

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.  
C12Q 1/68 (2006.01)



## [12] 发明专利申请公开说明书

[21] 申请号 200480033877.5

[43] 公开日 2006年12月20日

[11] 公开号 CN 1882700A

[22] 申请日 2004.9.20

[21] 申请号 200480033877.5

[30] 优先权

[32] 2003.9.22 [33] US [31] 60/504,211

[86] 国际申请 PCT/IL2004/000866 2004.9.20

[87] 国际公布 WO2005/028674 英 2005.3.31

[85] 进入国家阶段日期 2006.5.17

[71] 申请人 特莱索根生物科技有限合伙公司

地址 以色列佩塔提克瓦

[72] 发明人 D·哈勒

[74] 专利代理机构 中国专利代理(香港)有限公司  
代理人 权陆军 梁 谋

权利要求书 4 页 说明书 52 页 序列表 69 页  
附图 5 页

[54] 发明名称

用于检测基因座拷贝数改变的方法和试剂盒

[57] 摘要

提供了鉴定基因座拷贝数改变的方法。通过测定该基因座中至少一种基因的甲基化状态实现该方法，其中与所述至少一种基因的预定甲基化状态不同的甲基化状态表明所述基因座拷贝数的改变。

1. 鉴定基因座拷贝数的改变的方法，所述方法包括确定所述基因座中至少一种基因的甲基化状态，其中与所述至少一种基因的预定甲基化状态不同的甲基化状态表明所述基因座拷贝数的改变。

2. 权利要求 1 的方法，其中所述基因座拷贝数的改变由选自非整倍性和多倍性的染色体畸变引起。

3. 权利要求 1 的方法，其中通过

- (i) 限制酶消化甲基化检测；
- (ii) 基于硫酸氢盐的甲基化检测；
- (iii) 质谱分析法分析；
- (iv) 序列分析；和/或
- (v) 微阵列分析

实现所述确定所述至少一种基因的甲基化状态。

4. 权利要求 1 的方法，其中所述基因座位于选自 1 号染色体、2 号染色体、3 号染色体、4 号染色体、5 号染色体、6 号染色体、7 号染色体、8 号染色体、9 号染色体、10 号染色体、11 号染色体、12 号染色体、13 号染色体、14 号染色体、15 号染色体、16 号染色体、17 号染色体、18 号染色体、19 号染色体、20 号染色体、21 号染色体、22 号染色体、X 染色体和 Y 染色体的染色体上。

5. 鉴定受试者中基因座拷贝数的改变的方法，所述方法包括：

- (a) 得到受试者的染色体 DNA；和
- (b) 确定所述染色体 DNA 的基因座上的至少一种基因的甲基化状态，其中与所述至少一种基因的预定甲基化状态不同的甲基化状态表明所述基因座拷贝数的改变，从而鉴定所述受试者中基因座拷贝数的改变。

6. 权利要求 5 的方法，其中通过

- (i) 限制酶消化甲基化检测；和
- (ii) 基于硫酸氢盐的甲基化检测；
- (iii) 序列分析；和/或
- (iv) 微阵列分析

实现所述确定所述至少一种基因的甲基化状态。

7. 权利要求 6 的方法，其中所述基因座位于选自 1 号染色体、2

号染色体、3号染色体、4号染色体、5号染色体、6号染色体、7号染色体、8号染色体、9号染色体、10号染色体、11号染色体、12号染色体、13号染色体、14号染色体、15号染色体、16号染色体、17号染色体、18号染色体、19号染色体、20号染色体、21号染色体、22号染色体、X染色体和Y染色体的染色体上。

8. 产前鉴定基因座拷贝数改变的方法，该方法包括：

(a) 得到胎儿或者胚胎的包括所述基因座的染色体 DNA；和

(b) 确定所述胎儿或者胚胎的包括所述基因座的染色体 DNA 中至少一种基因的甲基化状态，其中与所述至少一种基因的预定甲基化状态不同的甲基化状态表明所述基因座拷贝数的改变，从而产前鉴定所述基因座拷贝数的改变。

9. 权利要求 8 的方法，其中通过：

(i) 羊膜腔穿刺术；

(ii) 胎儿活组织检查；

(iii) 绒毛膜绒毛取样；和/或

(iv) 母体活组织检查

实现得到所述胎儿或者胚胎染色体 DNA。

10. 权利要求 8 的方法，其中通过

(i) 限制酶消化甲基化检测；和

(ii) 基于硫酸氢盐的甲基化检测；

(iii) 质谱分析法分析；

(iv) 序列分析；和/或

(v) 微阵列分析

实现所述确定所述至少一种基因的甲基化状态。

11. 产前诊断唐氏综合征的方法，该方法包括：

(a) 得到胎儿或者胚胎 21 号染色体；和

(b) 确定所述胎儿或者胚胎 21 号染色体中至少一种基因的甲基化状态，其中选择的所述至少一种基因不在唐氏综合征中扩增，而与预定甲基化状态不同的所述甲基化状态表明所述胎儿或者胚胎 21 号染色体的扩增，从而产前诊断唐氏综合征。

12. 权利要求 11 的方法，其中所述至少一种基因选自 APP 和胱硫醚- $\beta$ -合酶。

13. 权利要求 11 的方法，其中通过：

- (i) 羊膜腔穿刺术；
- (ii) 胎儿活组织检查；
- (iii) 绒毛膜绒毛取样；和/或
- (iv) 母体活组织检查

实现得到所述胎儿或者胚胎 21 号染色体。

14. 权利要求 11 的方法，其中通过

- (i) 限制酶消化甲基化检测；和
- (ii) 基于硫酸氢盐的甲基化检测；
- (iii) 质谱分析法分析；
- (iv) 序列分析；和/或
- (v) 微阵列分析

实现所述确定所述至少一种基因的甲基化状态。

15. 鉴定“与生命相容的”基因的方法，所述方法包括：

- (a) 确定所扩增的染色体序列区中多种基因的甲基化状态；和
- (b) 鉴定所述多种基因中显示出与预定甲基化状态不同的甲基化状态的基因，从而鉴定“与生命相容的”基因。

16. 权利要求 15 的方法，其中通过

- (i) 限制酶消化甲基化检测；和
- (ii) 基于硫酸氢盐的甲基化检测；
- (iii) 质谱分析法分析；
- (iv) 序列分析；和/或
- (v) 微阵列分析

实现所述确定所述至少一种基因的甲基化状态。

17. 鉴定“与生命相容的”基因的方法，所述方法包括：

- (a) 确定所扩增的染色体序列区中多种基因的表达水平；和
- (b) 鉴定所述多种基因中显示出低于预定阈值的表达水平的基因，从而鉴定“与生命相容的”基因。

18. 权利要求 17 的方法，其中在 mRNA 水平上实现所述确定所述多种基因的表达水平。

19. 权利要求 17 的方法，其中在蛋白质水平上实现所述确定所述多种基因的表达水平。

20. 制成品，其包含包装材料和所述包装材料中含有的用于检测基因座拷贝数改变的试剂，其中所述试剂能够确定所述基因座中至少一种基因的甲基化状态，而与所述至少一种基因的预定甲基化状态不同的甲基化状态表明所述基因座拷贝数的所述改变。

21. 权利要求 20 的制成品，其中所述基因座拷贝数的所述改变由选自非整倍性和多倍性的染色体畸变引起。

22. 用于鉴定基因座拷贝数改变的试剂盒，所述试剂盒包含用于确定所述基因座中至少一种基因的甲基化状态的试剂，所述至少一种基因选自 APP 和胱硫醚- $\beta$ -合酶，其中与所述至少一种基因的预定甲基化状态不同的甲基化状态表明所述基因座拷贝数的改变。

23. 权利要求 22 的试剂盒，其中所述基因座拷贝数的改变由选自非整倍性和多倍性的染色体畸变引起。

## 用于检测基因座拷贝数改变的方法和试剂盒

### 发明领域和背景

本发明涉及用于检测基因座拷贝数异常（例如，扩增）的方法和试剂盒，所述异常导致染色体异常，如三体性。

在其中遗传成分相对于环境因素占优势的疾病状态称作遗传病并且通常分成如下三类之一：(i)特征是一个或多个染色体的缺失、过多或者异常排列的疾病；(ii)孟德尔或者简单遗传的疾病，其主要由单个突变的基因导致并且细分成常染色体显性、常染色体隐性或者 X-连锁的类型；和(iii)多个基因和环境因素相互作用导致的多因子疾病。

非整倍性 (Aneuploidias) 是在 50% 以上流产儿中发现的最常见的染色体异常 [McConnell HD, Carr DH. Recent advances in the cytogenetic study of human spontaneous abortions. *Obstet Gynecol.* 1975 May; 45 (5): 547-52]。三体性在胎儿或者胚胎状态是致命的，而常染色体三体性是允许胎儿在出生后存活的三体性。

唐氏综合征也称作第 21 对染色体三体性，是可以在产前诊断的最常见的遗传病之一。它导致患有该疾病儿童中智力迟钝和许多身体和生理异常。许多患者生下来就患有先天性心脏缺损和胃肠异常，其可以通过手术矫正。身体特征包括扁平后脑勺、和斜眼、鼻桥凹陷、手和脚小、出生时颈后部的过多皮肤、肌肉紧张度降低和手掌中猿褶 [Down syndrome, (1994) National Down Syndrome Congress. Atlanta, GA: NDSC]。

唐氏综合征的患病率占美国中每 10,000 个活产中的 9.2 例。尽管唐氏综合征发生的原因仍然还未充分理解，但是已经充分确定较高的母亲年龄起一定作用。从而，对于 35 岁以上的母亲，怀上具有第 21 对染色体三体性胚胎的危险性呈指数上升。由于在美国生育的母亲年龄增长，所以有危险怀有在子宫中诊断为唐氏综合征的儿童的母亲的患病率比以前更高。因此，认为 35 岁以上的所有母亲都可能具有唐氏综合征的高度危险并且应该为其提供试验。

当前唐氏综合征的产前筛选方法是多样的并且包括，血清筛选、超声、侵入性试验、遗传咨询和染色体研究。已经进行了大量研究来

改进唐氏综合征的产前诊断，特别是在首三月中的诊断，但是迄今还没有被证明在唐氏综合征的诊断中 100% 准确的试验。

下面概述了当前用于唐氏综合征的产前筛选和诊断的方法。

### 非侵入性试验

胎儿的超声成像 - 该试验在妊娠的第 12 - 18 周进行。它寻找颈项透明层 (nuchal translucency) (即, 增加的颈项 (nuchal) 增厚或者肿胀)、缩短的长骨的长度和第一和第二趾之间的趾间间隔明显 (sandal gap)。但是可以理解用于检测胎儿三体性状况的超声检查的灵敏性随着染色体异常的类型、超声检查时孕龄、治疗安排的理由、阳性超声检查发现的标准和超声检查的质量而变。作为估计, 可以在 50% 到 70% 具有第 21 对染色体三体性 (唐氏综合征) 的胎儿中鉴定一个或多个超声检查发现。从而, 超声检查标记的存在或不存在可以显著改变胎儿唐氏综合征的危险性并且是遗传超声图的基础。因为母体生物化学和超声检查标记在很大程度上是独立的, 所以组合的危险估计导致比单独的危险估计更高的检测率。

母体血清筛选 - 母体血清筛选也称作多重标记筛选测试, 包括三重标记测试, 其检查血清甲胎蛋白 (AFP, 其低水平表明唐氏综合征); 人绒毛膜促性腺激素 (hCG, 其高水平表明唐氏综合征); 和未缀合的雌三醇 (uE3, 其低水平表明唐氏综合征)。最近已经加入第四种标记抑制素 A, 其高水平表明唐氏综合征诊断 [Wald, Watt, 和 Hackshaw, (1999) *The New England Journal of Medicine*, vol. 341, no. 7.461-469]。现在三重标记测试与抑制素 A 的加入得到了四重标记测试。这些标记与母亲年龄参数可以用于诊断唐氏综合征, 其检测率为约 70%, 假阳性率为约 5%。这些标记可以与在首三月中使用的 AFP 测试和超声一起用于诊断妊娠期的次三月中唐氏综合征。

所述四重测试现在与颈项透明层超声波检查和测试妊娠有关的血浆蛋白质-A (PAPP-A) 一起使用。该方法可以增加检测率到 85%, 具有 5% 假阳性率, 从而提供了当前可利用的对唐氏综合征的最可靠的非侵入性检测 [Wald, Kennard, Hackshaw 和 McGuire, (1998) *Health Technology Assessment*, vol 2, no. 1.1-124.]。然而, 应该指出, 当前可利用的血清标记提供统计学结果, 其是不确定的并且通常难以解

释。

### 侵入性测试

羊膜腔穿刺术 - 羊膜腔穿刺术是侵入性方法，其中在次三月中抽出羊水用于检测胎儿异常。推荐该测试用于高母亲年龄的妇女，她们怀有患有诸如唐氏综合征的遗传异常的孩子的危险性更大。羊膜腔穿刺术的治疗安排可以包括非常低或高水平的 AFP。通常在次三月中进行羊膜腔穿刺术，但是其可以早在妊娠的第 11 周进行。在妊娠的约 16 周取羊水样品。由于仅 20% 羊水细胞适于测试，所以需要培养样品以得到足够的分裂细胞用于分裂中期分析。因此，1-3 周后可以得到结果，其可以导致母亲焦虑增加，和考虑次三月和第三个三月的终止。核型分析检测染色体病而不是唐氏综合征。然而，200 个妊娠中的约 1 个由于羊膜腔穿刺术导致流产。

绒毛膜取样 - 绒毛膜取样包括采集绒毛膜样品，其形成胎盘并且由胎儿形成，因此含有胎儿细胞。该测试可以在首三月结束时进行（即 10-12 周）。经子宫颈或者经腹进行该方法。两种方法一样安全和有效。该方法快速（在 24 小时以内得到结果）并且有很小或无疼痛。然后在显微镜下分析样品（即，未培养的样品），特别寻找染色体异常。CVS 的优点是在首三月内早期检测，和母亲细胞污染的危险减小。缺点是流产的危险增加和成本。仍然重要的是查看母体血清标记，尽管查看 AFP 时，太晚而不能进行 CVS。阳性结果以 60-70% 的比率检测遗传疾病，如唐氏综合征。将明白 1% CVS 显示出限制的胎盘镶嵌性，其中从直接或培养的 CVS 得到的结果与胎儿的结果不同。所培养的 CVS 从比直接 CVS 更接近胎儿系的细胞生长而来，所述直接 CVS 更接近胎盘。流产的危险性高于羊膜腔穿刺术的危险。此外，CVS 期间腿和手截断（amputation）的危险相对较高。

未培养的羊水细胞的分裂间期荧光原位杂交（FISH） - 可以使用荧光原位杂交（FISH）分析羊水的载玻片。该测试对未培养的分裂间期细胞进行并且可以检测许多染色体异常。结果在 24 小时内得到。来自染色体 21 标准区的探针用于诊断唐氏综合征。另一种探针用于测试倍性。对于一些易位，由于两种信号可以重叠，所以探针位置可以导致假阴性结果。



定量聚合酶链反应 (PCR) 诊断 - 该方法已经证明可用于唐氏综合征中不分离的研究。通常使用的是 21 号染色体内的多态性 (GT)<sub>n</sub> 重复序列和 Alu 序列。[Petersen (1991) *Am J Hum Genet*, 48: 65-71; Celi (1994); Messari (1996) *Hum Genet*, 97: 150-155]。从而, 例如, 可以使用通过子宫颈内管冲洗从妊娠的 7 到 9 周得到的经子宫颈细胞 (TCC) 样品得到的胎儿 DNA。滋养层回收对于 Y 染色体特异的 DNA 序列的 PCR 扩增和父亲 - 特异的微卫星等位基因的检测是足够的。该方法准确地预测胎儿性别。使用超氧化物歧化酶 - 1 (SOD 1) 的荧光原位杂交 (FISH) 和半定量 PCR 分析诊断 TCC 中的第 21 对染色体三体性胎儿。后来, 阐明了使用定量荧光聚合酶链反应 (PCR) 同时诊断第 21 对染色体和第 18 对染色体三体性以及检测来自 X 和 Y 染色体的 DNA 序列。通过定量荧光 PCR 扩增从羊水、胎儿血液或者组织提取的 DNA 样品以检测对 21 和 18 号染色体每一种上两个等位基因特异的多态性小串联重复序列 (STR)。扩增产物的定量分析允许诊断第 21 对染色体和第 18 对染色体三体性, 而使用来自 X 和 Y 染色体的 DNA 序列的 PCR 扩增同时进行性别鉴定。使用用于检测第 21 对染色体染色体三体性的两组 STR 标记证实了用于所选染色体异常的快速产前诊断的定量荧光多重 PCR 的有用性 [Pertl *Obstet Gynecol.* (2001) Sep; 98 (3): 483-90]。

在另一研究中, 从过多的羊水提取 DNA 并在基于荧光的 PCR 反应中扩增, 使用位于 21 号染色体上的三种小串联重复序列标记。在 DNA 测序仪上分析反应产物以鉴定 21 号染色体的两个或三个拷贝的存在。使用该方法, 用三种标记实现了共 99.6% 提供信息的结果 (Verma 1998)。还对非多态性靶基因通过荧光 PCR 产物进行了染色体定量分析。Rahil 等人 (2002) 设置了 DSCR1 (唐氏综合征关键区 1)、DCC (在结肠直肠癌中缺失) 和 RB1 (成视网膜细胞瘤 1) 的部分的共同扩增, 分别允许 21、18 和 13 号染色体的非整倍性的分子检测。在 400 种羊水的盲法前瞻性研究中进行了定量分析。对所有样品进行随访核型分析, 并且分子结果与细胞遗传学数据一致, 没有假阳性或者假阴性结果。从而, 使用 PCR 对胎儿 DNA 通过染色体定量进行非整倍性诊断是有效且可靠的方法。然而, 这些方法对胎儿 DNA 纯度非常敏感, 因为母亲 DNA 可能掩盖染色体定量。

单个细胞中非整倍性的检测 - 该方法用于植入前遗传诊断。从裂解的单个细胞得到 DNA 并使用简并寡核苷酸引发的 PCR(DOP-PCR) 进行扩增。将产物用切口平移进行标记并与正常参照基因组 DNA 一起杂交。用比较基因组杂交 (CGH) 荧光比率图确定非整倍性, 截断阈值为 0.75 和 1.25。使用该技术分析已知对 13、18 或 21 号染色体为三倍性的单个细胞[Voullaire 等人(1999), Tabet (2001), Rigola 等人(2001)]。

