

[19] 中华人民共和国国家知识产权局

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



# [12] 发明专利申请公布说明书

[21] 申请号 200580036116.X

[43] 公开日 2007 年 9 月 26 日

[11] 公开号 CN 101044250A

[22] 申请日 2005.9.23

[21] 申请号 200580036116.X

[30] 优先权

[32] 2004.10.1 [33] FR [31] 0452228

[86] 国际申请 PCT/FR2005/050775 2005.9.23

[87] 国际公布 WO2006/054003 法 2006.5.26

[85] 进入国家阶段日期 2007.4.20

[71] 申请人 拜奥默里克斯公司

地址 法国玛西 - 勒托勒

[72] 发明人 F·马莱 G·奥利奥尔  
J-P·皮雄

[74] 专利代理机构 中国国际贸易促进委员会专利商  
标事务所  
代理人 胡国群

权利要求书 2 页 说明书 27 页 序列表 16 页  
附图 4 页

## [54] 发明名称

产生转录物的方法

## [57] 摘要

本发明涉及由以下物质产生转录物的方法：·  
包含“引物”区和目的区的至少一种待扩增 RNA 序  
列，和·包含启动子区和能与待扩增 RNA 序列的所  
述“引物”区杂交的区域的扩增引物，所述方法在  
恒温下进行，并包括下列步骤：a) 将所述引物与待  
扩增 RNA 杂交，b) 通过逆转录酶的酶活性来延伸引  
物，以产生待扩增 RNA 的互补脱氧核糖核酸 (cD-  
NA) 序列，c) 通过具有核糖核酸酶 H 活性的酶来切  
割与所述 cDNA 杂交的待扩增 RNA，以获得与所述  
cDNA 杂交的待扩增 RNA 片段，d) 通过逆转录酶和  
链置换酶来延伸所述待扩增 RNA 片段的末端，以获  
得 RNA-DNA/DNA 杂交体，e) 通过具有 RNA 聚合  
酶活性的酶从在步骤 d) 中形成的 RNA-DNA/DNA  
杂交体获得 RNA 转录物。

1. 由以下物质产生转录物的方法:

- 包含“引物”区和目的区的至少一种待扩增 RNA 序列, 和
- 包含启动子区和能与待扩增 RNA 序列的所述“引物”区杂交的区域的扩增引物,

所述方法在恒温下进行, 并包括下列步骤:

- a) 将所述引物与待扩增 RNA 杂交,
- b) 通过逆转录酶的酶活性来延伸引物, 以产生待扩增 RNA 的互补脱氧核糖核酸 (cDNA) 序列,
- c) 通过具有核糖核酸酶 H 活性的酶来切割与所述 cDNA 杂交的待扩增 RNA, 以获得与所述 cDNA 杂交的待扩增 RNA 片段,
- d) 通过逆转录酶和链置换酶来延伸所述待扩增 RNA 片段的末端, 以获得 RNA-DNA/DNA 杂交体,
- e) 通过具有 RNA 聚合酶活性的酶从在步骤 d) 中形成的 RNA-DNA/DNA 杂交体获得 RNA 转录物。

2. 权利要求 1 的产生转录物的方法, 所述方法还包括以下步骤:

- f) 使用所述目的区的特异性杂交探针来检测产生的 RNA 转录物。

3. 权利要求 1 或 2 的产生转录物的方法, 根据所述方法, 施行步骤 a) 至 e) 的时间足以获得足够数量的转录物。

4. 权利要求 1 至 3 中任一项的产生转录物的方法, 其中具有 RNA 聚合酶活性的酶为噬菌体的 RNA 聚合酶, 优选 T7 噬菌体的 RNA 聚合酶。

5. 由以下物质产生转录物的方法:

- 至少第一待扩增 RNA 序列和第一扩增引物, 所述第一待扩增 RNA 序列包含“引物”区和目的区, 所述第一扩增引物包含启动子区和能与第一待扩增 RNA 序列的所述“引物”区杂交的区域,
- 至少第二待扩增 RNA 序列和第二扩增引物, 所述第二待扩增 RNA 序列与所述第一待扩增 RNA 序列不同, 并包含“引物”区和目的区, 所述第二扩增引物包含启动子区和能与第二待扩增 RNA 序列的所述“引物”

区杂交的区域，

所述方法在恒温下进行，并包括下列步骤：

a) 所述第一和第二引物与所述第一和第二待扩增 RNA 杂交，

b) 通过逆转录酶的酶活性来延伸所述第一和第二引物，以产生第一待扩增 RNA 的第一互补脱氧核糖核酸 (cDNA) 序列和第二待扩增 RNA 的第二互补脱氧核糖核酸 (cDNA) 序列，

c) 通过具有核糖核酸酶 H 活性的酶来切割分别与第一和第二 cDNA 杂交的第一和第二待扩增 RNA，以获得与第一 cDNA 杂交的第一待扩增 RNA 片段和与第二 cDNA 杂交的第二待扩增 RNA 片段，

d) 通过逆转录酶和链置换酶来延伸所述第一和第二待扩增 RNA 片段的末端，以获得 RNA-DNA/DNA 杂交体，

e) 通过具有 RNA 聚合酶活性的酶从在步骤 d) 中形成的 RNA-DNA/DNA 杂交体获得所述第一和第二待扩增 RNA 的转录物。

6. 权利要求 5 的产生转录物的方法，所述方法还包括以下步骤：

f) 使用第一杂交探针来检测第一待扩增 RNA 的转录物，和使用第二杂交探针来检测第二待扩增 RNA 的转录物。

7. 权利要求 5 或 6 的产生转录物的方法，根据所述方法，施行步骤 a) 至 e) 的时间足以获得足够数量的转录物。

8. 权利要求 5 至 7 中任一项的产生转录物的方法，其中具有 RNA 聚合酶活性的酶为噬菌体的 RNA 聚合酶，优选 T7 噬菌体的 RNA 聚合酶。

## 产生转录物的方法

本发明涉及扩增靶核酸序列的方法。

在与核酸和遗传物质相关的技术中，经常需要确定基因、基因部分或核苷酸序列在活生物体、该生物的细胞提取物或任何其他生物样品中是否存在。搜索特定的核苷酸序列具有重大意义，特别是对于检测病原生物、确定等位基因的存在、检测宿主基因组中损伤的存在和检测特定 mRNA 的存在或细胞宿主中修饰的存在。因此，可以通过分析甚至定量某些基因的表达来诊断许多遗传疾病。对这些基因的表达进行定量也是基础性的。

文献中描述了多种类型的用于检测核酸的方法。这些方法，特别是需要检测多核苷酸的方法，基于 DNA-DNA、DNA-RNA 和 RNA-RNA 双链体中核酸互补链的配对特性，所述配对通过双链 DNA 中腺嘌呤与胸腺嘧啶碱基 (A-T) 和鸟嘌呤与胞嘧啶碱基 (G-C) 之间氢键的建立，或者通过 DNA-RNA 或 RNA-RNA 双链体中腺嘌呤与尿嘧啶碱基 (A-U) 之间氢键的建立。核酸链的配对通常称为“核酸杂交”或简称为“杂交”。

一般地，在鉴定了需要分析的生物或疾病的特异性序列之后，应当从样品中提取核酸，并确定此序列（也称为“目的序列”）是否存在。对此目的已经开发了多种检测方法。

检测核酸样品中目的序列的存在的最直接方法是获得“探针”，其序列与靶核酸的一部分充分互补，从而与后者杂交。可以使如此合成的探针与含有核酸的样品接触，如果存在目的序列，则探针将会杂交并形成反应产物。在不存在目的序列并且避免所有非特异性杂交现象时，将没有反应产物形成。如果合成的探针与可检测标记物偶联，则可通过测定存在的标记物的量来检测反应产物。Southern 印迹法 (Southern E.M., J. Mol. Biol., 98, 503 (1975)) 或 Northern 印迹法或者斑点印迹或夹心杂交技术 (DUNN A. R. 和 HASSEL J. K., Cell,

12, 23 (1977) ) 组成了可使用方法的实例。

然而, 此方法的主要困难在于, 它不能直接应用于样品中存在的目的序列拷贝数低的情况。在这些条件下难于区分出强于反应背景噪声的显著信号(即区分开探针与其目的序列的特异性结合和探针与不同于目的序列的序列之间的非特异性结合)。对这一问题的一种解决方案为通过补充反应来增强检测信号。因此, 已经描述了多种方法以提高这些杂交技术的检测能力。这些称为“扩增”的方法可以分为三类: 靶、探针或信号的扩增。一方面, Lewis (1992, Genetic Engineering News 12:1-9) 的文章, 另一方面, Abramson 和 Myers (1993, Curer. Opus. Biotechnol. 4:4147) 的文章, 组成了这些方法的非常好的总括性综述。

靶扩增是特异性倍增样品中存在的核酸片段。它使得能够显著地提高待检测的靶核酸序列的拷贝数。

文献中描述的靶扩增技术主要基于重复进行体外 DNA 合成循环或者重复进行体外 RNA 合成循环, 前者通过 DNA 聚合酶来延伸与待扩增靶序列杂交的核苷酸引物来进行(“聚合酶链式反应”, 也称为 PCR: 参阅专利 US 4 683 195、US 4 683 202 和 US 4 800 159; EP 201 184; 称为“修复链反应”或 RCR 的方法: 参阅专利申请 WO 90/01069; 称为“链置换扩增”(SDA)的方法: 参阅专利 EP 0 497 272; “外切核酸酶介导的链置换扩增”方法: 参阅专利 EP 500 224), 后者通过使用 RNA 聚合酶的转录反应来进行。

已经描述了基于扩增转录物的这些靶扩增方法中的几种。专利申请 WO 88/10315 中描述的“TAS”方法由三步骤循环的重复组成。第一个步骤使得能够在逆转录酶和脱氧核苷酸引物存在下由 RNA 合成 cDNA, 所述脱氧核苷酸引物含有对于噬菌体 RNA 聚合酶启动子特异的序列。在使 RNA/cDNA 异源双链体热变性之后, 在反义寡核苷酸引物存在下通过逆转录酶来复制单链 cDNA 链。在第二个步骤中如此获得的 DNA 同源双链体含有噬菌体的 DNA 依赖性 RNA 聚合酶可以结合的双链启动子。第三个步骤是转录, 使每个模板产生 30 至 1000 个 RNA 分子,

所述 RNA 分子继而可以用作合成 cDNA 的模板,从而使扩增循环继续 (DAVIS 等认, 1990. J. Infect. Dis. 162:13-20)。

存在多种由 TAS 衍生的方法,包括专利申请 WO 90/06995 和专利 EP 0 373 960 中描述的“自动维持序列复制”(或 3SR),专利申请 WO 91/02818 和欧洲申请 EP 329 822 中描述的“基于核酸序列的扩增”(或 NASBA)方法,和专利 US 5 194 370 中描述的“单引物序列复制”(或 SPSR)方法。

这些方法都包括三种酶活性的组合:RNA-和 DNA-依赖性 DNA 聚合酶(逆转录酶)、核糖核酸酶 H(RNaseH, 大肠杆菌(*Escherichia coli*)的酶,和/或与逆转录酶相关的酶活性)和 DNA-依赖性 RNA 聚合酶(T7 噬菌体的 RNA 聚合酶)。这些方法基于相同的原理,并按照逆转录反应和转录反应的连续过程在固定的温度(37℃至 45℃)下进行,以通过 cDNA 来复制 RNA 靶。对于 TAS 的情况,通过用于逆转录步骤的引物将 RNA 聚合酶(T7 噬菌体)的结合位点引入 cDNA 中。然而, RNA/cDNA 异源双链体的变性通过 RNase H 活性对该异源双链体中 RNA 的特异性水解而等温地来进行。接着使用第二种寡核苷酸引物通过逆转录酶来复制游离的 cDNA。通过 T7 RNA 聚合酶将 DNA/DNA 同源双链体转录成 RNA,该 RNA 又可重新作为模板用于下一循环。

