 和/或杂交等技术检测该序列或相关序列的存在。

可利用类似的技术，如使用和末端重复序列相互补的 PCR 引物鉴定插入的转座子。

5

转座酶

有效的转座子移动依赖于转座元件自身向宿主细胞的有效输送和有效关联性转座酶的存在，以催化转座子跳跃。这里所用的“关联性”转座酶是能够有效活化转座子转座的任意转座酶，转座子的转座包括
10 转座子从第一整合位点的切除和/或在第二整合位点处的整合。优选的关联性转座酶是和自然界中体内情况下的转座子天然相关的转座酶。但是，本发明也包含修饰的转座酶，这些转座酶具有本发明范围内的优越改进活性。例如，编码转座酶的基因序列可以加以修饰，优化密码子利用率(usage)，从而增加转座频率。优化密码子利用率是本
15 领域所熟知的增加给定基因表达水平的方法。可替换的方案是，转座酶可以包含一个或多个氨基酸的插入、替换或缺失，以增在加宿主生物体中的活性。

编码转座酶的基因可以存在于第二生物体中，通过与基因组中包含转座子的第一生物体相杂交，产生用在本发明方法中的胚胎。在一个可代替的实施方案中，一个或多个拷贝的转座子和编码关联性转座
20 酶的基因存在于第一生物体中，将该生物体和第二生物体相杂交来产生胚胎，其中第二生物体包含一个或多个拷贝的、允许诱导性转座酶表达所需的调控元件。

将基因导入宿主细胞基因组中的许多方法都是本领域中已知的，
25 可以用在本发明中。例如，可以通过转基因手段将转座酶基因插入宿主细胞基因组。下面会进一步讨论这些方法。

转座酶表达的调控

根据需要，将编码转座酶的克隆序列可操作性地和调节转座酶表达的调控序列相连接。和编码转座酶的序列可操作性相连接的控制序
30 列包括启动子/增强子和其他表达调控信号。选择的这些控制序列和

要表达转座酶的宿主生物体相容。术语启动子是本领域所熟知的，从大小和复杂性方面来说所包含的核酸区域涵盖了最小启动子和包括上游元件和增强子的启动子。

启动子典型地选自在一些细胞类型中发挥作用的启动子，这些细胞类型和正被讨论的生物体，或生物体所属的属、科、目、界或其他分类同源，但是异源启动子也发挥作用，如一些原核启动子在真核细胞中具有功能。启动子可以来源于病毒或真核基因的启动子序列。例如，它可以是来源于将发生表达的细胞基因组的启动子。至于真核启动子，它们可以是以普遍(ubiquitous)的方式(如 α -肌动蛋白、 β -肌动蛋白、微管蛋白的启动子)，或者，可替换的方案是以组织特异性方式(如丙酮酸激酶基因的启动子)发挥作用的启动子。在胚系转座事件的产生过程中，启动子可以来源于在配子形成，卵子发生或精子发生过程中诱导表达的基因。可替换的方案是，对于诸如发育发育过程中转座作用之类的发育性调控转座事件来说，启动子可以来源于其表达受发育调控的基因。对于早期合子中的表达，可以使用来源于母性效应基因的启动子。它们还可以是对特定刺激物产生应答的启动子，例如和甾体类激素受体相结合的启动子。还可以使用病毒启动子，例如莫洛尼鼠白血病病毒长末端重复(MMLV LTR)启动子、肉瘤病毒(RSV) LTR启动子或人巨细胞病毒(CMV) IE启动子。

根据本发明，编码转座酶的基因位于一个或多个调控序列的控制之下，也就是说所获得的表达水平可以调控如使用启动子。例如，调控序列可以是诱导性的调控序列。基因表达的诱导系统是本领域中已知的，包括四环素、蜕皮素和雌激素诱导系统或lac操纵子-阻遏子系统。

广泛应用于哺乳动物细胞的这类系统是与可被四环素-多西环素阻遏的转录活化因子 tTA 联用的 tetO 启动子-操纵子，也称为 Tet-Off 基因表达系统(Gossen, M. & Bujard, H. (1992) 四环素应答启动子对哺乳动物细胞中基因表达的严紧控制. Proc. Natl. Acad. Sci. USA 89:5547-5551)，或多西环素诱导性 rtTA 转录活化因子，也称为 Tet-On 系统(Gossen, M., Freundlieb, S., Bender, G., Muller, G., Hillen, W. & Bujard, H. (1995)哺乳动物细胞中四环素引起的转录活化. Science

268:1766-1769)。

在 Tet-Off 系统中，四环素 (Tc) 或多西环素 (doxycycline, Dox; Tc 衍生物) 从培养基中去除时，基因开始表达。相比之下，在 Tet-On 系统中，通过加入 Dox 使基因开始表达。以前曾经描述了建立带有
5 转录活化基因和染色体中稳定整合有 Tet 调控基因的细胞系的方法。例如，参见 <http://www.clontech.com/techinfo/manuals/PDF/PT3001-1.pdf>。例如，可以采用 Tet-On 系统实现转基因动物中 Minos 转座酶的四环素诱导性表达。

可以通过常规的同源重组 ES 细胞技术产生双转基因动物。使用
10 了两个构建体：其一是包含位于组成性启动子控制之下的 rtTA 基因的构建体。这种构建体的一个例子是 pTet-On 质粒 (Clontech)，它包括编码 rtTA 活化因子的基因，该基因位于巨细胞病毒即早 (CMV) 启动子控制之下。该构建体编码的 rtTA 转录活化因子仅在多西环素存在的条件下活化。第二构建体包含位于四环素应答元件或 TRE 控制之下的 Minos 转座酶基因。TRE 有包含 tet 操纵子 (tetO) 的 42 bp
15 序列的七个正向重复组成，位于基础 CMV 启动子的即上游，该启动子缺少正常情况下和 CMV 即早启动子相关的增强子元件。因为没有这些增强子元件，在没有 rtTA 的结合时，没有转座酶从 TRE 的“泄漏(leaky)”表达。这种构建体的一个例子是 pTRE2 质粒 (Clontech)，
20 在其 MCS 中插入了编码 Minos 转座酶的基因。在稳定转化有两个构建体的细胞中，表达 rtTA，但是如果向动物体内施用多西环素，则不会活化 Minos 转座酶的转录。

因此，当根据本发明的方法，将包含 pTet-On 和 pTRE2 构建体的转基因动物和基因组中包含一个转座子的另一个动物相杂交时，在
25 不存在多西环素的条件下，所形成的、在基因组中包含转座子和编码转座酶的基因的胚胎中不发生转座移动。因此可以使用多西环素来诱导转座作用。

可以使用任意适当的方法来施用诱导物，如多西环素。在本发明的一个优选实施方案中，通过给亲本生物体喂食或喂水来施用转座酶
30 表达诱导物。

可替换的方案是，诱导系统包括他莫昔芬诱导的转座酶[和转座

酶相偶联的修饰的雌激素受体结构域 (Indra et al., Nucl Acid Res. 27, 4324-27, 1999), 在他莫昔芬加入培养物之前使转座酶滞留在细胞质中], RU418 诱导的转座酶 (和糖皮质激素受体的操纵原理相同; 参见 Tsujita et al., J. Neuroscience, 19, 10318-23, 1999), 或雌激素诱导系统。

蜕皮素诱导系统基于异源二聚化的果蝇蜕皮素受体, 它可以为昆虫激素、蜕皮素和其衍生物所诱导。在黑腹果蝇的蜕变过程中, 通常称作“蜕皮素”的甾体类激素 20-OH 蜕皮素通过蜕皮素受体引发级联的形态变化。通过稳定表达调控优化的蜕皮素应答启动子的修饰蜕皮素受体, 将蜕皮素应答性转移到哺乳动物细胞体内。在施用激素或其衍生物时, 表达修饰蜕皮素受体的转基因生物体, 如小鼠可以活化整合的蜕皮素应答性启动子。一旦受体与蜕皮素或蜕皮素类似物 muristerone 结合, 受体就会活化蜕皮素应答性启动子, 实现对目的基因表达的控制。据报道, 和基于四环素的系统相比, 基于蜕皮素的诱导系统表现出较低的基础活性和更高的诱导性。基于蜕皮素的诱导系统其进一步细节可以参考 US6,245,531 和 No D, Yao TP, Evans RM. 哺乳动物细胞和转基因小鼠中蜕皮素诱导性基因的表达, Proc Natl Acad Sci USA 1996 Apr 93:3346-51, 其内容在此引入作为参考。

最近研究表明 lac 操纵子-阻遏子系统在哺乳动物中, 尤其在小鼠中是具有功能的。Cronin et al., Genes and Development, 15, 1506-1517 (2001)介绍了 lac 阻遏子转基因的应用, 在密码子利用率和结构上 lac 阻遏子转基因和典型的哺乳动物基因相似, 在细胞系中普遍地表达 lac 阻遏子蛋白, 控制位于 lac 启动子控制下的报道基因的表达, 其内容在此引入作为参考。使用乳糖类似物 IPTG 可以逆转报道基因的表达。其中 IPTG 在小鼠或胚胎母体或喂养幼鼠的饮水中供应。可以通过将转基因置于 lac 启动子的控制之下, 使将 lac 操纵子-阻遏子系统适用于调控转座酶的表达。

此外, 可以通过加入其他的调控序列, 例如增强子序列来修饰这些启动子。还可以使用来自上述两个或多个不同启动子的序列元件的嵌合启动子。

也考虑使用基因座控制区。LCR 能够对转基因进行严紧调控的组

织特异性控制，从而增加转基因表达的忠实性。许多 LCR 都是本领域中已知的。它们包括 β -珠蛋白 LCR (Grosveld et al., (1987) Cell 51:975-985)； α -珠蛋白 (Hatton et al., (1990) Blood 76:221-227)；和 CD2 (Festenstein et al., (1996) Science 271:1123-1125)；T 细胞特异性 CD4 (Boyer et al., J Immunol 1997, 159:3383-3390) 和 TCR 基因座 (Diaz P, et al., Immunity 1994, 1:207-217; Ortiz et al., EMBO J 1997, 16:5037-5045; Hong et al., Mol Cell Biol 1997, 17:2151-2157)；B 细胞特异性 II E α 型 MHC (Carson et al., Nucleic Acids Res 1993, 21:2065-2072)；巨噬细胞特异性溶菌酶基因 (Bonifer et al EMBO J 1990, 9:2843-2848)；神经元特异性 S100 基因 (Friend et al., J Neurosci 1992, 12:4337-4346)；肝脏特异性 LAP 基因 (Talbot et al., Nucleic Acids Res 1994, 22:756-766)；人生长激素基因座 (Jones et al., Mol Cell Biol 1995, 15:7010-7021)；和免疫球蛋白、肌肉组织等等。Fraser, P. & Grosveld, F. (1998). Curr. Opin. Cell Biol. 10:361-365 和 Li, Q., Harju, S. & Peterson, K.R. (1999). Trends Genet. 15:403-408 中介绍了 LCR 的其他细节。

可替换的方案是，体内所有细胞中需要开放或打开的基因结构域；即蛋白质为所有细胞生存所需、普遍表达的基因结构域可以用来实现转座子和/或转座酶在任意组织中的表达。这种普遍活化的染色质开放元件 (ECO) 的例子包括人基因 TBP 和 hnRNPA2。使用这种 ECO 的其他细节可以参考 Antoniou, M. and Grosveld, F. (1999). (血珠蛋白病变治疗的遗传学方法. 在：血细胞生物化学中，第 8 卷：血细胞生成和基因治疗，Fairbairn 和 Testa 版本。Kluwer Academic/Plenum 出版社，纽约，219-242 页) 和 PCT/GB99/02357 (WO 0005393)，其内容在此引入作为参考。

还可以通过使用 ES 细胞来实现对转座酶和/或转座子表达的调节。使用转化的 ES 细胞构建嵌合胚胎，有可能产生仅在某些组织中包含转座酶基因或转座子元件的转基因胚胎。这可以提供进一步的调控水平。

使转座效率最大化

如上所述, 及如 WO 01/71019 和 WO 02/062991 中所述, 通过转座酶作用于整合转座子中的末端重复序列, 使转座子从“宿主”基因组的原始位置上切除, 重新插入到基因组的另一个位置上, 从而实现转座作用。

- 5 和大多数生化方法一样, 仅提高底物的浓度就可使本方法效率更高, 得到高水平的末端重复序列, 即拷贝数增加、转座酶水平升高。

通过在原始位点上产生多拷贝排列来增加拷贝数。例如, 通过常规转基因或使用 PAC 载体来产生 10 至 100 拷贝。可替换的方案是, 通过基因组不同位点上多个插入物的存在来产生多拷贝。

- 10 可以对转座酶加以修饰, 优化密码子利用率, 进而增加转座频率。优化密码子利用率是本领域中用来增加给定基因表达水平的熟知方法。

- 因此, 将蝇密码子利用率替换为哺乳动物密码子利用率, 使 mRNA 更有效地进行翻译, 增加转座酶的浓度, 进而增加哺乳动物细胞或动物体内蝇转座酶的效率。

15 测定转座酶效率的分析方法包括 Klinakis et al., *Insect Molecular Biology*, 9(3):269-275, 2000 中所述的常规转座分析。

- 还可以将诸如编码生长因子、珠蛋白、肌动蛋白或白蛋白的大量稳定 mRNA 中发现的 5' 和 3' 序列包含转座酶 mRNA 序列内, 来增加转座酶 mRNA 的浓度。

