



Title	Methods and compositions for high-throughput bisulphite DNA-sequencing and utilities
Inventor(s)	Lok, S
Citation	China Published Patent Application CN 101802223. Beijing, PRC: State Intellectual Property Office (SIPO) of the P.R.C., 2010
Issued Date	2010
URL	http://hdl.handle.net/10722/176630
Rights	Creative Commons: Attribution 3.0 Hong Kong License For Public Patent Documents



(12) 发明专利申请

(10) 申请公布号 CN 101802223 A

(43) 申请公布日 2010.08.11

(21) 申请号 200880103541.X

(22) 申请日 2008.08.08

(30) 优先权数据

60/935472 2007.08.15 US

60/935867 2007.09.05 US

(85) PCT申请进入国家阶段日

2010.02.08

(86) PCT申请的申请数据

PCT/CN2008/001435 2008.08.08

(87) PCT申请的公布数据

W02009/024019 EN 2009.02.26

(71) 申请人 香港大学

地址 中国香港薄扶林道

(72) 发明人 骆树恩

(74) 专利代理机构 中国专利代理(香港)有限公司
72001

代理人 梁谋 付磊

(51) Int. Cl.

C12Q 1/68(2006.01)

权利要求书 2 页 说明书 12 页

(54) 发明名称

用于高通量亚硫酸氢盐 DNA- 测序的方法和组合物及其用途

(57) 摘要

本文公开了生产适用于化学修饰和高通量 DNA 测序的 DNA 模板的新颖方法和组合物。本文也公开了 DNA 接头设计的方法,其中组分胞嘧啶被 5- 甲基胞嘧啶取代使得所得接头对亚硫酸氢盐介导的脱氨作用具有抗性。当所述接头连接至双链 DNA 模板时,后续的 DNA 变性和亚硫酸氢盐处理将模板 DNA 胞嘧啶差异脱氨形成尿嘧啶而所连接接头的 5- 甲基胞嘧啶对这一化学转化具有抗性使得接头序列仍然保持不变。因此可用与未改变接头序列杂交的单一引物组扩增亚硫酸氢盐处理的 DNA 两条链。也公开了生产具确定甲基化组成的对照模板以优化亚硫酸氢盐反应的条件的方法。在优选的实施方案中,本发明可用于生产适于使用传统的 Solexa™、SOLiD™或 454™- 型 DNA 测序平台进行全基因组亚硫酸氢盐 -DNA 测序以研究 DNA 甲基化的模板。

1. 测定靶 DNA 群的外遗基因组状态的方法,其包括:

将靶 DNA 群片段化至适当大小,将其与包含在一个或多个或所有位点取代其未经修饰的核苷酸类似物的经修饰核苷酸的组合物中的一个或多个不同的经修饰-接头连接,与该靶 DNA 连接得到包含下列的组合物:(1) 与靶 DNA 片段的两个末端连接的相同的经修饰-接头;或(2) 与靶 DNA 片段的每个末端连接的不同的经修饰接头;

之后,与经修饰接头连接的靶 DNA 经化学处理,使得靶 DNA 组合物和经修饰接头组合物在化学和功能上可区分,其中所连接的经修饰的一个或多个接头在实质和功能上均未发生变化;

之后,使用与经修饰接头序列互补的引物将化学处理的与经修饰接头连接的 DNA 扩增至少一个循环;

之后,对被扩增的 DNA 进行 DNA 测序。

2. 测定靶 DNA 群胞嘧啶甲基化状态的方法,其包括:

将靶 DNA 群片段化至适当大小,将其与包含在一个或多个或所有胞嘧啶位点取代胞嘧啶的经修饰核苷酸的组合物中的一个或多个不同的经修饰-接头连接,与该靶 DNA 连接得到包含以下的组合物:(1) 与靶 DNA 片段的两个末端连接的相同的经修饰-接头;或(2) 与靶 DNA 片段的每个末端连接的不同的经修饰接头;

之后,与经修饰接头连接的靶 DNA 经化学处理,使得靶 DNA 组合物和经修饰接头组合物在化学和功能上可区分,其中经修饰的一个或多个接头在实质和功能上均未发生变化;

之后,使用与经修饰接头序列互补的引物将化学处理的与经修饰接头连接的 DNA 扩增至少一个循环;

之后,对被扩增的 DNA 进行 DNA 测序。

3. 权利要求 1 的方法,其中所述靶 DNA 群的外遗基因组状态为 5- 碳位胞嘧啶的甲基化。

4. 权利要求 1 的方法,其中所述经修饰的一个或多个接头由一个或多个经修饰核苷酸组成,所述经修饰核苷酸选自任何经修饰的腺嘌呤、鸟嘌呤、胞嘧啶、尿嘧啶或胸腺嘧啶核苷酸。

5. 权利要求 1 和 2 的方法,其中所述经修饰的一个或多个接头由一个或多个选自任何甲基化核苷酸的经修饰核苷酸组成。

6. 权利要求 1 和 2 的方法,其中所述经修饰的一个或多个接头由一个或多个选自 5- 甲基胞嘧啶的经修饰核苷酸组成。

7. 权利要求 1 和 2 的方法,其中所述化学处理为亚硫酸氢盐介导的胞嘧啶脱氨作用。

8. 权利要求 1 和 2 的方法,其表征衍生自正常、疾病或其他表型来源的序列数据,其方式为通过作图或与一个或多个参比 DNA 序列比对以辨别可与任何表型相关的遗传或外遗基因组差异。

9. 权利要求 1 和 2 的方法,其中经修饰的一个或多个接头由一种或多种核苷酸组成,所述核苷酸选自与能够生成仪器或视觉检查可读取检测信号的部分缀合的任何核苷酸。

10. 权利要求 1 和 2 的方法,其中所述经修饰的一个或多个接头能够在固相载体上指导 DNA 扩增。

11. 权利要求 1 和 2 的方法,其中所述经修饰的一个或多个接头能够在固相载体上指导

等温 DNA 扩增。

12. 权利要求 1 和 2 的方法,其中所述经修饰接头与另一经修饰接头在功能上和空间上相连接。

13. 权利要求 1 和 2 的方法,其中所述经修饰的一个或多个接头由一种或多种核苷酸组成,所述核苷酸选自与亲和纯化标签缀合的任何核苷酸。

14. 权利要求 1 和 2 的方法,其中所述经修饰的一个或多个接头由一种或多种核苷酸组成,所述核苷酸选自与生物素部分缀合的任何核苷酸。

15. 权利要求 1 和 2 的方法,其中所述一个或多个 DNA 接头包含一个或多个可被能够与 DNA 形成三螺旋结构的寡核苷酸靶向的序列。

16. 权利要求 15 的方法,其中所述形成三螺旋的寡核苷酸与亲和纯化标签缀合。

17. 权利要求 1 和 2 的方法,其中所述靶 DNA 选自基因组 DNA、线粒体 DNA、叶绿体 DNA、质粒 DNA、cDNA、病毒 DNA、微生物 DNA、化学合成的 DNA、核酸扩增的 DNA 产物以及从 RNA 转录的 DNA。