指纹法系统是进行植入前遗传诊断的另一种方法。选择具有高杂合性、已知的等位基因大小范围和最小的 PCR 打滑假象的四核苷酸微卫星标记用于 X、13、18 和 21 号染色体并以多重荧光 (FL) - PCR 形式最优化[Katz 等人(2002) Hum Reprod. 17 (3): 752-9]。然而, 这些方法对于体外受精受限, 因为从母亲血清分离胎儿细胞纯的级分需要当前还不可利用的技术方法。

母体循环中的胎儿细胞 - 该技术的主要优点是它是非侵入性的并且因此该方法本身对妊娠无危险。可能可以比 CVS 更早进行, 因为在第 5 周就已经检测到胎儿 DNA。

在母体循环中仅发现少许胎儿细胞 (滋养层、淋巴细胞和具核红细胞), 因此需要选择和富集这些细胞。富集技术包括流式/磁分选, 和双密度离心。有约 1-2 个胎儿细胞/1 千万母体细胞, 并且 50% 胎儿细胞将不适于核型分析。特别地, 淋巴细胞不适于用于该技术, 因为此类细胞在母体循环中保持几年的持续时间, 并且因此结果可能受到以前妊娠的影响。与常规核型分析相比, 该方法仅检查单个染色体。

a) FISH 可以用于检查尽可能多细胞中的信号数/细胞以得到具有 3 种信号的细胞的比例。探针的杂交效率可以巨大地影响所见到的信号数(从而偏移结果)。

b) 引发的原位标记 (PRINS) 基于特定和未标记的 DNA 引物与互补基因组位点的原位退火和随后通过 PCR 掺入经标记的核苷酸进行的延伸。

诊断唐氏综合征的其他方法包括在 10 周采集并且需要培养和核型分析和子宫腔灌洗/经子宫颈细胞取样的体腔液。后者比羊膜腔穿刺术或者 CVS 的侵入性低。它在 7-9 周进行并且涉及收集从胎盘丧失的细胞, 从而类似于直接 CVS。然而, 该方法使母亲受到污染和感染。

从而，通常的和具体的唐氏综合征中染色体异常（即，三体性）的产前诊断是复杂的，需要杰出的技术技能，不完全有效并且可能导致妊娠失败。由于对于唐氏综合征没有确定的产前测试这一事实，所以结束健康胎儿妊娠的危险很高。

因此广泛认识到需要检测导致染色体异常的基因座扩增并且没有上面的局限的方法，并且拥有该方法将是非常有利的。

### 发明概述

根据本发明的一个方面，提供了鉴定基因座拷贝数改变的方法，该方法包括确定所述基因座中至少一种基因的甲基化状态，其中与所述至少一种基因的预定甲基化状态不同的甲基化状态表明基因座拷贝数的改变。

根据在下述本发明的优选实施方案中的其他特征，基因座拷贝数的改变由选自非整倍性和多倍性的染色体畸变导致。

根据本发明的另一方面，提供了鉴定受试者中基因座拷贝数改变的方法，该方法包括：(a)得到受试者的染色体 DNA；和 (b) 确定所述染色体 DNA 的基因座上至少一种基因的甲基化状态，其中与所述至少一种基因的预定甲基化状态不同的甲基化状态表明基因座拷贝数的改变，从而鉴定受试者中基因座拷贝数的改变。

根据本发明的另一方面，提供了产前鉴定基因座拷贝数改变的方法，该方法包括：(a) 得到胎儿或者胚胎的包括所述基因座的染色体 DNA；和 (b) 确定胎儿或者胚胎的包括所述基因座的染色体 DNA 中至少一种基因的甲基化状态，其中与所述至少一种基因的预定甲基化状态不同的甲基化状态表明基因座拷贝数的改变，从而产前鉴定基因座拷贝数的改变。

根据所述优选实施方案中的其他特征，得到胎儿或者胚胎染色体 DNA 通过：(i) 羊膜腔穿刺术；(ii) 胎儿活组织检查；(iii) 绒毛膜绒毛取样；和/或母体活组织检查来实现。根据所述优选实施方案中的其他特征，确定至少一种基因的甲基化状态通过：(i) 限制酶消化甲基化检测；(ii) 基于硫酸氢盐的甲基化检测；(iii) 质谱分析法分析；(iv) 序列分析；和/或(v) 微阵列分析来实现。

根据本发明的再一个方面，提供了产前诊断唐氏综合征的方法，

该方法包括：(a) 得到胎儿或者胚胎 21 号染色体；和 (b) 确定胎儿或者胚胎 21 号染色体中至少一种基因的甲基化状态，其中选择的所述至少一种基因不在唐氏综合征中扩增，而与预定甲基化状态不同的甲基化状态表明胎儿或者胚胎 21 号染色体的扩增，从而产前诊断唐氏综合征。

根据所述优选实施方案中的其他特征，所述至少一种基因选自 APP 和胱硫醚- $\beta$ -合酶。

根据本发明的额外方面，提供了鉴定“与生命相容的”基因的方法，该方法包括：(a) 确定所扩增的染色体序列区中多种基因的甲基化状态；和 (b) 鉴定所述多种基因中显示出与预定甲基化状态不同的甲基化状态的基因，从而鉴定“与生命相容的”基因。

根据本发明的再一个额外方面，提供了鉴定“与生命相容的”基因的方法，该方法包括：(a) 确定所扩增的染色体序列区中多种基因的表达水平；和 (b) 鉴定所述多种基因中显示出低于预定阈值的表达水平的基因，从而鉴定“与生命相容的”基因。

根据所述优选实施方案中的其他特征，在 mRNA 水平上实现确定多种基因的表达水平。

根据所述优选实施方案中的其他特征，在蛋白质水平上实现确定多种基因的表达水平。

根据本发明的额外方面，提供了制成品，其包含包装材料和所述包装材料中含有的用于检测基因座拷贝数改变的试剂，其中所述试剂能够确定所述基因座中至少一种基因的甲基化状态，而与所述至少一种基因的预定甲基化状态不同的甲基化状态表明基因座拷贝数的改变。

根据本发明的另一方面，提供了用于鉴定基因座拷贝数改变的试剂盒，所述试剂盒包含用于确定所述基因座中至少一种基因的甲基化状态的试剂，所述至少一种基因选自 APP 和胱硫醚- $\beta$ -合酶，其中与所述至少一种基因的预定甲基化状态不同的甲基化状态表明基因座拷贝数的改变。

根据所述优选实施方案中其他特征，选自非整倍性和多倍性的染色体畸变导致基因座拷贝数的改变。本发明通过提供用于鉴定基因座扩增的方法和试剂盒成功地解决了当前已知的结构的缺点。

除非另外定义，用于本文的所有技术和科学术语具有与本发明所述领域技术人员通常理解的相同的含义。尽管类似或者等价于本文描述的方法和材料的方法和材料可以用于实践或者检验本发明，但是下面描述适宜的方法和材料。对于有冲突的情况，以本专利说明书，包括定义为准。此外，所述材料、方法和实施例仅用于阐明而不意在限制。

### 附图简述

仅通过实例，参考附图描述本发明。现在具体地详细参考附图，要强调的是所示细节作为实例并且仅用于本发明优选实施方案的阐明性讨论的目的，并且是为了提供被认为是对于本发明的原理和概念方面最有用和最容易理解的说明而提供的。在这点上，仅仅为了对本发明的基本理解所需要而更详细地显示本发明的结构细节，说明书与附图一起使得本领域技术人员清楚怎样可以在实践中体现本发明的几种形式。

在附图中：

图 1a 是 APP 启动子的扩增产物的核苷酸序列，其从启动子区延伸到人 APP 区的第一个外显子。+1 指转录起始位点。用作引物 1 (SEQ ID NO: 1) 和 2 (SEQ ID NO: 2) 的序列加双下划线。9bp 长的富含 GC 的元件的 6 个拷贝加下划线。C 上的点指出所扩增启动子区 (-251 到 +22) 中 CpG 双联体中的胞嘧啶。

图 1b 是用于检测图 1a 中给出的 DNA 序列的甲基化状态的引物的核苷酸序列。引物 1 (a-b) - 表示分别在磺化作用之后或之前的引物 1 (SEQ ID NO: 1, APP-F) 的序列；引物 2c-e- 表示磺化作用后 (c)，反义取向的 (d) 或者磺化作用前 (e) 引物 2 (SEQ ID NO: 2, APP-R) 的序列。

图 2a-b 是雄激素受体外显子 1 的天然 (图 2a) 和亚硫酸氢盐修饰的 (图 2b) 序列的核苷酸序列。图 2a-# 表示正向引物的位置；# 表示反向引物的位置；\* 表示 HpaII 位点；\*\* 表示 HhaI 位点。图 2b- 绿色加亮区指出 CpG 岛；粉红色下划线 - 指出 CpG 位点；(#) 指出 AR-F-1 (SEQ ID NO: 60) 的位置；(\*) 指出 AR-F-34 引物 (SEQ ID NO: 61) 的位置；(\*\*) 指出 AR-R-282 引物 (SEQ ID NO: 62) 的位置。

图 3 是琼脂糖凝胶照片，显示了雄性、雌性和 Klinefelter 综合征侵袭的受试者中雄激素受体甲基化状态的基于限制酶分析的产物。泳道 1 - DNA 标记；泳道 2 - 阴性对照；泳道 3 - XX 未切割的；泳道 4 - XY 未切割的；泳道 5 - XY 未切割的；泳道 6 - X 染色体三体性未切割的；泳道 7 - XX 切割的；泳道 8 - XY 切割的；泳道 9 - XY 切割的；泳道 10 - X 染色体三体性切割的。

图 4a-b 是天然的 (图 4a) 和亚硫酸氢盐修饰的 (图 4b) DSCAM 启动子的核苷酸序列。( # ) - 表示正向引物的位置；( \$ ) 表示反向引物的位置；绿色加亮区指出 CpG 岛。

图 5a-b 是天然的 (图 5a) 和亚硫酸氢盐修饰的 (图 5b) IFNAR1 启动子的核苷酸序列。绿色加亮区指出 CpG 岛。( \* ) 指出 IFNR-f4-bis (SEQ ID NO: 247) 的位置；( \*\* ) 指出 IFNR-nes-f-bis (SEQ ID NO: 249) 的位置；( \*\*\* ) 指出 IFNR-r4-bis (SEQ ID NO: 248) 的位置。

### 优选实施方案描述

本发明为可以用于鉴定导致染色体异常的基因座拷贝数异常的方法和试剂盒。具体地，本发明可以用于产前检测基因座扩增，如三体性。

参考附图和所附描述可以更好地理解本发明的原理和操作。

在详细解释本发明的至少一个实施方案前，将理解本发明的应用不限于下面描述中给出或者通过实施例例证的细节。本发明能够有其他实施方案或者以多种方式实践或者实施。而且，将理解本文所用的措词和术语用于描述的目的并且不应认为是作为限定。

遗传病是最通常由染色体数的变异，如非整倍性、整倍性和多倍性导致的病理状况。染色体数或者其部分中此类变异通常对于胚胎或者胎儿是致死的。第 21 对染色体三体性 (唐氏综合征)、第 18 对染色体三体性 (Edward 氏综合征)、第 13 对染色体三体性 (Patau 综合征) 和性染色体三体性是仅有的活产常染色体三体性。与第 21 对染色体三体性相反，第 13 对染色体和第 18 对染色体三体性疾病倾向于具有更严重的临床表现并且很少有患病的婴儿活过一周岁。在患有三体性疾病的胎儿中存在多种异常，但是没有给定三体性典型的单一异常。相反，存在一组提示特定诊断的特征性临床发现。此外，因为

这些患者的一些可能对于三体性细胞系是镶嵌的，所以可能具有多种表型。

迄今为止，对于任何三体性疾病还没有特定治疗、疗法或者治愈。由于这些原因，通常非常需要染色体异常，且尤其三体性的早期产前诊断。

用于三体性的产前诊断的当前可利用的方法包括羊水细胞或者绒毛膜细胞的超声检查和细胞遗传学分析。尽管超声检查受到高假阳性率的限制，但是侵入性测试不完全有效，需要高技术技能并且可能导致妊娠失败。备选地，母体血浆中循环的胎儿 DNA 的诊断用途当前限于在胎儿而不是母亲中发现的基因或者突变。

如在下面的实施例部分中进一步描述的，在寻找染色体畸变的新的诊断方式时，本发明人发现由于某些基因的沉默，常染色体三体性或者单体性允许出生后存活，其中所述基因的超表达与生命不相容。

DNA 甲基化是一种可逆机制，通过 DNA 甲基化，基因表达在原核和真核生物中沉默。通过甲基转移酶将甲基加到特别是启动子序列区中胞嘧啶嘧啶环的 5 位碳的能力，实现了该水平的基因表达控制 [Adams (1995) *Bioessays* 17 (2): 139-45]。真核细胞中甲基化的序列通常是无活性的 [Gold 和 Pedersen (1994)]。

已经清楚地证明了异常 DNA 甲基化是癌症中的普遍现象，并且可能是肿瘤发生期间发生的最早改变 [Stirzaker (1997) *Cancer Res.* 57 (11): 2229-37]。已经表明 DNA 甲基化在基因印记、胚胎发育、X-染色体沉默和细胞周期调节中起重要作用 [Costello (2001) *J. Med. Genet.* 38 (5): 285-303]。建立基因甲基化的正常模式的失败是许多遗传病的病因，所述遗传病包括 Rett 综合征（智力迟钝的主要形式）、Prader-Willi 综合征、Angelman 氏综合征、ICF 综合征和 Beckwith-Wiedmann 综合征。

考虑到 DNA 甲基化在基因沉默中的重要作用，很容易想象相同机制应用于如下基因的沉默，所述基因的过表达是致死的（即不与生命相容），从而提示确定基因甲基化状态可以用于检测基因座扩增。

实际上，尽管负责唐氏综合征（即，智力迟钝、先天性心脏病等等）的临床表型的 21 号染色体上的基因在 DS 患者中以三体性水平表达，但是如通过微阵列分析的，来自唐氏综合征患者中 21 号染色体

的基因的总基因表达中没有显著差异[Gross SJ, Ferreira JC, Morrow B, Dar P, Funke B, Khabele D, Merkatz I. Gene expression profile of trisomy 21 placentas: a potential approach for designing noninvasive techniques of prenatal diagnosis. Am J Obstet Gynecol. 2002 Aug; 187 (2): 457-62].

这些发现提示 DNA 甲基化沉默 21 号染色体上额外拷贝上的重要基因。Kuramitsu 和同事的发现进一步证实了该假设，他们显示唐氏综合征患者中 21 号染色体的 h2-钙调蛋白基因由于 21 号染色体的三个拷贝之一中甲基化而没有超表达[Kuromitsu (1997) Mol. Cell Biol. 2: 707-12].

该新近鉴定的基因座拷贝数的改变和甲基化状态之间的连锁首次允许使用容易执行、节省成本并且对个体受试者造成最小或者无危险的分子生物学技术有效检测染色体畸变。

从而，根据本发明的一方面，提供了鉴定基因座拷贝数改变的方法。

本文所用术语“基因座”指染色体上基因的位置或定位。根据本发明的该方面的方法可以检测位于 1-22 号、X 和 Y 染色体上的基因座的获得、基因座扩增或者基因座的丧失。

本文所用短语“基因座扩增”指基因座拷贝数增加。根据本发明的该方面的基因座扩增和基因座缺乏可以由染色体结构的改变（例如，重复、倒位、易位、缺失、插入）和/或染色体数 ( $>2n$ ) 或者其部分（也称作染色体标记）的增加或减少引起。染色体数的改变可以是非整倍性性质的，包括一个或多个染色体的获得或丧失，但不是整个组的染色体的获得或丧失（例如，三体性和四体性）。备选地，基因座扩增可以由多倍性引起，其中存在三个或更多个完整的染色体组。

将明白仅在身体的某些细胞类型中发生的染色体数的改变（即，镶嵌性）也可以根据本发明的该方面进行检测[Modi D, Berde P, Bhartiya D. Down syndrome : a study of chromosomal mosaicism.Reprod Biomed Online. 2003 Jun; 6 (4): 499-503].