转基因生物体

- 本发明的方法可以使用一个或多个转基因生物体, 这些转基因生物体基因组中整合有转座子、编码关联性转座酶的基因, 或二者兼而有之。

- 25 可以使用任意可用的技术导入转座子或编码转座酶的基因, 这些技术包括转化/转染、病毒载体或非病毒载体的传递及微注射。每项技术都是本领域中已知的。可以使用相同或不同的方法插入转座酶和编码转座酶的基因。例如, 使用果蝇 P 元件将 Minos 转座酶构建体导入果蝇体内。

例如, 在一个优选的实施方案中, 可使用转基因技术将转座子或

编码转座酶的基因插入宿主细胞基因组中，产生包含转座子、编码关联性转座酶或二者兼而有之的转基因动物。转基因动物同时包含转座子和编码转座酶基因的情况下，可以使用相同或不同的方法来获得这两个构建体。构建体的传递通过病毒载体介导时，可以使用包含转座子
5 和编码转座酶基因的复合载体，其中转座子和编码转座酶的基因位于诸如 Tet 操纵子之类控制序列的控制之下。此外可两者分开的载体。

本发明中可以使用任意适当的转基因动物。这些动物包括刺胞动物门、栉水母门、扁形动物门、线虫纲、环节动物门、软体动物门、
10 螯肢动物门、单肢动物门、甲壳纲和脊索动物门的动物。单肢动物门动物包括六足动物亚门的动物，这类动物包括有翅昆虫之类的昆虫。脊索动物门的动物包括脊椎动物，如哺乳动物、鸟、鱼、爬行动物和两栖动物。具体的哺乳动物例子包括非人类的灵长类、猫、犬、有蹄动物，如牛、山羊、猪、绵羊和马，以及啮齿类动物，如小鼠、大鼠、
15 沙土鼠和仓鼠。

用于本发明方法中、用来产生转基因动物的技术是本领域中所熟知的。关于这一主题，有用的通用教材是 Houdebine, 转基因动物-产生和使用 (Harwood Academic, 1997) -用于产生转基因动物的技术泛论。

20 在一个优选的实施方案中，动物为昆虫。用在本发明方法中、用来产生转基因昆虫的方法是人们所熟知的（例如参见 Loukeris et al (1995), Science 270 2002-2005）。简要地说，将携带有目的基因的转座元件插入到前胚层胚胎中，例如可使用微注射方法。优选的是，新的转基因材料置于极性细胞质部位，卵在这一部位形成新生昆虫自身的卵细胞或精细胞。核物质分裂许多次之后，大部分被分隔至边缘，
25 在那里变成昆虫身体的核。少数核迁移到极性端，成熟时形成卵细胞。如果这些细胞包含转基因，则子代就是转基因的。Loukeris et al. (1995) Science 270 2002-2005 和 O'Brochta and Atkinson (1998) Scientific American 279 60-65 中提供了产生转基因昆虫进一步的细节。

30 在另一个优选的实施方案中，动物优选为哺乳动物。胚胎微操作技术的进展使得人们能够将异源 DNA 导入诸如受精的哺乳动物卵

中。例如，可以通过微注射、磷酸钙介导的沉淀、脂质体融合、逆转录病毒感染或其他手段转化全能的或多能的干细胞，然后将转化的细胞导入胚胎，胚胎进一步发育成转基因动物。在一个非常优选的方法中，用包含所需 DNA 的逆转录病毒感染发育的胚胎，从受感染的胚胎生成转基因动物。但是，在最优选的方法中，优选在单细胞阶段将适当的 DNA 同时注射到胚胎的原核或胞浆中，胚胎进而发育成成熟的转基因动物。这些技术是人们所熟知的（参见将异源 DNA 微注射到哺乳动物受精卵中的常规实验室方法的综述，包括 Hogan et al., 小鼠胚胎操作（冷泉港出版社 1986）；Krimpenfort et al., Bio/Technology 9:844(1991); Palmiter et al., Cell, 41, 41:343(1985); Kraemer et al., 小鼠胚胎的遗传操作，（冷泉港出版社 1985）；Hammer et al., Nature, 315:680(1985); Wagner et al., U.S. Pat. No. 5,175,385; Krimpenfort et al., U.S. Pat. No. 5,175,384，将其各自的内容在此引入作为参考）。

另一个用于产生转基因动物的方法涉及包括使用常规方法将核酸微注射到原核阶段的卵中。对注射的卵进行培养，然后转移到假孕受者的输卵管中。

还可以用 Schnieke, A.E. et al., 1997, Science, 278:2130 和 Cibelli, J.B. et al., 1998, Science, 280:1256 中所述的核转移技术来产生转基因动物。使用该方法，将质粒稳定转染来自供体动物的成纤维细胞，其中质粒包含位于调控序列控制之下的目的多肽编码序列。将稳定转染体与去核的卵母细胞相融合，培养并转移到雌性受者体内。

按照常规方法进行 PCR 或 Southern 印迹分析，对包含转基因序列的动物进行分析。

通过构建转基因动物，如牛的具体实施例中方法，使用诸如 U.S. Pat. No.4,873,191 中所述的技术将包含编码结合分子的序列导入来源于刚从哺乳动物体内摘除的卵巢的卵母细胞中。将卵母细胞从卵泡中吸出，沉降，然后使用冻融的精子进行受精，用肝素使精子获能，使用 Percoll 梯度进行预分级，分离能动级分。

将受精的卵母细胞离心，如 15,000 g 离心 8 分钟，使原核显现出来以进行注射，然后在输卵管组织条件培养基中从合子培养至桑葚胚或胚囊阶段。使用从输卵管上刮取的腔组织在培养基中加以稀释来

制备这种输卵管组织条件培养基。微注射 2 小时后必须将合子置于培养基中。

- 通过施用粪甾醇使受者动物，如牛体内的发情期同步化。2 天后产生发情期，发情期 5~7 天后将胚胎转移到受者体内。通过 Southern
5 印迹法评价后代体内的成功转移。

- 可替换的方案是，将所需的构建体导入胚胎干细胞（ES 细胞），进行细胞培养，确保转基因的修饰。然后将修饰的细胞注射到囊胚期的胚胎中，将囊胚置入假孕宿主体内。产生的后代对于 ES 和宿主细胞来说是嵌合的，可以使用传统的杂交技术获得只包含 ES 子代的非
10 嵌合株。WO91/10741 中就介绍了这种技术。

输送转座子和/或转座酶基因并使之稳定整合入宿主动物基因组中的可替换方法包括病毒载体，如腺病毒载体，逆转录病毒载体、杆状病毒载体和疱疹病毒载体。此外，本领域中 Zhang 等（Nucl. Ac. Res., 1998, 26:3687-3693）也对这种技术作了介绍。

- 15 适当的病毒载体可以是逆转录病毒载体，来源于或可来源于任意适当逆转录病毒载体。例如：鼠白血病毒（MLV）、人免疫缺陷病毒（HIV）、猿猴免疫缺陷病毒、人 T 细胞白血病毒（HTLV）、马感染性贫血病毒（EIAV）、小鼠乳腺癌病毒（MMTV）、劳氏肉瘤病毒（RSV）、Fujinami 肉瘤病毒（FuSV）、莫洛尼鼠白血病毒（Mo-MLV）、FBR 鼠骨肉瘤病毒（FBRMSV）、莫洛尼鼠肉
20 瘤病毒（Mo-MSV）、埃布尔森小鼠白血病毒（A-MLV）、禽髓细胞瘤病毒-29（MC29）和禽成红细胞增多病病毒（AEV）。逆转录病毒的详细列单可以参考 Coffin et al., 1997, “逆转录病毒”，冷泉港实验室出版社，JM Coffin, SM Hughes, HE Varmus, 758-763 页。

- 25 本领域中对一些逆转录病毒基因组的结构作细节了描述。例如，HIV 和 Mo-MLV 的详情可以 NCBI GenBank 获知（基因组登记号分别为 AF033819 和 AF033811）。

- 可以将逆转录病毒广泛地分为两类，即“简单型”和“复杂型”。还可以进一步将逆转录病毒分为七组。其中的五组为具有致癌潜力的
30 逆转录病毒。其余的两组是慢病毒和泡沫病毒。这些逆转录病毒的综述见 Coffin 等 1997 (ibid)。

不同的逆转录病毒具有不同的宿主范围和组织向性(tropism)。在某些情况下,这种特异性会限制重组逆转录病毒载体的转导潜力。由于这一原因,许多基因治疗实验都使用了 MLV。具有包膜蛋白的一个具体 MLV 是称作 4070A 的两性向性病毒,由于它的包膜蛋白和人类与小鼠之间保守的磷酸盐转运蛋白“停靠(docks)”在一起,它能够感染人细胞。磷酸盐转运蛋白是普遍的,因此这些病毒能够感染许多细胞类型。

联合使用包装或辅助细胞系和重组载体使具有复制缺陷的逆转录病毒载体典型地增殖,以制备适当滴度的逆转录病毒载体用于随后的转导。也就是说,以反式形式提供三个包装蛋白。

“包装细胞系”包含一个或多个逆转录病毒的 gag、pol 和 env 基因。包装细胞系能够产生包装逆转录病毒 DNA 所需的蛋白质,但是由于缺少 psi 区,不能够进行包装。辅助蛋白质能够包装 psi 阳性的重组载体,产生重组病毒原种。该病毒原种可以用于转导细胞,将载体导入靶细胞的基因组。现有包装细胞系的概览见 Coffin 等 1997 年 (ibid)。

慢病毒可以分为“灵长类”和“非灵长类”。灵长类慢病毒包括人免疫缺陷病毒(HIV)和猿猴免疫缺陷病毒(SIV)。非灵长类慢病毒包括原型“慢病毒”绵羊脱髓鞘性脑白质炎/绵羊肺腺瘤病病毒(VMV),以及相关的羊关节炎-脑炎病毒(CAEV)、马感染性贫血病毒(EIAV)和最近描述的猫免疫缺陷病毒(FIV)和牛免疫缺陷病毒(BIV)。参考 Rovira et al., Blood. 2000; 96:4111-4117; Reiser et al., J Virol. 2000 Nov; 74(Mulder, M.P et al. (1995). Hum Genet 96(2): 133-141): 10589; Lai et al., Proc Natl Acad Sci USA 2000 Oct 10; 97(Southern, E.M. (1975). J. Mol. Biol 98, 503-517):11297-302; 和 Saulnier et al., J Gene Med 2000 Sep-Oct; 2(5):317-25。

慢病毒家族与其他类型的逆转录病毒的区别之一在于慢病毒能够感染分裂和非分裂细胞。相比之下,其他逆转录病毒,如 MLV 不能感染非分裂细胞,如那些组成肌肉、大脑、肺和肝脏组织的细胞。

已经在逆转录病毒科的多个慢病毒亚家族成员的基础上构建了许多载体,其中许多都是专利申请的主题(WO-A-98/18815; WO-A-

97/12622)。优选的慢病毒载体是在 HIV、SIV 或 EIAV 的基础上构建的。从 HIV-1 构建的最简单载体具有完整的基因组,只是缺失了 env 编码区部分或置换了 nef 编码区。值得注意的是,这些载体表达 gag/pol 和所有的辅助基因,因此仅需要一个包膜蛋白即可生成具有感染性的病毒颗粒。在辅助基因中, vif、vpr、vpu 和 nef 是非必需的。

基于 HIV 的慢病毒载体的一个优选通用格式是: HIV 5'LTR 和先导序列,一些 gag 编码区序列(提供包装功能),报道序列盒、rev 应答元件(RRE)和 3'LTR。这些载体中 gag/pol 辅助基因产物和包膜蛋白功能来自单一质粒或两个或更多共转染质粒或使用 HIV 共感染包含载体的细胞。

腺病毒是双链线形 DNA 病毒,不经历 RNA 中间体。腺病毒共有 50 种不同的人血清亚型,根据基因序列同源性分为 6 个亚组,全都表现出相当的基因组结构。在腺病毒载体系统中,人腺病毒 C 组血清亚型 2 和血清亚型 5(具有 95%序列同源性)是最为常用的,和青年人中的上呼吸道感染正常相关。

用在本发明中的腺病毒/腺病毒载体可以是人源的或动物来源的。对于人源的腺病毒载体来说,优选的腺病毒是 C 组的腺病毒,尤其是 2 型腺病毒(Ad2)、5 型腺病毒(Ad5)、7 型腺病毒(Ad7)或 12 型腺病毒(Ad12)。在动物来源的各种腺病毒中,可以使用犬腺病毒、小鼠腺病毒或禽腺病毒,如 CELO 病毒(Cotton et al., 1993, J Virol 67:3777-3785)。

举例来说,HSV 载体可以来源于 HSV1 或 HSV2 株,或其衍生物。可以使用减毒株,如 1716 株(MacLean et al., 1991, J Gen Virol 72:632-639)、R3616 和 R4009 株(Chou and Roizman, 1992, PNAS 89:3266-3270)和 R930 株(Chou et al., 1994, J. Virol 68:8304-8311),这些病毒株都具有 ICP34.5 和 d27-1 中的突变(Rice and Knipe, 1990, J. Virol 64:1704-1715),以及 ICP27 中有缺失的 d27-1。具有 VMW65 中灭活突变的、缺失了 ICP4、ICP0、ICP22、ICP6、ICP47、vhs 或 gH 的其他病毒株,或上述病毒株的任意组合也可用于产生本发明的 HSV 株。