18. 权利要求 1 和 2 的方法,其中通过应用机械力或通过单独或联合使用一种或多种核酶完全或部分消化将所述靶 DNA 随机片段化。

19. 权利要求 18 的方法,其中所述核酶为选自 BshI 236 I、BstU I、CviJ I、FspB I、Hae III、Hha I、Hpa II、Mse I、Msp I、Sau3 AI、TaqI、Tsp509 I、它们的同裂酶和新裂酶的限制性核酸内切酶。

20. 生产半甲基化 DNA 对照模板以监测亚硫酸氢盐反应效率的方法,其包括:

使两条互补的 DNA 链 A 和 B 复性,其中 B 链的胞嘧啶在 5- 碳位甲基化,其中的 A 链胞嘧啶未甲基化;并且

A 链在包含引物 A 和 B 的扩增反应中生成,藉此引物 A 被生物素基团标记,引物 B 胞嘧啶被 5- 甲基胞嘧啶取代,在包含 dATP、dTTP、dGTP 和 5- 甲基 -dCTP 的三磷酸脱氧核糖核苷酸混合物存在下进行 DNA 扩增;并且

B 链在包含引物 A 和 B 的使用与 A 链相同的 DNA 模板的扩增反应中生成,藉此引物 B 被生物素基团标记并且在包含 dATP、dTTP、dGTP 和 dCTP 的三磷酸脱氧核糖核苷酸混合物存在下进行 DNA 扩增;并且

将等摩尔量的两种扩增产物混合、变性、使其再次复性,接着进行亲和素亲和色谱法移除不需要的产物。

21. 生产甲基化 DNA 对照模板的方法,其中两条链的胞嘧啶在 5- 碳位被甲基化以监测亚硫酸氢盐反应效率,该方法包括:

使用组分胞嘧啶被 5- 甲基胞嘧啶取代的引物将对照 DNA 模板进行 DNA 扩增并且 DNA 扩增在包含 dATP、dTTP、dGTP 和 5- 甲基 -dCTP 的三磷酸脱氧核糖核苷酸混合物存在下进行。

用于高通量亚硫酸氢盐 DNA- 测序的方法和组合物及其用途

与相关申请的交叉引用

[0001] 本申请要求基于 2007 年 8 月 15 日申请的美国临时专利申请序列号 60/935,472 和 2007 年 9 月 5 日申请的 60/935,867 的优先权。上述临时申请的全部内容通过引用结合到本文中。

本公开的技术领域和摘要

[0002] 本发明涉及生产适用于化学修饰和高通量 DNA 测序的 DNA 模板的新颖方法和组合物。本发明的一种方法涉及 DNA 接头设计,其中组分胞嘧啶被 5- 甲基胞嘧啶取代,使所得接头对亚硫酸氢盐介导的脱氨作用具有抗性。当所述接头与双链 DNA 模板连接时,后续 DNA 变性和亚硫酸氢盐处理将模板 DNA 胞嘧啶差异脱氨形成尿嘧啶,而所连接接头的 5- 甲基胞嘧啶对这一化学转化具有抗性使得接头序列保持不变。因此可使用与未改变接头序列杂交的单一引物组扩增经亚硫酸氢盐处理的 DNA 两条链。本发明也涉及生产具确定甲基化组成的对照模板以优化适用于亚硫酸氢盐反应条件的方法。在一个优选的实施方案中,本发明可用于生产适于使用传统的 Solexa™、SOLiD™ 或 454™- 型 DNA 测序平台进行全基因组亚硫酸氢盐 -DNA 测序以研究 DNA 甲基化的模板。

发明背景

[0003] 外遗调控 (epigenetic regulation) 的一个主要机制涉及 DNA 甲基化,藉此 S- 腺苷 - 甲硫氨酸的甲基被酶催化转移至胞嘧啶的 5- 碳位上得到 5- 甲基胞嘧啶 (综述: Caiafa 和 Zampiere, 2005 ;Novik 等, 2002 ;Bird, 2002 ;Costello 和 Plass, 2001 ;Laird 和 Jaenisch, 1996)。在人类中,绝大多数的胞嘧啶甲基化发生在 CpG 岛的 CpG 二核苷酸、G+C 等容线和 CpG 热点上,但是位于 CpNG、CC(a/t)GG、CpA 和 CpT 序列中的胞嘧啶也可被低频率甲基化 (Lorincz 和 Groudine, 2001 ;Woodcock 等, 1997 ;Clark 等, 1995)。调控区的 CpG 二核苷酸胞嘧啶甲基化促成基因沉默例如在 X- 染色体活化中并经常在癌症的肿瘤抑制基因沉默中发挥重要作用。已报导了致癌作用各种阶段以及许多其他疾病中不同基因组区域的低甲基化和超甲基化 (综述: Jones 和 Baylin, 2007 ;Brena 和 Costello, 2007 ;Esteller, 2007 ;Rodenhiser 和 Mann, 2006 ;Laird 和 Jaenisch, 1996)。因此,需要表征被定义为基因组甲基化状态的“甲基化组 (methylome)”的方法,以阐明可导致发现药物、药物作用靶或可用的疾病生物标记的调控网络。

[0004] 已研发了各种方法用于以不断增大的分辨率分析胞嘧啶甲基化以阐明调控网络并用于鉴定疾病的生物标记。用于测定胞嘧啶甲基化的最早方法是基于色谱法 (Hotchkiss, 1948) 但是其分辨率低,因为这一技术和其他相关技术只能观察到 DNA 中的整体甲基化变化。可分析甲基胞嘧啶分布更准确变化的分辨率提高的后续方法利用甲基 - 敏感性限制性核酸内切酶或亲和色谱法 (Zhang 等, 2006 ;Cross 等, 1994)。胞嘧啶化学修饰偶联 DNA 测序仍然是外遗研究中以单个核苷酸分辨率检测 5- 甲基胞嘧啶的精选方法,即