通过确定基因座中至少一种基因的甲基化状态（即，甲基化模式和/或水平）实现根据本发明的该方面的方法。与至少一种基因的预定



甲基化状态不同的甲基化状态表明基因座拷贝数的改变。

本文所用的“预定的甲基化状态”指从未扩增的基因座得到的相同基因的甲基化状态，优选为相同发育状态的。

从而，上述基因座中所述至少一种基因的至少一个等位基因的甲基化状态的改变（即，模式和/或增加的水平）表明根据本发明的该方面的基因座拷贝数的改变。

通常，人 DNA 的甲基化在包括相邻鸟嘌呤和胞嘧啶的二核苷酸序列上发生，其中胞嘧啶位于鸟嘌呤的 5'（也称作 CpG 二核苷酸序列）。CpG 二核苷酸内的大多数胞嘧啶在人基因组中甲基化，然而，一些在特定富含 CpG 二核苷酸的基因组区中保持未甲基化，称作 CpG 岛[见 Antequera, F. 等人, *Cell* 62: 503-514 (1990)]。“CpG 岛”是富含 CpG 二核苷酸的区，其中 CpG 二核苷酸组成 DNA 序列的至少 50%。

因此，根据本发明的该方面的甲基化状态通常在优选启动子区的 CpG 岛中确定。但是将明白人基因组中其他序列易受 DNA 甲基化，如 CpA 和 CpT[见, Ramsahoye (2000) *Proc. Natl. Acad. Sci. USA* 97: 5237-5242; Salmon 和 Kaye(1970)*Biochim. Biophys. Acta.* 204: 340-351; Grafstrom (1985) *Nucleic Acids Res.* 13: 2827-2842; Nyce (1986) *Nucleic Acids Res.* 14: 4353-4367; Woodcock (1987) *Biochem. Biophys. Res. Commun.* 145: 888-894]。

如上文提到的，确定基因座中至少一种基因的甲基化状态。实施例部分的实施例 1-4 中列出了可以用于确定 X、9 和 21 号染色体的扩增的许多基因。

优选地，根据基因的表达模式选择至少一种基因。从而，确定基因的甲基化，所述基因座经扩增但是没有显示出表达改变，即，与仅两个基因拷贝相容的表达模式。此类基因的实例在下面的表 1 中列出。

表 1

基因名称	染色体	位置
RASSF1-Ras 结合 (RalGDS/AF-6) 结构域家族 1	3	3p21.3
配对框 5; 配对框同源异型基因 5 (B-细胞谱系特异性激活蛋白); B-细胞谱系特异性激活蛋白	9	9p13
组织因子途径抑制剂 2	7	7q22
ARHI, ras 同系物 I	1	1p13
FHIT 脆性组氨酸三联体基因; 双 (5'-腺苷基) - 三磷酸酶; 二核苷三磷酸酶; 二腺苷 5',5'''-P1,P3 - 三磷酸水解酶; AP3A 水解酶	3	3p14.2
VHL	3	3p26-p25
OPCML 结合阿片样物质的细胞粘着分子前体; 结合阿片样物质的蛋白质/细胞粘着分子样; 结合阿片剂的细胞粘着分子	11	11q25
具有分叉头和环指结构域的 CHFR 关卡	12	12q24.33
Semaphorin 3B	3	3p21.3
MLH1 MutL 蛋白质同系物 1	3	3p21.3
COX2 前列腺素 - 内过氧化物 (endoperoxide) 合酶 - 2 前体; 前列腺素 G/H 合酶和环加氧酶	1	1q25.2-q25.3
MGMT O-6-甲基鸟嘌呤 - DNA 甲基转移酶	10	10q26
视黄酸受体 $\beta$	3	3p24.1
PTEN 磷酸酶和张力蛋白同系物; 在多种晚期癌症中的突变 1	10	10q23.3
RASSF1A	3	3p21.3
APC 腺瘤性结肠息肉病	5	5q21-q22
P15-CDKN2B	9	9p21
BLu 蛋白质	3	
CDH1 钙粘着蛋白 1, 1 型, E-钙粘着蛋白 (上皮的)	16	16q22.1

TIMP-3 金属蛋白酶的抑制剂 - 3	22	22q12.3
GSN-凝溶胶蛋白	9	9q33
p14-p14ARF-依赖细胞周期蛋白的激酶抑制剂 2A	9	9p21
CDKN1C-依赖细胞周期蛋白的激酶抑制剂 1C	11	11p15.5
LOT1-多态腺瘤基因样 1	6	6q24-25
PIK3CG-磷酸肌醇 - 3 - 激酶, 催化性, $\gamma$ 多肽	7	7q22.2
TSLC1-免疫球蛋白超家族, 成员 4	11	11q23.2
RBI-成视网膜细胞瘤 1	13	13q14.2
具有分叉头和环指结构域的 Chfr 关卡	12	12q24.33
HTERT-端粒酶逆转录酶	5	5p15.33
MYO18B-肌球蛋白 XVIII B	22	22q12.1
CASP8-天冬氨酸特异性半胱氨酸蛋白酶 - 8	2	2q33-q34
hSNF5/INI1-染色质的 SWI/SNF 相关的、基质结合的、肌动蛋白依赖性调节物, 亚家族 b, 成员 1; 蔗糖非发酵的, 酵母, 同系物 - 样 1; 整合酶 interactor 1; SWI/SNF 相关的、基质结合的、肌动蛋白依赖性调节物	22	22q11.23
HIC1-在癌症中超甲基化	17	17p13.3

确定基因表达的方法是本领域公知的。实例包括但不限于基于 RNA 的方法, 包括使用寡核苷酸的基于杂交的技术 (例如, RNA 印迹、PCR、RT-PCR、RNA 酶保护、原位杂交、引物延伸、微阵列分析和斑点印迹分析), 或者基于蛋白质的方法, 如层析法, 电泳, 免疫检测测定, 如 ELISA 和蛋白质印迹分析、免疫组织化学等等, 其可以使用特异抗体实现。进一步的技术细节见可以在 <http://www.protocol-online.org/> 得到的实验室参考书, 和在下面的实施例部分引用的其他参考文献。

用于确定基因甲基化的许多方法是本领域已知的, 包括基于限制酶消化的甲基化检测和基于硫酸氢盐的甲基化检测。一些此类方法在下文和在下面的实施例部分的实施例 1 中概述 (关于用于检测甲基化的进一步的技术细节公开在 Ahrendt (1999) J. Natl. Cancer Inst. 91:

332-9; Belinsky (1998) Proc. Natl. Acad. Sci. USA 95: 11891-96; Clark (1994) Nucleic Acids Res. 22: 2990-7; Herman (1996) Proc. Natl. Acad. Sci. USA 93: 9821-26; Xiong 和 Laird (1997) Nuc. Acids Res. 25: 2532-2534)。

#### 限制酶消化甲基化检测测定法

该测定法基于一些限制酶不能切割甲基化的 DNA。通常使用的是成对酶 HpaII-MspI, 其包括识别基序 CCGG, 和 SmaI-XmaI, 其具有较不寻常的识别基序 CCCGGG。从而, 例如, 当内部胞嘧啶甲基化时, HpaII 不能切割 DNA, 使得 HpaII-MspI 是快速甲基化分析的有用工具。该方法通常与 DNA 印迹分析结合进行。采取措施分析基因序列, 其将不会使得难以解释结果。从而, 首先选择侧翼是 CG 甲基化不敏感的酶 (例如, BamHI) 的限制位点的目的区域。选择的此类序列不包括超过 5-6 个 HpaII 的位点。用于 DNA 印迹或者 PCR 的探针应该位于该区域内并且完全或者部分覆盖该区域。该方法已经由 Buller 和同事 (1999) (Association between non random X-chromosome inactivation and BRCA1 mutation in germline DNA of patients with ovarian cancer J. Natl. Cancer Inst. 91 (4): 339-46) 成功地使用。

因为通过甲基化敏感酶 (例如, HpaII) 消化通常是部分的, 所以优选使用仅消化甲基化 CpG 位点的 MspI 或者其他酶进行补充分析 [Yamada 等人 Genome Research 14 247-266 2004] 以检测多种甲基化模式。

#### 基于硫酸氢盐的甲基化检测

基因组测序 - 使用从小于 100 个细胞分离的 DNA, 基因组测序技术 [Clark 等人, (1994) 上述] 能够检测任意靶序列的两条链上的每个甲基化胞嘧啶。在该方法中, 在 5-甲基胞嘧啶保持无反应性的条件下, 用亚硫酸氢钠将单链 DNA 中的胞嘧啶残基转化成尿嘧啶残基。用特异引物扩增所转化的 DNA 并测序。序列中剩余的所有胞嘧啶残基代表基因组中以前甲基化的胞嘧啶。该方法利用确定的步骤, 其使变性、亚硫酸氢盐转化和扩增的效率最大化, 以允许对容易从生殖细

胞和早期发育阶段得到的小量基因组 DNA 的单个基因进行甲基化作图。

甲基化 - 特异性 PCR (MSP) - 这是用于甲基化的灵敏检测的最广泛使用的测定法。简言之, 扩增前, 将 DNA 用亚硫酸氢钠处理将所有未甲基化的胞嘧啶转化成尿嘧啶。该亚硫酸氢盐反应将甲基化信息有效转化成序列差异。使用引物扩增 DNA, 所述引物匹配 DNA 的一种特定甲基化状态, 如其中 DNA 在所有 CpG 都甲基化的状态。如果该甲基化状态存在于 DNA 样品中, 那么可以在凝胶上显示所产生的 PCR 产物。

然而, 将明白该方法特异的引发需要引物结合位点中的所有 CpG 都被共同甲基化。从而, 当存在共同甲基化时, 在凝胶上观察到扩增的产物。当一个或多个 CpG 未被甲基化时, 没有产物。因此, 该方法不能区分部分甲基化水平和完全没有甲基化 [见美国专利号 5,786,146; Herman 等人, Proc. Natl. Acad. Sci. USA 93: 9821-9826 (1996) ]。用于检测指示 21 号染色体扩增的甲基化的示例性引物在下面的实施例部分的实施例 2 中提供。

实时荧光 MSP(MethyLight)-利用荧光探针结合 MSP, 使用实时 PCR 允许更高通量的均相反应。如果探针不含有 CpG, 那么该反应基本上是 MSP 的定量形式。然而, 通常设计荧光探针以与含有一个或多个 CpG 的位点退火, 且该第三种寡核苷酸增加用于完全甲基化靶链的测定法的特异性。因为实时发生扩增的检测, 所以不需要其次的电泳步骤。因为没有 PCR, 后样品的操作, 所以减少了污染的危险。MethyLight 探针可以是任意形式的, 包括但不限于 Taqman 探针或者 LightCycler 杂交探针对, 且如果使用多种报导染料, 那么可以同时几个探针 [Eads (1999) Cancer Res. 59: 2302-2306; Eads (2000) Nucleic Acids Res. 28: E32; Lo (1999) Cancer Res. 59: 3899-390]。通过 MethyLight 定量分析的优点在前列腺癌中用谷胱甘肽 - S-转移酶 P1(GSTP1)甲基化进行证明 [Jeronimo (2001) J. Natl. Cancer Inst. 93: 1747-1752]。使用该方法, 可能显示良性前列腺增生样品、前列腺 intraepithelial 肿瘤区和局部前列腺腺癌中的甲基化。

亚硫酸氢盐修饰的 DNA 的限制性分析 - 该定量技术也称作 COBRA (Xiong 等人, 1997, 上述), 可以用于确定小量基因组 DNA

中特定基因座中的 DNA 甲基化水平。用限制酶消化揭示亚硫酸氢钠处理的 DNA 的 PCR 产物中甲基化依赖性的序列差异。最初 DNA 样品中甲基化水平由消化和未消化的 PCR 产物的相对量表示, 所述甲基化水平在宽范围的 DNA 甲基化水平间为线性定量形式。该技术可以可靠地应用于从显微解剖的石蜡包埋的组织样品得到的 DNA。从而 COBRA 组合了容易使用、定量准确和与石蜡切片相容的有力特征。

差别甲基化杂交(DMH)-DMH 整合了高密度、基于微阵列的筛选策略来检测甲基化 CpG 二核苷酸基因组片段的存在或不存在[见, Schena 等人, *Science* 270: 467- 470 (1995) ]。当要分析单个区中许多(例如, >3 个)甲基化位点时, 使用基于阵列的技术。首先, 从基因组文库产生 CpG 二核苷酸核酸片段, 将其扩增并附着在固相支持体上以产生富含 CpG 二核苷酸的筛选阵列。通过用限制性内切核酸酶消化来自样品的 DNA 产生扩增子, 所述限制性内切核酸酶将 DNA 消化成片段但是保留甲基化的 CpG 岛完整。这些扩增子用于探测附着在筛选阵列上的富含 CpG 二核苷酸的片段以鉴定 DNA 样品的富含 CpG 二核苷酸区中的甲基化模式。不同于局限于一次分析一种基因的其他甲基化分析方法, 如 DNA 杂交, 亚硫酸氢盐 DNA 测序和甲基化特异性 PCR, DMH 利用多种富含 CpG 二核苷酸的基因组片段, 它们经特别设计成允许同时分析基因组中多种甲基化相关的基因(进一步细节见美国专利号 6,605,432)。

关于分析 DNA 甲基化的其他细节和额外方法(例如, 质谱分析法分析)见 Tost J, Schatz P, Schuster M, Berlin K, Gut IG. *Analysis and accurate quantification of CpG methylation by MALDI mass spectrometry. Nucleic Acids Res.* 2003 May 1; 31 (9): e50; Novik KL, Nimmrich I, Genc B, Maier S, Piepenbrock C, Olek A, Beck S. *Epigenomics: genome-wide study of methylation phenomena. Curr Issues Mol Biol.* 2002 Oct; 4 (4): 111-28. 综述; Beck S, Olek A, Walter J. *From genomics to epigenomics: a loftier view of life. Nat Biotechnol.* 1999 Dec; 17 (12): 1144; Fan (2002) *Oncology Reports* 9: 181-183; <http://www.methods-online.net/methods/DNAMethylation.html> ; Shi (2003) *J. Cell Biochem.*88(1): 138-43; Adoryan (2002) *Nucleic Acids*

**Res. 30 (5): e21.**

将明白可以用许多通过商业途径可获得的试剂盒检测基因的甲基化状态。实例包括,但不限于, EZ DNA methylation kit™ (可从 Zymo Research, 625 W Katella Ave, Orange, CA92867, 美国得到)。

通常,根据所选技术设计用于上述基于硫酸氢盐的甲基化检测方法的寡核苷酸。

如本文所用的术语“寡核苷酸”指核糖核酸(RNA)或者脱氧核糖核酸(DNA)的单链或者双链寡聚体或者多聚体或者其模拟物。该术语包括由天然存在的碱基、糖和共价核苷间键(例如,主链)组成的寡核苷酸以及具有功能类似于各自天然存在的部分的非天然存在的部分的寡核苷酸。(见,公开在美国专利号 4,469,863; 4,476,301; 5,023,243; 5,177,196; 5,188,897; 5,264,423; 5,276,019; 5,278,302; 5,286,717; 5,321,131; 5,399,676; 5,405,939; 5,453,496; 5,455,233; 5,466,677; 5,476,925; 5,519,126; 5,536,821; 5,541,306; 5,550,111; 5,563,253; 5,571,799; 5,587,361; 和 5,625,050)。

从而,例如,影响甲基化特异性 PCR 的特异性的最关键的参数由引物设计决定。因为亚硫酸氢盐对 DNA 的修饰破坏链互补性,所以每条链都可以作为随后 PCR 扩增的模板,并且可以确定每条链的甲基化模式。然而,将明白在实践中优选扩增一条链(例如,有义链)。设计引物来扩增长为 80-250 bp 的区域,其在最初链中整合足够的胞嘧啶以确保未经修饰的 DNA 不作为引物的模板。此外,CpG 二核苷酸内胞嘧啶的数目和位置决定了用于甲基化和未甲基化模板的引物的特异性。通常,在每种引物中包括 1-3 个 CpG 位点,其在每条引物的 3'区集中。这提供了最佳特异性并使由于错误引发导致的假阳性最小化。为了促供同时分析相同热循环仪(thermocycler)中给定基因的每条引物,调节引物的长度以得到几乎相同的解链/退火温度。

此外,因为 MSP 利用特异引物识别来区分甲基化和未甲基化的等位基因,所以在扩增期间保持严格退火条件。基本上,选择最大退火温度以允许退火和随后扩增。优选地,设计引物的退火温度低于计算的解链温度 5-8°C。进一步细节见 Herman 和 Baylin (1998) **Methylation Specific PCR, in Current Protocols in Human Genetics.**

根据本发明的教导设计的寡核苷酸可以根据本领域已知的任意寡

核苷酸合成方法如酶促合成或固相合成来产生。用于进行固相合成的设备和试剂可以通过商业途径从例如, Applied Biosystems 得到。可以使用用于此类合成的任意其他方法;寡核苷酸的实际合成是完全在本领域技术人员能力之内的,并且可以通过成熟的方法,使用固相化学,例如,氰乙基亚磷酰胺然后脱保护,脱盐和通过例如,自动化的含三苯甲基的方法或者 HPLC 纯化来完成,所述成熟方法在例如“Molecular Cloning: A laboratory Manual” Sambrook 等人, (1989); “Current Protocols in Molecular Biology” 卷 I-III Ausubel, R. M., 编者 (1994); Ausubel 等人, “Current Protocols in Molecular Biology”, John Wiley and Sons, Baltimore, Maryland (1989); Perbal, “A Practical Guide to Molecular Cloning”, John Wiley & Sons, New York (1988) 和 “Oligonucleotide Synthesis” Gait, M. J., 编著 (1984) 中详述。

本文上述方法可以用于检测与基因座拷贝数改变相关的病理(上述)。

此类病理的实例包括但不限于三体性,包括第 1 对染色体三体性、第 2 对染色体三体性、第 3 对染色体三体性、第 4 对染色体三体性、第 5 对染色体三体性、第 6 对染色体三体性、第 7 对染色体三体性、第 8 对染色体三体性、第 9 对染色体三体性、第 10 对染色体三体性、第 11 对染色体三体性、第 12 对染色体三体性、第 13 对染色体三体性 (Patau 氏综合征)、第 14 对染色体三体性、第 15 对染色体三体性、第 16 对染色体三体性、第 17 对染色体三体性、第 18 对染色体三体性 (Edward 氏综合征)、第 19 对染色体三体性、第 20 对染色体三体性、第 21 对染色体三体性 (唐氏综合征)、第 22 对染色体三体性、三 X 染色体综合征 (Klinefelter 综合征)、三 Y 染色体综合征、6q 染色体部分三体性、9p 染色体三体性、11q 染色体三体性、镶嵌性第 14 对染色体三体性、镶嵌性第 22 对染色体三体性;单体性,如第 1 对染色体单体性和 X 染色体单体性 (Turner 综合征);四体性,如 18p 染色体四体性;三倍性,如三倍体综合征(关于罕见疾病见国际组织 <http://www.rarediseases.org/>, 且关于染色体镶嵌性见 <http://www.medgen.ubc.ca/wrobinson/mosaic/contents.htm>);和癌症,如慢性髓细胞性白血病。



为了鉴定受试者中基因座拷贝数的改变，从个体受试者（即哺乳动物）得到 DNA 样品并如上文描述进行分析。根据本发明的该方面的优选受试者为任何发育阶段的人（例如，植入前胚胎受试者、胚胎受试者、胎儿受试者、新生儿受试者和出生后受试者）。

通常进行出生后检查以排除经典的染色体综合征和对如下个体进行基因型分析，所述个体为具有多发性先天性异常（MCA）的个体、具有染色体异常的个体的亲代或者同胞、具有平衡的或者结构染色体异常的个体的孩子、具有两次或更多次胎儿丧失病史的夫妇、具有不育问题的夫妇、具有不确定生殖器的个体、具有原发性经闭的女性、具有智力迟钝的个体和具有青春期失败（pubertal failure）的男性和女性。

从个体受试者（即，新生儿和出生后）的生物样品得到 DNA。如本文所用的，短语生物样品指从个体分离的组织或者流体的样品，包括但不限于，例如，血浆、血清、脊髓液、淋巴液、尿、皮肤、呼吸道、肠道和生殖泌尿道的外部切片、泪液、唾液、奶、血细胞、肿瘤、器官和体内细胞培养成分的样品。优选使用组织活组织检查、血液或者骨髓样品。

优选在肝素钠或者 EDTA-涂覆的管中收集血液。新生儿需要最少 1 - 2ml 血液，儿童或者成人需要最少 3 - 5ml 血液。对于白细胞分析，细胞必须超过 10,000，其具有 10% 未成熟细胞。