用来描述各种 HSV 基因的术语参见 Coffin and Latchman, 1996. 基

于疱疹病毒的载体。Latchman DS(版),神经系统的遗传操作,Academic Press: 伦敦, 99-114 页。

本发明中还可使用杆状病毒载体。杆状病毒苜蓿银纹夜蛾核型多角体病毒 (AcMNPV) 是一种 DNA 病毒, 仅在某些鳞翅目昆虫的细胞中复制, 已广泛用于昆虫细胞中重组蛋白的表达。最近已使用杆状病毒 AcMNPV 将异源 DNA 高效导入各种哺乳动物细胞中, 如肝细胞系及原代肝细胞和内皮细胞 (Boyce FM, Bucher NL (1996) 杆状病毒介导的向哺乳动物细胞中的基因转移. Proc Natl Acad Sci USA 93, 2348-52; Airenne KJ, Hiltunen MO, Turunen MP, Turunen AM, Laitinen OH, Kulomaa MS, YlaHerttuala S (2000)杆状病毒介导的向兔颈动脉中的基因转移.Gene Ther 7, 1499-1504)。另外, 用于基因转移的杆状病毒、将异源 DNA 导入其基因组的方法和昆虫细胞培养物中产生重组病毒的方法都可以从商用途获得; 而且, 杆状病毒不能在哺乳动物细胞中正常复制, 因此没有必要为了该用途将其加以改造。

可以采用常规的连接技术来构建用在本发明方法中的载体。分离的病毒载体、质粒或 DNA 片段可以按所需方式加以切割、加尾, 并重新连接, 产生所需的质粒。需要的话, 用已知的方式进行分析, 确认构建的质粒中的序列正确。举例来说, 可以通过常规的 Southern 印迹法、点杂交、PCR 或原位杂交, 使用根据转座子中存在的序列设计的适当标记探针, 直接测定转座子的存在和/或移动。本领域的熟练技术人员可以考虑如何根据需要来改进这些方法。本发明中用到的载体最好带有标记基因, 以便如上所述的转座子鉴定以及定位。

发明用途

本发明的方法能够产生转基因胚胎或生物体, 它们包括对一个或多个突变同质的一个或多个克隆细胞群。因此, 转基因胚胎和动物可以在细胞簇、一种组织或多种组织、一个器官或多个器官中产生, 这些组织和器官都具有相同的基因修饰。许多细胞、组织或器官中相同基因修饰的存在便于分析修饰的表型和基因型, 并使得通过和相应的野生型基因或相同生物体的同一类型细胞、组织或器官中的其他基因修饰相比较, 了解具体基因修饰的效应。

本发明还可用于监测修饰基因的基因表达模式，可在胚胎发育期间以及成年细胞和组织中监测。

可用序列分析来鉴定转座子及其从中切除的位点。例如，Minos 典型地整合在 TA 碱基对处，切除后留下一个足迹，该足迹由两个靶 TA 序列组成，位于转座子四个末端核苷酸的旁侧。可以通过测序、PCR 或杂交之类的技术来检测 Minos 序列、其足迹或相关序列的存在。

可用类似的技术，如使用和末端重复序列互补的 PCR 引物来鉴定插入的转座子。

10 本发明通过对发育的预定阶段进行精确的基因调控，并使用转座子进行检测，可以绘制基因组的功能谱。因此，本发明提供了发育早期阶段转基因生物体细胞中的外显子捕获功能性，实现了有效的转座子移动，以及向生物体的一个或多个细胞或组织基因组中的插入。在胚胎发育的这些阶段诱导转座子的移动能够产生转座基因为同质的细胞群或组织。这样就能够进一步快速、有效地检测表型的变化和/或鉴定修饰的基因。

图 12 中给出了一个适当的捕获构建体实例。

20 在一个有利的实施方案中，本发明可以将基因加以标记，用于细胞群或组织中的功能性遗传分析，或通过转座子插入将基因加以敲除，然后通过转座子“标签”进行特异性鉴定，不需要昂贵而耗时的遗传分析，且经常不需要大量的测序。

25 本发明的另一个实施方案产生诸如转基因小鼠的转基因生物体文库。根据一个或多个细胞、组织或器官的表型从表型上鉴定靶基因，并通过测定转座子插入位点进行遗传学鉴定。上述的诱导性表达系统可用于调控基因的部分敲除和反义诱导的完全敲除之间的开关。将带有转座子插入的体细胞永生化，如通过常规方法产生永生细胞系，或通过核植入方法产生转基因动物系。

这种文库可用于表型分析和基因相关性的鉴定。本发明的方法优于目前的方法。

30 在血珠蛋白病变、血友病、囊肿纤维症和肌肉失养症之类遗传性疾病中，单一基因或其调控元件中的明确突变产生异常的基因表达，

导致临床疾病，极大程度地影响受影响个体的生活方式和预期寿命。

但是，在其他疾病中，遗传因素更为复杂，对其知之甚少。尤其是和衰老相关的疾病，如痴呆和精神病；精神病学的紊乱，如精神分裂症和躁狂抑郁症；与骨和关节相关的炎症；肥胖、胰岛素抵抗、2
5 型糖尿病和相关的血管和心血管病情（参见 Lander and Schork, Science(1994) 265, 2037-2048）。

不同靶点中的多个突变决定疾病发生时间（如绝经期后的胰岛素抵抗）和疾病状态严重程度（如血管疾病和中风的事先倾向）之类的因素，因此对饲养的小鼠系进行随机烷基化，从发生单一基因缺失的
10 大的 ES 细胞文库费劲地形成转基因小鼠的策略必将会失败。

本发明提供了一个更具有吸引力的策略，该策略从已知易于发生某种疾病状态的背景出发，快速产生随机“标记”的小鼠系，因此有可能具有导致相关疾病之基因中的现有突变。使用在已知整合位点处带有多个“显性”转座子的建立者细胞系来选择背景。可调控的转座
15 酶活性能够快速产生具有不同遗传背景的小鼠文库，反映为不同的疾病模型。

例如，研究表明胰岛素受体中的突变导致轻至重度的高胰岛素血症，其程度取决于小鼠的遗传背景（参见 Kido(2000) Diabetes 49,589-596），说明胰岛素抵抗事先有就其他基因的参与。在“轻度
20 表型”小鼠背景下，使用体内基因转座技术产生的这种带有标签的小鼠文库应能使一些动物具有更严重的表型。对新的转座事件旁侧 DNA 进行测序分析，鉴定导致出现严重表型的新候选致病基因。

因此，候选疾病基因成为进一步研究的焦点，以明确它们在动物模型中的确切作用并证明其在人体内的疾病相关作用。

25 人类中靶点的证实将使用现有的临床数据库和基因数据库，这些数据库包括 DNA 和与相关患者及对照组相关的临床信息。

可以通过简单、快速的指标来分析所产生的转基因生物体的表型，测定指标包括尿、血液、脊髓液或组织中存在的代谢物、蛋白质（如胰岛素）、脂类或碳水化合物（当测定葡萄糖耐受时）的变化。
30 可以通过本领域的熟练技术人员已知的许多技术中的任意技术来实施体液中的指标测定，这些技术包括通过 NMR、ELISA、GMS 等等。

还可以使用光试验、声音试验、记忆试验和应激试验，测定行为模式或对外来刺激的反应性，对其他表型特征进行分析。

其他的可检测表型特征包括生长和衰老参数、肿瘤生长和肥胖等，这些特征可以通过诸如称量体重、脂肪含量和生长速度的方法来加以测定。也可以评价其他可检测特征的变化，如血压、心率、肺功能等的变化。

本发明转座子技术的优势在于能够生成转基因动物文库而不需要烦琐的保存。以前产生转基因动物的方法包括产生突变的化学诱变方法。这些方法涉及每个动物（如小鼠）进行多处突变。动物一旦产生，就分析其表型，并需要归档/保存以备后用。本发明能够产生基因组中插入不同转座子的起始细胞系或起始转基因生物体。然后使用这些细胞系或生物体和带有转座酶的转基因动物一起繁殖，产生一个新的文库。

在另一个实施方案中，可使用本发明的方法产生 ES 细胞文库，该文库具有分布在整个基因组上的不同转座子插入。可以对其进行测序和分析。ES 细胞便于保存，以备后用。

在一个可替换的实施方案中，可以使用本发明的方法“标记”表达受外来刺激物调节的基因。因此，表现出具有标记转座子的转座移动的一个胚胎、生物体或来源于二者之一的组织或细胞用外来刺激物处理，观察对标记基因表达的调节，其中刺激物可以是候选药物或其他受试试剂。标记基因超高表达或过低表达的细胞可能具有插入的转座子，且转座子插入到了受刺激物上调或下调的基因中或基因附近。因此本发明通过插入包含标记基因的转座子，可用于提供体内增强子捕获和外显子捕获功能，其中标记基因表达水平受增强子或外显子的接近调节。

本方法可用于研究药物发现和毒理学研究中药物对基因的调节作用。另外，可用于研究天然刺激物，如激素、细胞因子和生长因子对基因的调节；鉴定分子介入的新靶点，包括用语人类、植物或动物疾病治疗的靶点，开发杀虫剂、除草剂、抗真菌剂和抗菌剂。

下面的实施例对本发明作了进一步示例性说明。

实施例

A: 使用多西环素体内活化 Minos

实施例 1: 产生转座子携带小鼠和转座酶携带小鼠

产生了两个转基因小鼠系。转座子携带系 (MCG 系) 包含带有

5 Minos 转座子的片段串联体, 其中 Minos 转座子包含位于巨细胞病毒启动子控制之下的 GFP 基因。对转座子加以改造, 使内部至反向重复之间的几乎所有序列都置换为 CMV/GFP 序列盒。由于不包含编码转座酶的基因, 该转座子是非自主的, 只有当转座酶存在的条件下才能移动。第二转基因小鼠系包含在诱导性启动子控制之下表达的

10 Minos 转座酶基因。

如下所述构建转座子 MiCMVGFP: 用 *BamH*I 切割质粒 pMILRTetR (Klinakis et al. (2000) *Ins. Mol. Biol.* 9, 269-275 (2000b)), 然后重新连接, 以去除 Minos 末端之间的四环素抗性基因, 产生质粒 PMILR Δ *BamH*I。将来源于 pMILR Δ *BamH*I 的、包含 *D.hydei* 中 Minos 反

15 向重复序列和原始旁侧序列的 *Asp718/SacI* 片段克隆到质粒 pPolyIII-I-lox [将 *loxP* 寡核苷酸:

ATAACTTCGTATAGCATACATTATACGAAGTTAT

插入到载体 pPolyIII-I 中产生 (登记号 M18131) 中产生] 中, 产生质粒 ppolyMILR Δ *BamH*I。通过将来源于质粒 pBluescriptGFP 的 2.2

20 kb *SpeI* 片段插入到 pPolyMILR Δ *BamH*I 的 *SpeI* 位点上, 产生用于产生转基因小鼠的最终构建体 (pMiCMVGFP, 图 5), 其中质粒 pBluescriptGFP 包含由 CMV 启动子驱动的人源化 GFP 基因 (来源于 Clontech 质粒 pHGFP-S65T), 其后为 SV40 介入序列和腺苷酸化信号。

25 将来源于 pMiCMVGFP 质粒的 3.2 kb *XhoI* 片段微注射到 FVB X FVB 受精卵母细胞中, 构建携带转座子的 MCG 系。使用 GFP DNA 作为探针, 对来源于尾部活检样品的 DNA 进行 Southern 印迹分析, 鉴定转基因动物。

通过常规的同源重组 ES 细胞技术将转座酶基因导入胚胎干细胞,

30 产生转座酶表达系。用常规方法将 ES 细胞注射到囊胚中获得转基因动物 (小鼠胚胎操作, Hogan et al., 冷泉港出版社, 1994)。使

用了两个构建体。其中一个构建体包含有靶细胞中表达位于组成性启动子控制之下的 rtTA 基因。所用的构建体是 pTet-On 质粒 (Clontech), 它包含位于巨细胞即早 (CMV) 启动子控制之下、编码 rtTA 活化因子的基因。该构建体编码的 rtTA 转录活化因子仅在多西环素存在的条件下才能活化。第二构建体包含位于四环素应答元件或 TRE 控制之下的 Minos 转座酶基因。TRE 包含 7 个含有 tet 操纵子 (tetO) 的 42-bp 序列的正向重复, 就位于基础 CMV 启动子的上游, 该启动子缺少正常情况下和 CMV 即早启动子相关的增强子元件。因为没有这些增强子元件, 在没有 rtTA 的结合时, 没有转座酶从 TRE 的“泄漏”表达。使用的第二构建体是 pTRE2 质粒 (Clontech), 其多克隆位点 (MCS) 中插入了编码 Minos 转座酶的基因。在稳定转化有两个构建体的细胞中, rtTA 得表达, 但是只有往培养基中加入多西环素时才能活化 Minos 转座酶的转录。

使用转座酶 cDNA 片段作为探针对来源于尾部活检样品的 DNA 进行 Southern 印迹分析, 鉴定转基因动物。

实施例 2: 体内活化 Minos

转座子携带 MCG 系的转基因小鼠和转座酶携带系的转基因小鼠相杂交。基因组中包含转座子和编码转座酶基因的胚胎中, 转座子的移动仅在多西环素存在的条件下发生。将多西环素施用在母体生物体水中的胚胎上。为了限制转座事件的可能数目, 多西环素仅在限定的时间内施用 (妊娠 1~2 天时)。