专利 US 5 194 370 中描述的称为“连接激活转录”(或 LAT)的另一方法使用与 3SR、SPSR 和 NASBA 方法相同的酶活性,并以相同的循环进行运作。然而,它的区别在于安装启动子序列的方式,在这种情况下,通过在 DNA 连接酶存在下连接含有启动子的茎环结构而在 cDNA 末端引入所述启动子序列。

还存在两种其他基于扩增转录物的靶扩增方法,描述于专利申请 EP 369 775。一种方法与 LAT 的差异仅在于启动子由两种不同的寡核苷酸组成。另一种方法使用两种酶活性:连接酶和 RNA 聚合酶,所述 RNA 聚合酶能识别双链 DNA 启动子,能沿着 RNA 模板进行延伸。称为移动性启动子的启动子由两种寡核苷酸组成。一种带有有义启动子序列和使得能够在 RNA 靶的 3'末端上进行杂交的探针序列,另一种仅带

有反义启动子序列。反义启动子寡核苷酸的 5'末端通过杂交与靶的 3'末端并列在一起，然后与此末端连接。通过合适的 RNA 聚合酶进行的转录引起合成了多个转录物。这些转录物与第二个移动性启动子杂交并连接。RNA 模板的转录引起合成了互补转录物。可以重复该过程，导致起始靶序列的指数式扩增。

还可以提及专利申请 WO 99/4385 中描述的方法，其是核酸的非特异性扩增方法。这种扩增方法也基于转录，但仅需要一种引物。在 3'端封闭的所述引物使得能够加入位于有义 RNA 3'端的 T7 启动子序列。接着，此序列使得能够转录反义 cRNA。由于所使用的引物将与 mRNA 的 polyA 尾杂交，因此此方法是非特异性的。

一般地，为了得到扩增子，所述扩增方法例如 PCR 或 NASBA 需要两种引物。作为其结果，扩增反应的产物本身接着成为扩增的模板，使扩增成为指数型的。尽管这些方法使得能够得到非常好的扩增产率，但他们仍然不适合用于同时扩增多个靶（称为多重扩增），然而需要分析几种基因的表达谱时，这一步骤是必要的。此外，申请 WO 99/4385 中所描述的方法基于扩增所有的 mRNA，这在 DNA 芯片的分析中会促进背景噪声的产生。此外，此方法局限于基因的 3'区。因此，这种方法不适合用于研究突变或者研究位于基因中间或其 5'区的靶区域。

本发明提出通过提供特异性线性扩增方法来解决现有技术的所有缺点，所述方法使得能够同时扩增不同基因的靶区域，所述靶区域可以位于基因的 3'、5'或中间。

在这一方面中，本发明涉及由以下物质产生转录物的方法：

- 包含“引物”区和目的区的待扩增 RNA 序列，和
- 包含启动子区和能与待扩增 RNA 序列的所述“引物”区杂交的区域的扩增引物，

所述方法在恒温下进行，并包括下列步骤：

- a) 将所述引物与待扩增 RNA 杂交，
- b) 通过逆转录酶的酶活性来延伸引物，以产生待扩增 RNA 的互补

脱氧核糖核酸 (cDNA) 序列,

c) 通过具有核糖核酸酶 H 活性的酶来切割与所述 cDNA 杂交的待扩增 RNA, 以获得与所述 cDNA 杂交的待扩增 RNA 片段,

d) 通过逆转录酶和链置换酶来延伸所述待扩增 RNA 片段的末端, 以获得 RNA-DNA/DNA 杂交体,

e) 通过具有 RNA 聚合酶活性的酶从在步骤 d) 中形成的 RNA-DNA/DNA 杂交体获得 RNA 转录物。

在本发明的范围内, 术语“转录物”意指转录产物, 即在起始转录的启动子控制下的模板转录过程中新合成的 RNA。

术语“序列”意指含有天然碱基 (如果涉及脱氧核糖核苷酸性质的序列则为 A、C、G、T, 如果涉及核糖核苷酸性质的序列则为 A、C、G、U) 和/或一种或多种经修饰的碱基的核苷酸序列, 所述经修饰的碱基例如为肌苷、5-甲基-脱氧胞苷、脱氧尿苷、5-二甲基氨基-脱氧尿苷、2,6-二氨基-嘌呤、5-溴-脱氧尿苷或允许杂交反应的任何经修饰的碱基。也可以在核苷酸间键 (可包括如硫代磷酸酯键、Hphosphonate 键、烷基磷酸酯键) 的水平上或在骨架的水平上修饰多核苷酸序列, 例如对于  $\alpha$ -寡核苷酸 (法国专利号 2 607 507) 或 PNA (EGHOLM 等人, 1992, J. Am. Chem. Soc. 114: 1895-1897) 的情况。在经修饰的多核苷酸序列中, 如上文所指出的, 多种修饰可以组合存在。此序列可以是 DNA 或 RNA 性质或者 DNA/RNA 嵌合体性质, 并且“长度”通常为至少 5 个脱氧核糖核苷酸和/或核糖核苷酸, 其中可以任选地含有至少一个经修饰的核苷酸。序列可以含有各自与不同功能相关的多个不同的区域。

为了实施本发明的方法, 从待扩增 RNA 序列开始, 其可以被看作是待扩增 RNA 的第一个模型。在本发明的范围内, 待扩增 RNA 可以是任何核糖核酸序列。优选地, 待扩增 RNA 为信使 RNA。

特别地, 待扩增 RNA 包含目的区, 即能与所述目的区的特异性杂交探针杂交的多核苷酸。因此, 该探针使得能够如下文所述地检测目的区存在与否。

该待扩增 RNA 还包含“引物”区，即下文定义的扩增引物可以与其特异性杂交的多核苷酸。优选地，“引物”区位于待扩增 RNA 序列的目的区的 3' 区。

“能与另一序列/区域杂交的序列或区域”这种表述意指可以在杂交条件下与另一序列/区域杂交的序列或区域，所述杂交条件可以根据各种情况以已知方式确定。也称为互补序列/区域。与另一序列或区域严格互补的序列或区域是其中每个碱基都能与另一序列的碱基配对而没有错配的序列。术语“杂交”意指这样的过程，在该过程中，在适当条件下，具有充分互补性序列的两个核苷酸片段能够形成具有稳定且特异的氢键的双链。通过严紧性（即操作条件的严格性）来确定杂交条件。进行杂交的严紧性越高，特异性就越高。特别地，按照探针/靶双链体的碱基组成来定义严紧性，还通过两个核酸间的错配程度进行定义。严紧性还取决于反应参数，如杂交溶液中存在的离子种类的浓度和类型、变性剂的性质和浓度和/或杂交温度。进行杂交反应的条件严紧性主要取决于使用的杂交探针。所有这些资料都是熟知的，并且本领域技术人员可以确定合适的条件。

在本发明的范围内，术语“扩增引物”意指当与待扩增核酸（DNA 或 RNA 性质）杂交时使得能够引发延伸反应的序列。本发明使用的引物可以分成两个区域，它们是（在 5'-3' 方向上）：

- 可以被 RNA 聚合酶识别从而起始转录（即合成 RNA）的启动子区。该启动子区包含启动子和由 RNA 聚合酶起始转录的位点。特别地，该启动子区包含 RNA 聚合酶启动子的一条链。例如，可以提及 RNA 聚合酶的天然启动子、来自天然启动子并保留其功能性的截短的序列或能够起始转录的“环”结构，如 MOLLEGAARD N. E. 等人，1994, Proc. Natl. Acad. Sci. USA, 91, 3892-3895 所述。该启动子区还可以包含有利于起始转录的序列，其功能为限制形成无效的转录物。RNA 聚合酶可以是 DNA 依赖性的，这使得能够从 DNA 序列来转录 RNA；

- 能够与待扩增 RNA 的“引物”区杂交（即与待扩增 RNA 的所述“引物”区互补）的区域。该区域可以是 DNA 性质的。

因此，引物的 5'末端可以包含不与所述待扩增 RNA 杂交的核苷酸（启动子的反义序列、转录起始位点的反义序列等），而 3'末端必须具有允许在合适的温度和严紧性条件下起始引物延伸的长度和核苷酸组成。根据本发明的一个优选实施方案，引物为 DNA 性质的。引物的 3'不封闭，这使得能够稳定 DNA-RNA 杂交体。

本发明的方法为等温方法，即在多个步骤中不需要对温度进行任何改变。根据本发明的一个优选实施方案，该方法在 37℃ 和 50℃ 之间的温度下进行，优选 40℃ 和 45℃ 之间，更加优选 41℃。

在步骤 a) 中，在允许引物/待扩增 RNA 序列杂交的所述区域范围内使所述引物与待扩增 RNA 杂交。在合适的严紧性和温度条件下，通过待扩增 RNA 的预设区与引物的预设区之间的互补性来进行所述杂交。根据本发明的一个优选实施方案，在此步骤之前进行变性待扩增 RNA（优选在 65℃）的步骤。

在步骤 b) 中，通过逆转录酶的酶活性来延伸引物，以产生待扩增 RNA 的互补脱氧核糖核酸序列（cDNA）。此步骤需要存在脱氧核糖核苷酸，酶使用所述脱氧核糖核苷酸通过互补性来合成 cDNA。在与步骤 a) 中相当的合适的严紧性和温度条件下，在引物的 5'-3'方向进行所述延伸。这种逆转录酶优选为 AMV 逆转录酶。

在步骤 c) 中，用具有 RNase H 活性的酶切割与所述 cDNA 杂交的 RNA，从而获得与所述 cDNA 杂交的 RNA 片段。RNase H 活性使得能够水解 RNA-DNA 杂交体，而不水解 RNA-RNA 或 DNA-DNA 杂交体。RNase H 活性可以通过单独的酶（特别地，例如大肠杆菌的 RNase H 酶）而获得，或者通过逆转录酶（特别地，例如 AMV 逆转录酶）的 RNase H 活性而获得。

在步骤 d) 中，通过逆转录酶和链置换酶来延伸所述 RNA 片段的末端，从而获得 RNA-DNA/DNA 复合体。所述逆转录酶优选为 AMV 逆转录酶。

在步骤 e) 中，通过具有 RNA 聚合酶活性的酶从在步骤 d) 中形成的 RNA-DNA/DNA 复合体获得 RNA 转录物。通过在 5'-3'方向合成与转录的

## 核苷酸链反向平行且互补的 RNA 来进行转录

所使用的具有 RNA 聚合酶活性的酶是能够与引物的启动子结合并在体外从引物的转录起始位点开始起始 RNA 合成的酶。根据本发明的一个优选实施方案,具有 RNA 聚合酶活性的酶为噬菌体的 RNA 聚合酶,优选 T7 噬菌体的 RNA 聚合酶,但也可以使用其他聚合酶,如噬菌体 T3、SP6、gh-1 等的 RNA 聚合酶。当然,位于引物上的启动子序列必须适合于在步骤 e) 中使用的 RNA 聚合酶。