在骨髓运送培养基或者肝素钠管中收集骨髓(0.5-2 cc 骨髓)。

在无菌生理盐水或者无菌组织培养基中收集组织活组织检查 [3mm 样品，例如，胎盘、带、皮肤（通常用于测试镶嵌性的程度）]。

通常使用的是来自外周血的 DNA。由于正常循环的淋巴细胞在培养条件下不分裂，所以得到淋巴细胞并使其受到外部刺激因子（即，促分裂原）刺激以诱导细胞分裂（即，有丝分裂）。可以在温育 45 - 96 小时后的任何时间收获刺激的细胞。

一旦得到样品，就优选例如通过使用可以从 Qiagen(28159 Avenue Stanford Valencia CA 91355)得到的 QIAamp 血液试剂盒提取基因组 DNA 并如上述进行分析。

因为染色体异常是流产和出生缺陷的主要原因，所以优选使用上述方法鉴定未出生的婴儿中的基因座扩增。可以使用亚硫酸氢盐修饰

检测从母亲血液得到的胎儿 DNA 的甲基化是成熟的方法，从而允许在产前筛选中使用本发明的外遗传标记[见 Poon 等人(2002) *Clin. Chem.* 48: 35-41]。

从胚胎（即，从受孕到发育 8 周的正发育的婴儿）或者胎儿（即，从发育的第 9 周到出生的正发育的婴儿）细胞得到 DNA 的方法是本领域公知的。实例包括例如但不限于母体活组织检查（例如，子宫颈取样、羊膜腔穿刺术取样、血液取样）、胎儿活组织检查（例如，肝脏活组织检查）和绒毛膜绒毛取样（见，背景部分和美国专利号 6,331,395）。

优选使用根据本发明的该方面从母体血液分离胎儿 DNA，因为它是一种非侵入性方法，其不对正在发育的婴儿造成任何危险[见 Lo(1998) *Am. J. Hum. Genet.* 62(4): 768-75]。

可以从母体循环收集无细胞的胎儿 DNA 并如上述进行分析[见 Bauer (2002) *Am. J. Obstet. Gynecol.* 186: 117-20; Bauer (2001) *Ann. NY Acad. Sci.* 945: 161-3; Pertl (2001) *Obstet. Gynecol.* 98: 483-90; Samura (2000) *Hum. Genet.* 106: 45-9]。

备选地，可以使用抗体捕获技术从母体血液富集胎儿细胞，所述技术中固定化抗体结合胎儿细胞并捕获胎儿细胞以促进它们的富集 [Mueller 等人，“Isolation of fetal trophoblasts cells from peripheral blood of pregnant women”，*The Lancet* 336: 197- 200 (1990); Ganshirt-Ahlert 等人，“Magnetic cell sorting and the transferring receptor as potential means of prenatal diagnosis from maternal blood” *Am. J. Obstet. Gynecol.* 166: 1350-1355 (1992)]。

还可以用抗体和其他特异结合部分标记胎儿细胞以促进细胞分选方法，如流式细胞术 [Herzenberg 等人，“Fetal cells in the blood of pregnant women: Detection and enrichment by fluorescence-activated cell sorting”，*Proc. Natl Acad. Sci. (USA)*76: 1453-1455 (1979); Bianchi 等人，“Isolation of fetal DNA from nucleated erythrocytes in maternal blood” *Proc. Natl Acad. Sci. (USA)*87: 3279-3283 (1990); Bruch 等人，“Trophoblast-Like cells sorted from peripheral maternal blood using flow cytometry: a multiparametric study involving transmission electron microscopy and fetal DNA

amplification” Prenatal Diagnosis 11: 787-798 (1991). Price 等人  
“Prenatal diagnosis with fetal cells isolated from maternal blood by multiparameter flow cytometry” Am. J. Obstet. Gynecol 165: 1731-1737 (1991) ]。

PCR 技术通常结合使用以便增加胎儿 DNA 的相对量并从而允许分析 [Bianchi 等人, “Isolation of fetal DNA from nucleated erythrocytes in maternal blood”, Proc. Natl Acad. Sci (USA) 87: 3279-3283 (1990); Adkinson 等人, “Improved detection of fetal cells from maternal blood with polymerase chain reaction”, Am. J. Obstet. Gynecol. 170: 952-955 (1994); Takabayashi 等人, “Development of non-invasive fetal DNA diagnosis from maternal blood” Prenatal Diagnosis 15: 74-77 (1995)]。

在美国专利号 6,331,395 中公开了使用母体循环中胎儿细胞的产前诊断的特定构造。

例如, 可以从 8 - 20 周妊娠的孕妇得到血液 (50ml)。通过在 Ficoll-hypaque 上离心分离单核细胞 (MNC) 级分, 并将其在具有 10 % FCS 的  $\alpha$  培养基中以  $5 \cdot 10^6$  个/ml 培养 7 天, 其中使用 SCF 100ng/ml, IL-3 100ng/ml, 和 IL-6 100 u/ml。然后回收未贴壁细胞并在  $\alpha$  培养基中以  $3 \cdot 10^5$  个/ml 再次铺平板, 所述培养基含有 30 % FCS、1 % BSA、 $10^{-4}$ M  $\beta$ - 巯基乙醇, 和青霉素和链霉素, 以及 SCF 100ng/ml, IL-3 100 ng/ml, 和 IL-6 100mu/ml。所有温育都在具有 5 % CO<sub>2</sub>, 和室内空气或者 5 % 氧气的湿润培养箱中进行。21 天后, 回收细胞。离心细胞并使用标准方法提取 DNA, 用于如上述的甲基化分析。

用于从母体血液富集胎儿细胞的试剂盒可以从 AVIVA Biosciences Corporation (San Diego, CA, [http://www.avivabio.com/Technology/fetal\\_cell\\_isolation.html](http://www.avivabio.com/Technology/fetal_cell_isolation.html)) 得到。

将明白还可以在胎儿死亡或者流产后得到胚胎或者胎儿 DNA。在该情况中, 使用酶解离的细胞或者组织碎片 (外植块) 从胚胎或者胎儿组织开始培养。当将该组织置于合适的培养条件中时, 细胞贴壁到表面并以单层生长。

使用本方法所得染色体信息可以用本领域公知的许多细胞学 (例如, 吉姆萨染色) 和基于杂交的技术 (例如, FISH) 进一步验证 (见,

例如，美国专利号 5,906,919 和 5,580,724)。

用于确定如上述的基因座扩增的试剂可以在包装或者分配器装置，如诊断试剂盒中提供。包装可以例如包含金属或者塑料箔，如泡罩包装 (blister pack)。包装或者分配器装置可以附带有用于诊断的说明书。

将明白本发明还可以用于检测与异常 DNA 甲基化机制相关的病理，所述机制导致如上述的异常甲基化。实例包括，但不限于 Pradi-Willi、Angelman、Beckwith-Wiedemann、Rett 和 ICF 综合征。例如，ICF 综合征由称作 Dnmt3b 的 DNA 甲基转移酶的异常功能导致。类似地，识别和结合 mC(称作 MeCP2)的蛋白质之一的异常导致 Rett 综合征，其是影响年轻女性的一种形式的智力迟钝。

将还明白通过本发明也可以预期进行性别确定 (例如，产前)，因为女性 X 染色体的额外拷贝上的基因受到 DNA 甲基化抑制[Goto (1998) *Microbiol. Mol. Biol. Rev.* 62 (2): 362-78]。

将还明白本发明使得可以鉴定与生命相容的 (重要的) 基因。

从而，根据本发明的另一方面，提供了鉴定“与生命相容的”基因的方法。

通过如上述确定所扩增的染色体序列区中多种基因的甲基化状态可以实现所述方法。

随后，鉴定多种基因中显示处于与预定甲基化状态不同的甲基化状态的基因，从而鉴定“与生命相容的”基因。

可以有效使用此类方法来注释基因和鉴定新的治疗靶。

当考察下面的实施例后，本发明的额外目的、优点和新特征将对于本领域技术人员是显而易见的，所述实施例不意在作为限制。此外，如上述和如权利要求书中要求保护的本发明的多种实施方案和方面的每一种将在下面的实施例中找到实验支持。

## 实施例

现在参考下面的实施例，这些实施例与上面的描述一起以非限定的方式阐明本发明。

通常，用于此处的命名法和用于本发明的实验方法包括分子、生物化学、微生物学和重组 DNA 技术。此类技术在文献中详细解释。

见, 例如, “Molecular Cloning: A laboratory Manual” Sambrook 等人, (1989); “Current Protocols in Molecular Biology” 卷 I-III Ausubel, R. M., 编著 (1994); Ausubel 等人, “Current Protocols in Molecular Biology”, John Wiley and Sons, Baltimore, Maryland (1989); Perbal, “A Practical Guide to Molecular Cloning”, John Wiley & Sons, New York(1988); Watson 等人, “Recombinant DNA”, Scientific American Books, New York; Birren 等人(eds) “Genome Analysis: A Laboratory Manual Series”, 卷 1-4, Cold Spring Harbor Laboratory Press, New York (1998); 如美国专利号 4,666,828; 4,683,202; 4,801,531; 5,192,659 和 5,272,057 中陈述的方法; “Cell Biology: A Laboratory Handbook”, 卷 I-III Cellis, J. E., 编著(1994); “Current Protocols in Immunology” Volumes I-III Coligan J. E., 编著(1994); Stites 等人(eds), “Basic and Clinical Immunology” (第八版), Appleton & Lange, Norwalk, CT (1994); Mishell 和 Shiigi (编著), “Selected Methods in Cellular Immunology”, W. H. Freeman and Co., New York(1980); 可利用的免疫测定法在专利和科学文献中详细描述, 见, 例如, 美国专利号 3,791,932; 3,839,153; 3,850,752; 3,850,578; 3,853,987; 3,867,517; 3,879,262; 3,901,654; 3,935,074; 3,984,533; 3,996,345; 4,034,074; 4,098,876; 4,879,219; 5,011,771 和 5,281,521; “Oligonucleotide Synthesis” Gait, M. J., 编著 (1984); “Nucleic Acid Hybridization” Hames, B. D., 和 Higgins S. J., 编著 (1985); “Transcription and Translation” Hames, B. D., 和 Higgins S. J., 编著(1984); “Animal Cell Culture” Freshney, R. I., 编著(1986); “Immobilized Cells and Enzymes” IRL Press, (1986); “A Practical Guide to Molecular Cloning” Perbal, B., (1984) 和 “Methods in Enzymology” 卷 1-317, Academic Press; “PCR Protocols: A Guide To Methods And Applications”, Academic Press, San Diego, CA (1990); Marshak 等人, “Strategies for Protein Purification and Characterization-A Laboratory Course Manual” CSHL Press (1996); 将所有这些参考文献完整引入本文作为参考。在该文件全文提供了其他一般参考文献。认为其中提供的方法是本领域公知的并且为了读者的方便而提供。其中包含的所有信息都引入作为参考。

### 材料和实验步骤

DNA 提取 - 使用 QIAamp Blood 试剂盒(Qiagen, 28159 Avenue Stanford Valencia CA 91355)从血浆和羊水样品提取 DNA。每个柱子用 800 $\mu$ l 血浆或者羊水进行 DNA 提取。用 50 - 110 $\mu$ l 洗脱缓冲液洗脱 DNA。用 Nucleon DNA 提取试剂盒(Scotlabs Woburn, MA)根据生产商的说明书从白细胞的棕黄层提取 DNA。

DNA 的亚硫酸氢盐处理 - 将 DNA(最多 2 $\mu$ g)在 50 $\mu$ l 蒸馏水中稀释并向其中加入 5.5 $\mu$ l 2M NaOH。然后向反应混合物中加入 5 $\mu$ g 鲑精 DNA。

将溶液在 50 $^{\circ}$ C 温育 10 分钟以从而产生单链 DNA。向每管加入氢醌[30 $\mu$ l 10 mM 氢醌(Sigma), 通过向 50ml 水中加入 55 mg 氢醌新鲜配制]。之后, 向该溶液加入 520 $\mu$ l 新鲜配制的 3M 亚硫酸氢钠[Sigma S-8890, 通过向每 5ml 水加入 1.88 gm 亚硫酸氢钠并用 NaOH 调节 pH 到 5.0 制备]。采取措施以确保 DNA 溶液均匀混合。将 DNA 溶液用矿物油分层并允许其在 50 $^{\circ}$ C 温育 16 小时或者在 70 $^{\circ}$ C 温育 1 - 2 小时。防止更长的温育时间以避免甲基化胞嘧啶转化成胸苷。一旦温育结束, 除去油。向每管加入 1ml DNA Wizard cleanup (Promega A7280) 溶液并将溶液应用于试剂盒中的小量制备柱。应用真空并将柱子用 2 ml 80% 异丙醇洗涤。通过加入 50 $\mu$ l 温水(即 80 $^{\circ}$ C)将 DNA 洗脱到清洁、标记的 1.5ml 管中。将管离心 1 分钟并向每管加入 5.5 $\mu$ l 3M NaOH。在室温温育磺化 DNA 溶液 5 分钟。

加入作为载体的 1 $\mu$ l 糖原(Boehringer Ingelheim GmbH, 德国), 33 $\mu$ l 10 M NH<sub>4</sub>Ac, 和 3 体积乙醇。将 DNA 在 -20 $^{\circ}$ C 沉淀至少 1 小时至过夜, 并用 70% 乙醇洗涤。将干燥的沉淀重悬浮在 20 $\mu$ l 水中并在 -20 $^{\circ}$ C 保存。

扩增反应 - 向含有 1 $\times$  GC 缓冲液 2(TaKaRa, Shuzo, Kyoto, 日本)、2.5 mM 每种 dNTP、5U TaKaRa LA Taq (TaKaRa, Shuzo, Kyoto, 日本)和 50 pmol 反义引物的 50 $\mu$ l PCR 反应混合物中加入磺化 DNA 溶液的 1 $\mu$ l 等分试样。使反应混合物在 94 $^{\circ}$ C 的温度温育 5 分钟。

预热后, 将亚硫酸氢盐处理的 DNA 的有义序列的互补链如下延

伸 2 轮：94°C 1 分钟、60°C 3 分钟和 72°C 3 分钟。

之后，将 50 pmol 有义引物加入并加热混合物到 94°C。

将 DNA 在扩增 8 轮，每轮为 94°C 1 分钟、60°C 1.5 分钟和 72°C 2 分钟。

通过 30 轮的 94°C 1 分钟、55°C 1.5 分钟和 72°C 2 分钟实现进一步扩增。

通过限制酶分析或者直接测序检测所得 PCR 产物中的甲基化。

通过限制酶检测甲基化位点 - 如上述甲基胞嘧啶向胸腺嘧啶的化学修饰改变 HpaII 和 HahI 限制酶的位点，从而这些酶不能在它们的各自位点消化 DNA。胞嘧啶甲基化不允许甲基化敏感酶在这些位点消化，而甲基化耐受性的其他酶，如 MspI 将产生常规限制模式。对硫酸氢盐处理的 DNA 使用这种酶允许使用特异性限制酶区分甲基化位点（进一步细节见下面的实施例 3）。

用 McrBC 进行甲基化 - 从 New England Biolabs 得到 McrBC。将该酶加入到 5µg 基因组 DNA 中并根据生产商的说明在 37°C 过夜温育反应物。通过在 65°C 温育 20 分钟灭活该酶。将 50µg 所消化的 DNA 用作 PCR 反应的模板。使用启动子特异性引物。通过琼脂糖凝胶分离分析产物。

PCR 产物的直接测序 - 使用通过商业途径可得到的试剂盒（例如，Geneclean 等）纯化所得 PCR 产物并通过商业途径可得到的自动化测序仪测序。

等位基因特异的寡核苷酸杂交 - 在该测定中，扩增含有目的基因上所有候选甲基化位点的大片段。PCR 产物含有通过荧光素或者其他荧光团（Cy3, Cy5）标记的一个核苷酸。标记产物的第二种方法是通过放射性核苷酸（<sup>32</sup>P、<sup>33</sup>P、<sup>35</sup>S、<sup>3</sup>H 或 <sup>14</sup>C），其整合入 PCR 产物。然后将 PCR 产物与用于甲基化胞嘧啶（即，胸腺嘧啶）对胞嘧啶的特异寡核苷酸杂交。与寡核苷酸的杂交可以使用微阵列方法在玻璃或者硝化纤维素上进行。

用于检测突变（或者 SNP）的商业试剂盒 - 可以通过商业途径可得到的“Gamidagene”公司的“Pronto”试剂盒进行甲基化位点的检测。设计这些试剂盒用于与特异性探针结合检测突变和/或单核苷酸多态性（SNP），所述探针经设计和构造成识别目的甲基化位点。类似地，

也可以使用可以识别核苷酸序列中突变的其他方法。例如，扩增失效突变系统 - ARMS。在该方法中，使用两种补充反应，一种含有对正常等位基因特异的引物，另一种含有突变等位基因（两种都具有公共的第二种引物）。因为 PCR 引物与变体 DNA 完美匹配，所以鉴定完美匹配的等位基因基因型分析的优先扩增。如上面描述的，通过该方法可以检测通过亚硫酸氢盐转化成胸腺嘧啶的甲基胞嘧啶。

### 实施例 1

#### 21 号染色体的基因

下面的表 2 列出了由 Gardiner 和 Davisson *Genome Biology* 2000 1 (2): 综述 0002.1 - 0002.9 注释的 21 号染色体的 122 种基因的给定功能。其大多数具有完整或者大概完整的 cDNA 序列。基于直接实验的文献报导或者基于从与其他蛋白质的相似性的推断给定功能注释。具有仅部分结构信息的基因注释是基于其中的特定功能结构域的，并且通过 (\*) 指出 (Gardiner K. <http://genomebiology.com/2000/1/2/reviews/0002.1>)。

选择的功能类别将是概括地描述性的，每种基因仅在一种类别中出现。

表 2

功能类别	基因数	功能注释
转录因子、调节物和调节剂	17	GABPA、BACH1、RUNX1、SIM2、ERG、ETS2(转录因子); ZNF294、ZNF295、Pred65、ZNF298、APECED(锌指); KIAA0136(亮氨酸拉链); GCFC(富含 GC 的结合蛋白质); SON(DNA 结合结构域); PKNOX1(同源框); HSF2BP(热激转录因子结合蛋白); NRIP1(雌激素对转录激活的调节剂)
染色质结构	4	H2BFS(组蛋白 2B)、HMG14(高迁移组)、CHAF1B(染色质装配因子)、PCNT(中心粒周围蛋白 (pericentrin), 中心体的中心粒旁基质的整合组分)