出生时, 分离从胚胎发育而来的转基因后代, 各种细胞和组织用于基因型分析。

实施例 3: 转座的检测

使用能够与非移动性 *Drosophila hydei* 序列相杂交的引物对转座子进行 PCR 分析, 检测小鼠组织中 Minos 转座酶所活化的转座, 其中 *Drosophila hydei* 序列位于构建体中 Minos 转座子的旁侧 (Klinakis et al., (2000) *Ins. Mol. Biol.* 9, 269-275)。在果蝇细胞中, 转座酶介导的 Minos 切除后, 染色单体进行修复, 通常留下特征性的 6-碱基对足迹

(Arca et al. (1997) Genetics 145, 267-279)。使用特异引物进行 PCR 分析,产生诊断性的 167 bp 的 PCR 片段(Catteruccia F. et al. (2000) Proc. Natl. Acad. Sci. USA 97, 2157-2162)。

使用 Dneasy 组织试剂盒 (QIAGEN), 根据生产商的说明从不同组织中分离基因组 DNA。使用引物 11DML:

(5'-AAGTGTAAGTGCTTGAAATGC-3')

和 GOUM67

(5'-GCATCAAATTGAGTTTTGCTC-3')

进行 PCR 反应。PCR 条件如下: 每 25 μ L 反应体系中包含 10 mmol/L Tris-HCl (pH 8.8)、50 mmol/L KCl、1.5 mmol/L $MgCl_2$ 、0.001% 明胶; 1.2 单位 Taq 2000™ DNA 聚合酶 (STRATAGENE), 200 g 模板 DNA 和 10 pmol 各引物。94℃ 30 s, 59℃ 30s, 72℃ 30s 进行 43 或 60 个循环。将 PCR 产物克隆到 PCR II TA 克隆载体 (Invitrogen) 中, 并用 T7 引物进行测序。

诊断性条带存在于转基因后代的某些组织中。通过使用对扩增序列特异的标记 DNA 探针进行 Southern 印迹分析, 证实片段的相同性 (数据未列出)。细胞簇对于相同的转座基因来说是同质的。

B:使用位于 ZP3 启动子控制之下的转座酶在体内活化 Minos

20

实施例 4

在该实施例中, 说明了将编码转座酶的基因的表达置于特定发育阶段产生的基因调控信号的控制之下, 使转座作用在胚胎发育的这一阶段得以诱导。在该实施例中, 编码转座酶的基因位于 ZP3 启动子的控制之下, 从而使转座酶仅在 2 至 3 周卵发生阶段中生长的卵母细胞中表达。在生长的卵母细胞中表达的 Minos 转座酶催化修饰的、非自主的 Minos 转座子的切除, 并促进其向基因组中新位点上的重新整合。

图 11 显示可替换使用的构建体 (敲入构建体, 用于常规的转基因), 其中转座酶插入到内源精子特异性 H1t 基因, 以在精子中发生

转座。通过将 H1t 序列置换为 ZP3（卵特异性表达）或 hnRNP（普遍表达）基因起始位点旁侧的相应序列，构建卵特异性表达和普遍表达的可替换构建体。

5 C: Minos 序列的哺乳动物密码子利用率

实施例 5

改进转座酶

- 10 提高蝇转座酶哺乳动物细胞或动物体内的效率的一个方法是将蝇密码子利用率置换为哺乳动物密码子利用率，使 mRNA 更有效翻译，导致转座酶浓度升高。为了这一目的，我们将蝇 Minos 转座酶的编码序列置换成了序列：

-SEQ 新：1026 bp;

- 15 组成：321 A; 235 C; 261 G; 209 T; 0 其他
百分比：31% A; 23% C; 25% G; 20% T; 0%其他
分子量（kDa）：ssDNA: 317.79 dsDNA: 632.5
来源

```

1  ATGGTGC GCG GTAAGCCTAT CTCTAAGGAG ATCAGAGTAC TGATCAGGGA CTATTTTAAG
61  TCTGGGAAGA CACTCACTGA GATAAGCAAG CAGTTAAACT TGCCTAAGAG CTCTGTGCAT
121 GGGGTGATAC AGATTTTCAA GAAAAATGGG AACATTGAGA ATAACATCGC GAATAGAGGC
181 CGAACATCCG CAATAACCCC CCGCGACAAG AGACAGCTGG CCAAAAATTGT GAAGGCTGAC
241 CGCCGCCAAT CCCTGAGAAA CTGGCTTCC AAGTGGTCGC AGACCATTGG CAAGACTGTC
301 AAGCGGGAGT GGACCCGGCA GCAATTAAAG AGTATTGGCT ACGGTTTTTA TAAGGCCAAG
361 GAAAAACCCC TGCTTACGCT TCGGCAAAAA AAGAAGCGTC TGCAATGGGC TCGGGAAAGG
421 ATGTCTTGGA CTCAAAGGCA GTGGGATACC ATCATCTTCA GCGATGAGGC TAAATTTGAT
481 GTGAGTGTGCG GCGACACGAG AAAGCGCGTC ATCCGTAAGA GGTCCGAGAC ATACCATAAG
541 GACTGCCTGA AAAGAACAAC CAAGTTTCCT GCAAGCACTA TGGTATGGGG ATGTATGTCT
601 GCCAAAGGAC TCGGAAAGCT TCACTTCATC GAAGGGACCG TTAATGCCGA AAAATACATT
661 AACATTCTCC AGGATAGTTT GCTGCCCTCA ATACCAAAAC TATCCGATTG TGGTGAATTC
721 ACTTTTCAGC AGGACGGAGC ATCATCGCAC ACCGCCAAGC GGACCAAAAA CTGGCTGCAG
781 TACAATCAGA TGGAGGTGCT CGATTGGCCC TCAAATAGTC CGGATCTAAG CCAATCGAA
841 AATATCTGGT GGCTAATGAA AAACCAGCTG CGAAACGAGC CACAGAGGAA CATTTCGAC
901 TTGAAAATCA AGCTGCAAGA GATGTGGGAC TCAATCTCTC AGGAGCACTG CAAAAACCTG
961 CTCAGCAGCA TGCCTAAACG AGTGAAATGC GTGATGCAGG CCAAGGGCGA CGTTACACAG
1021 TTCTGA

```

- 该序列对应于正常的哺乳动物密码子利用率，翻译后产生和蝇转座酶蛋白序列相同的蛋白质序列。以寡聚核苷酸（上方正义链和下方反义链）形式合成的基因（cDNA）分为三部分，对各部分进行克隆，然后连接成一条 cDNA。

-部分 A 和起始密码子前面的 Kozak
接头 A Kozak: CCACCATGG

```

      AatII   NcoI
-15  CCCCAGCGTCCCACCATGGTGC GCGG GTAAGCCTAT CTCTAAGGAG ATCAGAGTAC TGATCAGGGA CTATTTTAAG
      TACCACGCGC CATTCGGATA GAGATTCCTC TAGTCTCATG ACTAGTCCTT GATAAAATTC

61      TCTGGGAAGA CACTCACTGA GATAAGCAAG CAGTTAAACT TGCCTAAGAG CTCTGTGCAT
      AGACCCCTCT GTGAGTGACT CTATTCGTTT GTCAATTGA ACGGATTCTC GAGACACGTA

121     GGGGTGATAC AGATTTTCAA GAAAAATGGG AACATTGAGA ATAACATCGC GAATAGAGGC
      CCCCACTATG TCTAAAAGTT CTTTTTACCC TTGTAACTCT TATGTAGCG CTTATCTCCG

181     CGAACATCCG CAATAACCCC CCGCGACAAG AGACAGCTGG CCAAAATTGT GAAGGCTGAC
      GCTTGTAGGC GTTATTGGGG GCGCTGTTT TCTGTGACC GGTTTTAACA CTTCCGACTG

241     CGCCGCCAAT CCCTGAGAAA CTTGGCTTCC AAGTGGTCGC AGACAATTGG CAAGACTGTC
      GCGGCGGTTA GGGACTCTTT GAACCGAAGG TTCACCAGCG TCTGTTAACC TGATCAGGGG
                                     MfeI   SpeI

```

接头 B

```

                                     MfeI
                                     GGGGCAATTGG CAAGACTGTC
GCGGCGGTTA GGGACTCTTT GAACCGAAGG TTCACCAGCG TCTGTTAACC GTTCTGACAG

301  AAGCGGGAGT GGACCCGGCA GCAATTAAAG AGTATTGGCT ACGGTTTTTA TAAGGCCAAG
      TTCGCCCTCA CCTGGGCCGT CGTTAATTTT TCATAACCGA TGCCAAAAAT ATTCGGGTTT

361  GAAAAACCCC TGCTTACGCT TCGGCAAAAA AAGAAGCGTC TGCAATGGGC TCGGGAAAGG
      CTTTTTGGGG ACGAATGCGA AGCCGTTTTT TTCTTCGCAG ACGTTACCGG AGCCCTTTCC

421  ATGTCTTGGA CTCAAAGGCA GTGGGATACC ATCATCTTCA GCGATGAGGC TAAATTGAT
      TACAGAACCT GAGTTTCCGT CACCCTATGG TAGTAGAAGT CGCTACTCCG ATTTAACTA

481  GTGAGTGTCT GCGACACGAG AAAACGCGTC ATCCGTAAGA GGTCCGAGAC ATACCATAAG
      CACTCACAGC CGCTGTGCTC TTTTGCAGC TAGGCATTCT CCAGGCTCTG TATGGTATTC

541  GACTGCCTGA AAAGAACAAC CAAGTTTCCT GCAAGCACTA TGGTATGGGG ATGTATGTCT
      CTGACGGACT TTTCTTGTG GTTCAAAGGA CGTTCGTGAT ACCATACCCC TACATACAGA

601  GCCAAAGGAC TCGGAAAGCT TCACCTTCATC GAAGGGACCG TTAATGCCGA AAAATACATT
      CGGTTTCCTG AGCCTTTGGA ACCCCTACGTACCCC
      HindIII   NsiI

```

接头 C

```

HindIII
601      CCCCAGCT TCACTTCATC GAAGGGACCG TTAATGCCGA AAAATACATT
      TTCGA AGTGAAGTAG CTTCCTGGC AATTACGGCT TTTTATGTAA

661      AACATTCTCC AGGATAGTTT GCTGCCCTCA ATACCAAAC TATCCGATTG TGGTGAATTC
      TTGTAAGAGG TCCTATCAAA CGACGGGAGT TATGGTTTTC ATAGGCTAAC ACCACTTAAG

721      ACTTTTCAGC AGGACGGAGC ATCATCGCAC ACCGCCAAGC GGACCAAAAA CTGGCTGCAG
      TGAAAAGTCG TCCTGCCTCG TAGTAGCGTG TGGCGGTTTC CCTGGTTTTT GACCGACGTC

781      TACAATCAGA TGGAGGTGCT CGATTGGCCC TCAAATAGTC CGGATCTAAG CCCAATCGAA
      ATGTTAGTCT ACCTCCACGA GCTAACCGGG AGTTTATCAG GCCTAGATTC GGGTTAGCTT

841      AATATCTGGT GGCTAATGAA AAACCAAGCTG CGAAACGAGC CACAGAGGAA CATTTCCGAC
      TTATAGACCA CCGATTACTT TTTGGTCGAC GCTTTGCTCG GTGTCTCCTT GTAAAGGCTG

901      TTGAAAATCA AGCTGCAAGA GATGTGGGAC TCAATCTCTC AGGAGCACTG CAAAAACCTG
      AACTTTTAGT TCGACGTTCT CTACACCTG AGTTAGAGAG TCCTCGTGAC GTTTTGGAC

961      CTCAGCAGCA TGCCTAAACG AGTGAAATGC GTGATGCAGG CCAAGGGCGA CGTTACACAG
      GAGTCGTCGT ACGGATTGTC TCACCTTACG CACTACGTCC GGTTCCTCGT GCAATGTGTC

1021     TTCTGAGGAT CC
      AAGACTCCTA GGCCCCGGGG AGATCTCCTA CGTAGGGG
      BamHI      XbaI      NsiI

```

材料和方法

5 质粒构建

实施例 1 中描述了用于产生转基因小鼠的修饰型 Minos 转座子 pMiCMVGFP (图 5) 的构建。简要地说, 包含由 CMV 启动子驱动的人源化 GFP 基因以及下游介入序列和 SV40 多聚腺苷酸化信号的 2.2 kb 片段置于 Minos 反向重复之间。需要的话, 左侧反向重复前面可以包括 lox P 位点, 用于产生单拷贝转基因动物。

以载体 Pev3 (Clare Gooding, Biotechnology Dept, Zeneca, Macclesfield, UK; Needham et al., Nucl. Acids Res., 20, 997-1003, 1992 中进一步介绍了 Pev3) 中 1 kb *Clal/NotI* 片段的形式克隆 Minos 转座酶 cDNA。将获得质粒中的 3.8 kb 的 *Clal/Asp718* 片段 (包含 Minos 转座酶 cDNA, 及后面的人 β 珠蛋白基因内含子和多聚腺苷酸化信号) 克隆到 pBluescript SK+ (Stratagene, La Jolla, Ca, USA) 中, 产生质粒 pBlue/ILMi/3' β 。来自质粒 ZP3/6.5 Luc (Lira, S. et al (1990) Proc. Natl. Acad. Sci. USA 87, 7215-7219) 的 6.5 kb 平末端 *Asp718* 片段包含透明带 3 (ZP3) 基因的 5' 旁侧区和启动子, 将该片段克隆到 pBlue/ILMi/3' β 的 *EcoRV* 位点中, 产生质粒 ZP3/ILMi (图 6), 该智力可用于产生在发育的卵中表达转座酶的转基因小鼠。