本发明的扩增方法可以通过在适当条件下加入实施该方法所需的所有化合物(即待扩增 RNA、酶、扩增引物和脱氧核糖核苷酸)来进行。该反应介质不含有可能干扰扩增过程的物质,如可以抑制所需酶的活性的物质、可以干扰引物与待扩增 RNA 杂交的物质或者可以降解扩增产物的物质。

因此,根据本发明的一个优选实施方案,施行步骤 a) 至 e) 的时间足以获得足够数量的转录物。优选地,施行步骤 a) 至 e) 的时间为 30 分钟至 3 小时 30 分钟,更加优选 1 小时 30 分钟至 2 小时 30 分钟。

然后,可以通过本领域技术人员已知的任何技术来检测如此获得的转录物。

根据本发明的一个优选实施方案,本发明的产生转录物的方法还包括以下步骤:

f) 使用所述目的区的特异性杂交探针来检测产生的 RNA 转录物。

术语“检测”意指通过物理方法的直接检测或使用标记物的检测方法。存在多种用于检测核酸的检测方法[参阅如 Kricka 等人, *Clinical Chemistry*, 1999, No. 45(4), 第 453-458 页;或 Keller G. H. 等人, *DNA Probes*, 第二版, Stockton Press, 1993, 第 5 和 6 部分, 第 173-249 页]。术语“标记物”意指能产生信号的示踪物。这些示踪物的非限制性列表包括产生可检测信号(如通过比色法、荧光或发光而可检测的)的酶例如辣根过氧化物酶、碱性磷酸酶、 $\beta$ -半乳糖苷酶或葡萄糖-6-磷酸脱氢酶;生色团例如荧光、发光或染料化合物;可以通过电子显微术或其电特性(如电导率)、可以通过电流分析法或伏安

法、或者通过阻抗测量而可检测的高电子密度基团；可以通过光学方法如衍射、表面等离子体共振（*résonance plasmon de surface*）或接触角变化，或者通过物理方法如原子力光谱法、隧道效应等而可检测的基团；放射性分子如  $^{32}\text{P}$ 、 $^{35}\text{S}$  或  $^{125}\text{I}$ 。因此，可以在本发明的扩增步骤中标记多核苷酸，如通过将标记的核苷三磷酸用于扩增反应。还可以使用间接系统，例如能与抗配体反应的配体。配体/抗配体对是本领域技术人员所熟知的，这例如是以下对的情况：生物素/链霉抗生物素蛋白、半抗原/抗体、抗原/抗体、肽/抗体、糖/凝集素、多核苷酸/该多核苷酸的互补序列、组氨酸连续序列（称为“标签”，用于金属如镍）。在这种情况下，配体带有结合剂。抗配体可以通过如以上段落所述的标记物直接进行检测，或者其自身可以通过配体/抗配体进行检测。

经标记的核苷酸将是核糖核苷酸。还可以在扩增步骤之后标记多核苷酸，例如按照文献 WO-A-91/19812 中所述的夹心杂交技术与经标记的探针杂交。这些探针还可用于 OLISA 技术，如申请 WO 03/098217 所述。

还可以使用识别根据本发明扩增出的 RNA 的所述目的区的杂交探针来进行该检测。术语“杂交探针”意指包含 5 至 100 个核苷酸基序（特别是 6 至 35 个核苷酸基序）的核苷酸片段，其具有在给定的条件下与目的区形成杂交复合体的杂交特异性。这种杂交探针可以是捕获探针。在这种情况下，可以用标记物对靶核苷酸片段进行预先标记。可以通过任何适当的手段（即直接或间接地，例如通过共价或吸附）使捕获探针固定在或者可固定在固体支持物上。接着进行所述探针与经标记的靶核苷酸片段之间的杂交反应。

杂交探针还可以是检测探针。在这种情况下，可以用标记物对杂交探针进行标记。接着进行所述检测探针与靶核苷酸片段之间的杂交反应。夹心测试还可以使用固定在固体支持物上的捕获探针，靶核苷酸片段与所述捕获探针杂交，检测探针与所述靶核苷酸片段杂交。

无论使用捕获探针还是检测探针，杂交反应都可以在固体支持物

上进行, 所述固体支持物包括可以在其上固定核酸的所有材料。可以使用合成材料或天然材料(其任选地进行化学修饰)作为固体支持物, 特别是多糖如基于纤维素的材料, 例如纸, 纤维素衍生物如乙酸纤维素和硝酸纤维素或葡聚糖, 聚合物, 共聚物(特别是基于苯乙烯型单体的), 天然纤维如棉花, 和合成纤维如尼龙; 无机材料如二氧化硅、石英、玻璃、陶瓷; 胶乳; 磁性颗粒; 金属衍生物、凝胶等。固体支持物可以是微量滴定板形式、申请 WO-A-94/12670 所述的膜形式、颗粒形式或生物芯片形式。术语“生物芯片”意指尺寸很小的固体支持物, 在所述支持物的预设位置上固定了大批捕获探针。生物芯片(更准确地, DNA 芯片)的概念出现于 20 世纪 90 年代初。由于已经开始开发蛋白质芯片, 因此目前这一概念已经得到了扩展。它基于整合了微电子学、核酸化学、图像分析和信息学的多学科技术。工作原理基于分子生物学: 杂交现象, 即两个 DNA 和/或 RNA 序列的碱基通过互补性而配对。

生物芯片方法基于使用固定在固体支持物上的捕获探针, 在所述捕获探针上让用荧光染料直接或间接标记的靶核苷酸片段样品发生作用。捕获探针以特定的方式安置在支持物或芯片上, 每个杂交给出关于该靶核苷酸片段的一条特定信息。获得的信息是累加的, 并使得能够例如定量一个或多个靶基因的表达水平。为了分析靶基因的表达, 可以制备带有非常大量的探针的生物芯片, 所述探针相应于转录成 mRNA 的靶基因的全部或部分。接着使例如来自需要分析的靶基因的 mRNA 的互补 DNA 进行杂交。杂交后, 洗涤支持物或芯片, 用例如扫描仪读数并通过信息学对荧光分析结果进行处理。为说明起见, 可以提及 Affymetrix 公司开发的用于分子诊断的 DNA 芯片("Accessing Genetic Information with High-Density DNA arrays", M. Chee 等人, Science, 1996, 274, 610-614. "Light-generated oligonucleotide arrays for rapid DNA sequence analysis", A. Caviani Pease 等人, Proc. Natl. Acad. Sci. USA, 1994, 91, 5022-5026)。在这一技术中, 捕获探针一般较小, 约为 20 个核苷酸。

在出版物 G. Ramsay, Nature Biotechnology, 1998, No. 16, 第 40-44 页; F. Ginot, Human Mutation, 1997, No. 10, 第 1-10 页; J. Cheng 等人, Molecular diagnosis, 1996, No. 1(3), 第 183-200 页; T. Livache 等人, Nucleic Acids Research, 1994, No. 22(15), 第 2915-2921 页; J. Cheng 等人, Nature Biotechnology, 1998, No. 16, 第 541-546 页中或者在专利 US-A-4,981,783、US-A-5,700,637、US-A-5,445,934、US-A-5,744,305 和 US-A-5,807,522 中给出了生物芯片的其他实例。还可以提及 Apibio 公司开发的低密度芯片,如申请 WO 03/098217 所述。固体支持物的主要特征必须是保留捕获探针与靶核苷酸杂交的特征,同时对检测方法产生最小的背景噪声。本发明特别适用于这些低密度芯片系统。

可以实施本发明的方法以产生来源于多种不同基因的 RNA 转录物。使得能够同时扩增不同基因的靶区域,这正是本发明的优点之一,所述靶区域可以位于基因的 3'、5'或中间。

在这一方面中,本发明还涉及由以下物质产生转录物的方法:

- 至少第一待扩增 RNA 序列和第一扩增引物,所述第一待扩增 RNA 序列包含“引物”区和目的区,所述第一扩增引物包含启动子区和能与第一待扩增 RNA 序列的所述“引物”区杂交的区域,
- 至少第二待扩增 RNA 序列和第二扩增引物,所述第二待扩增 RNA 序列与所述第一待扩增 RNA 序列不同,并包含“引物”区和目的区,所述第二扩增引物包含启动子区和能与第二待扩增 RNA 序列的所述“引物”区杂交的区域,

所述方法在恒温下进行,并包括下列步骤:

- a) 所述第一和第二引物与所述第一和第二待扩增 RNA 杂交,
- b) 通过逆转录酶的酶活性来延伸所述第一和第二引物,以产生第一待扩增 RNA 的第一互补脱氧核糖核酸(cDNA)序列和第二待扩增 RNA 的第二互补脱氧核糖核酸(cDNA)序列,
- c) 通过具有核糖核酸酶 H 活性的酶来切割分别与第一和第二 cDNA 杂交的第一和第二待扩增 RNA,以获得与第一 cDNA 杂交的第一待

扩增 RNA 片段和与第二 cDNA 杂交的第二待扩增 RNA 片段,

d) 通过逆转录酶和链置换酶来延伸所述第一和第二待扩增 RNA 片段的末端, 以获得 RNA-DNA/DNA 杂交体,

e) 通过具有 RNA 聚合酶活性的酶从在步骤 d) 中形成的 RNA-DNA/DNA 杂交体获得所述第一和第二待扩增 RNA 的转录物。

可以以相当的方式使用此方法从来源于 3、4、5、10 个或更多个不同基因的待扩增 RNA 开始来产生转录物。

按照与上文所述相当的方法施行步骤 a) 至 e), 并且其施行时间足以获得足够数量的转录物, 如上文所定义。

为了检测转录物, 优选在步骤 f) 中使用第一杂交探针来检测第一待扩增 RNA 的转录物, 和使用第二杂交探针来检测第二待扩增 RNA 的转录物。

优选地并如上文所定义, 具有 RNA 聚合酶活性的酶为噬菌体的 RNA 聚合酶, 更加优选 T7 噬菌体的 RNA 聚合酶。

所附的图用于作为说明性示例, 而不具有任何限制性的性质。它们使得能够更清楚地理解本发明。

图 1 示意性地显示了本发明方法的步骤。

图 2 显示了杂交探针发射的信号, 所述杂交探针与如实施例 2 所述根据本发明获得的转录物杂交。

图 3 显示了按照实施例 3 所述方案扩增一系列转录物 ( $1 \times 10^7$ 、 $1 \times 10^8$ 、 $1 \times 10^9$  和  $1 \times 10^{10}$  个拷贝) 之后得到的线性回归。

图 4 显示了使按照实施例 4 所述方案获得的扩增产物杂交并显影之后, 与每个特异性探针相关的所发射出的信号 (3 次重复的平均值) 及其标准偏差。

在图 1 中, 以波浪线代表核糖核酸性质的序列, 而直线代表脱氧核糖核酸性质的序列。

在步骤 a) 中, DNA 性质的引物 (以长虚线代表) 与 RNA 模板的“引物”区 (以粗线代表) 杂交, 引物 5' 端的启动子序列不与模板 RNA 杂交。为了更清楚地理解本发明, 以短虚线代表需要在扩增阶段后检测

的目的序列，其位于“引物”区的 5'侧。然而，目的序列和“引物”序列可以是单一的相同序列。

在步骤 b) 中，使具有逆转录酶(RT)活性的酶发挥作用，以按 5'-3' 方向延伸与 RNA 模板杂交的引物。获得了序列与 RNA 模板序列互补的 DNA (cDNA，以虚线代表)。由此在靶区域与模板 RNA 的 5'末端之间获得了模板 RNA/cDNA 双链杂交体序列。由于其缺乏互补性，因此引物的启动子序列和位于靶 RNA 的靶区域的 3'侧的序列仍保持为未杂交(单链)的序列。