蛋白酶和蛋白酶抑制剂	6	BACE( $\beta$ -位点 APP 切割酶); TMPRSS2、TMPRSS3(跨膜丝氨酸蛋白酶); ADAMTS1、ADAMTS5 (金属蛋白酶); CSTB(蛋白酶抑制剂)
遍在蛋白途径	4	USP25、USP16(遍在蛋白蛋白酶); UBE2G2(遍在蛋白缀合酶); SMT3A(遍在蛋白样)
干扰素和免疫应答	9	IFNAR1、IFNAR2、IL10RB、IFNGR2(受体/辅助因子); MX1、MX2(干扰素诱导的); CCT8(T-复合体亚基)、TIAM1(T-淋巴瘤侵入和转移诱导蛋白)、TCP10L(T-复合体蛋白 10 样)
激酶	8	ENK(肠激酶); MAKV、MNB、KID2(丝氨酸/苏氨酸); PHK(吡哆醛激酶); PFKL(果糖磷酸激酶); ANKRD3(具有激酶结构域的锚蛋白样); PRKCBP2(蛋白质激酶 C 结合蛋白质)
磷酸酶	2	SYNJ1(聚磷酸肌醇磷酸酶); PDE9A(环磷酸二酯酶)
RNA 加工	5	rA4(SR 蛋白质)、U2AF35(剪接因子)、RED1 (编辑酶(editase))、PCBP3(poly(C)-结合蛋白); RBM11(RNA-结合基序)
粘着分子	4	NCAM2(神经细胞)、DSCAM; ITGB2(淋巴细胞); c21orf43(类似于内皮紧密连接分子)
通道	7	GRIK1(谷氨酸受体, 钙通道); KCNE1、KCNE2、KNCJ6、KCNJ15(钾); CLIC11(氯化物); TRPC7(钙)
受体	5	CXADR(柯萨奇病毒和腺病毒); Claudins 8、14、17(Claustidia); Pred12(甘露糖)
转运蛋白	2	SLC5A3(Na-肌醇); ABCG1(ATP-结合盒)
能量代谢	4	ATP50(赋予 ATP 合酶寡霉素 - 敏感性的蛋白质); ATP5A(ATP 酶偶联因子 6); NDUFV3(NADH-泛醌氧化还原酶亚基前

		体); CRYZL1(醌氧化还原酶)
结构的	4	CRYA(晶状体蛋白); COL18、COL6A1、COL6A2(胶原)
甲基转移酶	3	DNMT3L(胞嘧啶甲基转移酶)、HRMTIII(蛋白质精氨酸甲基转移酶); Pred28(AF139682)(N6-DNA 甲基转移酶)
SH3 结构域	3	ITSN、SH3BGR、UBASH3A
一碳代谢	4	GART(嘌呤生物合成)、CBS(胱硫醚-β-合成酶)、FTCD(亚胺甲基转移酶环化脱氨酶)、SLC19A1(还原的叶酸载体)
氧代谢	3	SOD1(超氧化物歧化酶); CBR1、CBR3(碳酰还原酶)
其他	28	HLCS(羧化全酶合酶); LSS(羊毛固醇合成酶); B3GALT5(半乳糖基转移酶); AGPAT3(酰基转移酶); STCH(微粒体应激蛋白); ANA/BTG3(细胞周期控制); MCM3(DNA 复制相关因子); APP(阿尔茨海默氏病淀粉状蛋白前体); WDR4、WDR9(含有 WD 重复序列的蛋白质); TFF1、2、3(三叶蛋白); UMODL1(尿调节素); *Pred5(脂肪酶); *Pred3(角质形成细胞生长因子); KIAA0653、*IgSF5(Ig 结构域); TMEM1、*Pred44(跨膜结构域); TRPD(含有三十四肽 (tetra-trico-peptide) 重复序列); S100b(Ca 结合); PWP2(周期性色氨酸蛋白质); DSCR1(富含脯氨酸); DSCR2(富含亮氨酸); WRB(富含色氨酸的蛋白质); Pred22(tRNA 合成酶); SCL37A1(磷酸甘油通透酶)

## 实施例 2

第 21 对染色体三体性的基因和可用于检测其甲基化状态的引物背景

原纤维淀粉状蛋白作为神经原纤维缠结在神经元内的沉积、作为斑块在细胞外以及在血管中的沉积是阿尔茨海默氏病 (AD) 和老年唐氏综合征患者的特征。在这些沉积物中发现的主要蛋白质是小的、不溶性和高度聚集的多肽 a4, 认为其来自它的前体 - 淀粉状蛋白前体的异常分解代谢, 所述前体定位于 21 号染色体(21q21.2)。

#### 实验步骤

为了检测 APP (GenBank 检索号 X127522) 的扩增, 通过亚硫酸氢盐测序确定 APP 启动子区的甲基化。

表 3

引物 ID	寡核苷酸序列 (5'-3')/SEQ ID NO:	X127522 中的位置	PCR 产物大小 (bp)
APP-F	tggtttagattttttttattg (1)	3449-3473	272
APP-R	acctaccactaccaaaaaactaac (2)	3696-3721	

下面表 4 列出了优选的 PCR 条件。

表 4

温度 (°C)	时间	循环数
95	10 分钟	35
94	30 秒	
62	30 秒	
72	30 秒	
72	10 分钟	

对所得 PCR 产物进行测序以鉴定胞嘧啶向胸苷的替代。从 APP 启动子扩增的 PCR 产物 (使用引物 APP-F 和 APP-R, 图 1b) 在图 1a 中显示。

备选地, 所得 PCR 产物可以与寡核苷酸微阵列杂交。

下面的表 5 和 6 列出了一些寡核苷酸结构, 其可以用于鉴定用亚硫酸氢盐处理 DNA 后, 人 21 号染色体上甲基化的 DNA 部分, 如上述。

表 5 - 淀粉状蛋白前体蛋白质(APP)基因(GenBank 检索号 X127522)  
21 号染色体

WT 探针 (5'-3')/SEQ ID NO:	甲基化探针 (5'-3')/SEQ ID NO:	位置 (gi35230)
gagggggtgtgtggg/(5)	gagggggcgtgtggg/(6)	3509-3523
gtaagggtgtgtat/(7)	gtaaggcgtgtat/(8)	3535-3549
ttgtgggtgtgggt/(9)	ttgtggcgtgggt/(10)	3550-3563
ttttggtgtgagt/(11)	ttttggcgtgagt/(12)	3573-3591
gagtgggttagt/(13)	gagtggcgttagt/(14)	3583-3597
tttgggtgtgtta/(15)	tttggcgtgtta/(16)	3598-3613
ggtgtgtgtttggg/(17)	ggtgttcgtttggg/(18)	3677-3692
tgtgtgtgggagt/(19)	tgtgttcgggagt/(20)	3492-3506
ttttttggtgtga/(21)	tttttcggtgtga/(22)	3570-3584
agttttggtgtg/(23)	agttttcgtgtg/(24)	3592-3606
ggtgggttgattag/(25)	ggtggcgttgattag/(26)	3639-3653
tggggagtgagggg/(27)	tggggagcggagggg/(28)	3500-3514
ttttggcgtgagt/(29)	ttttggcgtgagt/(30)	3572-3586
gggggtgtgtggg/(31)	gggggtcgtgtggg/(32)	3511-3525
gtgtagggtgtta/(33)	gtgtaggcgtgtta/(34)	3523-3537
tttgggtgagtggg/(35)	tttggcgtgagtggg/(36)	3574-3588

表 6-H2-钙调理蛋白基因(GenBank 检索号 gi: 4758017),  
21 号染色体 [Kuromitsu J, 等人 Mol Cell Biol. (1997) 17 (2): 707-12]

WT 探针 (5'-3')/SEQ ID NO:	甲基化探针 (5'-3')/SEQ ID NO:	位置 (gi4758017)
aatgtgttttta/(37)	aatgttcgttttta/(38)	966501-966515
atattgcgttttg/(39)	atattgcgttttg/(40)	966528-966542
tgtgtttgggttaa/(41)	tgtgttcgggttaa/(42)	966533-966547
ggtgtgtgtgtga/(43)	ggtgttcgtgtgtga/(44)	966559-966573
tgtggcgtgtggagt/(45)	tgtggcgtgtggagt/(46)	966561-966575
tggagtttgggtgt/(47)	tggagttcgggtgt/(48)	966570-966584
agtttgggtgttt/(49)	agtttggcgtgttt/(50)	966572-966586
aatgttcgttagt/(51)	aatgttcgttagt/(52)	966588-966602
gttagtttgggtgt/(53)	gttagttcgggtgt/(54)	966596-966610

### 实施例 3

#### X 染色体三体性的基因和可用于检测其甲基化状态的引物

在女性中，重复 X 染色体的多数基因的一组被沉默。通常通过此类基因的启动子的 CpG 甲基化发生沉默。几种甲基化分析方法用于检测男性、女性和 Klinefelter 综合征影响的受试者中雄激素受体 (GenBank 检索号 NM\_00044) 的甲基化状态。

### 实验步骤

细胞 - 从 Coriell Institute NJ 得到男性、女性和 Klinefelter 综合征影响的胚胎的培养 12 天的羊水细胞。Klinefelter 细胞目录号 GMO3102。正常细胞目录号。

DNA 提取 - 将细胞以 2500rpm 离心 10 分钟。将细胞沉淀重悬浮在包括 75 mM NaCl 和 25 mM EDTA 的裂解缓冲液中并充分涡旋以分裂质膜。之后，向混合物中加入 10 % SDS 溶液 (1/10 终体积) 并通过倒转混合溶液。在蛋白酶 K (10 mg/ml, 1/10 终体积) 存在下在 55°C 过夜温育溶液。向溶液加入等体积酚:氯仿 (1:1)，通过倒转充分混合 (5 分钟) 并以 14,000xg 离心 15 分钟以达到相分离。向上层相中加入氯仿，通过倒转 5 分钟充分混合溶液，以 14,000xg 离心 5 分钟以达到相分离，收集上面的相，向其中加入 3M 乙酸钠 (1/10 终体积) 并通过倒转充分混合。用乙醇 (70 %) 过夜沉淀 DNA 并之后测定其浓度和纯度。

### 基于限制酶的分析

用 HpaII (30 单位, NEB Enzyme, New England Biolabs. Inc. Beverly MA 01915-5599 美国) 消化 0.5µg DNA 分子 (即, 亚硫酸氢盐处理的或者未处理的)。为了确保完全消化, 允许温育过夜进行, 包括温育 8 小时后再次加入新鲜酶。

消化后, 将来自消化混合物的 2µl DNA 用作 PCR 的模板, 该 PCR 使用下面表 7 中列出的引物并且在下面表 8-9 中描述的条件下进行。

表 7

引物名称	序列 (5'-3')/SEQ ID NO:	NM000044 中的位置	PCR 产物 (bp)
AR-f	TCCAGAATCTGTTCCAGAGCGTGC/55	1183-1207	~300*
AR-r	GCTGTGAAGGTTGCTGTTCCAT/56	1447-1470	

\*-PCR 产物的大小取决于从患者回收的 DNA 中 CAG 重复序列数。

表 8

	反应混合物
缓冲液 10X	0.1 终体积 (20 $\mu$ l)
dNTPs 2mM	0.1 终体积 (20 $\mu$ l)
AR-f 10pmol/ $\mu$ l	0.1 终体积 (20 $\mu$ l)
AR-r 10pmol/ $\mu$ l	0.1 终体积 (20 $\mu$ l)
水	配齐至终体积 (20 $\mu$ l)
酶*	1 单位
DNA	0.1-0.15 终体积 (20 $\mu$ l)

\*NEB 酶 - Taq DNA 聚合酶目录号 M0267; New England Biolabs. Inc. Beverly MA 01915-5599 美国

表 9

温度	时间	循环数
94 °C	4 分钟	35
94 °C	45 秒	
59 °C	45 秒	
72 °C	1 分钟	
72 °C	7 分钟	

将所得约 300bp 的 PCR 产物在 2.5% 琼脂糖凝胶上分离和显现。

甲基化特异性 PCR(MSP)

如上文实验步骤中描述的用亚硫酸氢盐处理 DNA。

引物和 PCR 条件分别在表 10、11 和 12 中列出。

表 10

引物名称	序列 (5'-3')/SEQ ID NO:	NM000044 中的位置	PCR 产物 (bp)
AR-U	tagaatttggttttagagtgtgtgt/57	1185-1208	
AR-M	tttggttttagagcgtgcg/58	1189-1207	~225
AR-R	aaaaccatcctcaccctact/59	1385-1404	

表 11

	未甲基化的混合液	甲基化的混合液
缓冲液10X DS*	0.1 终体积	
dNTPs 2mM	0.1 终体积	0.1 终体积
AR-U 10pmol/ $\mu$ l	0.1 终体积	
AR-M 10pmol/ $\mu$ l		0.1 终体积
AR-U 10pmol/ $\mu$ l	0.1 终体积	0.1 终体积
水	配齐至终体积	配齐至终体积
酶*	1 单位	1 单位
DNA	0.1-0.15 终体积	0.1-0.15 终体积

\*缓冲液 DS 10X-166 mM 硫酸铵, 670 mM Trizma; 67mM 氯化镁; 100 mM 巯基乙醇; 1% DMSO; 硫酸铵-Sigma A 4418; Trizma-Sigma T 5753; DMSO- Trizma D 8414; MgCl<sub>2</sub>-Sigma M-1028; 巯基乙醇-Sigma M 3148.

表 12

温度	时间	循环数
94 <sup>o</sup> C	4 分钟	35
94 <sup>o</sup> C	45 秒	
59 <sup>o</sup> C	45 秒	
72 <sup>o</sup> C	1 分钟	
72 <sup>o</sup> C	7 分钟	

通过两步嵌套 PCR 反应证实产物同一性, 其引物在下面的表 13 中列出, PCR 条件在下面的表 14-16 中列出。第一次 PCR 的 DNA 模板用于亚硫酸氢盐修饰 (~50ng)。将 PCR 产物用作第二次 PCR 反应的模板 (1/20 终体积)。对反应产物进行测序。

表 13

引物名称	序列 (5'-3')	NM000044 中的位置	PCR 产物 (bp)
AR-F-1	agatttagttaagtttaaggatggaagt/60	1096-1124	
AR-F-34	gggttgggaagggtttatttt/61	1131-1151	~280*
AR-R-282	aaaaacctcctcaccctactactac/62	1379-1404	

\*PCR 产物的大小与从患者所得 DNA 中 CAG 重复序列数线性相关。

表 14 - 步骤 I

	混合液1
缓冲液 10X	0.1 终体积
dNTPs 2mM	0.1 终体积
AR-F-1 10pmol/ $\mu$ l	0.1 终体积
AR-R-282 10pmol/ $\mu$ l	
水	配齐至终体积
DNA	0.1-0.15 终体积
酶*	1 单位

\*NEB 酶

在下面的表 15 和 16 中列出了用于从亚硫酸氢盐修饰的 DNA 扩增雄激素受体的外显子 1 的 PCR 条件。

表 15

温度	时间	循环数
94 <sup>0</sup> C	4 分钟	35
94 <sup>0</sup> C	45 秒	
59 <sup>0</sup> C	45 秒	
72 <sup>0</sup> C	1 分钟	
72 <sup>0</sup> C	7 分钟	

表 16-步骤 II

	混合液1
缓冲液 10XNEB	0.1 终体积
dNTPs 2mM	0.1 终体积
AR-F-34 10pmol/ $\mu$ l	0.1 终体积
AR-R-282 10pmol/ $\mu$ l	0.1 终体积
水	配齐至终体积
DNA (步骤I的产物)	0.05 终体积
酶*	1 单位

\*NEB 酶

将步骤 II 的 PCR 产物在 2.5% 琼脂糖凝胶上分离并用通过商业途径可获得的试剂盒 (Amersham Bioscience Piscataway Bioscience NJ 08855-美国的 GFX PCR 目录号 27-9602-01) 纯化, 然后亚克隆到 pGEM 质粒 (pGEM-T Easy Vector Vector System I 目录号 A1360 或者



pGEM-T Vector Vector System I 目录号 A3600 Promega Corporation Madison WI 美国)。通过对每种 PCR 产物的 5-10 个克隆测序进行准确测序。以 ABI 测序仪实现测序。

### 结果

在图 2a 中给出了雄激素受体的外显子 1 的天然序列以及 HpaII 和 HhaI 限制位点。亚硫酸氢盐修饰后得到的推定序列在图 2b 中显示。

图 3 和 4 描绘了如通过基于限制酶的分析 and 通过甲基化特异性 PCR (MSP) 确定的, 男性、女性和 Klinefelter 综合征影响的受试者中雄激素受体甲基化状态的结果。

如图 3 中所示, 从 XY (即, 男性) 受试者得到的用 HpaII 处理的 DNA 样品的 PCR 扩增没有得到产物。然而, 使用从 XX 和 XXY 受试者得到的 HpaII 处理的 DNA 样品的相同反应产生清晰的 280bp 带, 其是雄激素受体外显子 1 的产物 exon1。

对男性和 Klinefelter 综合征影响的受试者的雄激素受体的外显子 1 的甲基化状态的 MSP 分析表明使用甲基化引物的 DNA 扩增仅在 Klinefelter 影响的受试者 (即, 46XXY) 的 DNA 中发生。