所谓“黄金标准”。“亚硫酸氢盐-DNA 测序”化学的发展使得能够对胞嘧啶甲基化进行直接阳性检测。亚硫酸氢盐-DNA 测序法来自 20 世纪 70 年代描述的化学,通过该方法亚硫酸氢钠催化胞嘧啶有效脱氨得到尿嘧啶 (Shapiro 等,1973 ;Shapiro 等,1970 ;Hayatsu 等,1970),其功能上等价于在测序和 DNA 扩增反应中的胸腺嘧啶。根据反应条件,5-甲基胞嘧啶脱氨形成胸腺嘧啶的速率可比胞嘧啶转化为尿嘧啶的速率慢大约两个数量级 (Hayatsu 和 Shiragami,1979 ;Wang 等,1980)。Frommer 等 (1992) 利用甲基胞嘧啶和胞嘧啶之间的选择性化学差异以及聚合酶链式反应 (PCR),来提供位于单个 DNA 链上 5-甲基胞嘧啶的阳性展示。从这一最初的报告和其他报告开始,亚硫酸氢盐-DNA 测序法迅速成为并且仍然是用于在单核苷酸水平分辨率上研究靶基因座中胞嘧啶甲基化的精选方法 (Ehrich 等,2007 ;Grunau 等,2001 ;Eads 等,2000 ;Paulin 等,1998 ;Clark 等,1994 ;Raizis 等,1995 ;Feil 等,1994 ;Frommer 等,1992)。

[0005] 近些年来尝试采用大规模全基因组 DNA 甲基化检测的方法。这些尝试的第一个为“限制性标记基因组扫描”,其中通过连续分级分离用甲基-敏感性限制酶消化的末端标记 DNA 检测差异 DNA 甲基化 (Hatada 等,1991)。这一方法受到可获得性和基因组中限制性酶切位点分布的限制并且分辨率低。Olek 等,(1998) (美国专利第 6,214,556 号) 描述了另一种方法,藉此经亚硫酸氢盐处理的 DNA 与通过使用 PCR 的选择性完整基因组 DNA 扩增偶联。通过引物延伸测定检测被扩增的产物以得到可用于评价细胞甲基化状态的复杂 DNA 甲基化指纹图谱。所进行的引物延伸测定的数量指明这一方法的分辨率和基因组覆盖范围。另一策略是基于使用抗-甲基胞嘧啶抗体或甲基-CpG 结合蛋白的甲基化 DNA 片段亲和层析 (Zhang 等,2006 ;Cross 等,1994)。甲基化拟南芥 DNA 的免疫沉淀和亲和色谱法与被捕获标记产物和基因组寡核苷酸嵌合芯片 (tiling array) 杂交相偶联已产生了第一个全基因组甲基化图谱 (Zhang 等,2006)。所得甲基化图谱具有 35-碱基分辨率,对应于嵌合芯片上寡核苷酸的长度。使用低分辨率阵列在人类癌细胞系上进行的相似研究已揭示大量的差异甲基化基因 (Keshet 等,2006 ;Weber 等,2005)。尽管可用于全基因组扫描,但是这一方法受到阵列分辨率和在其可被亲和纯化捕获前 DNA 片段上甲基-CpG 的最低阈值浓度的限制。因此,需要相对大量的原料,因此排除了其在许多临床应用中的用途。在本领域中明显需要只需更少量的原料并且在单个核苷酸水平上具有更高分辨率的更敏感的检测方法。

[0006] 直到最近,技术难题一直阻碍亚硫酸氢盐-DNA 测序法在全基因组范围作图绘制甲基化变化的应用。在原理性试验的小规模验证中,Meissner 等 (2005) 证实了使用亚硫酸氢盐-DNA 测序法以单核苷酸分辨率作图绘制基因组 DNA 文库插入物 5-甲基胞嘧啶的实际可行性。用亚硫酸氢盐处理大小经过选择的随机片段化基因组 DNA 片段 (装有接头),经 PCR 扩增并克隆至用于测序的载体中。所得序列数据显示胞嘧啶至尿嘧啶的转化率大于 99.9%,表明基因组文库插入物的随机鸟枪法亚硫酸氢盐-DNA 测序可应用于全基因组范围。然而,这一方法的用途是有限的,主要受到高成本和传统桑格双脱氧测序法低通量的限制。因此,本领域仍然存在对较低成本和更高通量的经改进的亚硫酸氢盐-DNA 测序法的需要。

[0007] 下一代大规模并行 DNA 测序技术提供了高出若干个数量级的通量并且成本降低,但是这些平台仍然不适用于亚硫酸氢盐-DNA 测序,做不到低成本全基因组地研究 DNA 甲基化。目前有三种市售的用于高通量 DNA 测序的系统:基因组测序仪 (The Genome

SequencerFLX™) 系统 (通常称作 454™- 测序仪) (Roche Diagnostics, Indianapolis, IN); Solexa™ (Illumina, San Diego, CA); 和 SOLiD™ 系统 (AppliedBioSystems, Foster City, CA)。

[0008] 454- 技术基于传统的焦磷酸测序化学, 其在经克隆扩增的 DNA 模板上进行, 这些模板在单个装载至高密度光学板刻蚀孔的微珠上 (Margulies 等, 2005)。各次碱基延伸生成的信号被精密的光学纤维捕获。

[0009] Solexa 测序模板被固定至专有流动池表面, 在那里它们被原位克隆扩增以形成密度高达一千万集簇每平方厘米的分离的序列模板集簇。使用引物介导的 DNA 合成以分步方式在四种专有的经修饰核苷酸 (具有可逆 3' 双 - 脱氧核苷酸部分和可裂解的荧光生色团 (chromofluor)) 存在下进行 Solexa 测序。在各延伸循环之前化学移除 3' 双 - 脱氧核苷酸部分和荧光生色团用于连续的碱基判定 (basecalling)。来自各模板集簇的分步式核苷酸添加循环由激光激发检测, 之后进行成像从而完成碱基判定。

[0010] 用于大规模并行 DNA 测序的 Applied Biosystems' SOLiD 方法是基于连续的 DNA 连接循环, 这是由哈佛大学 George Church 创新的策略 (Shendure 等, 2005)。通过这一方法, 在微珠上克隆扩增固定化 DNA 模板 (乳滴 PCR), 这些微珠以高密度置于玻璃流动池的表面。通过短的确定的标记探针连接至与固定化模板杂交的一系列引物的连续循环完成序列测定。

[0011] 这些新仪器每次运行的通量可超过数十亿碱基判定, 是目前基于 96- 道毛细管电泳测序仪器的约一万五千或更多倍。因此, 仍然存在对使亚硫酸氢盐 -DNA 测序化学适应 454-、Solexa 或 SOLiD 测序平台以便低成本全基因组地研究 DNA 甲基化的方法和组合物的需要。在他们的初步研究中, Meissner 等 (2005) 提出新一代 454-DNA 测序仪可提供经济的解决方案以使亚硫酸氢盐 -DNA 测序的全基因组应用得以进行, 但是他们没有讨论关键问题也没有公开任何付诸实施的方法。更重要的是, 研究者没有认识到将他们的方法应用至 Solexa 或 SOLiD 平台 (典型的序列读取长度仅为 35-50 个碱基) 的巨大困难。本发明提供这些和其他实质性益处。

发明详述

[0012] 本发明提供在下一代 DNA 测序仪中使用的用于亚硫酸氢盐 -DNA 测序的新颖改进方法和可用组合物, 以使大规模高通量全基因组研究胞嘧啶甲基化类型的改变得以实现并用于其它优选的应用。