在步骤 c) 中，使具有 RNase H 活性的酶发挥作用，从而诱导模板 RNA 内的切割。这些切割是非特异性的，使得能够释放与 cDNA 杂交的 RNA 片段的 3' OH 末端(以叉号代表)。

在步骤 d) 中，通过具有逆转录酶活性和链置换活性的酶延伸如此释放出的 3' OH 末端。延伸终止于合成与引物的启动子序列互补的链。这样获得了多种不同的 RNA-DNA/DNA 双链杂交体序列，其包含在 DNA 依赖性 RNA 聚合酶作用下能起始转录的双链启动子区(以圆圈代表)。如此，以说明性的方式显示了从两个不同切割位点开始获得的两种不同杂交体。在步骤 e) 中，DNA 依赖性 RNA 聚合酶的作用使得能够获得与在步骤 d) 中获得的多种不同杂交体相对应的 RNA 转录物。从而获得一群大小不同但都包含需要检测的目的区的转录物。

给出以下实施例用于举例说明，但不具有任何限制性的性质。它们使得能够更清楚地理解本发明。

#### 实施例 1: 获得待扩增 RNA 序列

按照供应商推荐的方案，在双启动子载体 pCR-II-TOPO (Invitrogen) 中，从来自 HML-4 内源性逆转录病毒的 *pol* 基因区域的克隆序列开始产生第一待扩增 RNA 序列，其相应于 SEQ ID No. 1:

GUCACUGUCA AGCCUCCAGA UGCCAUCCCU UUGACUUGAA AACUCAGCCA  
GUUUGGGUGG AUCAGUGGCC GCUCCCAAAA AAUAAGCUGG AGGCGCUCCA

UAAUUUAGUC CUGGAACAGU UAGAAUUGGG ACACAUUGAG GAAUCUUUCU  
 CUCCAUGGAA UUCACUUGUC UUUGUUAUCC AAAAGAAAUC UGGGGAAAAC  
 AGAGAAUGCU CACUCAUCUU AGGGCAGUUA AUGCUGUACU UCAACCUCUG  
 GGGACAUIIAC AAUCUGGCUU ACCCUCCCCG UCUAUGCUCG CUGAGUAUUG  
 GCCUCUAAUC CUCAUAGAUC UUAAGAUUG CUUUUUUAAC AUUCCACUGG  
 CCUCUCAGGA CUUUGAAAAG UUUGCUUUUA UGGUCCCUUC CCUCAACAAU  
 GUCGCUCAGG CUACAUGCUA CUAUUGGAAA GUCCUACCAC AAGGCAUGCU  
 UAAUAGUCCC ACUAUUUGUC AGUAUUUUGU GGGGCGUGUG CUUCAACCUG  
 UCAGGGAUCA GUUUCCTCCGA UGUUACAUCG UUCACUACAU GGAUGAUCUC  
 CUCUGCACAG CCCCCCAUA CACCAUUUUG AUUUCUGCU UUUCUGUGAU  
 UCAACAGGCC AUUUCAGAAG CAGGUUUGAC UAUUGCACCA GAAAAAUUC  
 AACUACCUC UCAUUUUCAA UAUUUGGGCA UGCAGUUGGA AGACAAGCUG  
 AUUACACCAC AAAAAGUUCA GCUUAGGAGA GACGCCU .

使用 Ampliscribe T7 Flash 试剂盒 (Epicentre) 并按照供应商推荐的方案进行转录。

以相似的方式从来自 HERV-4.1 内源性逆转录病毒的 *pol* 基因区域的克隆序列开始产生第二待扩增 RNA 序列, 其相应于 SEQ ID No. 2:

GGCCAAGAGA UAAGACCAG CUCUGGCUAA GCGGUGGCCA AGAGUGUGGG  
 CGGAAGACAA CCCUCCAGGG UUGGCAGUCA ACCAAGCCCC CGUGCUUAUA  
 GAAGUUAAGC CUGGGGUCCA GCCGGUAGG CAAAAACAGU ACCCGGUCCU  
 CAGAGAAGCU CUUGAAGGUA UCCAGGUCCA UCUCAAGUGC CUAAGAACCU  
 UUAGAAUUAU AGUUCCUUGU CAGUCUCCA UGGAACACUCC CCUCCUGCCU  
 GUUCCCAAGC CUGGGACCAA GGACUACAGG CCGGUACAGG AUUUGCGCUU  
 GGUUAAUCAG GCUACAGUGA CUUUACAUC AACAGUACCU AACCUGUACA  
 CAUUGCUGGG GUUGCUGCCA GCUGAGGACA GCUGGUUCAC CUGCUUGGAC  
 CUGAAAGAUG CCUUCUUUAG CAUCAGAUUA GCCCCUGAGA GACAGAAGCU  
 GUUUGCCUUU CAGUGGGAAG AUCCAGAGUC AGGUGUCACU ACUCAAUACA  
 CUUGGACCCA GCCUCCCCAA AGGUUCAAGA ACUCCCCAC CAUCUUUGGG  
 GAGGCGUUGG CUCGAGACCU CCAGAAGUUU CCCACCAGAG ACCUAGGCUG

CGUGUUGCUC CAGUACGUUG AUGACCUUUU GCUGGGACAC CCCACGGCAG  
UCGGGUGCGC CAAGGGAACA GAUGCUCUAC UCCGGCACCU GGAGGACUGU  
GGGUAUAAGG UGUCCAAGAA AAAAAGCUCA GAUCUGCCGA CAGCAGGUAU  
GUUACUUGGG AUUUACUAUC CAACAGGGGG AGCACAGCCU GGGAUCAGAA  
AGAAAGCAGG UCAUUUGUAA UCUACCGGAG CCUAAGACCA GAAGGCAGGU  
GAGAGAAU

实施例 2: 通过本发明方法从实施例 1 中定义的 RNA 序列产生转录物  
为了扩增 SEQ ID No. 1 的 RNA 序列, 使用 SEQ ID No. 3 的引物序  
列:

SEQ ID No. 3: *TAATACGACT* **CACTATAGGG** **AGATCTAGCA** GCATAGGGGG  
GGCTGTGCAG AGGAGATCAT CC,

其包含:

- 包含 T7 RNA 聚合酶的启动子 (在 SEQ ID No. 3 中以斜体标出) 和转录起始位点 (在 SEQ ID No. 3 中以粗体标出) 的启动子区,
- 能与待扩增 RNA 序列的引物区杂交的区域 (在 SEQ ID No. 3 中以下划线标出)。

该引物在 3'端未封闭。

应该指出, 能与引物区杂交的区域可以更短, 特别是可以包含 15 至 30 个碱基对。还应该指出, 在起始位点和能与引物区杂交的区域之间包含了促进起始的序列, 其限制了无效转录物的形成。

为了扩增 SEQ ID No. 2 的 RNA 序列, 使用 SEQ ID No. 4 的引物序列:

SEQ ID No. 4: *TAATACGACT* **CACTATAGGG** **AGATCTAGCA** GCATACAGCA  
AAAGGTCATC AACGTACTGG AG

其中包含:

- 包含 T7 RNA 聚合酶的启动子 (在 SEQ ID No. 4 中以斜体标出) 和转录起始位点 (在 SEQ ID No. 4 中以粗体标出) 的启动子区,
- 能与待扩增 RNA 序列的引物区杂交的区域 (在 SEQ ID No. 4 中

以下划线标出)。

该引物在 3'端未封闭。

应该指出,能与引物区杂交的区域可以更短,特别是可以包含 15 至 30 个碱基对。还应该指出,在起始位点和能与引物区杂交的区域之间包含了促进起始的序列,其限制了无效转录物的形成。

转录物的产生在反应混合物(NucliSens Basic Kit, bioMérieux)中进行,所述反应混合物包含具有逆转录酶活性的酶(AMV-RT)、具有 RNase H 活性的酶(RNase H)和具有 RNA 聚合酶活性的酶(T7 RNA 聚合酶)。

在以下引物存在下进行扩增:

- SEQ ID No. 3 的引物序列 (0.5  $\mu$ M), 用于扩增 SEQ ID No. 1 的 RNA 序列,
- SEQ ID No. 4 的引物序列 (0.5  $\mu$ M), 用于扩增 SEQ ID No. 2 的 RNA 序列,
- SEQ ID No. 3 的引物序列 (0.5  $\mu$ M) 和 SEQ ID No. 4 的引物序列 (0.5  $\mu$ M), 用于同时扩增 SEQ ID No. 1 的 RNA 序列和 SEQ ID No. 2 的 RNA 序列。

因此,以整合有生物素的方式对于 SEQ ID No. 1 和 SEQ ID No. 2 的 RNA 序列进行了扩增:

- 彼此独立地进行:

为此,将水溶液中已知量的 SEQ ID No. 1 的 RNA 或 SEQ ID No. 2 的 RNA (5  $\mu$ l H<sub>2</sub>O 中有  $1 \times 10^9$  或  $1 \times 10^{10}$  个拷贝)与 10  $\mu$ l 反应混合物(NucliSens Basic Kit)混合,在所述反应混合物中已经预先加入了 UTP-生物素 (0.6 mM) 以及使得能够扩增 SEQ ID No. 1 或 SEQ ID No. 2 的引物,

- 同时进行:

为此,将水溶液中已知量的 SEQ ID No. 1 的 RNA 和 SEQ ID No. 2 的 RNA 与 10  $\mu$ l 反应混合物(NucliSens Basic Kit)混合,在所述反应混合物中已经预先加入了 UTP-生物素以及使得能够扩增 SEQ ID

No. 1 和 SEQ ID No. 2 的引物。

对于每个扩增反应，得到的体积均为 15  $\mu$ l（足够量的无核酸酶的 H<sub>2</sub>O）。然后将混合物于 65℃ 孵育 5 分钟，接着于 41℃ 孵育 5 分钟，以使引物与待扩增 RNA 杂交。

然后加入体积为 5  $\mu$ l 的酶混合物（NucliSens Basic Kit）。将体积为 20  $\mu$ l 的混合物于 41℃ 孵育 90 分钟，以使：

1. 引物与待扩增 RNA 杂交，
2. 通过逆转录酶活性来延伸引物，以产生待扩增 RNA 的互补脱氧核糖核酸（cDNA）序列，
3. 通过具有核糖核酸酶 H 活性的酶来切割与所述 cDNA 杂交的待扩增 RNA，从而获得与所述 cDNA 杂交的待扩增 RNA 片段，
4. 通过逆转录酶和链置换酶来延伸所述待扩增 RNA 片段的末端，从而获得 RNA-DNA/DNA 杂交体，
5. 通过具有 RNA 聚合酶活性的酶从在步骤 d) 中形成的 RNA-DNA/DNA 杂交体获得 RNA 转录物。

通过 ELOSA 方法分析获得的 RNA 转录物（Mallet F. 等人，1993）。此方法允许通过核酸序列与移植到微量滴定板孔底的寡核苷酸探针（称为检测探针）的杂交来特异性地检测核酸序列。