这些结果一起显然支持 DNA 甲基化介导 Klinefelter 综合征影响的受试者中雄激素受体的基因沉默并且提示它是该病理的有价值的诊断工具。

将明白因为 MSP 不有效检测部分甲基化 (即, 不是给定等位基因中的所有甲基化位点都实际上被甲基化), 所以使用寡核苷酸微阵列将是有利的。

可以有效用于此类微阵列中的寡核苷酸在下面的表 17 中列出。

表 17

WT 探针 (5'-3')/SEQ ID NO:	甲基化探针 (5'-3')/SEQ ID NO:	位置 (M00044)
ggtttatttttggttggtggt/63	ggtttattttcggttggtggt/66	1142-1162
	ggtttatttttggtcggtggt/67	
	ggtttatttttggttggtggt/68	
	ggtttattttcggtcggtggt/69	
	ggtttatttttggtcggtggt/70	
	ggtttattttcggttggtggt/71	
	ggtttattttcggtcggtggt/72	
tatttttggttggtggttaag/64	tattttcggttggtggttaag/73	1146-1166
	tatttttggtcggtggttaag/74	
	tatttttggttggtggttaag/75	
	tattttcggtcggtggttaag/76	
	tatttttggtcggtggttaag/77	
	tattttcggtcggtggttaag/78	
	tattttcggtcggtggttaag/79	
ttttggttggtggttaagattt/65	tttcggttggtggttaagattt/80	1150-1170
	ttttggtcggtggttaagattt/81	
	ttttggttggtcggttaagattt/82	
	tttcggtcggtggttaagattt/83	
	ttttggtcggtcggttaagattt/84	
	tttcggttggtcggttaagattt/85	
	tttcggtcggtcggttaagattt/86	

taagatttattgaggagtttt/89	taagatttatcgaggagtttt/88	1162-1182
tgtttttagag tgtgtgtgaag/90	tgtttttagag cgtgtgtgaag/91	1192-1212
	tgtttttagag tgtgctgaag/92	
	tgtttttagag tgtgtgcaag/93	
	tgtttttagag cgtgctgaag/94	
	tgtttttagag cgtgtgcaag/95	
	tgtttttagag tgtgctgaag/96	
	tgtttttagag cgtgctgaag/97	
ttagagtgtgtgtgaagtgat/98	ttagagcgtgtgtgaagtgat/99	1196-1216
	ttagagtgtgctgaagtgat/100	
	ttagagtgtgtgcaagtgat/101	
	ttagagcgtgctgaagtgat/102	
	ttagagcgtgtgcaagtgat/103	
	ttagagtgtgctgaagtgat/104	
	ttagagcgtgctgaagtgat/105	
agagtgtgtgtgaagtgattt/106	agagcgtgtgtgaagtgattt/107	1198-1218
	agagtgtgctgaagtgattt/108	
	agagtgtgtgcaagtgattt/109	
	agagcgtgctgaagtgattt/110	
	agagcgtgtgcaagtgattt/111	
	agagtgtgctgaagtgattt/112	
	agagcgtgctgaagtgattt/113	
atttagaatttgggttttagg/114	atttagaattcgggttttagg/115	
atttagaggttgtgagtgtag/116	atttagaggtcgcgagcgtag/117	
	atttagaggtcgtgagtgtag/118	
	atttagaggttgcgagtgtag/119	
	atttagaggttgtgagcgtag/120	
	atttagaggtcgcgagtgtag/121	
	atttagaggtcgtgagcgtag/122	
	atttagaggttgcgagcgtag/123	
	atttagaggtcgcgagcgtag/124	
atttagaggttgtgagtgtag/125	Ttagaggtcgcgagcgtagta/126	
	Ttagaggtcgtgagtgtagta/127	
	Ttagaggttgcgagtgtagta/128	
	Ttagaggttgtgagcgtagta/129	
	Ttagaggtcgcgagtgtagta/130	
	Ttagaggtcgtgagcgtagta/131	
	Ttagaggttgcgagcgtagta/132	
	ttagaggtcgcgagcgtagta/133	
aggttgtgagtgtagtatttt/134	Aggtcgcgagcgtagtatttt/135	
	Aggtcgtgagtgtagtatttt/136	
	Aggttgcgagtgtagtatttt/137	
	Aggttgtgagcgtagtatttt/138	
	Aggtcgcgagtgtagtatttt/139	
	Aggtcgtgagcgtagtatttt/140	
	Aggttgcgagcgtagtatttt/141	
	aggcgcgagcgtagtatttt/142	
tagtatttttgggttagtt/143	tagtatttttggcgttagtt/144	
	tagtatttttccgttagtt/145	
	tagtatttttccgcttagtt/146	
tagtatttttgggttagtttgt/147	tagtatttttggcgttagtttgt/148	
	tagtatttttccgttagtttgt/149	
	tagtatttttccgcttagtttgt/150	

#### 实施例 4

根据本发明教导鉴定的第 21 对染色体常染色体三体性的推定标

记

如上所述，位于扩增的染色体或染色体区的基因通常可能由于上游启动子区域的甲基化而不超表达，所述甲基化导致特异性沉默。

下面的表 18 显示了在来自唐氏综合征影响的受试者的羊水细胞相对于从正常受试者得到的羊水细胞中 21 号染色体基因表达的比率。X<1.5 的比率表明基因沉默([www.hgu.mrc.ac.uk/Research/Cellgen/Supplements/Unigene/t21all.htm](http://www.hgu.mrc.ac.uk/Research/Cellgen/Supplements/Unigene/t21all.htm))。

表 18

基因名称	检索号	比率	定位	CpG 岛	符号
淀粉状蛋白 $\beta$ (A4)前体蛋白质(蛋白酶连接蛋白-II, 阿尔茨海默氏病)	M28373	1.38	21q21.3	Y	APP
ATP-结合盒, 亚家族 G (WHITE), 成员 1	AF038175	1.23	21q22.3	Y	ABCG1
自身免疫调节剂 (自身免疫多内分泌腺病白假丝酵母病外胚层营养不良)	AJ009610	1.12	21q22.3	Y	AIRE
BTB 和 CNC 同源物 1, 碱性亮氨酸拉链转录因子 1	AI830904	1.02	21q22.11	Y	BACH1
BTG 家族, 成员 3	BE896159	1.83	21q21.1-q21.2	Y	BTG3
碳酸还原酶 1	AP000688	1.28	21q22.13		CBR1
碳酸还原酶 3	AB003151	1.06	21q22.2	Y	CBR3
染色质装配因子 1, 亚基 B (p60)	NM_005441	0.97	21q22.13	Y	CHAF1B
21 号染色体可读框 18	AB004853	1.08	21q22.12	Y	C21orf18
21 号染色体可读框 18	AA984919	0.99	21q22.12	Y	C21orf18
21 号染色体可读框 2	AF001754	0.89	21q22.3	Y	C21orf2
胶原, VI 型, $\alpha$ 1	X99135	1.58	21q22.3	Y	COL6A1
胶原, VI 型, $\alpha$ 2	AI635289	1.23	21q22.3	Y	COL6A2
胶原, XVIII 型, $\alpha$ 1	AF018081	1.17	21q22.3	Y	COL18A1
柯萨奇病毒和腺病毒受体	AI557255	1.23	21q21.1	Y	CXADR
半胱氨酸蛋白酶抑制剂 B(stefin B)	BF341232	1.94	21q22.3		
21 号染色体上的 DNA 片段(独特的)2056 表达的序列	AL137757	1.07	21q22.3	Y	D21S2056E
唐氏综合征细胞粘着分子	AF217525	0.9	21q22.2	Y	DSCAM

唐氏综合征关键区基因 1	U85267	0.82	21q22.12	Y	DSCR1
唐氏综合征关键区基因 3	D87343	1.19	21q22.2	Y	DSCR3
f-box 和 WD-40 结构域蛋白质 1B	AA436684	0.9	21q22.11		
谷氨酸受体, 离子化作用的, 红藻氨酸 1	NM_000830	0.96	21q21.3	Y	GRIK1
HMT1(hnRNP 甲基转移酶, 啤酒糖酵母 ( <i>S. cerevisiae</i> ))-样 1	NM_001535	1.15	21q22.3	Y	HRMT1L1
羧化全酶合成酶 (生物素-[丙酰辅酶 A-羧化酶 (ATP-水解)]连接酶)	D87328	0.95	21q22.13	Y	HLCS
整联蛋白, $\beta 2$ (抗原 CD18(p95), 淋巴细胞功能相关的抗原 1; 巨噬细胞抗原 1 (mac-1) $\beta$ 亚基)	X64072	0.83	21q22.3	Y	ITGB2
干扰素 ( $\alpha$ , $\beta$ 和 $\omega$ ) 受体 1	AU137565	0.94	21q22.11	Y	IFNAR1
干扰素 ( $\alpha$ , $\beta$ 和 $\omega$ ) 受体 2	L41943	1.28	21q22.11	Y	IFNAR2
干扰素 $\gamma$ 受体 2 (干扰素 $\gamma$ 转导物 1)	U05875	1.41	21q22.11	Y	IFN
干扰素 $\gamma$ 受体 2 (干扰素 $\gamma$ 转导物 1)	U05875	1.41	21q22.11	Y	
白细胞介素 10 受体, $\beta$	Z17227	0.97	21q22.11	Y	IL10RB
Intersectin 1 (SH3 结构域蛋白质)	AI033970	0.95	21q22.11	Y	
KIAA0653 蛋白质	AI421115	1.32			
微型染色体维持缺陷 (啤酒糖酵母) 3-相关的蛋白质	AB011144	1.14	21q22.3	Y	MCM3AP
粘液病毒 (流感) 抗性 1, 鼠同系物 (干扰素-可诱导的蛋白质 78)	NM_002462	1.19	21q22.3	Y	MX1
粘液病毒 (流感) 抗性 2, 鼠同系物	M30818	1.03	21q22.3	N	MX2
神经细胞粘着分子 2	U75330	1.07	21q21.1	N	NCAM2
核受体相互作用蛋白 1	AF248484	1.3	21q11.2	N	NRIP1
PBX/有结的 1 同源框 (hoemobox) 1	Y13613	0.96	21q22.3	Y	PKNOX1
中心粒周围蛋白	AB007862	0.93	21q22.3	Y	PCNT2
果糖磷酸激酶, 肝脏	AL041002	1.29	21q22.3	Y	PFKL
磷酸核糖甘氨酸酰胺甲酰基转移酶, 磷酸核糖甘氨酸酰胺合成酶, 氨基咪唑核苷酸合成酶	AA436452	0.98	21q22.11		
垂体肿瘤转化 1 相互作用蛋白	BE795643	1.58	21q22.3	Y	PITGIP
内向整流钾通道, 亚家族 J, 成员 15	U73191	1.42	21q22.13	N	KCNJ15
蛋白酶, 丝氨酸, 7 (肠激酶)	U09860	0.87	21q21.1		
PWP2(周期性色氨酸蛋白质, 酵母)同系物	AP001753	0.95	21q22.3	Y	PWP2H

吡哆醛 (吡哆醇, 维生素 B6) 激酶	BE742236	1.5	21q22.3	Y	PDXK
runt 相关的转录因子 1 (急性髓细胞样白血病 1, aml1 癌基因)	D43968	0.89	21q22.12	Y	RUNX1
S100 钙结合蛋白, $\beta$ (神经的)	AV701741	1.21	21q22.3	Y	S100B
SH3 结构域结合性富谷氨酸蛋白	BE501723	0.96	21q22.2	N	SH3BGR
Single-minded (果蝇 ( <i>Drosophila</i> )) 同系物 2	U80456	1.32	21q22.13	Y	SIM2
SMT3(mif two 3 的抑制因子, 酵母) 同系物 1	W55901	1.34	21q22.3	Y	SMT3H1
SON DNA 结合蛋白	X63071	0.92	21q22.11		SON
超氧化物歧化酶 1, 可溶性的 (肌萎缩性侧索硬化 1 (成人))	AI421041	1.04	21q22.11	Y	SOD1
synaptojanin 1	NM_003895	0.97	21q22.11	Y	SYNJ1
三十四肽重复结构域 3	D84294	1.23	21q22.13	N	TTC3
瞬时型感受器电位通道 7	AB001535	0.97	21q22.3	Y	TRPM2
跨膜蛋白酶, 丝氨酸 2	U75329	1.16	21q22.3	Y	TMPRSS2
跨膜蛋白质 1	U61500	0.95	21q22.3	Y	TMEM1
富色氨酸碱性蛋白质	NM_004627	1.39	21q22.3	Y	WRB
遍在蛋白缀合酶 E2G 2(与酵母 UBC7 同源)	AL163300	0.99	21q22.3	Y	
v-ets 鸟类成红细胞增生病毒 E26 癌基因同系物 2	AF017257	0.98	21q22.2	Y	ETS-2

根据上面的表 18, 显然在三体性状态中 21 号染色体的基因的表达存在变化; 一些基因高度超表达 (即, 比率  $X=1.5$ ; 例如, PDXK), 而其他基因表达不足 (即, 比率  $X<1$ ; 例如, DSCAM)。该变化的原因可以是甲基化的 CpG 位点数。从而, 例如, 1.2 的比率提示过量等位基因中不是所有 CpG 位点都被甲基化, 而那些仍被甲基化的位点防止了 1.5 倍超表达 (即, 三个等位基因的最大超表达)。

### 实施例 5

在第 21 对染色体三体性中 21 号染色体的 DSCAM 和 IFNAR1 基因被部分甲基化

### 实施例 5a

### DSCAM

选择唐氏综合征细胞粘着分子 (DSCAM) 基因 (GenBank 检索号 AF217525) 以显示第 21 对染色体三体性中部分沉默的基因的甲基化模式 (即,  $X < 1.5$ )。

假设 DSCAM 外显子 1 上游的 CpG 岛的甲基化可以抑制 DS 患者中该基因的超表达。在图 4a 中给出了 DSCAM 启动子的天然序列。亚硫酸氢盐处理后所得推定的序列在图 4b 中显示。

#### 实验步骤

细胞 - 从 Coriell Institute NJ 得到来自健康和 DS 影响的胚胎的培养 12 天的羊水细胞。DS 细胞目录号 GMO2067。正常细胞目录号。

DNA 提取 - 见上面的实施例 3。

DSCAM 甲基化的基于测序的分析 - 下面的表 19 - 21 列出了用于从来自健康受试者和唐氏综合征影响的受试者的组织和细胞扩增 DSCAM 的引物和 PCR 条件。

使用下面表 19 中列出的引物和上面表 14 中描述的反应混合物试剂和浓度进行 PCR 反应。

表 19

引物名称	引物序列(5'-3')/SEQ ID NO:	位置 (AL163283)
DSCAM-fl-bis	GTTATATGGATTTTTTGTAAATTTTTTTT/87	333350-333379
DSCAM-r1-bis	TCTCTACTACTACTTTAAACTACAAAAC/151	333456-333481
DSCAM-nes-fl-bis	GGTTTTAGTTATATGGATTTTTTGTAAAT/152	333344-333373

表 20-步骤 1

温度	时间	循环数
94°C	4 分钟	35
94°C	45 秒	
52°C	45 秒	
72°C	1 分钟	
72°C	7 分钟	

\*使用引物 DSCAM-nes-fl-bis 和 DSCAM-r1-bis 在缓冲液 B-DS 中实现反应。

将所得 PCR 产物为 142bp。

将 PCR 产物用作第二次 PCR 反应的模板 (1/20 终体积)。

表 21 - 步骤 2

温度	时间	循环数
94°C	4 分钟	35
94°C	45 秒	
53°C	45 秒	
72°C	1 分钟	
72°C	7 分钟	

使用引物 DSCAM-f1-bis 和 DSCAM-r1-bis 在缓冲液 NEB 中实现反应。

所得 PCR 产物为 135bp。将 PCR 反应混合物加样到 3% 琼脂糖凝胶上并如上面实施例 3 中描述的纯化 135bp 产物。通过测序证实产物的序列同一性也如上述进行。

#### 结果

来自 DS 胚胎的羊水细胞中 DSCAM 甲基化状态的序列分析表明两种情况中 25% 的克隆显示出 CpG 位点上的甲基化。这些结果表明 DSCAM 的仅部分甲基化，提示优选使用寡核苷酸微阵列用于检测 DSCAM 甲基化。适于检测 DSCAM 甲基化的寡核苷酸在下面的表 22 中列出。

表 22



<b>WT 探针 (5'-3') / SEQ ID NO:</b>	<b>甲基化探针 (5'-3') / SEQ ID NO:</b>	<b>在染色体 (UCSC No.) 和 AL163283 克隆中的位置</b>
tttttgtttgtgagtcgggtg/246	ttttgtttgcgagttgggtg/153	41139457- 41139477 333376-333396
	ttttgtttgtgagtcgggtg/154	
	ttttgtttgtgagttgggcg/155	
	ttttgtttgcgagtcgggtg/156	
	ttttgtttgcgagttgggcg/157	
	ttttgtttgtgagtcgggcg/158	
	ttttgtttgcgagtcgggcg/159	
gtttgtgagttgggtgagtga/160	gtttgcgagttgggtgagtga/161	41139462- 41139482 333381-333401
	gtttgtgagtcgggtgagtga/162	
	gtttgtgagttgggcgagtga/163	
	gtttgcgagtcgggtgagtga/164	

	gtttgcgagttgggagtg/165	
	gtttgtgagtcgggagtg/166	
	gtttgcgagtcgggagtg/167	
gtgagttgggtgagtg/168	gagtcgggtgagtg/169	41139475- 41139495 333394-333414
	gtgagtcgggtgagtg/170	
	gtgagttgggagtg/171	
	gtgagttgggtgagtcg/172	
	gtgagttgggagtg/173	
	gtgagtcgggtgagtcg/174	
	gagtcgggtgagtcg/175	
	gtgagtcgggagtg/176	
	gagtcgggagtg/177	
	gagtcgggagtcg/178	
	gagtcgggtgagtg/179	
	gagtcgggagtg/180	
	gagtcgggtgagtcg/181	
	gagtcgggagtcg/182	
	gagtcgggagtcg/183	
tgagtgagttgagtg/184	gagtgagttgagtg/185	41139476- 41139496 333395-333415
	tgagtgagtcgagtg/186	
	tgagtgagttgagtg/187	
	tgagtgagttgagtcg/188	
	gagtgagtcgagtg/189	
	tgagtgagttgagtcg/190	
	tgagtgagtcgagtcg/191	
	gagtgagttgagtcg/192	
	tgagtgagtcgagtcg/193	
	gagtgagttgagtcg/194	
	gagtgagtcgagtcg/195	
	tgagtgagtcgagtcg/196	
	gagtgagttgagtcg/197	
	gagtgagtcgagtcg/198	
	gagtgagtcgagtcg/199	
tgaagttgagtg/200	tgaagtcgagtg/201	41139480- 41139500 333399-333419
	tgaagttgagtcgagtg/202	
	tgaagttgagtg/203	
	tgaagttgagtcgagtg/204	
	tgaagttgagtcgagtg/205	
	tgaagtcgagtg/206	
	tgaagttgagtcgagtg/207	
	tgaagtcgagtcgagtg/208	
	tgaagtcgagtcgagtg/209	
	tgaagtcgagtcgagtg/210	
	tgaagtcgagtcgagtg/211	
	tgaagtcgagtcgagtg/212	
	tgaagttgagtcgagtg/213	
	tgaagtcgagtcgagtg/214	
aagttgagtg/215	aagtcgagtg/216	41139481- 41139501 333401-333421
	aagttgagtcgagtg/217	
	aagttgagtg/218	
	aagttgagtcgagtg/219	
	aagttgagtcgagtg/220	
	aagtcgagtg/221	
	aagttgagtcgagtg/222	
	aagtcgagtcgagtg/223	

	aagtcgagcgtggaggtgagt/224	
	aagtcgagcgcggaggtgagt/225	
	aagtcgagcgtggagggcagt/226	
	aagtcgagtgccggagggcagt/227	
	aagttgagcgcggagggcagt/228	
	aagtcgagcgcggagggcagt/229	
agtgtggaggtgagttaggat/230	gtgcggaggtgagttaggat/231	41139488- 41139508 333407-333427
	gcgtggaggtgagttaggat/232	
	gtgtggagggcagtaggat/233	
	gtgcggagggcagtaggat/234	
	gcgtggagggcagtaggat/235	
	gcgcggaggtgagttaggat/236	
	gcgcggagggcagtaggat/237	
tgtttttggttggtggggtgt/238	tgtttttggtcgttggggtgt/239	41139517- 41139537 /333436-333456
	tgtttttggttggtggggcgt/240	
	tgtttttggtcgttggggcgt/241	
gttgttggggtgttttgtagt/242	gtcgttggggtgttttgtagt/243	41139525- 41139545 333444-333464
	gttgttggggcgtttttagt/244	
	gtcgttggggcgtttttagt/245	