[0013] Meissner 等 (2005) 的先导研究描述了亚硫酸氢盐 -DNA 测序方法, 通过该方法首先将短 DNA 接头与多个大小经选择的随机片段化基因组 DNA 片段的各个末端相连接。将与接头连接的 DNA 变性, 得到对亚硫酸氢盐处理敏感的单链形式, 其中原有 (resident) 胞嘧啶转化成尿嘧啶但是 5- 甲基胞嘧啶不发生改变。使用接头区的引物将已转化的 DNA 扩增重新生成 DNA 链并产生足够质量的经亚硫酸氢盐转化的 DNA 产物用于有效克隆至用于通过传统毛细管电泳进行测序分析的载体中。该研究显示该方法提供待测基因组 DNA 的无偏倚表现并且具有大规模应用的可行性。然而, Meissner 等的亚硫酸氢盐处理靶 DNA 的重要后果是已连接接头中的所有胞嘧啶也被转化成尿嘧啶。因此, 为了进行 DNA 扩增, 将 PCR 引物设计成不与接头序列杂交而是设计成与经亚硫酸氢盐转化的接头序列杂交, 这一策略是

被称作“甲基化-特异 PCR”方法的基础 (Cottrell, 2004 ;Li 和 Dahlya (2002) ;Herman 和 Baylin (1997) (美国专利第 6, 017, 704 号) ;Herman 等, 1996)。适用于扩增亚硫酸氢盐处理样品的本领域已知的其他适当 PCR 引物设计包括使用可从亚硫酸氢盐修饰位点扩增 DNA 的简并引物或使用靶向 DNA 无胞嘧啶区域的 DNA 的非常短的引物 (Olek 等, 1998 美国专利第 6, 214, 556 号)。

[0014] 如本领域所述的甲基化-特异 PCR 对引物设计的限制与现有 ABI SOLiD 或 Illumina Solexa 高通量测序平台的使用不相容。这些平台必须使用紧邻样品 DNA 插入物之后的经优化并经验证的专有接头序列。这些专有接头的功能为介导 DNA 测序模板的克隆固相扩增和测序引物的结合。Solexa 和 SOLiD 测序仪的阅读长度仅为 35 个碱基 (在 2008 年末延伸至 50 个碱基或以上)。位于专有接头和 DNA 插入物之间的外来序列例如甲基化-特异 PCR 所要求的那些将样品 DNA 本已较短的阅读长度降低至不可接受的水平。结果, 如所述目前的 Solexa 和 SOLiD 平台不能测序甲基化-特异 PCR 法生成的产物 (Meissner 等, 2006 ;Cottrell, 2004 ;Li 和 Dahlya, 2002 ;Herman 和 Baylin, 1997, 美国专利第 6, 017, 704 号 ;Herman 等, 1996)。尽管可能形式上生成其中经亚硫酸氢盐转化的接头序列可介导固相载体上克隆扩增的接头设计和结合至 Solexa 和 SOLiD 平台的测序引物, 但是技术和经济上的难题是难以克服的。在接头上将胞嘧啶亚硫酸氢盐转化成尿嘧啶将有效减少遗传密码至仅三个碱基, 因此严重限制了可有效并特异性用于平台所需固相扩增和高通量 DNA 测序的特异引物的设计。此外, 亚硫酸氢盐转化使得接头的两条链不互补, 因此需要生产和验证另一组用于其它样品 DNA 链的固相扩增引物和测序引物。已消耗了大量的公司经费、时间和资源研发并验证 SOLiD 和 Solexa 测序平台的现有接头和引物设计 ;对市场上现有产品的主要设计改变将产生不可接受的财务负担。454- 测序仪的读取长度为数百个碱基并且在样品 DNA 模板中加入甲基化-特异 PCR 引物将减少读取长度。然而, 去除 454- 模板中的外来序列将增加该平台的效率。

[0015] 本发明提供将现有基于 SOLiD、Solexa 或 454- 的 DNA 测序平台适用于测序亚硫酸氢盐处理的 DNA 样品以研究 DNA 甲基化的新颖、简便、有效并且低成本的方法。本发明的一方面创造了新颖的接头组合物, 其中组分胞嘧啶被 5- 甲基胞嘧啶取代使得所述接头对所结合模板 DNA 的亚硫酸氢盐处理过程中的脱氨作用具有抗性。当本发明接头与模板 DNA 连接时, DNA 变性并且亚硫酸氢盐处理将模板 DNA 胞嘧啶转化成尿嘧啶, 接头序列保持不变。因此可使用与原来未改变的接头序列互补的单个引物组扩增经亚硫酸氢盐处理的 DNA 两条链。相反, 传统接头的胞嘧啶经亚硫酸氢盐处理转化成尿嘧啶使得必须使用与经亚硫酸氢盐转化的接头序列杂交的 PCR 引物扩增经亚硫酸氢盐处理的模板。亚硫酸氢盐处理也使得两个 DNA 模板链不互补。亚硫酸氢盐处理也使传统接头的两条链不互补, 致使需要分离的引物组来扩增各 DNA 链。本发明的接头组合物不存在这一问题, 两个接头链保持互补并且单一引物组足以扩增经亚硫酸氢盐处理的两条链用于制备在 Solexa、SOLiD 或 454- 测序平台上测序的模板。这些已建立的平台采用本发明预期引起很少或没有材料消耗, 因为这些平台的专有接头的一级序列没有改变, 因此, 所有的下游操作例如固相 DNA 扩增和测序引物结合均不受影响。在优选的实施方案中, 本发明的一方面可用于生产试剂盒或试剂盒组件, 用于制备在 SOLiD、Solexa、454- 或其他测序平台上高通量亚硫酸氢盐-DNA 测序以进行甲基化研究的 DNA 模板。试剂盒组件与传统测序供应商提供的那些基本相同, 除了简单