产生转基因小鼠

为了产生 Minos 转座酶表达系，从 pZP3/ILMi 中切除一个 10.3 kb 的 *SmaI/Asp718* 片段，用凝胶电泳从质粒序列中分离出来 (Sambrook, J 等, 分子克隆: 实验指南, 冷泉港实验室出版社, NY(1989)), 使用 ELUTIP-d 柱进行纯化和浓缩 (Schleicher & Schuell GmbH, Dassel, Germany), 并以 4 ng/ μ L 的浓度注射到受精的卵母细胞中。将注射了的卵转移到假孕的小鼠体内, 通过对尾部 DNA 进行 Southern 印迹分析鉴定转基因后代 (Southern, E.M. (1975), J. Mol. Biol 98, 503-517)。

如上所述和 (Zagoraiou, L et al. (2001) Proc. Natl. Acad. Sci. USA 98, 11474-11478) 中所述产生转座子携带 (MCG) 系。

RT-PCR

对于 RT-PCR 分析来说, 使用 Ultraspec RNA 分离系统 (Biotech Laboratories, Houston, TX, USA) 从 ZP3/ILMi 转基因的不同器官中分离总 RNA。在 20 μ L 反应体系中, 使用逆转录酶 (Super RT; HT Biotechnology, Cambridge, UK) 和 oligo (dT) 引物, 从 1 μ g 总 RNA 合成 cDNA。在体积为 50 μ L 的 PCR 缓冲液中进行 PCR 反应, PCR 缓冲液包括 1.5 mmol/L $MgCl_2$ 、100 ng 各引物、0.2 mmol/L dNTPs 和 2.5 U Taq DNA 聚合酶 (Pharmacia)。94 $^{\circ}$ C 45 秒, 55 $^{\circ}$ C 30 秒, 72 $^{\circ}$ C 45 秒, 共进行 25 个循环。在 2% 琼脂糖凝胶上进行电泳观察 PCR 产物。Minos 转座酶特异性引物为 Minos1: 5'-CAGCTTCGAAATGAGCCAC-3' 和 β -EX: 5'-TGGACAGCAAGAAAGCGAG-3'。鼠次黄嘌呤磷酸核糖转移酶 (HPRT) 特异引物为: 5'-CACAGGACTAGAACACCTGC-3' 和 5'-GCTGGTGAAAAGGACCTCT-3'。

育种(breeding)程序

携带转座子 (MCG) 的雌性动物和 ZP3/ILMi 系中 15 个雄性动物相杂交。通过这些杂交获得的双阳性雌性动物和野生型 (WT) 雄

性动物相杂交, 通过 Southern 印迹分析, 分析后代中的可能转座事件。使用 *EcoRV* 或 *BglII* 消化基因组 DNA, 在 0.7% 或 1% 的琼脂糖凝胶 (Sigma, Steinheim, Germany) 上进行分离, 印到尼龙膜 (Hybond-N⁺, Amersham Pharmacia, Buckinghamshire England) 上, 用 ³²P 标记的、
5 来源于 pMiCMGFP 的 737 bp *SacI/NotI* GFP 片段进行探测。

DNA 的荧光原位杂交 (FISH) 分析

根据常规方法从腹腔白细胞 (Mulder, M.P et al. (1995). Hum Genet 96(2):133-141) 制备小鼠细胞分裂中期的涂片。使用 pMiCMVGFP 构
10 建体的 737-bp *SacI/NotI* GFP 片段作为探针。探针可以使用生物素 (Boehringer Mannheim) 进行标记并直接用 FITC 进行免疫化学检测, 或使用基于 Tyramide 的步骤改进信号检测 (Raap, A.K. et al (1995) Human Molecular Genetics 4, 529-534)。使用 DAPI 对 DNA 进行复染。
15

克隆插入位点

使用 *EcoRV* 或 *BglII*, 切割从带有新的 Minos 插入位点的动物获得的小鼠 DNA, 并在 0.7% 琼脂糖凝胶上进行电泳。将包含转座事件的凝胶区切下, 分离 DNA。根据片段大小, 直接在自身连接的片段
20 上进行反向 PCR, 使用 Minos 引物 IMio1 (5'-AAGAGAATAAAATTCTCTTTGAGACG-3') 进行第一次 PCR, 使用 IMio2 (5'-GATAATATAGTGTGTAAACATTGCGC-3') 进行巢式 PCR (Klinakis, A.G., Zagoraiou, L., Vassilatis, D.K & Savakis, C. (2000) EMBO Report 11, 416-421.); 或者使用 *AluI* 对获得的 *EcoRV* 或 *BglII*
25 片段进一步进行消化, 并加以环化。如前所述使用 IMio1 和 IMii1 (5'-CAAAAATATGAGTAATTTATTCAAACGG-3') 进行反向 PCR, 使用引物 IMio2 和 IMii2 (5'-GCTTAAGAGATAAGAAAAA AGTGACC-3') 进行巢式 PCR (Klinakis, A.G., Zagoraiou, L., Vassilatis, D.K & Savakis, C. (2000) EMBO Report 11, 416-421.)。用这种方式分
30 别扩增左侧和右侧序列。直接对 PCR 片段进行测序, 或者克隆到 pGEM T easy 载体 (Promega) 或 PCR II (Invitrogen) 上进行测序。使用获

得的序列，对当时 Celera (www.celera.com) 数据库中能够得到的小鼠基因组序列进行 BLAST 检索。

结果

5 产生了携带转座子的转基因小鼠系 (MCG)。它包含 6 个拷贝的、整合在小鼠第 14 号染色体中的 Minos 转座子 MiCMVGFP(图 5)。由于该转座子缺少转座酶基因，因此是非自主的，即它自身不能转座。平行产生了两个表达转座酶的小鼠系。由于使用 ZP3 基因的 6.5 kb 5' 旁侧区和启动子(图 6)，它们在生长点卵母细胞中特异性地表达 Minos
10 转座酶。在两个 Minos 转座酶 (ZP3/ILMi) 系中，转基因以串联体形式整合。在该实施例中，我们使用了整合有较高拷贝数的 ZP3/ILMi 系 15 (数据未列出)。与预期的一样，该系中转座酶的表达仅限于卵巢(图 7)。对来源于转基因 ZP3/ILMi 系不同组织的 RNA 样品进行 RT-PCR，结果仅从卵巢中扩增出了~360 bp 的片段，和正确剪接
15 的转座酶 RNA 相对应。通过使用跨越一个外显子/内含子连接部位的引物，防止扩增到相似大小片段的杂质 DNA (β珠蛋白 IVSII-图 6)。

既然 Minos 转座酶从转基因的表达受 ZP3 启动子的驱动，转座酶应该仅在 2 至 3 周卵发生阶段生长的卵母细胞中表达。

正常情况下，在原始卵母细胞 (10~15 μm) 中不能检测到透明带
20 转录物，在 50 μm 直径的卵母细胞中能够观察到最高水平。当卵母细胞到达最大 (70~80 μm) 时，ZP3 转录物的水平开始下降。排出的卵中包含的透明带转录物少于总透明带转录物峰值水平的 5% (Millar, et al (1991) *Molecular and Cellular Biology* 11, 6197-6204; Liu, C et al (1996) *Proc. Natl. Acad. Sci. USA.* 93, 5431-5436)。为了排除
25 残留在成熟卵母细胞中的一些转座酶活性介导亲本转座子转基因的可能性，将 Zp3/ILMi 雄性动物 (不表达转座酶) 和多拷贝携带 MCG 系的雌性动物相交配。选择对两个转基因阳性的雌性子代，用于进一步研究。我们用 Southern 印迹分析法分析了双阳性雌性动物和野生型雄性动物产生的 307 个子代。将 *EcoRV* 和/或 *BglII* 消化的尾部 DNA
30 印到膜上，并和 GFP 探针相杂交。由于这些酶都不在转座子内部切割，因此在 MCG 系中存在有和转座子探针相杂交的单一一条带。如果

发生了转座，转座子插入到其最初所在的基因组 *EcoRV* 或 *BglII* 片段之外，和转座子探针杂交后会检测到一条新的带。307 只小鼠中，有 146 只小鼠是转座子（GFP）阳性的。在这 146 只小鼠中，印迹分析观察到有 12 个转座事件产生了新的限制性片段，转座频率为 8.2%。

5 在两只小鼠中发现了两个独立的转座事件（图 8 泳道 2 和 5）。观察到一些后代带有 Minos 介导的插入，但是没有遗传转座酶转基因，有力地说明了转座作用发生在母体减数分裂 II 期之前的胚系细胞中。

为了证明观察到的转座事件确实发生在胚系中，进一步使发生了转座事件的动物和野生型小鼠杂交。对其子代的分析说明这些小鼠

10 稳定地转移了重新插入的 Minos 元件（图 8）。转座子串联体的分离以及向 F1 代两个不同系中的重新插入明确地说明了转座作用发生在另外一个染色体上（图 8，泳道 2 对泳道 3 和 4，泳道 8 对泳道 9）。所有的转座事件中都移动了单拷贝的转座子，只有一个转座事件除外。

15 以前曾经注意到转座子的大小影响其转座频率；较长的元件转座频率较低（Lampe, D.J. et al. (1998). *Genetics* 149, 179-187; Fischer, S.E., et al. (1999) *Mol. Gen. Genet.* 262, 268-274.），说明当转座酶结合位点（反向重复）彼此靠近时，可以更有效地识别。

Minos 转座特征为元件的精确整合而不移动旁侧 DNA。在果蝇和

20 HeLa 细胞中，转座子插入到 TA 双核苷酸中，导致插入时靶位点复制（17、19）。为了研究小鼠基因组中插入物的结构，我们从 5 个不同的转座事件中克隆了旁侧区（如材料和方法中所述）。和果蝇中观察到的相同，Minos 末端旁侧为 TA 双核苷酸，之后是和建立者小鼠系 MCG 中元件旁侧序列不相关的序列（图 9）。用所得序列在 Celera

25 小鼠基因组数据库中序列进行 BLAST 检索，结果说明所有新的 5 个旁侧序列（一种情况下仅获得了一个旁侧序列）都对应于广泛分布的基因组位置（图 9）。所分析的 5 个转座事件中，只有一个位于 14 号染色体上。它是整合到中心粒区域的单拷贝转座子，14 号染色体顶端不存在转座子串联体。因此，这是发生在不同染色体上的转座事件（参见图 10）。对来源于腹腔白细胞的分裂中期涂片进行 DNA 荧

30 光原位杂交（FISH 分析），验证了旁侧序列的测序结果（数据未列

出)。对于 Southern 印迹分析可以清楚地加以检测但并不靠近同一染色体上原始位置的转座事件来说, FISH 可能检测不到。通过 FISH 检测单拷贝 (3 kb) 转座作用确实非常困难, 这可以部分地解释了在使用 FISH 作为仅有的检测方法的情况下为什么我们以前获得的转座作用频率较低 (0.61%), 其中的转座作用使用包含相同转座子的转基因系 (MCG) 和 T 细胞中特异性表达的 Minos 转座酶获得 (Zagoraiou, L., et al (2001) *Proc. Natl. Acad. Sci. USA* 98, 11474-11478)。既然靠近元件原始部位的“局部插入物”至少对于 P 元件来说是相当频繁的 (Zhang, P. & Spradling, A.C. (1993) *Genetics* 133, 361-373) 和 Sleeping Beauty (Luo, G et al (1998) *Proc. Natl. Acad. Sci. USA* 95, 10769-10773; Fisher, S.E.J., et al (2001) *Proc. Natl. Acad. Sci. USA* 98, 6759-6764.), 对于 Minos 元件来说也不是不可能。

为了确定 EcoRV/ BglII 旁侧区的大小, 我们使用 loxP/Cre 系统从 MCG 系中产生了单拷贝转座系 (数据未列出)。在 Southern 印迹分析数据的基础上, EcoRV 和 BglII 转座子诊断片段中转座子旁侧大约有 10 kb 基因组序列。发生在该区域的局部重新插入 (但是对于诱变目的并没有意义) 会逃脱检测, 不包括在我们对转座频率的估算之内。因此, 频率 8.2% 是“有用”转座频率, 而非实际转座频率。据报道, Sleeping Beauty 在小鼠雄性胚系中的转座频率约为 20%, 但是仅有 2% 转座到了不同的染色体上 (Fisher, S.E.J., Wienholds, E. & Plsterk, R.H.A. (2001) *Proc. Natl. Acad. Sci. USA* 98, 6759-6764)。相比之下, 我们发现 Minos 大约有 6% 转座作用发生在不同的染色体上 (146 个后代中, 分析到的 11 个转座作用中有 8 个, 其中 1 个未知), 而 11 个转座作用中仅有 3 个发生在相同的染色体上 (如图 10B)。该结果说明 Minos 系统具有向不同染色体转座的倾向 (前提是靠近原始位点处不发生未能检测到的许多转座作用)。值得注意的是, 对至今发表的不同转座子系统进行直接比较并不是十分有效的。转座子的大小、拷贝数和初始染色体位置都会影响转座效率, 这些在不同的系统中无法比较。

30

综上所述, 这些结果表明可以通过选择诱导时间在胚胎发育的预

定阶段诱导转座子的移动在小鼠胚系中实现转座作用。另外，结果表明使用 Minos 的系统是在生物体，如小鼠中实现插入性基因失活、基因标签、增强子捕获和外显子捕获的有力工具。