使得能够证实 SEQ ID No. 1 的探针为序列 SEQ ID No. 5：  
CAACCTGTCAGGGATCAGTTTC。

使得能够证实 SEQ ID No. 2 的探针为序列 SEQ ID No. 6：  
CTCGAGACCTCCAGAAGTTTCC。

用生物素标记 RNA 转录物使得能够由于链霉抗生物素蛋白-酶缀合物与相应底物的作用而揭示出杂交，并通过光谱法读出 OD。OD 与通过本发明方法产生的转录物的量成比例。

图 2 中显示了得到的结果。字母 A 至 D 相应于彼此独立地扩增 SEQ ID No. 1 和 SEQ ID No. 2 的条件（A:  $1 \times 10^9$  个拷贝的 SEQ ID No. 1; B:  $1 \times 10^9$  个拷贝的 SEQ ID No. 2; C:  $1 \times 10^{10}$  个拷贝的 SEQ ID No. 1; D:  $1 \times 10^{10}$  个拷贝的 SEQ ID No. 2），而字母 E 至 H 相应于同时扩增 SEQ ID

No. 1 和 SEQ ID No. 2 (E:  $1 \times 10^9$  个拷贝的 SEQ ID No. 1 和  $1 \times 10^9$  个拷贝的 SEQ ID No. 2; F:  $1 \times 10^{10}$  个拷贝的 SEQ ID No. 1 和  $1 \times 10^{10}$  个拷贝的 SEQ ID No. 2; G:  $1 \times 10^9$  个拷贝的 SEQ ID No. 1 和  $1 \times 10^{10}$  个拷贝的 SEQ ID No. 2; H:  $1 \times 10^{10}$  个拷贝的 SEQ ID No. 1 和  $1 \times 10^9$  个拷贝的 SEQ ID No. 2)。通过 SEQ ID No. 5 的检测探针(灰色柱)的存在或通过 SEQ ID No. 6 的检测探针(黑色柱)的存在来检测得到的转录物。

结果显示,在独立扩增与同时进行的扩增之间检测到的信号相当。

因此,无论待扩增 RNA 为 SEQ ID No. 1 还是 SEQ ID No. 2,均产生数量相当的转录物(当按照条件 A 和条件 B 进行扩增时,OD 读数相当)。此外,当待扩增 RNA 的量提高时,产生的转录物的数量更多(按照条件 A 进行扩增时的 OD 读数比按照条件 C 进行扩增时低;类似地,按照条件 B 进行扩增时的 OD 读数比按照条件 D 进行扩增时低)。

最后,同时扩增 SEQ ID No. 1 和 SEQ ID No. 2 时产生的转录物的数量与彼此独立地扩增 SEQ ID No. 1 和 SEQ ID No. 2 时产生的转录物的数量相当。甚至在以相差 1 个对数级的转录物量进行同时扩增时,信号仍然与彼此独立地扩增 SEQ ID No. 1 和 SEQ ID No. 2 时得到的信号相同。

通过在 OLISA 芯片上分析产生的转录物得到相似的结果。OLISA 芯片是排列在标准形式 96 孔平板上的低密度 DNA 芯片(每个芯片至多 128 个探针)。将冠状分布的寡核苷酸检测探针移植到孔底。使用链霉抗生物素蛋白-酶缀合物进行探针/靶杂交体的比色法揭示,其中包括用生物素标记靶。

额外的用 RNase 进行消化的步骤确认,检测的产物确实是 cRNA。

因此,这些结果证明,本发明的方法特别适合于扩增用于在 DNA 芯片上对其进行定量分析的转录物。这是因为,此方法是转录方法,并且扩增产物不会成为扩增模板,这一事实使该扩增具有线性性质,所述线性性质非常适合于在生物芯片上对转录物进行定量。

### 实施例 3: 评估本发明方法的扩增的线性

基于用作待扩增 RNA 序列的 SEQ ID No. 2 和用作扩增引物的 SEQ ID No. 4 来进行本实施例。

根据实施例 2 中所述的条件进行该产生转录物的方法。在这种情况下，以 0.5  $\mu\text{M}$  的浓度使用 SEQ ID No. 2 的引物。按照与实施例 2 中所阐述的相似的原理，扩增出一系列转录物 ( $1 \times 10^7$ 、 $1 \times 10^8$ 、 $1 \times 10^9$  和  $1 \times 10^{10}$  个拷贝)，并使用 SEQ ID No. 6 的探针在 OLISA 芯片上分析反应产物。

如图 3 所示，采用由每个点的至少 4 次测量得到的平均值而进行的线性回归导致获得相关系数  $R^2 = 0.9781$  的直线。因此，这证明了本发明方法的扩增的线性。

#### 实施例 4：多重扩增

对于这个实验，如上文实施例 1 和 2 中所述实施扩增方案。

按照供应商推荐的方案，在双启动子载体 pCR-II-TOPO (Invitrogen) 中，从来自 HML-2 内源性逆转录病毒的 *pol* 基因区域的克隆序列开始产生第一待扩增 RNA 序列，其相应于 SEQ ID No. 7:

```
CACUGUAGAG CCUCCUAAAC CCAUACCAUU AACUUGGAAA
ACAGAAAAAC UGGUGUGGGU AAAUCAGUGG CCACUACCAA
AACAAAAACU GGAGGCUUUA CAUUUAUUAG CAAAUGAACA
GUUAGAAAAG GGUCAUAUUG AGCCUUCAUU CUCGCCUUGG
AAUUCUCCUG UGUUUGUAAU UCAGAAGAAA UCAGGCAAU
GGCAUAUGUU AACUGACUUA AGGGUCGUAA ACGCCGUAAU
UCAACCCAUG GGGCCUCUCC AACCCGGGUU GCCCUCUCCG
GCCAUGAUCC CAAAAGAUUG GCCUUUAGUU AUAAUUGAUC
UAAAGGAUUG CUUUUUUACC AUCCCUCUGG CGGAGCAGGA
UUGCGAAAAA UUUGCCUUUA CUAUACCAGC CAUAAAUAAU
AAAGAACCAG CCACCAGGUU UCAGUGGAAA UUGUUACCUC
AGGGAAUGCU UAAUAGUCCA ACUAUUUGUC AGACUUUUGU
AGGUCGAGCU CUUCAACCAG UUAGAGACAA GUUUUCAGAC
```

UGUUAUAUUA UUCAUUAUUA UGAUGAUUU UUAUGUGCUA  
 CAGAAACGAG AGAUAAAUAU AUUGACUGUU AUACAUUUCU  
 GCAAGCAGAG GUUGCCAAUG CAGGACUGGC AAUAGCAUCU  
 GAUAAGAUGC AAACCUCUAC UCCUUUUCAU UAUUUAGGGA  
 UGCAGAUAGA AAAUAGAAAA AUUAAGCCAC AAAAAUAGA  
 AAUAAGAAAA GACACAUUAA AACACUAAA UGAUUUUCAA  
 AAUUGCUGG GAGAUAUUAA UUGGAUUCGG CCAACUCUAG  
 GCAUUCCUAC UUAUGCCAUG UCAAAUUU

使用 Ampliscribe T7 Flash 试剂盒 (Epicentre) 并按照供应商推荐的方案进行转录。

以相似的方式, 按照供应商推荐的方案, 在双启动子载体 pCR-II-TOP0 (Invitrogen) 中, 从来自 HML-5 内源性逆转录病毒的 *pol* 基因区域的克隆序列开始产生第二待扩增 RNA 序列, 其相应于 SEQ ID No. 8:

UAUCCAAUUU GGGUAGACCA GUGGCCUUUA AAGGGAGAGA  
 AAUUGCAAAG AGCCCAUGAG UUUGUUGAAG AGCAAUUA  
 AGCCAGCCAU AUAGAACCAU CAAAAGACC UUGGAAUUCA  
 CCCAUUUUCA UCAUUCCCAA AAAGUCUGGU AAUAGAGAC  
 UUUUGCAUGA CUAACAUGCU AUCAAUGCUA AUUUGAAACC  
 UAUGGGACCC CUUGAGCAGG UGUUCUCCUC CCCUUCAGUG  
 AUUCCUUGAG AUUGGCAUUA AAUAGUUAUU GGCUUAUAAU  
 ACUGCAUUUG UACUAUUCU UUUGCAGAAC AGGACAGAGA  
 AAAAUUUGUG UUAUAUAUAC CAGCUAUAAA UAACGAAAGG  
 CCAGCUCCCC GAUUUCACUG GAAAGUGCUU CCUGAAGGGA  
 UGCUAAACAG UCCUACCAUG UGUCAGUAUC AUGUAAAUCA  
 AGCUUUGCUC CUCAGUAGAA AAGAAUUUCC UAAUUGCAAG  
 ACUAUUCAUU UUAUGGAUGA UAUUUUACUA GCAGCCCCAA  
 UGGAUCCAGC ACUUUUAAAGU UUAUGUGCCU CUGUUGUAAA  
 GAAUACACAG UUAAGAGGUU UAAUCAUAGC ACCUGAAAAA

GUACAGAUGU CCUCUCCUUG GAAUAUUCUU GGGUACAUAC  
 UAACUUCGUG GUCAGUAAGA CCUCAAAAGG UUAAAUUAAA  
 UACUAGCAAC UUACACACCU UAAAUGAUUA UCAGAAAUUA  
 CUAGGCAAUA UUAACUGGCU UUGCCCCACC UUGGGCAUAA  
 CUACUGAUAA GUUACAGAAC CUGUUUUCUA UCUUAAAAGA  
 CAAUGCUGCU CUAGACUCUC CCAGGUUUU AACUCCUGCA  
 GCACAAAGGG AAUUGAGGA AGUAGAGCAC

使用 Ampliscribe T7 Flash 试剂盒 (Epicentre) 并按照供应商推荐的方案进行转录。

以相似的方式, 按照供应商推荐的方案, 在双启动子载体 pCR-II-TOPO (Invitrogen) 中, 从来自 HERV-H 内源性逆转录病毒的 *pol* 基因区域的克隆序列开始产生第三待扩增 RNA 序列, 其相应于 SEQ ID No. 9:

CACCCUUACC CUGCUCUAAUG CCAAUAUCCC AUCCCACAGC  
 AUGCUUUAAA AGGAUUAAG CCUGUUAUCA CUCGCCUGCU  
 ACAGCAUGGG CUUCUAGAAC CUAUAAACUC UCUUUUCAAU  
 UCCCCCAUUU UACCUGUCCA AAAACCGGAU AAGUCUUACA  
 GACUAGUUCA GAAUCUGCGU CUUAUCAACC AAUUGUUUU  
 GCCUAUCCAC CCUGUGGUGC CCAACCUGUA CACUCUUUUG  
 UCCUCAAUAC CUUCCUCCAC AACUCACUAU UCCGUGCUUG  
 AUCUUAAGA UGGUUUUUUC ACUAUUCUCC UGCACUCCUC  
 GUCCCAGCCU CUCUUUGCUU UCACCUGGAC UGACCCUGAC  
 ACCCAUCAGU CCCAGCAGCU UACCUGGACU GUGCUGCCGC  
 AAGGUUUCAG GGACAGCCCU CGUUACUUCA GCCAAGCUCU  
 UUCUCAUGAU CUACUUUCCU UCCACCCUC CGCUUCUCAC  
 CUUAUUCAAU AUAUUGAUGA GCUUCUUCUU UGUAGCCCCU  
 CCUUUGAAUC UUCUCAAUAA GACACACUUC UGCUCCUUCA  
 GCAUUUAUUC UCCAAAGGAU AUCGGGUAUC CCCUCCAAA  
 GCUCAAAUUU CUUCUCCAUC CGUUACCUAC CUCGGCAUAA