并低成本地在接头中用 5- 甲基胞嘧啶替换胞嘧啶。

[0016] 通常, 一个接头包括两个短的互补 DNA 寡核苷酸链, 其包含天然的或使用本领域已知的合成路径通过化学或酶辅助合成生产的经修饰的寡核苷酸 (综述: Verma 和 Eckstein, 1998 ; Goodchild, 1990)。包含例如甲基与胞嘧啶的 5- 碳位结合生成 5- 甲基胞嘧啶的经修饰碱基的寡核苷酸可从许多供应商获得, 包括 Operon (Cologne, Germany) ; Sigma-Proligo (Paris, France) ; 和 Genosys (St. Louis, MO)。除了化学合成, 可能使用甲基转移酶将接头 DNA 酶促甲基化, 只要胞嘧啶在酶的识别位点内。也可能通过使用填充反应 (fill-in reaction) 中的 DNA 聚合酶或通过 PCR 将 5- 甲基 -dCTP 掺入接头 DNA 中。本领域技术人员知道优化的接头设计和合成方法。可操作地, 将接头的两条 DNA 链复性形成双链分子。一般而言, 接头序列可在 10 至 100 个碱基对 (bp) 或更长范围内变化, 通常为 15 至 30bp。接头的序列组成是可变的, 但是通常不含可能干扰潜在引物结合和其他功能的反向重复序列和类似序列。在一些应用中, 接头可空间上结合在一起使得已结合的接头可连接至多于一个的靶 DNA 末端。这一应用的典型为当需要有不同的接头与模板 DNA 每个末端连接时, 如克隆扩增和接着在下一代 Solexa、SOLiD 或 454-DNA 测序仪上测序的情况。所结合接头和靶 DNA 之间的分子内连接之后为分子间连接得到环形分子, 藉此靶 DNA 被两个不同的接头侧接。已描述了一定片段长度范围的 DNA 区段得到环形分子的分子内连接的条件 (Collins 和 Weissman, 1984 ; Dugaiczky 等, 1975 ; Wang 和 Davidson, 1966)。接头可经工程化具有不同的末端结构以辅助与 DNA 的连接。经常使用平端, 同样可使用用于和具有配偶体互补末端的 DNA 片段连接的特异粘性互补末端。接头与 DNA 连接以及使用靶向所连接接头的引物进行全基因组 DNA 扩增的操作为本领域已知 (Hughes 等, 2005 ; Klein 等, 1999 ; Lucito 等, 1989 ; Ludecke 等, 1989 ; Kinzler 和 Vogelstein, 1989)。除了之前所述用 5- 甲基胞嘧啶取代胞嘧啶之外, 接头可包括其他经修饰或缀合的核苷酸。可使接头分子对亚硫酸氢盐处理或对可将基因组胞嘧啶与经修饰的接头胞嘧啶区分开来的其他差异化学处理具有抗性的胞嘧啶其他化学修饰均视作属于本发明范围和原理。同样视作属于本发明范围和原理的为其他接头碱基修饰, 其中有可将经修饰的接头 DNA 与基因组 DNA 区分开来用于研究其他细胞外遗 DNA 修饰的化学反应。同样视作属于本发明领域和原理的为向接头中掺入表位或纯化标签, 例如含生物素的部分或可被形成三螺旋的寡核苷酸靶向的 DNA 序列 (综述: Vasquez 和 Glazer, 2002 ; Sun 等, 1996) 等, 使得可在化学处理各步骤之前、之后或过程中便捷地亲和纯化与接头连接的 DNA。

[0017] 根据本发明用于分析的 DNA 可衍生自任何细胞、组织或器官。在一些实施方案中, DNA 衍生自临床治疗不同时间点或阶段的肿瘤或具有疾病表型的其他细胞以评价疾病状态中甲基化类型的整体改变。这样, 本发明可用于鉴定疾病或疾病易感性或疾病后果的基因组诊断或预后甲基化生物标记。Ordway 等 (2006)、Sova 等 (2006) 和 Shames 等提供了这些生物标记的说明性实例。其他应用包括阐明导致可鉴定用于治疗干预的药物或药物靶标的调控网络。

[0018] 用于全基因组甲基化研究的 DNA 可通过随机片段化生成以提供基因组的无偏倚分析。适当大小的 DNA 可在 100 至 500bp 或更长范围内变化, 通常优选 100 至 250bp。生成随机 DNA 片段的方法包括: (1) 牛胰脏脱氧核糖核酸酶 I (DNase I), 其可在锰离子存在下在 DNA 中形成随机双链裂解 (Melgar 和 Goldthwait, 1968) ; (2) 物理剪切 (Shriefer

等,1990);和(3)声裂法(Deininger,1983)。在一些实施方案中,可用优先靶向消化 CpG 岛序列的酶消化基因组 DNA, CpG 岛序列为与基因组中的基因相关的富含 GC 的区域(Kato 和 Sasaki,1998)。大部分的甲基化发生在 CpG 序列内,因此用例如 MspI (CCGG)、Hae III (GGCC)、Taq I (TCGA) 等酶消化基因组 DNA 将优先将亚硫酸氢盐 -DNA 测序靶向基因组的这些区域。在宽松条件下使用在 GC 二核苷酸位置裂解 DNA 的限制性核酸内切酶 CviJI (Fitzgerald 等,1992) 在部分消化条件下尤其适用于生产可用的 DNA 片段大小的连续区。

[0019] 计算机模拟分析表明给定的随机 50- 碱基读取对人类基因组参比集合明确分配(unambiguous assignment) 的概率为约 93%。对于被 Msp I (CCGG) 或 Hae III (GGCC) 位点侧接的 50-bp 片段和其他具有富含 G+C 识别位点的其他酶,对基因组集合的明确分配大于 99%,因为观察到基因组中绝大多数重复 DNA 元件具有较低的 GC 含量并且这些酶的位点在这些基因组区域中低表达(under represented)。计算机模型也显示在通过 Msp I、Hae III 和 Taq I 消化生成的片段中存在高度重叠。在 50-400bp 片段大小范围内,绝大多数 CpG 岛序列可被由这三种酶单独消化构建的基因组文库的重叠 50bp 读取覆盖。经亚硫酸氢盐处理的 DNA 通常对参比序列的明确分配率较低,这归因于胞嘧啶转化成尿嘧啶(胸腺嘧啶),其可有效减少对三碱基遗传密码的原始查询(raw query)。使用 Solexa 和 SOLiD 测序仪的成对末端(pair-end) 读取能力延伸序列长度以及使用对立 DNA 链通过一致性比对和建立重叠群可解决这一问题。除了鉴定甲基化状态的改变,当序列数据与参比序列相比时本发明同时也可鉴定 SNPs 和其他遗传和体细胞改变。本领域具有用于甲基化数据集簇分析的信息工具(Wang 等,2007;Segal,2006;Siegmond,2004;Virmani 等,2002;Model 等,2001;Eads 等,2000)。