上面说明书中提到的所有出版物都在此引用以作参考。对于本领域中的熟练技术人员来说，显然可以在不违背本发明的范围和精神的条件下对本发明所述的方法和系统加以不同的改动和变通。虽然本发明是联系具体的优选实施方案进行介绍的，可以理解的是，要求保护的本发明并不局限于具体的实施方案。事实上，对用于实施本发明的所述模式所作的各种改动对于分子生物学和相关领域中熟练的技术人员来说是显而易见的，属于下面权利要求书的范围。

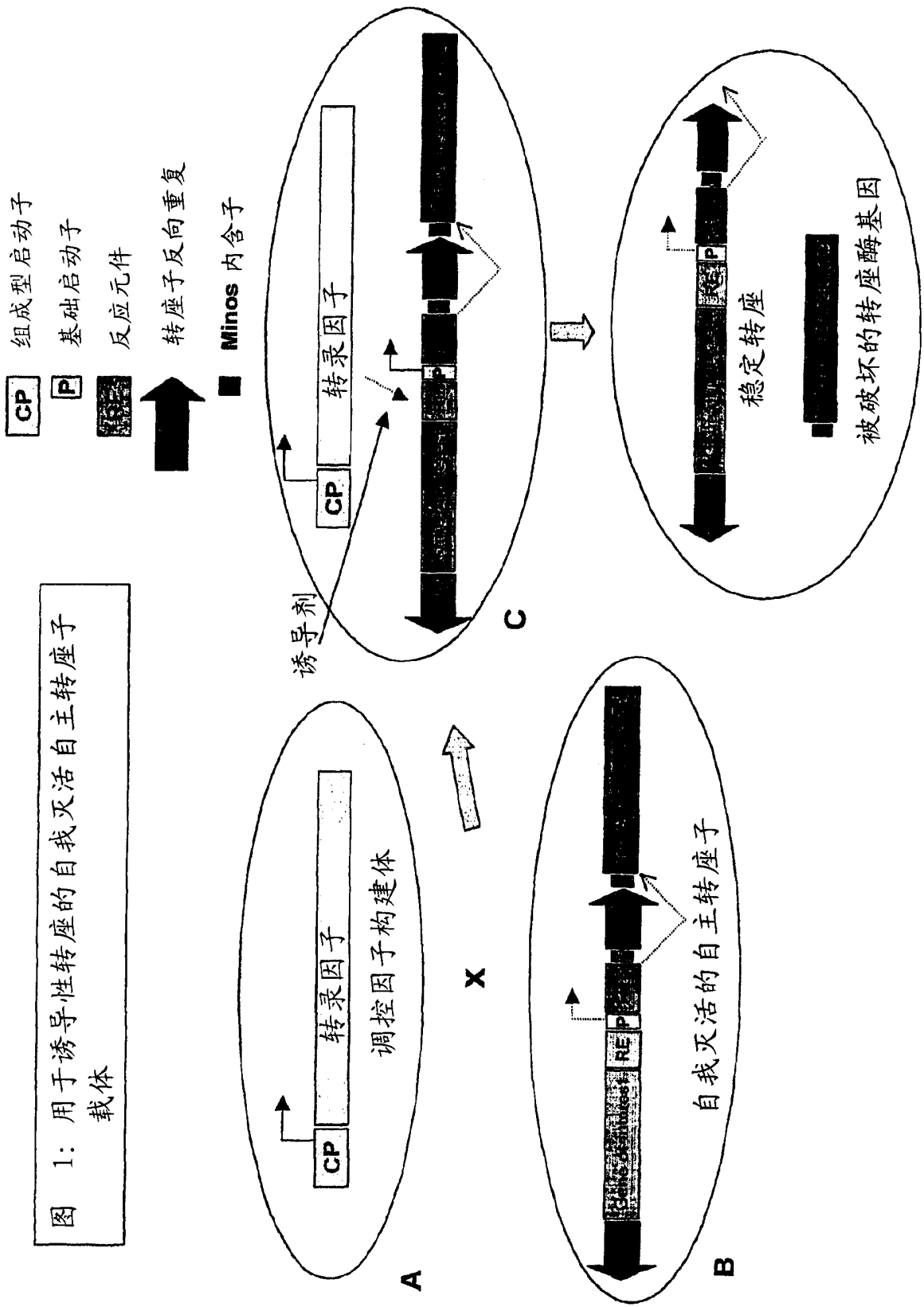
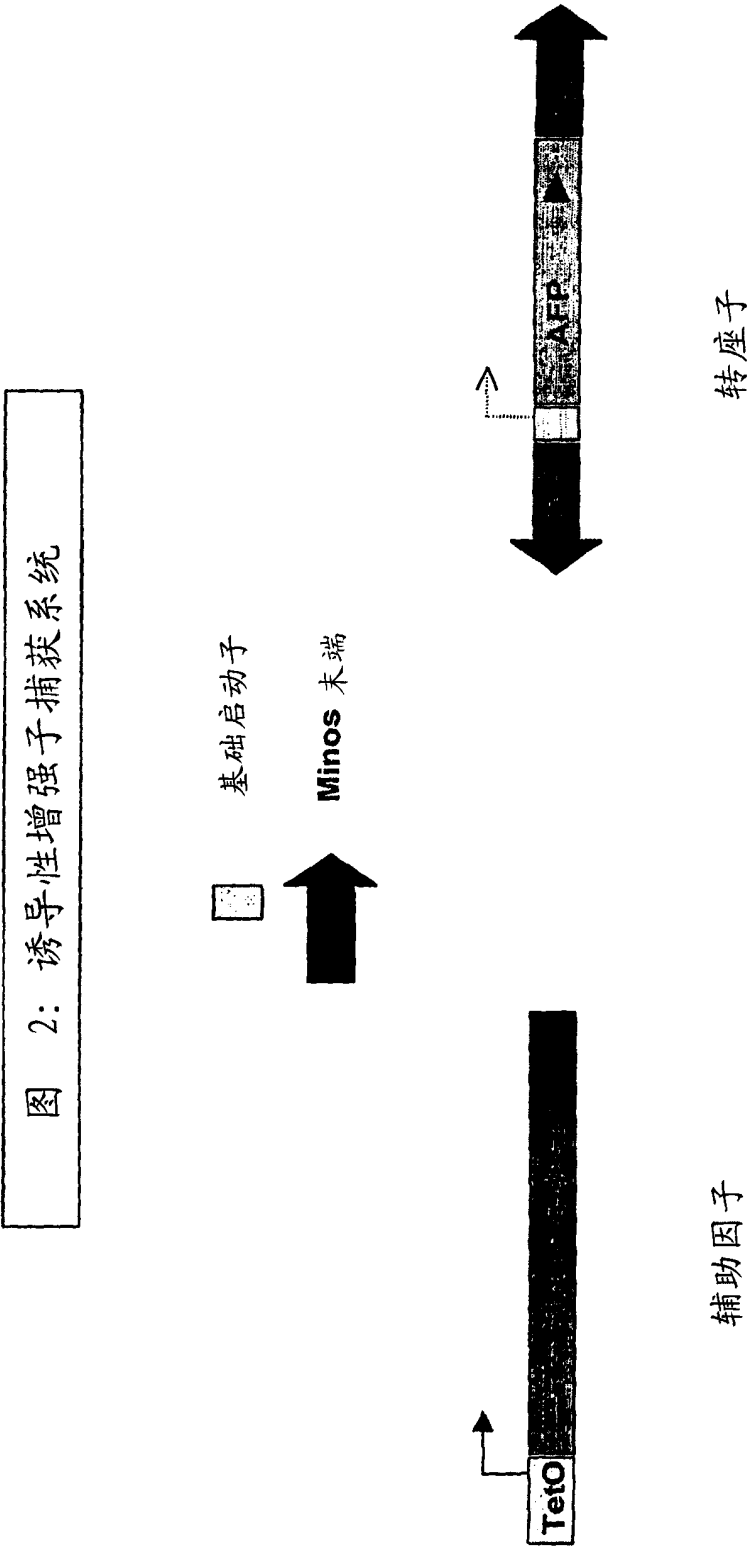
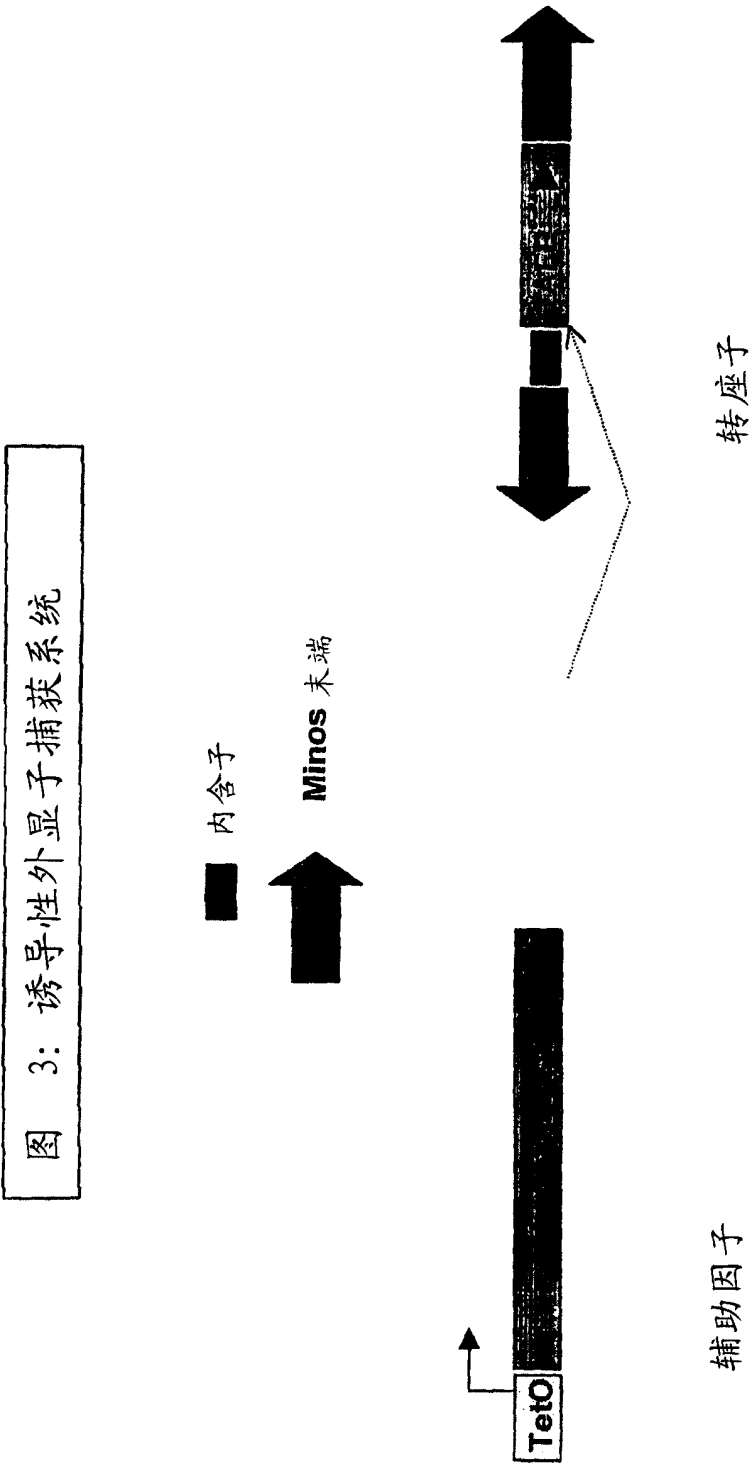
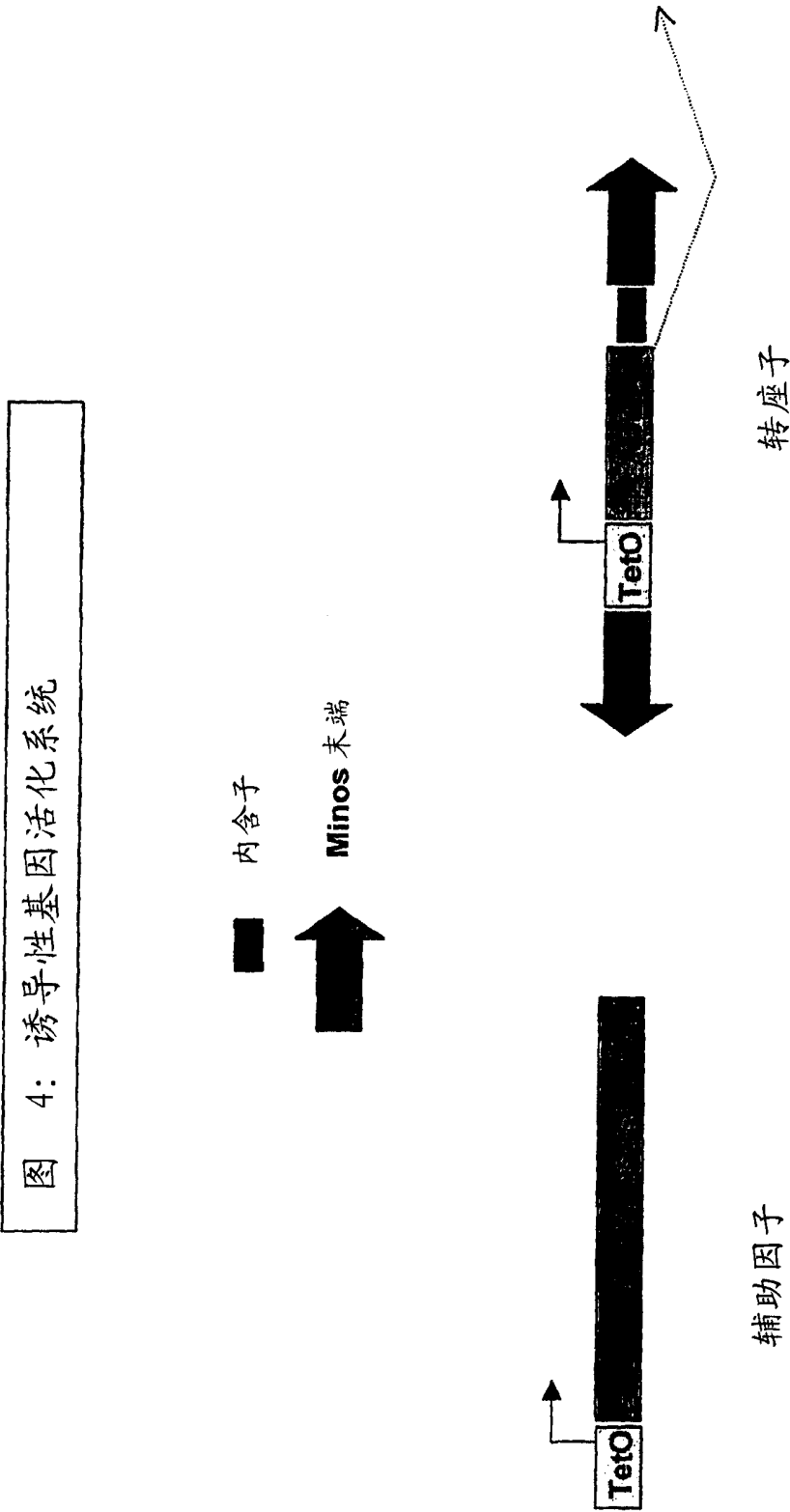
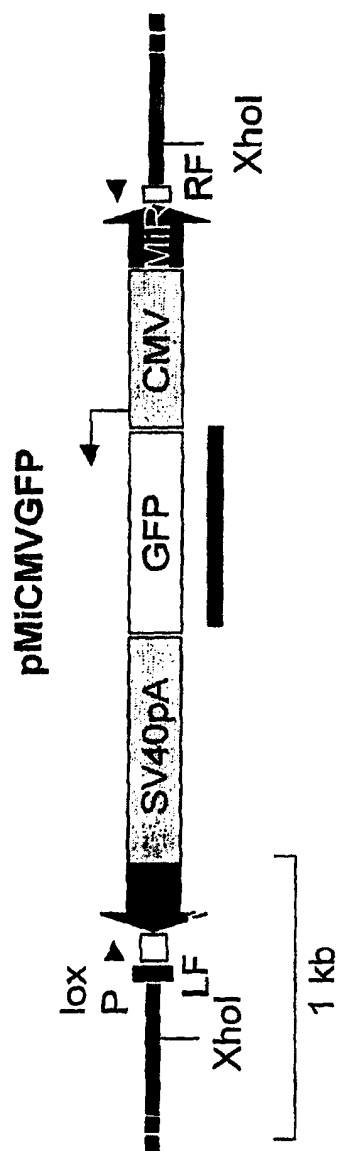