UUCUUCAUAA GAACACACGU GCUCUCCCUG CCGACCGUGU  
 CUGACUAAUC UCUCAAGCCC CAACCCCUUC UACAAAACAA  
 CAACUCCUUU CCUUCCUGGG CGUGGUUGGA UACUUUCGCC  
 UUUGGAUACC UGGUUUUGCC AUCCUAACAA AACCAUUAUA  
 UAAACUCACA AAAGGAAACU UAGCUGACCC CAUAGAUCCU  
 AAUCCUUUC CCCACUCCUC UUUCUG

使用 Ampliscribe T7 Flash 试剂盒 (Epicentre) 并按照供应商推荐的方案进行转录。

以相似的方式, 按照供应商推荐的方案, 在双启动子载体 pCR-II-TOPO (Invitrogen) 中, 从来自 ERV-9 内源性逆转录病毒的 *pol* 基因区域的克隆序列开始产生第四待扩增 RNA 序列, 其相应于 SEQ ID No. 10:

CUGAAGCUCU UAAAGGAUUA CAGGAUAUUG UUAACAUAUU  
 AAAAGCUCAA GGCUUAGUAA GGAAUUGCAG CAGUCCCUGC  
 AACACCCCAA UUCUAGGAGU ACAAAAACCA AAUGGUCACU  
 GGAGACUAGU GCAAGAUCUU AGACUCAUCA AUGAGGCAGU  
 AAUUCCUCUA UAUCCAGUUG UACCCAACCC CUAUACCCUG  
 CUUUCUCAA UACCAGAGGA AGCAGAAUGG UUCAUGGUUC  
 UGGACCUCAA GGAUGCCUUC UUCUGUUUCC CCUGCACUCU  
 GACUCCCAGU UUCUGUUUGC CUUUGAGGAU CCCACAGACC  
 ACACGUCCCA ACUUACAUGG AUGGUCUUGC CACAAGGGUU  
 UAGGGAUAGC CCUCACCUGU UUGGUCAGGC ACUGGCCCAA  
 GAUCUAGGCC ACUUCUCAAG UCCAGGCACU UUGGUCCUUC  
 AGUAUGUGGA UGAUUUACUU UUGGCUACCA GUUCAGAAGC  
 CUCAUGCCAG CAGGCUACUC UAGAUCUCUU GAACUUUCUA  
 CCUAAUCAAG GGUACAAGGC AUCUAGGUCA AAGGUGCAGC  
 UUUGCUUACA GCAGGCUAAA UAUCUAGGCC UAAUCUUAGC  
 CAGAGGGACC AGGGCCCUCA GCAAGGAAUG AAUACAGCCU  
 AUACUGGCUU AUCCUUGCCC UAAGACAUUA AAACAGUUGC

AGGGGUUCCU UGGAAUCACC GGCUUUUGCC GACUAUGGAU  
 CCCUGGAUAC AGUGAGAUAG UCAGGCCCCU CCAUACUCUA  
 AUCAAGGAGA CCC

使用 Ampliscribe T7 Flash 试剂盒 (Epicentre) 并按照供应商推荐的方案进行转录。

以相似的方式, 按照供应商推荐的方案, 在双启动子载体 pCR-II-TOPO (Invitrogen) 中, 从来自 HERV-W 内源性逆转录病毒的 *pol* 基因区域的克隆序列开始产生第五待扩增 RNA 序列, 其相应于 SEQ ID No. 11:

GACCCAAGGC CCAACAAGGA CUCCAAAAGA UUGUUAAGGA  
 CCUAAAAGCC CAAGGCCUAG UAAAACCAUG CAGUAACCCC  
 UGCAGUACUC CAAUUUUAGG AGUACAGAAA CCAACAGAC  
 AGUGGAGGUU AGUGCAAGAU CUCAGGAUUA UCAAUGAGGC  
 UGUUGUCCU CUAUAGCCAG CUGUACCUAG CCCUUAUACU  
 CUGCUUUCCC AAUACCAGA GGAAGCAGAG UGGUUUACAG  
 UCCUGGACCU UCAGGAUGCC UUCUUCUGCA UCCUGUACA  
 UCCUGACUCU CAUUCUUGU UUGCCUUUGA AGAUACUUCA  
 AACCCAACAU CUCAACUCAC CUGGACUUUU UACCCAAGG  
 UUCAGGGAUA GUCCCCAUCU AUUUGCCAGG CAUAGCCCA  
 AGACUUGAGC CAUCCUCAU ACCUGGACAC UUGUCUUCGG  
 UAGGUGGAUG AUUUACUUUU GGCCGCCCAU UCAGAAACCU  
 UGUGCCAUCA AGCCACCCAA GCGCUCUUCA AUUCCUCGC  
 UACCUGUGGC UACAUGGUUU CCAAACCAA GGCUCAACUC  
 UGCUCACAGC AGGUUACUUA GGGCUAAAAU UAUCCAAAGG  
 CACCAGGGCC CUCAGUGAGG AACACAUCCA GCCUAUACUG  
 GCUUAUCCUC AUCCCAAAC CCUAAAGCAA CUAAGGGGAU  
 UCCUUGGCGU AAUAGGUUUC UGCCGAAAAU GGAUUCCAG  
 GUAUGGCGAA AUAGCCAGGU CAUUAAAUAC ACUAAUUAAG  
 GAAACUCAGA AAGCCAUAC CCAUUUAGUA AGAUGGACAA

CUGAAGUAGA AGUGG

使用 Ampliscribe T7 Flash 试剂盒 (Epicentre) 并按照供应商推荐的方案进行转录。

以相似的方式, 按照供应商推荐的方案, 在双启动子载体 pCR-II-TOPO (Invitrogen) 中, 从来自 HERV-R 内源性逆转录病毒的 *pol* 基因区域的克隆序列开始产生第六待扩增 RNA 序列, 其相应于 SEQ ID No. 12:

```

AUCAUAAUCC AAUGCCAGUC ACCCUGGAAC ACUCCACUUU
UGCUGGUACA AAAACCAUUG CCUGGACCAG GAUCUGAUGA
GUAUAUACUG GUGCAGGGCU UGCAUGCUGU AAACCAAGCC
ACGGUGACCA UACAUCCAGU AGUACCAAAC CUGUAUACUU
UAAUGGGACU UAUUCUAGCA AGUGCCACCU GGUUUACAGU
CCUGGACUUA AAGGAUGCUU UCCUCUGUCU CUACCUGGCA
CCAGUUAGUC AGCCCAUCUU UGCAUUUUAA UGGGACAAUU
CAGUCACAGG CACAGGGGGA CAGCUCGCCU GGACUAGUCU
CCCACAAGGG UUCAAGAAUU CUCCCACAAU CUUUGGGGAA
GCACUGGCCU CAGACCUCAA GGCAUACACC CCACCAAUG
ACAACUGCGC CUUGUUGCAG UACAUAGACA ACCCUUCUUU
UGGCAGCCCC AACCCCAAGA GGACUGUAAC UGGGGGAACC
CAGGACCUCU UCCAUCUCUU AUGGGAAAGC AGGGUAUAGA
GUAUCCAAAA AGAAGGCCCA AAUUUGCCAU GAAAAGGUUA
AAUAUUUAGG CUUCAUAGUA AGCCAAGGGG AACGCUGGCU
UGGCCAUGGA UGAAAGCAAG CCAUUUGUGC ACUUCCAACU
CCAACCACCU GGCGCCAAAU AGGGGAAUUC UUAGGGGCAG
CAGGGUUCUG CCAUAUCUGG AUCCCAAUU UCUCACUUAU
AGCCAGGCCC UUAUAUGAAG CCACAAGGGA GGGUGAAAAG
GAACCCCUCC UCUGAAAGGC UGACCAGAAG AAGGUGUUUA
AACAAAUCAA AGAAGCUCUA ACUCAGGCUC CAGCCUUAGG
ACUGCCAGAU ACUACU

```

使用 Ampliscribe T7 Flash 试剂盒 (Epicentre) 并按照供应商推荐的方案进行转录。

以相似的方式, 按照供应商推荐的方案, 在双启动子载体 pCR-II-TOPO (Invitrogen) 中, 从来自HERV-L内源性逆转录病毒的 *pol* 基因区域的克隆序列开始产生第七待扩增RNA序列, 其相应于SEQ ID No. 13:

```

UUA AUGGAGU UAUGGCUCAG GUCUGACUUA CAGCAGGUCC
AGUGGGUCCC UGGACUCAUC CUGUGUUCAU UUUCCCAGUG
CCAGAAUGCA UAAUUGGCAU UGUCAUACUU AGAGGCUGGC
AGAACCCCCA CAUUGAUUCU GUGACUGGUA AGGUGAGGGC
UAUUAUGGUG GGAAAGGCCA AAUGGAAGCC AUUGGAGCUG
CCUUUACCUA GGAAAUAAGU AAAUCAAAAA CAUUAUCACC
ACCCUGGAGG GAUUGCAGAG AUUAGUGCCA CCAUCAAGGA
CUUGAAAAAU GCAGGGGUGG CGAUUCCCAU AACAUCCUUG
UUCAACUCUC CUUUUUGGCC UGUGCAGAAG ACAGAUGGAU
CUUGGAGAAU GAGAGUGGAU UAUCAUAAGC UUAACCAAGU
GGUGACUCCA AUUGCAGCUG CUAUACAAGA UGUGGUUUUC
AUUGCUCUAG CAAAUUAAUA CAUCUCCUGG UACCUUGUAU
GCAGCCAUUG ACUUGGCAAU UGGCCUUUUA CCCAUUCCAU
AAGCCCCACC AGAAGCGAUU UGCCUUCAGC UGGCAAGGCC
AGCAAUGUAU CUUUACUGUC CUACUUCAGG GGUUAUAUCAA
CUCUCCGGCU UUGUGUCAUA AUCUUAUUCA GAGUGAUCUU
GAUCACUUUU CACUGCCACA AGAUAUACA CUGGUCCAUA
ACAUUGAUGG CAUUAUGUUG AUUGGAUCCA AUGAGCAAGA
AGUAGCAAAC ACACUGGACU UAUUGGUGAG ACAUUUGCAU
GCCAUAGGAU GGGAAAUAAA UCCAAAUAAA AUUCACGCAC
CCUCUACCUC AGUAAAAUU

```

使用 Ampliscribe T7 Flash 试剂盒 (Epicentre) 并按照供应商推荐的方案进行转录。

如实施例 2 中所述, 在以下引物存在下进行扩增:

- SEQ ID No. 14 的引物序列 AATTCTAATA CGACTCACTA TAGGGAGACA CATAAAATAT CATCAACATA (0.75  $\mu$ M), 用于扩增 SEQ ID No. 7 的 RNA 序列,

- SEQ ID No. 15 的引物序列 AATTCTAATA CGACTCACTA TAGGGAGATG CAGAGGAGAT CATCCATGTA (0.75  $\mu$ M), 用于扩增 SEQ ID No. 1 的 RNA 序列,

- SEQ ID No. 16 的引物序列 AATTCTAATA CGACTCACTA TAGGGAGACT AGTAAAATAT CATCCATAAA (0.75  $\mu$ M), 用于扩增 SEQ ID No. 8 的 RNA 序列,

- SEQ ID No. 17 的引物序列 AATTCTAATA CGACTCACTA TAGGGAGAAA AGTAGAAGGT CATCAATATA (0.75  $\mu$ M), 用于扩增 SEQ ID No. 9 的 RNA 序列,