[0020] 尽管亚硫酸氢盐测序法的可用性和大范围使用,本发明倾向于处理误差和亚硫酸氢盐处理固有的竞争性和不利化学反应引起的问题。这些问题在全基因组应用中更明显。强效的亚硫酸氢盐处理方案(即延长温育时间、高温或高亚硫酸氢盐浓度)确保胞嘧啶完全转化成尿嘧啶,但是存在脱嘌呤作用引起的不可接受的 DNA 片段化以及最终将 5- 甲基胞嘧啶转化成胸腺嘧啶的风险(Hayatsu 和 Shiragami,1979;Wang 等,1980)。较低强度处理的风险在于过高估计甲基化水平,归因于胞嘧啶向尿嘧啶的不完全转化。因此,对于主要的实验条件例如温度、pH、反应时间、亚硫酸氢盐浓度、DNA 变性效率等仍需大量的工作以继续优化亚硫酸氢盐转化过程(Ehrich 等,2007;Hayatsu 等,2006;Grunau 等,2001;Eads 等,2000;Paulin 等,1998;Clark 等,1994;Raizis 等,1995;Feil 等,1994;Frommer 等,1992)。优化这一过程的主要限制步骤为缺乏便捷和综合性的对照模板以监测亚硫酸氢盐转化固有的复杂和竞争性反应。用于评价亚硫酸氢盐转化效率的现有方法利用高效液相色谱法(HPLC)、凝胶电泳和质谱法检测处理后 DNA 的质量(Ehrich 等,2007)。亚硫酸氢盐反应优化实验中胞嘧啶转化成尿嘧啶的速率通常由甲基化 -PCR 测定法测定,接着对衍生自一个或多个基因组待测基因座或待测对照模板的克隆产物(Frommer 等,1992) 测序,从而通过使用甲基转移酶甲基化两条 DNA 链的确定位点。使用甲基转移酶体外甲基化衍生的对照模板受累于潜在的不完全酶促作用,使得难以辨别特定位点胸腺嘧啶的存在是否由不完全体外甲基化引起或由过度的亚硫酸氢盐转化(其中甲基胞嘧啶被转化成胸腺嘧啶)引起(Hayatsu 和 Shiragami,1979;Wang 等,1980)。此外,仅可评价给定甲基转移酶识别位点内的胞嘧啶。因此,需要便捷、强效和综合性的测定法以监测亚硫酸氢盐 - 转化过程中的复杂

和竞争性反应,尤其当亚硫酸氢盐测序法在全基因组范围内进行时。

[0021] 本发明另一方面提供生产具精确限定的胞嘧啶甲基化组成的合成对照模板的方法以优化亚硫酸氢盐反应的条件。在本发明的一方面,该对照模板包含两个互补的复性 DNA 链 A 和 B,其中 B 链的胞嘧啶在 5- 碳位甲基化并且其中 A 链的胞嘧啶没有甲基化。通过将衍生自通用 DNA 模板的两个独立扩增反应的产物复性构建所得半-甲基化 DNA 分子。第一个反应包含扩增引物 A 和 B,藉此引物 A 被生物素部分标记,引物 B 胞嘧啶被 5- 甲基胞嘧啶取代并且在包含 dATP、dTTP、dGTP 和 5- 甲基 -dCTP 的三磷酸脱氧核糖核苷酸混合物(典型浓度为每核苷酸 10mM)存在下进行扩增。第二个扩增反应包含引物 A 和 B,藉此引物 B 被生物素部分标记并在包含 dATP、dTTP、dGTP 和 dCTP 的三磷酸脱氧核糖核苷酸混合物存在下进行扩增。将等摩尔量的两种扩增产物混合、变性、使其再次复性并接着进行亲和素亲和色谱法移除被生物素标记的 DNA 分子。因此没有被亲和色谱法捕获的种类包含甲基化胞嘧啶 A 链和未甲基化胞嘧啶 B 链的双链半甲基化分子。所得半甲基化对照模板(HM- 对照模板)可用于优化亚硫酸氢盐反应条件。因为已知两条 DNA 链每一条的 HM- 对照模板绝对精确的甲基化状态,所以与期望序列或亚硫酸氢盐处理后两个对照模板链结果的任何偏差是不完全或过度亚硫酸氢盐处理程度的定量测定。此外,对照模板可经工程化以包含一些特征,例如发夹结构、反向重复序列等,已知这些对亚硫酸氢盐处理、对所得实验条件的抗性更强从而影响它们的转化。在本发明的另一方面,也可通过将两个化学合成的寡核苷酸复性产生 HM- 对照模板,其中一条链包含胞嘧啶位点的 5- 甲基胞嘧啶取代并且互补链包含胞嘧啶。在本发明的另一方面,也可通过在包含 dATP、dTTP、dGTP 和 5- 甲基 -dCTP 的三磷酸脱氧核糖核苷酸混合物存在下进行 PCR 生成对照模板。所得对照模板在两条 DNA 链上具有 5- 甲基胞嘧啶完全取代胞嘧啶并可监测过度亚硫酸氢盐处理的可用对照模板。在优选的实施方案中,具有二级结构或同聚物束严重度不断增强区域的对照模板可用于监测在不同温育时间、温度、pH 和亚硫酸氢盐浓度的实验条件下亚硫酸氢盐处理的效率。在另一个优选的实施方案中,将对照模板加入基因组 DNA 以验证复杂 DNA 混合物存在下的试验条件。在另一个优选的实施方案中,可将微量的对照模板加入基因组 DNA 样品以提供在 Solexa、SOLiD 或 454- 平台上高通量亚硫酸氢盐 -DNA 测序的内部对照。在又另一优选的实施方案中,本发明的对照模板可用于提供基于 Solexa、SOLiD、454- 或其他测序平台进行高通量亚硫酸氢盐 -DNA 测序的试剂盒或试剂盒组件。

[0022] 应理解的是根据本文的公开内容在不偏离本发明范围和精神的条件下对本领域技术人员而言其他修改是显而易见的并可轻而易举地实现。

[0023] 参考文献

[0024] Bird A,2002. DNA methylation patterns and epigenetic memory(DNA 甲基化模式和外遗传记忆). *Genes Dev* 16 :6-21.

[0025] Brena RM 和 Costello JF,2007. Genome-epigenome interactions in cancer(癌症中的基因组 - 外遗传基因组相互作用). *Hum Mol Genetics* 16 :R96-R105.

[0026] Caiafa P 和 Zampieri M,2005. DNA methylation and chromatin structure :The puzzling CpG islands(DNA 甲基化和染色质结构 :令人迷惑的 CpG 岛). *J Cell Biochem* 94 :257-265.

[0027] Clark SJ 等,1995. CpNpG methylation in mammalian cells(哺乳动物细胞中的

CpNpG 甲基化). *Nat Genet* 10 :20-27.

[0028] Clark SJ等,1994.High sensitivity mapping of methylated cytosines(甲基化胞嘧啶的高灵敏度作图绘制). *Nuc Acids Res* 22 :2990-2997.

[0029] Collins FS 和 Weissman SM,1984.Direction cloning of DNA fragments at a large distance from an initial probe: A circularization method(远离起始探针的 DNA 片段的定向克隆:一种环化方法). *Proc Natl Acad Sci (USA)* 81 :6812-6816.

[0030] Cottrell SE,2004.Molecular diagnostic application of DNA methylation technology(DNA 甲基化技术的分子诊断应用). *Clin Biochem* 37 :595-604.

[0031] Costello JF 和 Plass C,2001.Methylation matters(甲基化物质). *J Med Genet* 38 :285-303.

[0032] Cross SH 等,1994.Purification of CpG islands using a methylated DNA binding column(采用甲基化 DNA 结合柱纯化 CpG 岛). *Nature Genet* 6 :236-244.