## 实施例 5b

### IFNAR1

从上面的表 18, 显然在第 21 对染色体三体性中, 干扰素 ( $\alpha$ 、 $\beta$  和  $\omega$ ) 受体 1 (IFNAR1, GenBank 检索号 AU137565) 被部分沉默。如实施例 5a 中描述的检查了细胞和组织中 IFNAR1 的甲基化模式。在图 5a 中给出 IFNAR1 启动子的天然序列。在图 5b 中显示了亚硫酸氢盐处理后得到的推定序列。

#### 实验步骤

细胞 - 见上面。

DNA 提取 - 见上面的实施例 3。

DSCAM 甲基化的基于测序的分析 - 下面的表 23 - 25 列出了用于从来自健康受试者和唐氏综合征影响的受试者的组织和细胞扩增 IFNAR1 的引物和 PCR 条件。

使用下面表 23 中列出的引物和上面表 14 中描述的反应混合物试剂和浓度实现 PCR 反应。

表 23

引物名称	(AY654286) 中的位置	序列 (5'-3')/SEQ ID NO:
IFNR-f4-bis	1327-1351	TTTTAGTTTTATTGGTTTTAGGT/247
IFNR-r4-bis	1372-1396	AAAAACCTTAACCTTCACAAAATC/248
IFNR-nes-f-bis	1533-1557	ATTGTTTAAGATTTAGGGTTAGTA/ 249

表 24-步骤 1

温度	时间	循环数
94°C	4 分钟	35
94°C	45 秒	
54°C	45 秒	
72°C	1 分钟	
72°C	7 分钟	

使用引物 IFNR-f4-bis 和 IFNR-r4-bis 在缓冲液 B-DS 中实现反应。

所得 PCR 产物为 231 bp。

将 PCR 产物用作第二次 PCR 反应的模板 (1/20 终体积)。

表 25 - 步骤 2

温度	时间	循环数
94°C	4 分钟	35
94°C	45 秒	
56°C	45 秒	
72°C	1 分钟	
72°C	7 分钟	

使用引物 IFNR-nes-f-bis 和 IFNR-r4-bis 在缓冲液 NEB 中实现反应。

所得 PCR 产物为 186 bp。

将 PCR 反应产物在 2.5% 琼脂糖凝胶上分离并如上面实施例 3 中描述的纯化 186 bp 产物。通过如上文描述的测序证实产物的序列同一性。

## 结果

看到 DS 样品中 IFNAR1 等位基因的甲基化。

根据上面的描述，可以设想 DSCAM 和 IFNAR1 甲基化状态可以作为第 21 对染色体三体性的有价值的诊断标记。这些结果还表明在第 21 对染色体三体性中没有上调的其他基因也可以作为染色体扩增的标记。

## 实施例 6

### 第 13 对染色体常染色体三体性的推定的标记

下面的表 26 显示了从第 13 对染色体三体性基因型的受试者得到的羊水细胞相对于从正常受试者得到的羊水细胞中 13 号染色体基因表达的比率 ([www.hgu.mrc.ac.uk/Research/Cellgen/Supplements/Unigene/t13all.htm](http://www.hgu.mrc.ac.uk/Research/Cellgen/Supplements/Unigene/t13all.htm))。有趣地，与第 21 对染色体三体性（其中多数基因被沉默（RNA 为不稳定状态水平））相反，该基因表达谱不在 13 号染色体中发生，从而解释了 21 号染色体扩增的活力。

表 26

基因名称	检索号	比率	位置
ADP-核糖基转移酶 (NAD <sup>+</sup> ; 聚 (ADP-核糖) 聚合酶) - 样 1	NM_006437	1.6	13q34
ATP 酶, H <sup>+</sup> /K <sup>+</sup> 交换, $\beta$ 多肽	NM_000705	1.11	13q34
羧肽酶, B2(血浆)	NM_001872	0.92	13q14.13
CDC16(细胞分裂周期 16, 啤酒糖酵母, 同系物)	NM_003903	1.08	13q34
蜡样脂 - 褐质症, 神经元的 5	NM_006493	1.5	13q22.3
凝固因子 X	AL521984	1.09	13q34
胶原, IV 型, $\alpha$ 1	XM_007094	0.43	13q34
Cullin 4A	AI638597	1.58	13q34
细胞周期蛋白 A1	NM_003914	1.06	13q13.3
依赖细胞周期蛋白的激酶 8	BE467537	1.59	13q12
Dachshund(果蝇)同系物	NM_004392	1.69	13q21.33
DnaJ(Hsp40)同系物, 亚家族 C, 成员 3	AW772531	1.22	13q32.1
doublecortin 和 CaM 激酶样 - 1	NM_004734	1.4	13q13.3
内皮缩血管肽受体型 B	BE837728	1	13q22.3
切割修复交叉互补啮齿动物修复缺陷, 互补群 5 (干皮病色素变性, 互补群 G(Cockayne 综合征))	NM_000123	1.12	13q33.1
FERM, RhoGEF(ARHGEF)和普列克底物蛋白结构域蛋白 1 (来自软骨细胞)	BF793662	1.19	13q32.2
成纤维细胞生长因子 9 (神经胶质激活因子)	AI869879	0.98	13q12.11
fms-相关酪氨酸激酶 1 (血管内皮生长因子/血管通透因子受体)	NM_002019	1.73	13q12.3
fms-相关酪氨酸激酶 1 (血管内皮生长因子/血管通透因子受体)	NM_002019	1.73	13q12.3
fms-相关酪氨酸激酶 3	NM_004119	1.3	13q12.2
分叉头框 O1A (横纹肌肉瘤)	NM_002015	1.17	13q14.11
生长抑制特异的 6	NM_000820	0.76	13q34
人 BRCA2 区, mRNA 序列 CG011	U50536	0.67	13q13.1
生长抑制剂 1 家族, 成员 1	AF181850	0.99	13q34

整联蛋白, $\beta$ -样 1 (具有 EGF-样重复结构域)	NM_004791	0.68	13q33.1
karyopherin $\alpha$ 3 (输入蛋白 $\alpha$ 4)	NM_002267	1.11	13q14.2
Klotho	NM_004795	1.44	13q13.1
连接酶 IV, DNA, ATP-依赖性的	NM_002312	1.3	13q33.3
脂肪瘤 HMGIC 融合配偶体	N67270	1.05	13q13.3
淋巴细胞胞质蛋白质 1 (L-肌动蛋白)	BF035921	0.98	13q14.13
线粒体中间肽酶	AA524277	0.88	13q12.12
线粒体翻译释放因子 1	AI884353	0.99	13q14.11
肌管素 (myotubularin) 相关蛋白质 6	AW205652	1.63	13q12.13
成骨细胞特异因子 2 (成束蛋白 1 样)	N71912	1.91	13q13.3
peroxiredoxin 2	AL523978	1.11	
丙酰辅酶 A 羧化酶, $\alpha$ 多肽	NM_000282	1.22	13q32.3
蛋白质磷酸酶 1, 调节 (抑制剂) 亚基 2	AI141349	1.57	
嘌呤能受体 (家族 A, 组 5)	AI823889	1.25	13q14.2
复制因子 C(活化剂 1)3 (38kD)	AA907044	0.96	13q13.2
ret 指蛋白 2	AL526890	1.25	13q14.2
成视网膜细胞瘤 1 (包括骨肉瘤)	NM_000321	1.65	13q14.2
sciellin	AK025320	0.95	13q22.3
丝氨酸/苏氨酸激酶 24 (Ste20, 酵母同系物)	NM_003576	1.12	
丝氨酸/苏氨酸激酶 24 (Ste20, 酵母同系物)	AU146392	1.06	13q32.2
溶质载体家族 10 (钠/胆汁酸协同转运蛋白家族), 成员 2	NM_000452	1.32	13q33.1
溶质载体家族 25 (线粒体载体; 鸟氨酸转运蛋白) 成员 15	AI382550	0.87	13q14.11
溶质载体家族 7 (阳离子氨基酸转运蛋白, $y^+$ 系统), 成员 1	X57303	1.09	13q12.3
Charlevoix-Saguenay 的痉挛性共济失调 (sacsin)	AB018273	0.97	13q12.12
Sprouty (果蝇) 同系物 2	NM_005842	1.54	13q31.1
转录因子 Dp-1	NM_007111	1.23	13q34
跨膜 9 超家族成员 2	AU131084	0.97	13q32.3
三肽基肽酶 II	NM_003291	1.1	13q33.1
肿瘤坏死因子 (配体) 超家族, 成员 11	AF053712	1.14	13q14.11
Zic 家族成员 2 (奇数配对果蝇同系物)	AF188733	1.9	13q32.3
锌指蛋白 198	AL138688	1.45	13q12.11

## 实施例 7

第 9 对染色体三体性的基因和可以用于检测其甲基化状态的引物

第 9 对染色体三体性是罕见的染色体疾病。特征性特征包括胎儿生长延缓、出生时存在心脏缺陷、面部异常（例如，低位和/或畸形耳）、异常的小头、肾和/或生殖器异常、骨骼异常（例如，固定和/或脱位的关节），和/或脑的畸形。

9 号染色体上 p16 在细胞周期和许多病理（包括黑素瘤、膀胱癌和肺癌）中起重要作用。肿瘤抑制基因 p16 的表达在多种癌症中受到阻遏，所述癌症包括膀胱癌、结肠癌和成视网膜细胞瘤。已经表明 p16 启动子中 CpG 岛的甲基化负责某些情况中该基因的失活 [Sharpless (2003) *Oncogene*. 22 (20): 3092-8; Virmani (2003) *Methods Mol Biol*. 2003; 222: 97-115]。

将 CpG WIZ®p16 Amplification Kit (Chemicon International, Inc.) 用于通过甲基化特异性 PCR (MSP) 确定 p16 启动子的甲基化状态。该试剂盒含有靶定启动子区的引物，所述区中的序列在亚硫酸氢盐处理后最趋异。已经鉴定了 PCR 参数，从而使得试剂盒中的所有引物组都在相同条件下扩增。还包括 p16 的对照基因组 DNA 样品（甲基化的和未甲基化的）。

### 实验步骤

使用 CpGenome DNA Modification Kit (Intergen, New York, NY) 进行亚硫酸氢盐转化。根据生产商的推荐用亚硫酸氢钠处理 1 $\mu$ g DNA。转化后，将亚硫酸氢盐处理的 DNA 以 25 $\mu$ l 的总体积重悬浮。

下面的表 27 概述了用于检测上述基因的甲基化状态的方法。

表 27

方法	实例	三体性
DNA 测序	*APP, AR, p16, DSCAM BACH1 ETS2 INFAR1	21, X, 9
限制酶	雄激素受体	X
MSP	雄激素受体	X
微阵列	APP	21, X
用于突变检测的商业试剂盒	P16	9



将理解本发明的某些特征为了清楚而在分开的实施方案的上下文中描述，但是所述特征也可以与单个实施方案组合提供。相反地，为了简明而在单个实施方案的上下文中描述的本发明的多种特征也可以分开或者以任意适宜的亚组合提供。

尽管已经结合本发明的特定实施方案描述了本发明，但是显然许多备选方案、修改和变形对本领域技术人员而言是显而易见的。因此，意在包括所附权利要求的精神和宽范围内的所有此类备选方案、修改和变形。该说明书中提到的所有出版物、专利和专利申请在此完整引入本说明书中作为参考，就像每个单独的出版物、专利和专利申请具体地且分别指出引入本文作为参考一样。此外，不应将本申请中对任何参考文献的引用或者鉴别理解为承认此类参考文献可以作为本发明的现有技术得到。

<110> Trisogen Biotechnology Limited Partnership

<120> 用于检测基因座拷贝数改变的方法和试剂盒

<130>

<160> 249

<170> PatentIn version 3.2

<210> 1

<211> 25

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 1

tggttttaga tttttttttt tattg

<210> 2

<211> 25

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 2

acctaccact accaaaaaaaa ctaac

<210> 3

<211> 15

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 3  
tattttttgg tgta

<210> 4  
<211> 15  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 4  
tatttttcgg tgta

<210> 5  
<211> 15  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 5  
gagggggtgt gtggg

<210> 6  
<211> 15  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 6  
gaggggCGT gtggg

<210> 7  
<211> 15  
<212> DNA  
<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 7

gttaaggtgt tgtat

<210> 8

<211> 15

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 8

gttaaggcgt tgtat

<210> 9

<211> 15

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 9

ttgtgggtgt ggggt

<210> 10

<211> 15

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 10

ttgtgggcgt ggggt

<210> 11  
<211> 15  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 11  
tttttggtgt gagtg

<210> 12  
<211> 15  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 12  
tttttgccgt gagtg

<210> 13  
<211> 15  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 13  
gagtgggtgt agttt

<210> 14  
<211> 15  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 14  
gagtgggcgt agttt

<210> 15  
<211> 16  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 15  
tttgggtggtg ttgtta

<210> 16  
<211> 16  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 16  
tttgggtggcg ttgtta

<210> 17  
<211> 16  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 17  
ggttgttgtg tttggg

<210> 18  
<211> 16  
<212> DNA  
<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 18

ggttggtgcg tttggg

<210> 19

<211> 15

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 19

tgttggttg ggagt

<210> 20

<211> 15

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 20

tgttggtcgg ggagt

<210> 21

<211> 15

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 21

ttttttttgg tgtga

---

<210> 22

<211> 15

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 22

tttttttcgg tgtga

<210> 23

<211> 15

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 23

agttttttgg tggtg

<210> 24

<211> 15

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 24

agtttttcgg tggtg

<210> 25

<211> 15

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸



---

<400> 25  
ggtgggttgg attag

<210> 26  
<211> 15  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 26  
ggtgggtcgg attag

<210> 27  
<211> 15  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 27  
tggggagtgg agggg

<210> 28  
<211> 15  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 28  
gggggagcgg agggg

<210> 29  
<211> 15  
<212> DNA  
<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 29

tttttggcgt gagtg

<210> 30

<211> 15

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 30

tttttggcgt gagtg

<210> 31

<211> 15

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 31

gggggtgtgt ggggt

<210> 32

<211> 15

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 32

gggggtgcgt ggggt

---

<210> 33  
<211> 15  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 33  
gtgtaggtgg tgtta

<210> 34  
<211> 15  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 34  
gtgtaggcgg tgtta

<210> 35  
<211> 15  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 35  
tttggtgtga gtggg

<210> 36  
<211> 15  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

---

<400> 36  
tttggtagca gtggg

<210> 37  
<211> 15  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 37  
aatttggtgt ttta

<210> 38  
<211> 15  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 38  
aatttgccgt ttta

<210> 39  
<211> 15  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 39  
atatttgct tttgg

<210> 40  
<211> 15  
<212> DNA  
<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 40

atatttgcgt tttag

<210> 41

<211> 15

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 41

tgtgttttgg gttaa

<210> 42

<211> 15

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 42

tgtgtttcgg gttaa

<210> 43

<211> 15

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 43

ggtgttgtgt gtgga

---

<210> 44  
<211> 15  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 44  
ggtgtggcgt gtgga

<210> 45  
<211> 15  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 45  
tgtggcgtgt ggagt

<210> 46  
<211> 15  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 46  
tgtggcgcgt ggagt

<210> 47  
<211> 15  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

---

<400> 47  
tggagtttgg tgtgt

<210> 48  
<211> 15  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 48  
tggagttcgg tgtgt

<210> 49  
<211> 15  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 49  
agtttggtgt gtttt

<210> 50  
<211> 15  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 50  
agtttggcgt gtttt

<210> 51  
<211> 15  
<212> DNA  
<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 51

aat t t t t g c g t t a g t t

<210> 52

<211> 15

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 52

aat t t t t g c g t t a g t t

<210> 53

<211> 15

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 53

g t t a g t t t g g t g g t t

<210> 54

<211> 15

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 54

g t t a g t t c g g t g g t t



---

<210> 55  
<211> 24  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 55  
tccagaatct gttccagagc gtgc

<210> 56  
<211> 24  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 56  
gctgtgaagg ttgctgttcc tcat

<210> 57  
<211> 24  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 57  
tagaatttgt tttagagtgt gtgt

<210> 58  
<211> 18  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

---

<400> 58

tttgtttag agcgtgcg

<210> 59

<211> 20

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 59

aaaaccatcc tcaccctact

<210> 60

<211> 29

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 60

agatttagtt aagttaagg atggaagtg

<210> 61

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 61

gggttgggaa gggtttattt t

<210> 62

<211> 26

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 62

aaaaaccatc ctcaccctac tactac

<210> 63

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 63

ggtttatttt tggttgttgt t

<210> 64

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 64

tatttttggg tggttgtaaa g

<210> 65

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 65

ttttggttgt tgtaagatt t

<210> 66

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 66

ggtttatattt cggttggtt t

<210> 67

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 67

ggtttatattt tggtcggtt t

<210> 68

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 68

ggtttatattt tggttgct t

<210> 69

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 69

ggtttatattt cggtcgttgt t

<210> 70

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 70

ggtttatattt tggtcgtcgt t

<210> 71

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 71

ggtttatattt cggltgtcgt t

<210> 72

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 72

ggtttatattt cggtcgtcgt t

<210> 73

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 73

tatttttcggt tgttgtttaa g

<210> 74

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 74

tatttttggg cgggtgttaa g

<210> 75

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 75

tatttttggg tgcgtttaa g

<210> 76

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 76

tattttcggg cgttgtttaa g

<210> 77

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 77

tatTTTTGGT cGtcGTTTaa g

<210> 78

<211> 20

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 78

tatTTTcGGT gTcGTTTaa g

<210> 79

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 79

tatTTTcGGT cGtcGTTTaa g

<210> 80

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

---

<400> 80  
tttcggttgt tgtaagatt t

<210> 81  
<211> 21  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 81  
ttttggtcgt tgtaagatt t