- SEQ ID No. 18 的引物序列 AATTCTAATA CGACTCACTA TAGGGAGAAG TAAATCATCC ACATACTGAA G (0.75  $\mu$ M), 用于扩增 SEQ ID No. 10 的 RNA 序列,

- SEQ ID No. 19 的引物序列 AATTCTAATA CGACTCACTA TAGGGAGAAA GTAAATCATC CACGTACCGA (0.75  $\mu$ M), 用于扩增 SEQ ID No. 11 的 RNA 序列,

- SEQ ID No. 20 的引物序列 AATTCTAATA CGACTCACTA TAGGGAGACC AGCAAAAGGT CATCAACGTA (0.75  $\mu$ M), 用于扩增 SEQ ID No. 2 的 RNA 序列,

- SEQ ID No. 21 的引物序列 AATTCTAATA CGACTCACTA TAGGGAGACC AAAAGAAGGT TGTCTATGTA (0.75  $\mu$ M), 用于扩增 SEQ ID No. 12 的 RNA 序列,

- SEQ ID No. 22 的引物序列 AATTCTAATA CGACTCACTA TAGGGAGATC ARCATAATGY CATCAATGTA (0.75  $\mu$ M), 用于扩增 SEQ ID No. 13 的 RNA 序列。

如实施例 2 中所述, 在 OLISA 芯片上分析得到的 RNA 转录物。

使得能够证实 SEQ ID No. 1 的探针为序列 SEQ ID No. 5:  
CAACCTGTCAGGGATCAGTTTC。

使得能够证实 SEQ ID No. 2 的探针为序列 SEQ ID No. 6:  
CTCGAGACCTCCAGAAGTTTCC。

使得能够证实 SEQ ID No. 7 的探针为序列 SEQ ID No. 23:  
CAACCAGTTA GAGACAAGTT TTCA。

使得能够证实 SEQ ID No. 8 的探针为序列 SEQ ID No. 24:  
TTGCTCCCCA GTAGAAAAGA ATT。

使得能够证实 SEQ ID No. 9 的探针为序列 SEQ ID No. 25:  
GACAGCCCCC ATTACTTCAG TCAAG。

使得能够证实 SEQ ID No. 10 的探针为序列 SEQ ID No. 26:  
CCAAGATCTA GGCCACTTCT CA。

使得能够证实 SEQ ID No. 11 的探针为序列 SEQ ID No. 27:  
CATCTATTTG GCCAGGCATT A。

使得能够证实 SEQ ID No. 12 的探针为序列 SEQ ID No. 28:  
AGACCTCAAG GCATACACCC。

使得能够证实 SEQ ID No. 13 的探针为序列 SEQ ID No. 29:  
GCTTTGTGTC ATAATCTTAT TCG。

从四个独立的扩增中计算平均值和标准偏差，并显示于图 4 中。

本发明的技术使得能够混合地扩增转录物（在此，多达 9 种不同的）。此外，应该注意到，在测试范围内（每种转录物有  $10^6$  至  $10^8$  个拷贝），在测量信号和引入反应体积中的转录物的量之间存在线性关系。这表明，此技术使得能够同时并且特异性地扩增多种转录物。因此，对通过此方法得到的扩增产物的 DNA 芯片分析（或通过其他方法的分析）允许对初始样品的转录物进行定量。

<110> bioMérieux SA

<120> 产生转录物的方法

<130> unknown

<160> 29

<170> PatentIn version 3.1

<210> 1

<211> 737

<212> RNA

<213> HML. 4

<400> 1

|                                                                   |     |
|-------------------------------------------------------------------|-----|
| gucacuguca agccuccaga ugccaucucu uugacuugaa aacucagcca guuugggugg | 60  |
| aucaguggcc gcucccaaaa aaauagcugg aggcgcucca uauuuuaguc cuggaacagu | 120 |
| uagaauuggg acacauugag gaauuuucu cuccauggaa uucacuuguc uuuguuaucc  | 180 |
| aaaagaaauc ugaggaaaac agagaauucu cacucaucu agggcaguua augcuguacu  | 240 |
| ucaaccucug gggacauuac aaucuggcuu accucuccgc ucuauugcug cugaguauug | 300 |
| gccucuaauc cucauagauc uuaaagauug cuuuuuuac auuccacugg ccucucagga  | 360 |
| cuuugaaaag uuugcuuuua uggucccuuc ccucaacaau gucgcucagg cuacaugcua | 420 |
| cuauuggaaa guccuaccac aaggcaugcu uaauagucc acuaauuguc aguauuuugu  | 480 |
| ggggcgugug cuucaaccug ucagggauga guuuccccga uguuacaucg uucacuacau | 540 |
| ggaugaucuc cucugcacag ccccccaua caccuuuug auuuccugcu uuucugugau   | 600 |
| ucaacaggcc auuucagaag cagguuugac uauugcacca gaaaaauuc aaacuaccuc  | 660 |

---

|                                                                  |     |
|------------------------------------------------------------------|-----|
| ucauuucaa uauuugggca ugcaguugga agacaagcug auuacaccac aaaaaguuca | 720 |
| gcuuaggaga gacgccu                                               | 737 |

&lt;210&gt; 2

&lt;211&gt; 857

&lt;212&gt; RNA

&lt;213&gt; HERV. E4. 1

&lt;400&gt; 2

|                                                                   |     |
|-------------------------------------------------------------------|-----|
| ggccaagaga uaagaccagc ucuggcuaag cgguggccaa gagugugggc ggaagacaac | 60  |
| ccuccagggg uggcagucua ccaagcccc gugcuuauag aaguuaagcc ugggguccag  | 120 |
| ccgguuaggc aaaaacagua cccgguccuc agagaagcuc uugaagguau ccagguccau | 180 |
| cucaagugcc uaagaaccuu uagaauuaua guuccuuguc agucuccaug gaacacuccc | 240 |
| cuccugccug uuccaagcc ugggaccaag gacuacaggc cgguacagga uuugcgcuug  | 300 |
| guuaaucagg cuacagugac uuuaucua acaguaccua accuguacac auugcugggg   | 360 |
| uugcugccag cugaggacag cugguucacc ugcuuggacc ugaaagaugc cuucuuuagc | 420 |
| aucagauuag cccugagag acagaagcug uuugccuuuc agugggaaga uccagaguca  | 480 |
| ggugucacua cucaauacac ugggaccag ccuccccaaa gguucaagaa cuccccacc   | 540 |
| aucuuugggg aggcguugc ucgagaccuc cagaaguuc ccaccagaga ccuaggcugc   | 600 |
| guguugcucc aguacguuga ugaccuuuug cugggacacc ccacggcagu cgggugcgcc | 660 |
| aagggaacag augcucuacu ccggcaccug gaggacugug gguauaaggu guccaagaaa | 720 |
| aaaagcucag aucugccgac agcagguaug uuacuuggga uuacuaucc aacaggggga  | 780 |
| gcacagccug ggaucagaaa gaaagcaggu cauuguaau cuaccggagc cuaagaccag  | 840 |

aaggcaggug agagaau 857

<210> 3

<211> 62

<212> DNA

<213> 人工

<400> 3

taatacgact cactataggg agatctagca gcataggggg ggctgtgcag aggagatcat 60

cc 62

<210> 4

<211> 62

<212> DNA

<213> HERV-E4

<400> 4

taatacgact cactataggg agatctagca gcatacagca aaaggtcatc aacgtactgg 60

ag 62

<210> 5

<211> 22

<212> DNA

<213> HML. 4

<400> 5

---

|                                                                   |     |
|-------------------------------------------------------------------|-----|
| caacctgtca gggatcagtt tc                                          | 22  |
| <210> 6                                                           |     |
| <211> 22                                                          |     |
| <212> DNA                                                         |     |
| <213> HERV. E4. 1                                                 |     |
| <400> 6                                                           |     |
| ctcgagacct ccagaagttt cc                                          | 22  |
| <210> 7                                                           |     |
| <211> 828                                                         |     |
| <212> RNA                                                         |     |
| <213> HML-2                                                       |     |
| <400> 7                                                           |     |
| cacuguagag ccuccuaaac ccauaccuu aacuuggaaa acagaaaaac uggugugggu  | 60  |
| aaaucagugg ccacuaccaa aacaaaaacu ggaggcuua cauuuuuag caaaugaaca   | 120 |
| guuagaaaag ggucuuuug agccuucuu cucgccuugg aaucuccug uguuuuaau     | 180 |
| ucagaagaaa ucaggcaau ggcauuguu aacugacuua agggucgua acgccguaau    | 240 |
| ucaaccaug gggccucucc aaccggguu gccucuccg gccaugaucc caaaagauug    | 300 |
| gccuuuagu auaauugauc uaaaggauug cuuuuuuacc auccucug cggagcagga    | 360 |
| uugcgaaaa uuugccuuu cuauaccagc cauaaauau aaagaaccag ccaccagguu    | 420 |
| ucaguggaaa uuguuaccuc agggaaugcu uaauagucca acuuuuuguc agacuuuugu | 480 |
| aggucgagcu cuucaaccag uuagagaca guuuucagac uguuauauu uucauuauu    | 540 |

---

|                                                                  |     |
|------------------------------------------------------------------|-----|
| ugaugauuu uuaugugcua cagaaacgag agauaaaaua auugacuguu auacauuucu | 600 |
| gcaagcagag guugccaug caggacuggc aaugcaucu gauaagauc aaaccucuac   | 660 |
| uccuuuucu uuuuagga ugcagauaga aaauagaaa auuaagccac aaaaauaga     | 720 |
| aaauagaaa gacacauuaa aaacacuaaa ugauuuuca aaauugcugg gagauuuua   | 780 |
| uuggauucgg ccaacucuag gcauuccuac uuaugccaug ucaaaaa              | 828 |