[0033] Deininger PL,1983.Random subcloning of sonicated DNA: Application to shotgun DNA sequence analysis(随机亚克隆经声处理的 DNA:应用于鸟枪法 DNA 序列分析). *Analyt Biochem* 129 :216-223.

[0034] Dugaiczyk A 等,1975.Ligation of Eco RI endonuclease-generated DNA fragments into linear and circular structures(Eco RI 核酸内切酶生成的 DNA 片段连接成线性和环状结构). *J Mol Biol* 96 :171-178.

[0035] Eads CA 等,2000.MethyLight: A high-throughput assay to measure DNA methylation(MethyLight:一种检测 DNA 甲基化的高通量测定法). *Nuc Acids Res* 28 :e32.

[0036] Ehrich M 等,2007.A new method for accurate assessment of DNA quality after bisulphite treatment(一种在亚硫酸氢盐处理后准确评价 DNA 质量的新方法). *Nuc Acids Res* 35 :e29.

[0037] Esteller M,2007.Epigenetic gene silencing in cancer: The DNA hypermethylome(癌症中的外遗传基因沉默:DNA 超甲基组). *Hum Mol Gen* 16 :R50-59.

[0038] Feil R 等,1994.Methylation analysis on individual chromosomes: Improved protocol for bisulphite genomic sequencing(单个染色体上的甲基化分析:亚硫酸氢盐基因组测序的改进方案). *Nuc Acids Res* 22 :695-696.

[0039] Fitzgerald MC 等,1992.Rapid shotgun cloning utilizing the two base recognition endonuclease CviJ I(应用二碱基识别核酸内切酶 CviJ I 的快速鸟枪法克隆). *Nuc Acid Res* 20 :3753-3762.

[0040] Frommer M 等,1992.A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands(可得到单个 DNA 链中 5-甲基胞嘧啶残基阳性展示的基因组测序方案). *Proc Natl Acad Sci (USA)* 89 :1827-1831.

[0041] Goodchild J,1990.Conjugates of oligonucleotides and modified oligonucleotides: A review of their synthesis and properties(寡核苷酸和经修饰寡核苷酸的缀合:它们的合成和性质的综述). *Bioconjugate Chem* 1 :165-187.

- [0042] Grunau C等,2001. Bisulphite genomic sequencing: Systematic investigation of critical experimental parameters(亚硫酸氢盐基因组测序法:关键实验参数的系统研究). *Nuc Acids Res* 29 :e65.
- [0043] Hatada I等,1991. A genomic scanning method for higher organisms using restriction sites as landmarks(使用限制性酶切位点作为标志用于更高级生物体的基因组扫描法). *Proc Natl Acad Sci(USA)* 88 :9523-9527.
- [0044] Hayatsu H 和 Shiragami M,1979. Reaction of bisulphite with the 5-Hydroxymethyl groups in pyrimidines and in phage DNAs(亚硫酸氢盐与嘧啶和噬菌体 DNAs 中 5-羟甲基基团的反应). *Biochem* 18 :632-637.
- [0045] Hayatsu H等,2006. Does urea promote the bisulphite-mediated deamination of cytosine in DNA? Investigation aiming at speeding-up the procedure for DNA methylation analysis(尿素促进亚硫酸氢盐-介导的 DNA 胞嘧啶的脱氨作用吗? 针对加速 DNA 甲基化分析操作的研究). *Nucleic Acids Symposium Series No 50* :69-70.
- [0046] Hayatsu H 等,1970. Reaction of sodium bisulphite with uracil, cytosine, and their derivatives(亚硫酸氢钠与尿嘧啶、胞嘧啶及它们的衍生物的反应). *Biochem* 9 :2858-2865.
- [0047] Herman JG 和 Baylin SB,1997. Method of detection of methylated nucleic acid using agents which modify unmethylated cytosine and distinguishing modified methylated and unmethylated nucleic acids(通过使用修饰未甲基化胞嘧啶的物质并区分经修饰的甲基化核酸和未甲基化核酸检测甲基化核酸的方法). U. S. Patent No. 6,017,704(2000年1月25日出版).
- [0048] Herman JG 等,1996. Methylation-specific PCR: A novel PCR assay for methylation status of CpG islands(甲基化-特异性 PCR: CpG 岛甲基化状态的新颖 PCR 测定法). *Proc Natl Acad Sci(USA)* 93 :9821-9826.
- [0049] Hotchkiss RD,1948. The quantitative separation of purine, pyrimidines, and nucleosides by paper chromatography(纸色谱法定量分离嘌呤、嘧啶和核苷). *J Biol Chem* 175 :315-332.
- [0050] Hughes S等,2005. The use of whole genome amplification in the study of human disease(全基因组扩增在研究人类疾病中的用途). *Prog Biophys Mol Biol* 88 :173-189.
- [0051] Jones P 和 Baylin SB,2007. The epigenomics of cancer(癌症的外显基因组学). *Cell* 128 :683-692.
- [0052] Kato R 和 Hiroyuki S,1998. Quick identification and localization of CpG islands in large genomic fragments by partial digestion of Hpa II and Hha I(通过 Hpa II 和 Hha I 部分消化在大基因组片段中快速鉴别和定位 CpG 岛). *DNA Res* 5 :287-295.
- [0053] Keshet I 等,2006. Evidence for an instructive mechanism of de novo methylation in cancer cells(癌细胞中从头甲基化指导性机制的证据). *Nat Genet* 38 :149-153.
- [0054] Kinzler KW 和 Vogelstein B,1989. Whole genome PCR: Application in the

identification of sequences bound by gene regulatory proteins(全基因组PCR:在鉴别基因调控蛋白结合序列中的应用). *Nuc Acids Res* 17 :3645-3653.

[0055] Klein CA 等,1999. Comparative genomic hybridization, loss of heterozygosity, and DNA sequence analysis of a single cell(单个细胞的竞争性基因组杂交、杂合性丢失和DNA序列分析). *Proc Natl Acad Sci(USA)* 96 :4494-4499.

[0056] Laird PW 和 Jaenisch R,1996. The role of DNA methylation in cancer genetics and epigenetics(DNA甲基化在癌症基因组学和外遗基因组学中的作用). *Ann Rev Genet* 30 :441-464.

[0057] Li L-C 和 Dahly R,2002. MethPrimer: Designing primers for methylation PCR(MethPrimer:设计甲基化PCRs的引物). *Bioinformatics* 18 :1427-1431.

[0058] Lorincz MC 和 Groudine M,2001. CpC(a/t)GG methylation: A new epigenetic mark in mammalian DNA? (CpC(a/t)GG甲基化:是哺乳动物DNA中的新外遗传标记吗?) *Proc Natl Acad Sci(USA)* 98 :11034-11036.