<210> 82  
<211> 21  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 82  
ttttggttgt cgtaagatt t

<210> 83  
<211> 21  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 83  
tttcggtcgt tgtaagatt t

<210> 84  
<211> 21  
<212> DNA  
<213> 人工序列



<220>

<223> 单链 DNA 寡核苷酸

<400> 84

ttttggtcgt cgttaagatt t

<210> 85

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 85

tttcggttgt cgttaagatt t

<210> 86

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 86

tttcggtcgt cgttaagatt t

<210> 87

<211> 30

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 87

gttatatgga ttttttggtt aatTTTTTTT

---

<210> 88

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 88

taagatttat cgaggagttt t

<210> 89

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 89

taagatttat tgaggagttt t

<210> 90

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 90

tgtttttagag tgtgtgtgaa g

<210> 91

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

---

<400> 91  
tgtttttagag cgtgtgtgaa g

<210> 92  
<211> 21  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 92  
tgtttttagag tgtgcgtgaa g

<210> 93  
<211> 21  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 93  
tgtttttagag tgtgtgcgaa g

<210> 94  
<211> 21  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 94  
tgtttttagag cgtgcgtgaa g

<210> 95  
<211> 21  
<212> DNA  
<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 95

tgtttttagag cgtgtgcgaa g

<210> 96

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 96

tgtttttagag tgtgcgcgaa g

<210> 97

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 97

tgtttttagag cgtgcgcgaa g

<210> 98

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 98

ttagagtgtg tgtgaagtga t

<210> 99

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 99

ttagagcgtg tgtgaagtga t

<210> 100

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 100

ttagagtgtg cgtgaagtga t

<210> 101

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 101

ttagagtgtg tgcgaagtga t

<210> 102

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

---

<400> 102  
ttagagcgtg cgtgaagtga t

<210> 103  
<211> 21  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 103  
ttagagcgtg tgcgaagtga t

<210> 104  
<211> 21  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 104  
ttagagtgtg cgcgaagtga t

<210> 105  
<211> 21  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 105  
ttagagcgtg cgcgaagtga t

<210> 106  
<211> 21  
<212> DNA  
<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 106

agagtgtgtg tgaagtgatt t

<210> 107

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 107

agagcgtgtg tgaagtgatt t

<210> 108

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 108

agagtgtgcg tgaagtgatt t

<210> 109

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 109

agagtgtgtg cgaagtgatt t

---

<210> 110  
<211> 21  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 110  
agagcgtgcg tgaagtgatt t

<210> 111  
<211> 21  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 111  
agagcgtgtg cgaagtgatt t

<210> 112  
<211> 21  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 112  
agagtgtgcg cgaagtgatt t

<210> 113  
<211> 22  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸



---

<400> 113  
agagcgtgcg cgcaagtgat tt

<210> 114  
<211> 21  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 114  
attdagaatt tgggttttag g

<210> 115  
<211> 21  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 115  
attdagaatt cgggttttag g

<210> 116  
<211> 21  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 116  
attdagaggt tdtgagtgta g

<210> 117  
<211> 21  
<212> DNA  
<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 117

atttagaggt cgcgagcgta g

<210> 118

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 118

atttagaggt cgtgagtgta g

<210> 119

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 119

atttagaggt tgcgagtgta g

<210> 120

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 120

atttagaggt tgtgagcgta g

---

<210> 121  
<211> 21  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 121  
atttagaggt cgcgagtgta g

<210> 122  
<211> 21  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 122  
atttagaggt cgtgagcgta g

<210> 123  
<211> 21  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 123  
atttagaggt tgcgagcgta g

<210> 124  
<211> 21  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

---

<400> 124  
atttagaggt cgcgagcgta g

<210> 125  
<211> 21  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 125  
atttagaggt tgtgagtgta g

<210> 126  
<211> 21  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 126  
ttagaggtcg cgagcgtagt a

<210> 127  
<211> 21  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 127  
ttagaggtcg tgagtgtagt a

<210> 128  
<211> 21  
<212> DNA  
<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 128

ttagaggttg cgagtgtagt a

<210> 129

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 129

ttagaggttg tgagcgtagt a

<210> 130

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 130

ttagaggtcg cgagtgtagt a

<210> 131

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 131

ttagaggtcg tgagcgtagt a

---

<210> 132

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 132

ttagaggttg cgagcgtagt a

<210> 133

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 133

ttagaggtcg cgagcgtagt a

<210> 134

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 134

aggttgtagag ttagtagtattt t

<210> 135

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

---

<400> 135  
aggtcgag cgtagtattt t

<210> 136  
<211> 21  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 136  
aggtcgtgag ttagtattt t

<210> 137  
<211> 21  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 137  
aggttgag ttagtattt t

<210> 138  
<211> 21  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 138  
aggttgtgag cgtagtattt t

<210> 139  
<211> 21  
<212> DNA  
<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 139

aggtcgag tgtagtattt t

<210> 140

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 140

aggtcgtgag cgtagtattt t

<210> 141

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 141

aggttgcgag cgtagtattt t

<210> 142

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 142

aggtcgag cgtagtattt t



---

<210> 143

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 143

tagtattttt tgggttagt t

<210> 144

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 144

tagtattttt tggcgtagt t

<210> 145

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 145

tagtattttt cgggttagt t

<210> 146

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

---

<400> 146  
tagtattttt cggcgtagt t

<210> 147  
<211> 24  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 147  
tagtattttt tggtagtagt ttgt

<210> 148  
<211> 24  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 148  
tagtattttt tggcgtagt ttgt

<210> 149  
<211> 24  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 149  
tagtattttt cggtagtagt ttgt

<210> 150  
<211> 24  
<212> DNA  
<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 150

tagtattttt cggcgtagt ttgt

<210> 151

<211> 29

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 151

tctctactac tactttaaaa ctacaaaac

<210> 152

<211> 30

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 152

ggttttagtt atatggattt ttttgtaat

<210> 153

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 153

tttttgttg cgagttgggt g

---

<210> 154  
<211> 20  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 154  
ttttgtttgt ggtcgggtg

<210> 155  
<211> 20  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 155  
ttttgtttgt ggttgggctg

<210> 156  
<211> 20  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 156  
ttttgtttgc ggtcgggtg

<210> 157  
<211> 20  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

---

<400> 157

ttttgtttgc gaggtagggcg

<210> 158

<211> 20

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 158

ttttgtttgt gaggtagggcg

<210> 159

<211> 20

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 159

ttttgtttgc gaggtagggcg

<210> 160

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 160

gtttgtgagt tgggtgagtg a

<210> 161

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 161

gtttgcgagt tgggtgagtg a

<210> 162

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 162

gtttgtgagt cgggtgagtg a

<210> 163

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 163

gtttgtgagt tgggcgagtg a

<210> 164

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 164

gtttgcgagt cgggtgagtg a

---

<210> 165  
<211> 21  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 165  
gtttgagagt tggcgagtg a

<210> 166  
<211> 21  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 166  
gtttgtgagt cggcgagtg a

<210> 167  
<211> 21  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 167  
gtttgagagt cggcgagtg a

<210> 168  
<211> 22  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

---

<400> 168  
gtgagttggg tgagtgaagt tg

<210> 169  
<211> 22  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 169  
gcgagttggg tgagtgaagt tg

<210> 170  
<211> 22  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 170  
gtgagtcggg tgagtgaagt tg

<210> 171  
<211> 22  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 171  
gtgagttggg cgagtgaagt tg

<210> 172  
<211> 22  
<212> DNA  
<213> 人工序列



<220>

<223> 单链 DNA 寡核苷酸

<400> 172

gtgagttggg tgagtgaagt cg

<210> 173

<211> 22

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 173

gtgagttggg cgagtgaagt cg

<210> 174

<211> 22

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 174

gtgagtcggg tgagtgaagt cg

<210> 175

<211> 22

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 175

gcgagttggg tgagtgaagt cg

---

<210> 176

<211> 22

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 176

gtgagtcggg cgagtgaagt tg

<210> 177

<211> 22

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 177

gcgagttggg cgagtgaagt tg

<210> 178

<211> 22

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 178

gcgagtcggg cgagtgaagt cg

<210> 179

<211> 22

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

---

<400> 179

gcgagtcggg tgagtgaagt tg

<210> 180

<211> 22

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 180

gcgagtcggg cgagtgaagt tg

<210> 181

<211> 22

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 181

gcgagtcggg tgagtgaagt cg

<210> 182

<211> 22

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 182

gcgagttggg cgagtgaagt cg

<210> 183

<211> 22

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 183

gcgagtcggg cgagtgaagt cg

<210> 184

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 184

tgagtgaagt tgagtgtgga g

<210> 185

<211> 22

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 185

cgagtgaagt tgagtgtctgg ag

<210> 186

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 186

tgagtgaagt cgagtgtgga g

---

<210> 187

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 187

tgagtgaagt tgagcgtgga g

<210> 188

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 188

tgagtgaagt tgagtgcgga g

<210> 189

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 189

cgagtgaagt cgagtgtgga g

<210> 190

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

---

<400> 190  
tgagtgaagt tgagcgcgga g

<210> 191  
<211> 21  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 191  
tgagtgaagt cgagtgcgga g

<210> 192  
<211> 21  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 192  
cgagtgaagt tgagcgtgga g

<210> 193  
<211> 21  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 193  
tgagtgaagt cgagcgtgga g

<210> 194  
<211> 21  
<212> DNA  
<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 194

cgagtgaagt tgagcgtgga g

<210> 195

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 195

cgagtgaagt cgagcgtgga g

<210> 196

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 196

tgagtgaagt cgagcgcgga g

<210> 197

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 197

cgagtgaagt tgagcgcgga g

---

<210> 198

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 198

cgagtgaagt cgagtgcgga g

<210> 199

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 199

cgagtgaagt cgagcgcgga g

<210> 200

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 200

tgaagttgag tgtggaggtg a

<210> 201

<211> 22

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸



---

<400> 201  
tgaagtcgag tgctggaggt ga

<210> 202  
<211> 21  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 202  
tgaagttgag cgtggaggtg a

<210> 203  
<211> 21  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 203  
tgaagttgag tgtggaggcg a

<210> 204  
<211> 21  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 204  
tgaagttgag tgcggaggcg a

<210> 205  
<211> 21  
<212> DNA  
<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 205

tgaagttgag cgtggaggcg a

<210> 206

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 206

tgaagtcgag tgtggaggcg a

<210> 207

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 207

tgaagttgag cgcggagggtg a

<210> 208

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 208

tgaagtcgag tgcggagggtg a

---

<210> 209

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 209

tgaagtcgag cgtggaggtg a

<210> 210

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 210

tgaagtcgag cgcggaggtg a

<210> 211

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 211

tgaagtcgag cgtggaggcg a

<210> 212

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

---

<400> 212  
tgaagtcgag tgcggaggcg a

<210> 213  
<211> 21  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 213  
tgaagttgag cgcggaggcg a

<210> 214  
<211> 21  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 214  
tgaagtcgag cgcggaggcg a

<210> 215  
<211> 21  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 215  
aagttgagtg tggaggtgag t

<210> 216  
<211> 22  
<212> DNA  
<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 216

aagtcgagtg ctggaggtga gt

<210> 217

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 217

aagttgagcg tggaggtgag t

<210> 218

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 218

aagttgagtg tggaggcgag t

<210> 219

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 219

aagttgagtg cggaggcgag t

---

<210> 220

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 220

aagttgagcg tggaggcgag t

<210> 221

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 221

aagtcgagtg tggaggcgag t

<210> 222

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 222

aagttgagcg cggaggtgag t

<210> 223

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

---

<400> 223  
aagtcgagtg cggaggtgag t

<210> 224  
<211> 21  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 224  
aagtcgagcg tggaggtgag t

<210> 225  
<211> 21  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 225  
aagtcgagcg cggaggtgag t

<210> 226  
<211> 21  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 226  
aagtcgagcg tggaggcgag t

<210> 227  
<211> 21  
<212> DNA  
<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 227

aagtcgagtg cggaggcgag t

<210> 228

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 228

aagttgagcg cggaggcgag t

<210> 229

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 229

aagtcgagcg cggaggcgag t

<210> 230

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 230

agtgtggagg tgagtaggga t



---

<210> 231  
<211> 20  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 231  
gtgCGgaggt gagtaggat

<210> 232  
<211> 20  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 232  
gcgtggaggt gagtaggat

<210> 233  
<211> 20  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 233  
gtgtggaggc gagtaggat

<210> 234  
<211> 20  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

---

<400> 234  
gtgCGgaggc gagtagggat

<210> 235  
<211> 20  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 235  
gcgtggaggc gagtagggat

<210> 236  
<211> 20  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 236  
gCGcggaggt gagtagggat

<210> 237  
<211> 20  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 237  
gCGcggaggc gagtagggat

<210> 238  
<211> 21  
<212> DNA  
<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 238

tgtttttggg t gttggggg t g

<210> 239

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 239

tgtttttggg t cgttggggg t g

<210> 240

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 240

tgtttttggg t gttggggcg t

<210> 241

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 241

tgtttttggg t cgttggggcg t

<210> 242

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 242

gttgttgggg tgtttttag t

<210> 243

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 243

gtcgttgggg tgtttttag t

<210> 244

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 244

gttgttgggg cgtttttag t

<210> 245

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

---

<400> 245  
gtcgttgggg cgttttgtag t

<210> 246  
<211> 21  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 246  
tttttgtttg tgagtcgggt g

<210> 247  
<211> 25  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 247  
tttttagtttt atttggtttt taggt

<210> 248  
<211> 25  
<212> DNA  
<213> 人工序列

<220>  
<223> 单链 DNA 寡核苷酸

<400> 248  
aaaaaacctt aaccttcaca aaatc

<210> 249  
<211> 25  
<212> DNA  
<213> 人工序列

<220>

<223> 单链 DNA 寡核苷酸

<400> 249

attgtttaag attttaggt tagta







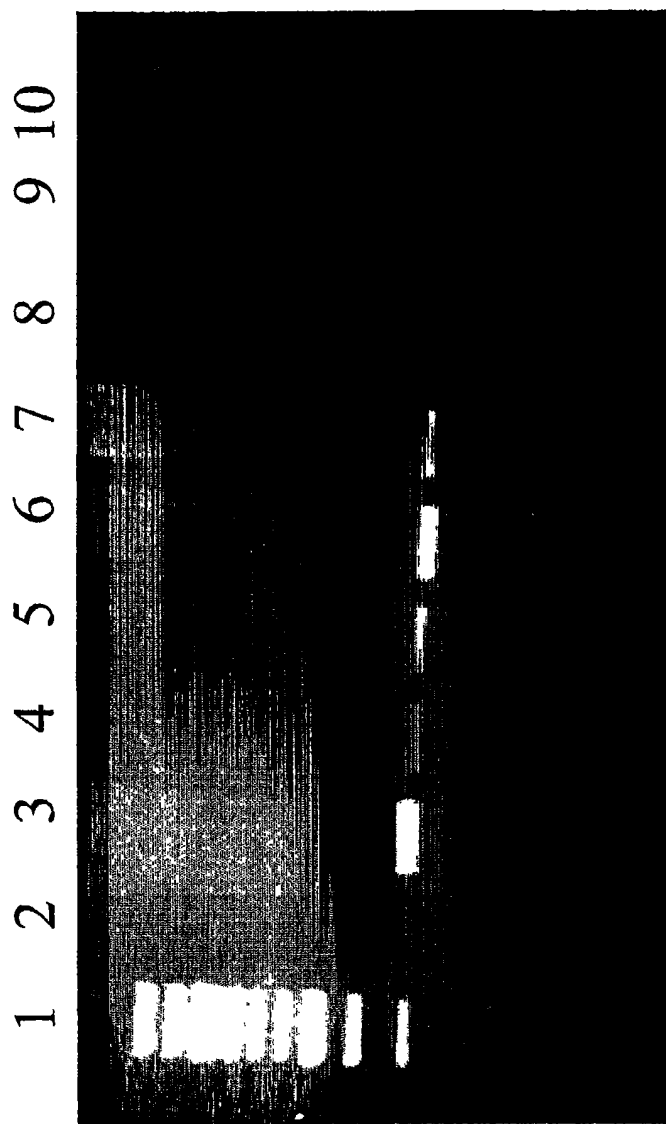


图 3

**██████████GTCAC<sup>#</sup>ATGGATCCCTCTGCCAACCTTCCCT<sup>#</sup>GCCTGCCGAGCCGGGCGAGTGA**  
**AGCCGAGCGCGGAGGCGAGCAGGGACC CCTCCCC**  
**TGCCTCTGCCCGCTGGGGC<sup>#</sup>GCTCTGCAGTCTTAAAGCAGCAGCAGAGA**

图 4a

**AGTTATATGGATTTTTTGTAAATTTTTTGTTCGAGTCGGGCCAGTGAAGTCGAGC**  
**GCGGAGGCGAGTAGGGATTTTTTTTTTGTTCGAGTCGGGCCAGTGAAGTCGAGC**  
**AGTAGTAGTAGAGA**

图 4b

TTCCAGCCTCATCTGGTTCCAGGCCGCTGGGGACTCCCAACGCC   
 AGCGCCCGGGCGGAGAAGGGCGAGGA  
CGAAGAGCGCCGGGCCGACCCAGGAGCCCACCCGCGCCCTCCGACTG  
CAGACATGGGGAAGAGACCGCGGAACTCCAAAGTCGCTGGGTCTGCCG  
AGGTGTGTGCCG GATCCTGTGAAGGTCAAGGCCTCCT

图 5a

\*TTTAGTTTTATTGGTTTTAGGTCGTTGGGGATTTTAAACGTT\*\*   
 AGCGTTTCGGGCGGAGAAGGGCGAGGA  
CGAAGAGCGTCGGGTTGCCATTAGGAGTTTATTCGCGTTTTTCGATTGT  
AGATATGGGGAAGAGACCGCGGAAATTTAAAGTCGTTGGGTTTGCGTAG  
GTGTGTGCCG\*\*\*GATTTTGTGAAGGTTAAGGTTTTT

图 5b