图 1: 用于诱导性转座的自我灭活自主转座子载体

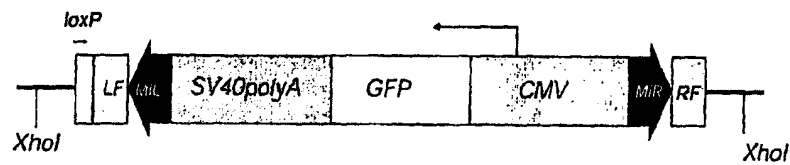









A



B

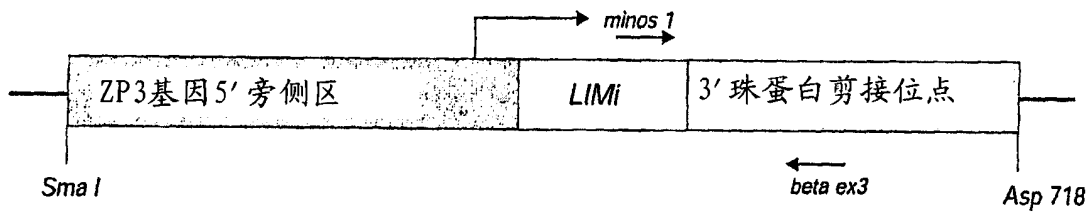


图 6

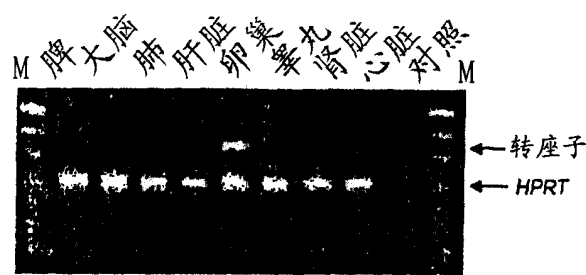


图 7

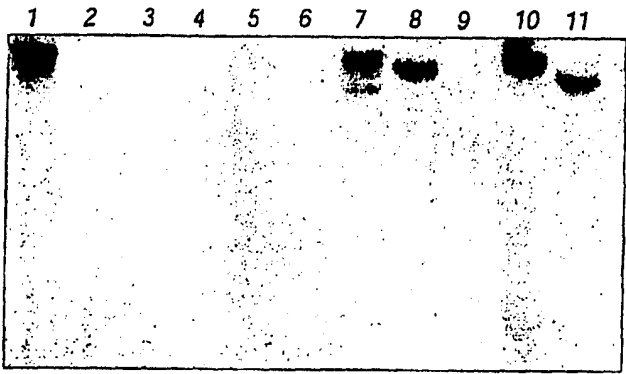


图 8

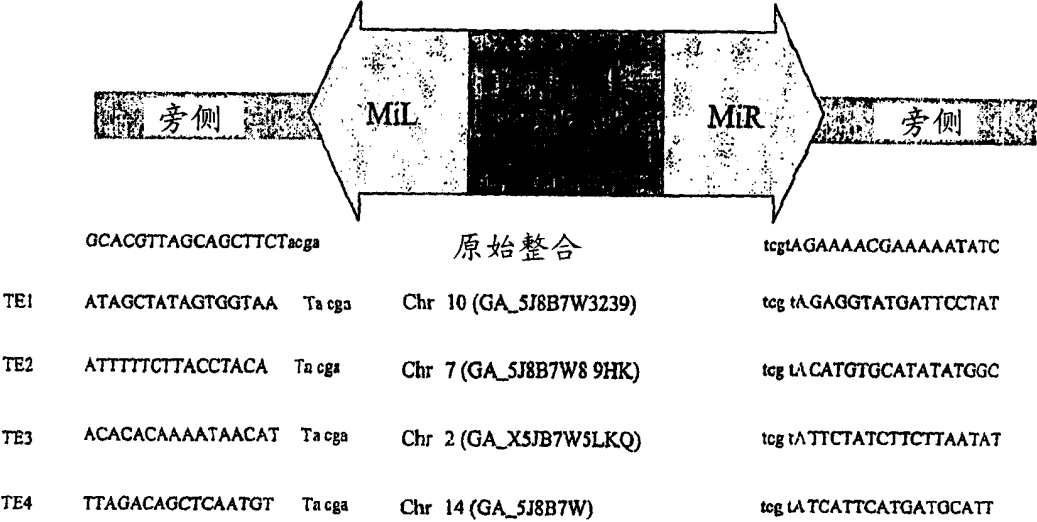


图 9

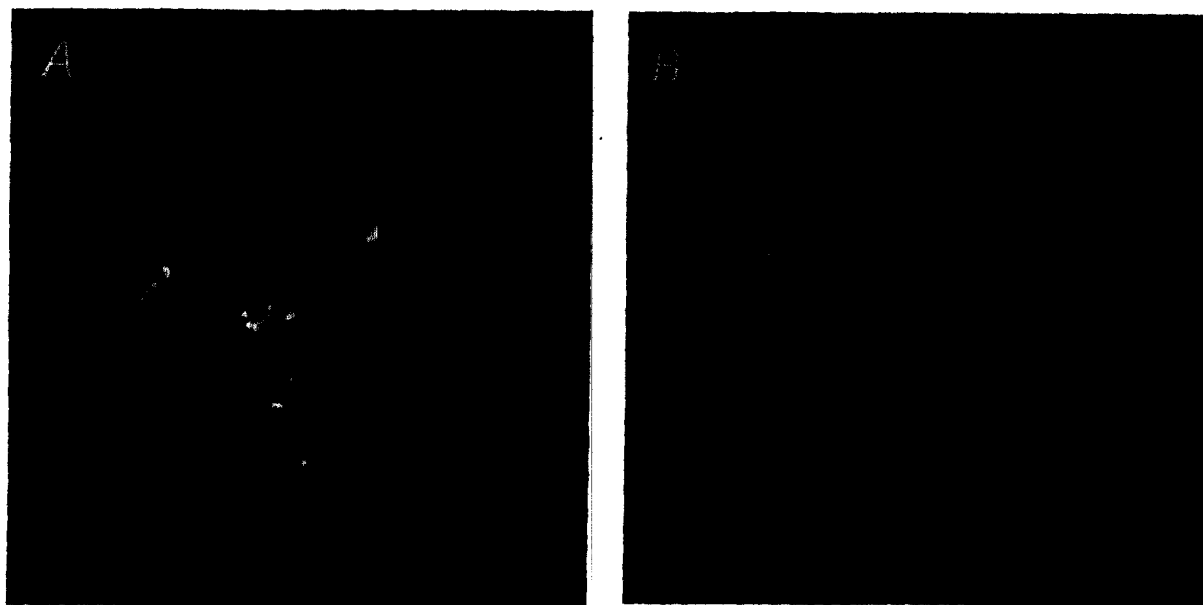


图 10

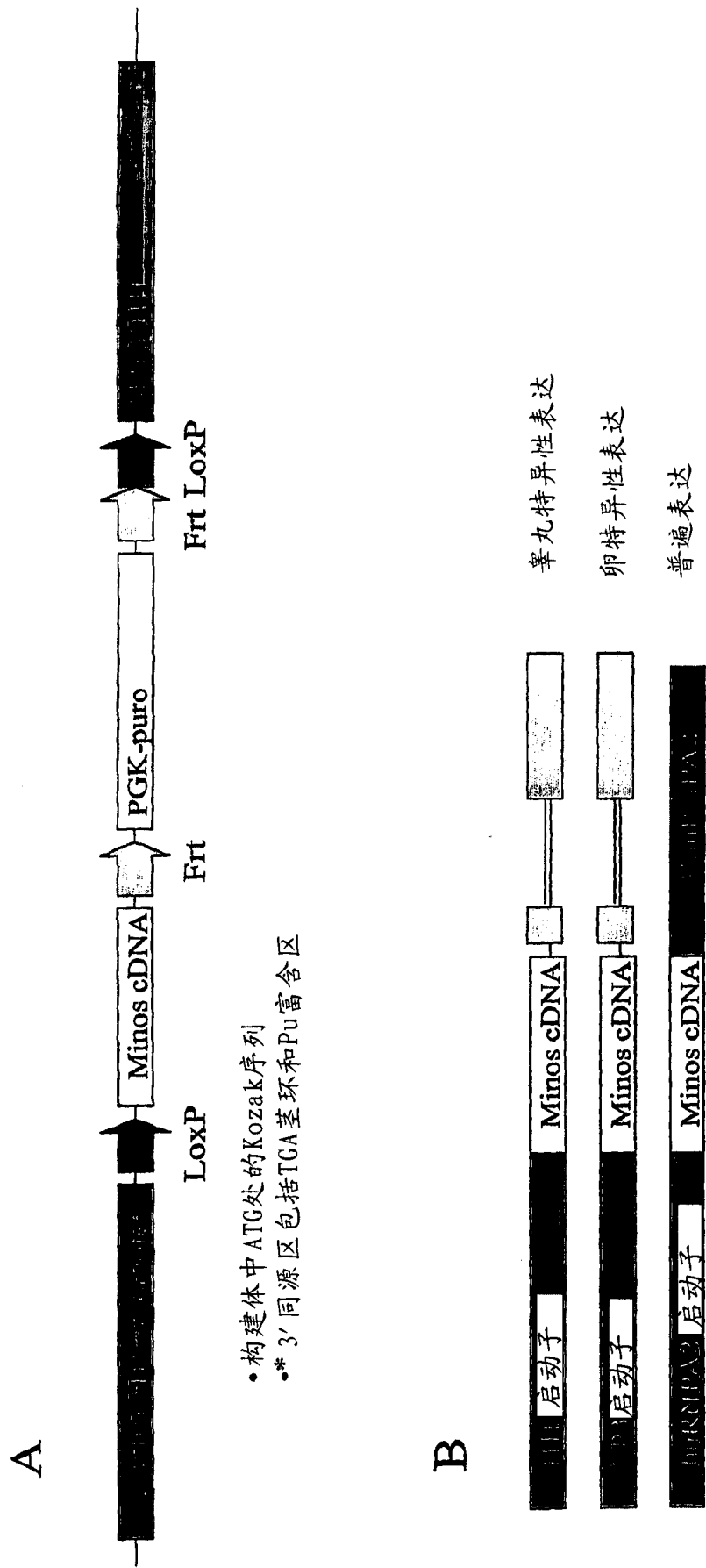