[0059] Ludecke H 等,1989. Cloning defined regions of the human genome by microdissection of banded chromosomes and enzymatic amplification(通过显微切割显带染色体和酶促扩增克隆人基因组的确定区域). *Nature* 338 :248-350.

[0060] Lucito R 等,1998. Genetic analysis using genomics representations(采用基因组代表进行遗传分析). *Proc Natl Acad Sci(USA)* 95 :4487-4492.

[0061] Margulies M 等,2005. Genome sequencing in microfabricated high-density picolitre reactors(在微结构高密度皮升反应器中的基因组测序). *Nature* 437 :376-380.

[0062] Meissner A 等,2005. Reduced representation bisulphite sequencing for comparative high-resolution DNA methylation analysis(用于竞争性高分辨率DNA甲基化分析的reduced representation亚硫酸氢盐测序法). *Nuc Acids Res* 33 :5868-5877.

[0063] Melgar E 和 Goldthwait DA,1968. Deoxyribonucleic acid nucleases: II. The effect of metals on the mechanism of action of deoxyribonuclease I(脱氧核糖核酸酶: II. 金属对脱氧核糖核酸酶I作用机制的作用). *J Biol Chem* 243 :4409-4416.

[0064] Model F 等,2001. Feature selection for DNA methylation based classification(基于DNA甲基化的分类的特征选择). *Bioinformatics* 17 :S157-S164.

[0065] Novik, KL 等,2002. Epigenomics: Genome-wide study of methylation phenomena(外遗基因组学:全基因组研究甲基化现象). *Curr Issues Mol Biol* 4 :111-128.

[0066] Paulin R 等,1998. Urea improves efficiency of bisulphite-mediated sequencing of 5' -methylcytosine in genomic DNA(尿素提高基因组DNA中5' -甲基胞嘧啶的亚硫酸氢盐-介导测序的效率). *Nuc Acids Res* 26 :5009-5010.

[0067] Olek 等,1998. Method for producing complex DNA

methylationfingerprints(产生复杂DNA甲基化指纹图谱的方法).U.S.Patent No6,214,556(issued April 10,2001).

[0068] Olek A 等,1996.A modified and improved method for bisulphite basedcytosine methylation analysis(一种基于亚硫酸氢盐的胞嘧啶甲基化分析的经修改和改进的方法).Nuc Acids Res 24 :5064-5066.

[0069] Ordway JM 等,2006.Comprehensive DNA methylation profiling inhuman cancer genome identifies novel epigenetic targets(人类癌症基因组的综合DNA甲基化谱式分析鉴别新的外遗传靶点).Carcinogenesis 27 :2409-2423.

[0070] Raizis AM 等,1995.A bisulphite method of 5-methylcytosine mappingthat minimizes template degradation(使模板降解最小化的作图绘制5-甲基胞嘧啶的亚硫酸氢盐法).Anal Biochem226 :161-166.

[0071] Rodenhiser D 和 Mann M,2006.Epigentics and human disease :Translating basic biology into clinical applications(外遗传学和人类疾病 :将基础生物学应用至临床应用).CMAJ 174 :341-348.

[0072] Schriefer LA 等,1990.Low pressure DNA shearing :A method forrandom DNA sequence analysis(低压DNA切变 :一种用于随机DNA测序分析的方法).Nuc Acids Res 18 :7455.

[0073] Segal MR,2006.Validation in genomics :CpG island methylationrevisited(基因组学中的验证 :再谈CpG岛甲基化).StatisticalApplications in Genetics and Molecular Biology 5 :article 29.

[0074] Siegmund KD 等,2004.A comparison of cluster analysis methods usingDNA methylation data(采用DNA甲基化数据比较聚类分析法).Bioinformatics 20 :1896-1904.

[0075] Shames DS 等,2006.A genome-wide screen for promoter methylationin lung cancer identifies novel methylation markers for multiplemalignancies(全基因组扫描肺癌中的启动子甲基化鉴别多种恶性肿瘤的新甲基化标记).PLOS Medicine 3 :e486.

[0076] Shapiro R 等,1970.Reactions of uracil and cytosine derivatives withsodium bisulphite :A specific deamination method(尿嘧啶和胞嘧啶与亚硫酸氢盐的反应 :一种特异性脱氨法).J Am Chem Soc92 :422-424.

[0077] Shapiro R等,1973.Nucleic acid reactivity and conformation II :Reaction of cytosine and uracil with sodium bisulphite(核酸反应性和构象II :胞嘧啶和尿嘧啶与亚硫酸氢钠的反应).J BiolChem 248 :4060-4064.

[0078] Shendure J等,2005.Accurate multiplex polony sequencing of anevolved bacterial genome(一个已进化细菌基因组的准确多路克隆阵列测序).Science 309 :1728-1732.

[0079] Sova P 等,2006.Discovery of novel methylation biomarkers incervical carcinoma by global demethylation and microarry analysis(通过全局脱甲基作用和微阵列分析发现新的甲基化生物标己).Cancer Epidemiol Biomarkers Rev 11 :291-297.

[0080] Verma S 和 Eckstein,1998.Modified oligonucleotides :Synthesis

andstrategy for users(经修饰的寡核苷酸:合成及使用策略).AnnRev Biochem 67:99-134.

[0081] Virmani AK 等,2002.Hierarchical clustering of lung cancer cell linesusing DNA methylation markers(使用DNA甲基化标记层次聚类肺癌细胞系).Cancer Epidemiol Biomarkers Rev 11:291-297.

[0082] Wang RY 等,1980.Comparison of bisulphite modification of5-methyldeoxycytidine and deoxycytidine residues(比较5-甲基脱氧胞苷和脱氧胞苷酸残基的亚硫酸氢盐修饰).Nuc Acids Res8:4777-4790.

[0083] Wang Z 等,2007.Heritable cluster and pathway discovery in breastcancer integrating epigenetic and phenotype data(通过整合外遗传和表型数据发现乳腺癌中的遗传聚类和途径).BMCBioinformatics8:38.

[0084] Wang JC 和 Davidson N,1966.On the probability of ring closure oflambda DNA(对 λ DNA 环合可能性的研究).J Mol Biol 19:469-482.

[0085] Weber M 等,2005.Chromosome-wide and promoter-specific analysesidentift sites of differential DNA methylation in normal andtransformed human cells(全染色体和启动子-特异性分析鉴别正常和转化人类细胞中的差异DNA甲基化).Nat Genet 37:853-862.

[0086] Woodcock DM 等,1997.Asymmetric methylation in the hypermethylatedCpG promoter region of the human L1 retrotransposon(人L1反转座子超甲基化CpG启动子区的不对称甲基化).J Biol Chem272:7810-7816.

[0087] ZhangX 等,2006.Genome-wide high-resolutioin mapping andfunctional analysis of DNA methylation in Arabidopsis(拟南芥中DNA甲基化的全基因组高分辨率作图绘制和功能分析).Cell1126:1189-1201.