图 11

用于转基因小鼠的捕获构建体和/或向逆转录病毒质粒中的克隆

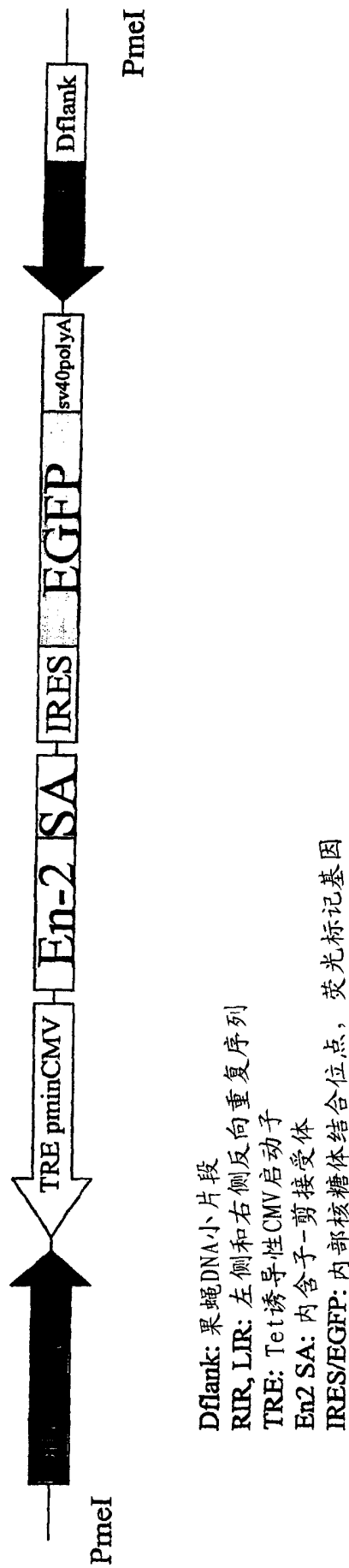


图 12

雄性动物提供转座子，雌性动物提供转座酶的体内转座技术。
转座发生在卵或早期胚胎中

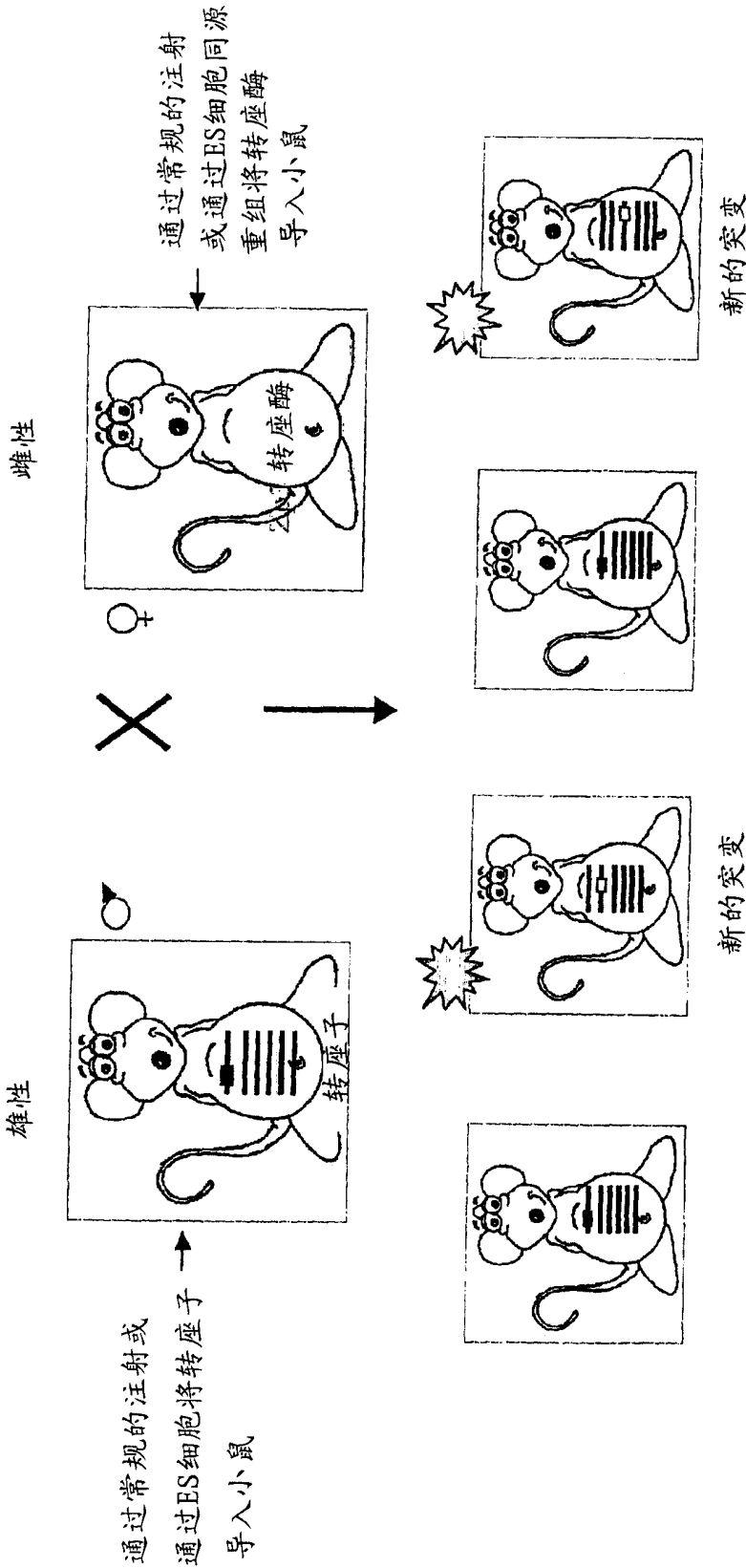


图 13

雌性动物提供转座子，雄性动物提供转座酶的体内转座技术。
转座发生在精子中

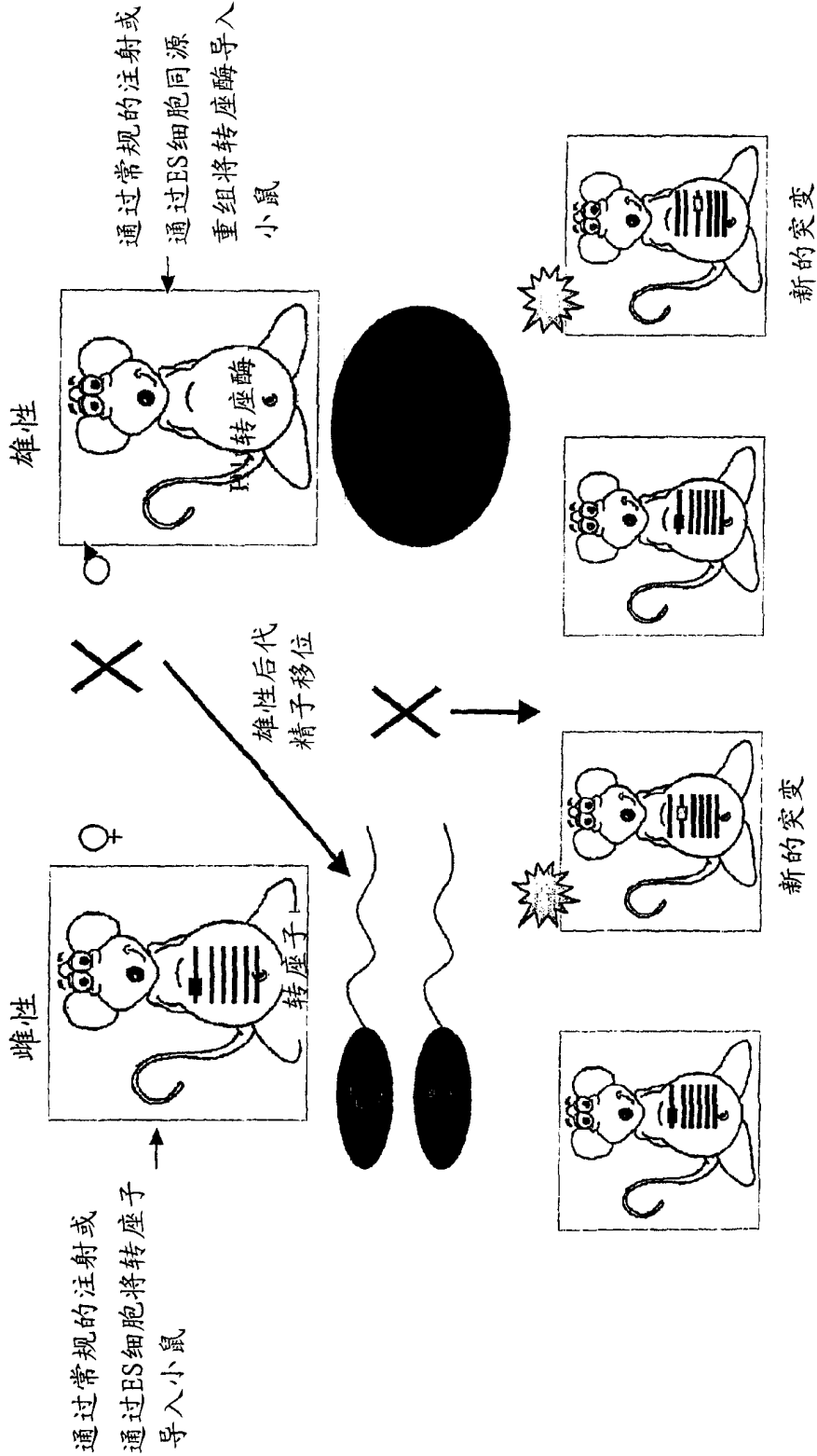


图 14

利用普遍表达的体内移位。
雄性动物和雌性动物在提供转座子或转座酶方面是可以相互交换的。

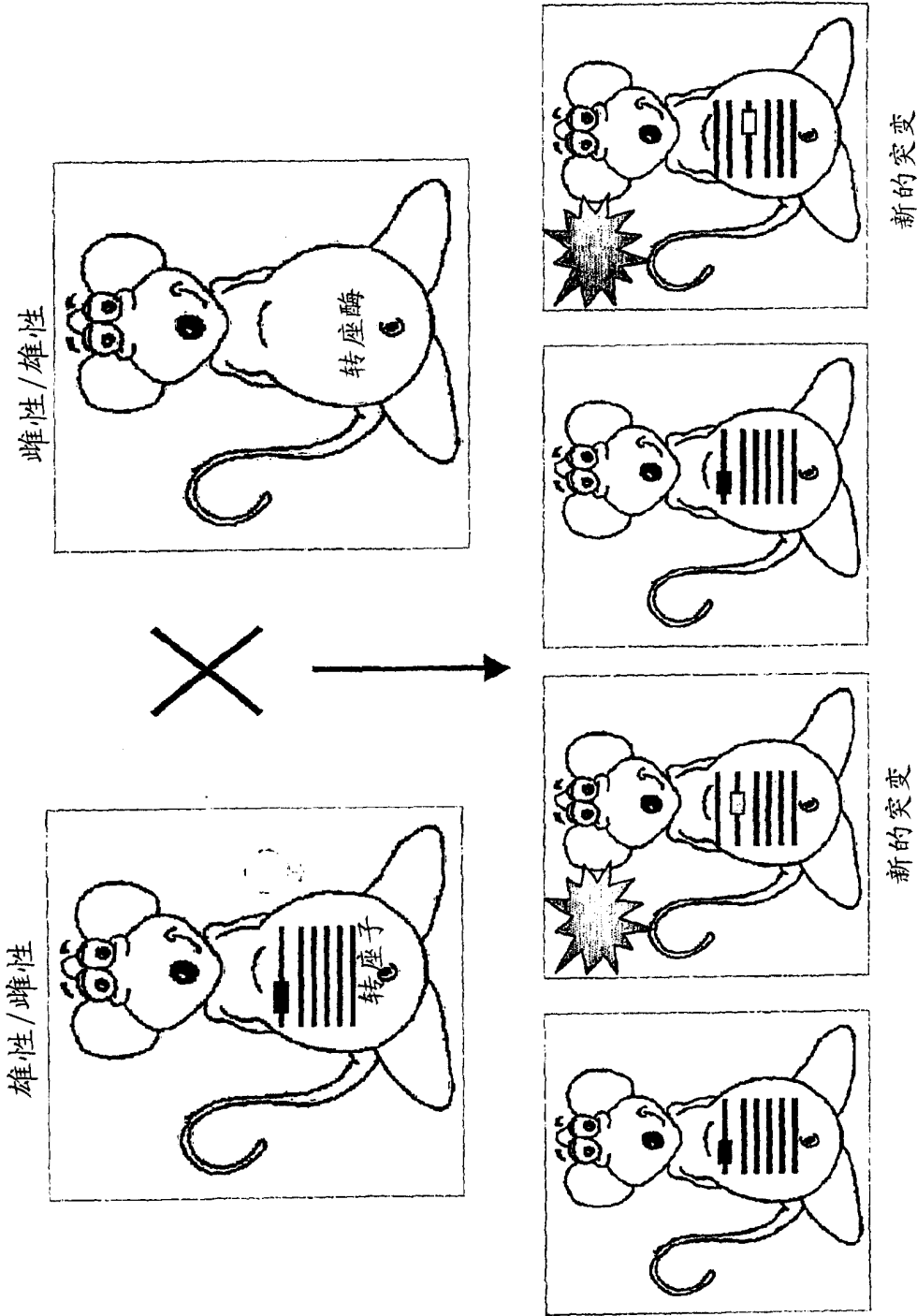


图 15