&lt;210&gt; 8

&lt;211&gt; 870

&lt;212&gt; RNA

&lt;213&gt; HML-5

&lt;400&gt; 8

|                                                                   |     |
|-------------------------------------------------------------------|-----|
| uauccaaaa gguagacca guggccuuu aaggagaga aaugcaaag agcccaugag      | 60  |
| uuuguugaag agcauuuaa agccagccau auagaaccu caaaaagacc uuggaaauca   | 120 |
| cccauuuua uauucccaa aaagucuggu aaauagagac uuugcauga cuaacaugcu    | 180 |
| aucaaugcua auuugaaacc uaugggacc cuugagcagg uguucuccuc ccuucagug   | 240 |
| auuccuugag auuggcuuu aauguuuuu ggcuuaaaau acugcauuug uacuauuccc   | 300 |
| uuugcagaac aggcagaga aaaaauugug uuuaauaac cagcuauaaa uaacgaaagg   | 360 |
| ccagcucucc gauuucacug gaaagugcuu ccugaaggga ugcuaaacag uccuaccaug | 420 |
| ugucaguauc auguaauca agcuuugcuc cucaguaga aagaauuucc uauugcaag    | 480 |
| acuauucauu uuauggauga uuuuuuacua gcagcccaa uggauccagc acuuuaagu   | 540 |
| uuaugugccu cuguuguaaa gaauacacag uuaagagguu uaucauagc accugaaaa   | 600 |
| guacagaugu ccucuccuug gaaauaucu ggguacauac uacuucgug gucaguaaga   | 660 |
| ccucaaagg uaaaauuua uacuagcaac uuacacaccu uaaugauua ucagaaaua     | 720 |

cuaggcaaua uuaacuggcu uugccccacc uugggcatau cuacugauaa guuacagaac 780

cuguuuucua ucuuaaaaga caaugcugcu cuagacucuc ccagguauuu aacuccugca 840

gcacaaaggg aaauugagga aguagagcac 870

<210> 9

<211> 866

<212> RNA

<213> HERV-H

<400> 9

cacccuuacc cugcucaaug ccaauauccc aucccacagc augcuuuaaa aggauuaaag 60

ccuguuauca cugccugcu acagcauggg cuucuagaac cuauaaacuc ucuuuucaau 120

ucucccauuu uaccugucca aaaaccggau aagucuuaa gacuaguca gaucugcg 180

cuuaucaacc aaauuguuuu gccuauccac ccuguggugc ccaaccugua cacucuuuug 240

uccucaauac cuuccuccac aacucacua uccgugcuug aucuuaaaga ugguuuuuuc 300

acuauucucc ugcacuccuc gucccagccu cucuuugcu ucaccuggac ugaccugac 360

acccaucagu cccagcagcu uaccuggacu gugcugccgc aagguuucag ggacagcccu 420

cguuacuua gccaaagcucu uucucaugau cuacuuccu uccaccccuc cgcuuucac 480

cuuauucaau auauugauga gcuucuucu uguagcccu ccuuugauc uucucaauaa 540

gacacacuuc ugcuccuua gcauuuuuuc uccaaaggau aucggguauc cccuccaaa 600

gcuaaaauu cuucucca cguuaccuac cucggcauaa uucuucuaa gaacacag 660

gcucuccug ccgaccgugu cugacuauc ucucaagccc caaccccuuc uacaaaaca 720

caacuccuuu ccuuccuggg cgugguugga uacuucgcc uuuggauacc ugguuuugcc 780

---

auccuaacaa aaccauuuaa uaaacucaca aaaggaaacu uagcugaccc cauagaucuu 840

aaauccuuuc cccacuccuc uuucug 866

<210> 10

<211> 773

<212> RNA

<213> ERV-9

<400> 10

cugaagcucu uaaaggauua caggauauug uaaaacauuu aaaagcucaa ggcuuaguaa 60

ggaaaugcag cagucccugc aacaccccaa uucuaggagu acaaaaacca aauggucacu 120

ggagacuagu gcaagaucuu agacucauca augaggcagu aaauccucua uauccaguug 180

uacccaaccc cuauaccug cuuucucaaa uaccagagga agcagaauagg uucaugguuc 240

uggaccucaa ggaugccuuc uucuguuucc ccugcacucu gacucccagu uucuguuugc 300

cuuugaggau cccacagacc acacguccca acuuacaugg auggucuugc cacaaggguu 360

uagggauagc ccucaccugu uuggucaggc acuggcccaa gaucuaggcc acuucucaag 420

uccaggcacu uugguccuuc aguaugugga ugauuuacuu uuggcuacca guucagaagc 480

cucaugccag caggcuacuc uagaucucuu gaacuuucua ccuaaucaag gguacaaggc 540

aucuagguca aaggugcagc uuugcuuaca gcaggcuaaa uaucuaggcc uaauuuagc 600

cagagggacc agggcccuca gcaaggaaug aaucagccu auacuggcuu auccuugccc 660

uaagacauua aaacaguugc agggguuccu uggaauccacc ggcuuuugcc gacuaggau 720

cccuggauac agugagauag ucaggccccc ccauacucua aucaaggaga ccc 773

<210> 11

<211> 815

<212> RNA

<213> HERV-W

<400> 11

|                                                                   |     |
|-------------------------------------------------------------------|-----|
| gacccaaggc ccaacaagga cuccaaaaga uuguuaagga ccuaaaagcc caaggccuag | 60  |
| uaaaaccaug caguaacccc ugcaguacuc caauuuuagg aguacagaaa cccaacagac | 120 |
| aguggagguu agugcaagau cucaggauua ucaaugaggc uguuguuccu cuauagccag | 180 |
| cuguaccuag ccuuauacu cugcuuuccc aaauaccaga ggaagcagag ugguuuacag  | 240 |
| uccuggaccu ucaggauGCC uucuucugca ucccuguaca uccugacucu caauucuugu | 300 |
| uugccuuuga agauacuua aaccaacau cucaacucac cuggacuuu uacccaagg     | 360 |
| uucagggaua gucccaucu auuugccagg cauugccca agacuugagc caauccuau    | 420 |
| accuggacac uugucuucgg uagguggaug auuuacuuu ggccgcccau ucagaaaccu  | 480 |
| ugugccauca agccacccaa ggcucucuua auuuccucg uaccuguggc uacaugguuu  | 540 |
| ccaaaccaa ggcucaacuc ugcucacagc agguuacuua gggcuaaaau uauccaaagg  | 600 |
| caccagggcc cucagugagg aacacaucca gccuauacug gcuuauccuc aucccaaac  | 660 |
| ccuaagcaa cuaaggggau uccuuggcgu aaauagguu ugcgcgaaa ggauucccag    | 720 |
| guauggcgaa auagccaggu cauuaaauc acuaauuaag gaaacucaga aagccaauac  | 780 |
| ccauuuagua agauggacaa cugaaguaga agugg                            | 815 |

<210> 12

<211> 856

<212> RNA

<213> HERV-R

## &lt;400&gt; 12

|                                                                   |     |
|-------------------------------------------------------------------|-----|
| aucauaaucc aaugccaguc acccuggaac acuccacuuu ugcugguaca aaaaccuug  | 60  |
| ccuggaccag gaucugauga guauauacug gugcagggcu ugcaugcugu aaaccaagcc | 120 |
| acggugacca uacauccagu aguaccaaac cuguauacuu uaaugggacu uauucuagca | 180 |
| agugccaccu gguuuacagu ccuggacuua aaggauvcuu uccucugucu cuaccuggca | 240 |
| ccaguuaguc agcccaucuu ugcauuuuua ugggacaauu cagucacagg cacaggggga | 300 |
| cagcucgccu ggacuagucu cccacaaggg uucaagaauu cuccacaa cuuuggggaa   | 360 |
| gcacuggccu cagaccucaa ggcauacacc ccacaaaug acaacugcg cuuguugcag   | 420 |
| uacauagaca acccuucuuu uggcagcccc aacccaaga ggacuguaac ugggggaacc  | 480 |
| caggaccucc uccaucucu augggaaagc aggguauga guaucaaaa agaaggcca     | 540 |
| aaauugccau gaaaagguua aaauuuagg cuucauagua agccaagggg aacgcugcu   | 600 |
| uggccaugga ugaaagcaag ccauuuguc acuuccaacu ccaaccaccu ggcgcaau    | 660 |
| aggggaauuc uuaggggcag caggguucug ccuaucugg aucccaauu ucucacuau    | 720 |
| agccagggcc uuauaugaag ccacaaggga gggugaaaag gaacccucc ucugaaaggc  | 780 |
| ugaccagaag aagguguua aacaaauca agaagcucua acucaggcuc cagccuag     | 840 |
| acugccagau acuacu                                                 | 856 |

## &lt;210&gt; 13

## &lt;211&gt; 819

## &lt;212&gt; RNA

## &lt;213&gt; HERV-L

## &lt;400&gt; 13

---

|                                                                     |     |
|---------------------------------------------------------------------|-----|
| uuauggagu uauggcucag gucugacuua cagcaggucc aguggguccc uggacucauc    | 60  |
| cuguguucau uuucccagug ccagaaugca uauuuggcau ugucauacuu agaggcuggc   | 120 |
| agaacccccca cauugauucu gugacuggua aggugagggc uauuauuggug ggaaaggcca | 180 |
| aauggaagcc auuggagcug ccuuuaccua ggaaaauagu aaaucaaaaa cauuaucacc   | 240 |
| accugggagg gauugcagag auuagugcca ccaucaagga cuugaaaaau gcaggggugg   | 300 |
| cgauucccau aacauccuug uucaacucuc cuuuuuggcc ugugcagaag acagauggau   | 360 |
| cuuggagaau gagaguggau uaucuaaagc uuaaccaagu ggugacucca auugcagcug   | 420 |
| cuauacaaga ugugguuuuc auugcucaag caaaauaaua caucuccugg uaccuuguau   | 480 |
| gcagccauug acuuggcaaa ugGCCUUUUA ccCAUuccau aagccccacc agaagcgauu   | 540 |
| ugccuucagc uggcaaggcc agcaauguau cuuuacuguc cuacuucagg gguauaucaa   | 600 |
| cucuccggcu uugugucaua aucuuauuca gagugaucuu gaucacuuuu cacugccaca   | 660 |
| agauaucaca cugguccauu acauugaugg cauuauguug auuggaucca augagcaaga   | 720 |
| aguagcaaac acacuggacu uauuggugag acauuugcau gccauaggau gggaaauaaa   | 780 |
| uccaaauaaa auucacgcac ccucuaccuc aguaaaauu                          | 819 |

<210> 14

<211> 50

<212> DNA

<213> HML-2

<400> 14

|                                                        |    |
|--------------------------------------------------------|----|
| aattctaata cgactcacta tagggagaca cataaaatat catcaacata | 50 |
|--------------------------------------------------------|----|

<210> 15

<211> 50

<212> DNA

<213> HML4

<400> 15

aattctaata cgactcacta tagggagatg cagaggagat catccatgta 50

<210> 16

<211> 50

<212> DNA

<213> HML-5

<400> 16

aattctaata cgactcacta tagggagact agtaaaatat catccataaa 50

<210> 17

<211> 50

<212> DNA

<213> HERV-H

<400> 17

aattctaata cgactcacta tagggagaaa agtagaaggt catcaatata 50

<210> 18

<211> 51

<212> DNA

<213> ERV-9

<400> 18

aattctaata cgactcacta tagggagaag taaatcatcc acatactgaa g 51

<210> 19

<211> 50

<212> DNA

<213> HERV-W

<400> 19

aattctaata cgactcacta tagggagaaa gtaaatacgc cacgtaccga 50

<210> 20

<211> 50

<212> DNA

<213> HERV E4.1

<400> 20

---

aattctaata cgactcacta tagggagacc agcaaaaggt catcaacgta 50

<210> 21

<211> 50

<212> DNA

<213> HERV-R

<400> 21

aattctaata cgactcacta tagggagacc aaaagaaggt tgtctatgta 50

<210> 22

<211> 50

<212> DNA

<213> HERV-L

<400> 22

aattctaata cgactcacta tagggagatc arcataatgy catcaatgta 50

<210> 23

<211> 24

<212> DNA

<213> HML-2

<400> 23

caaccagtta gagacaagtt ttca 24

<210> 24

<211> 23

<212> DNA

<213> HML-5

<400> 24

ttgctcccca gtagaaaaga att 23

<210> 25

<211> 25

<212> DNA

<213> HERV-H

<400> 25

gacagccccc attacttcag tcaag 25

<210> 26

<211> 22

<212> DNA

<213> ERV-9

<400> 26

ccaagatcta ggccacttct ca 22

<210> 27

<211> 21

<212> DNA

<213> HERV-W

<400> 27

catctatttg gccaggcatt a 21

<210> 28

<211> 20

<212> DNA

<213> HERV-R

<400> 28

agacctcaag gcatacaccc 20

<210> 29

<211> 23

<212> DNA

<213> HER-L

<400> 29

---

|                           |    |
|---------------------------|----|
| gccttgtgtc ataatcttat tcg | 23 |
|---------------------------|----|

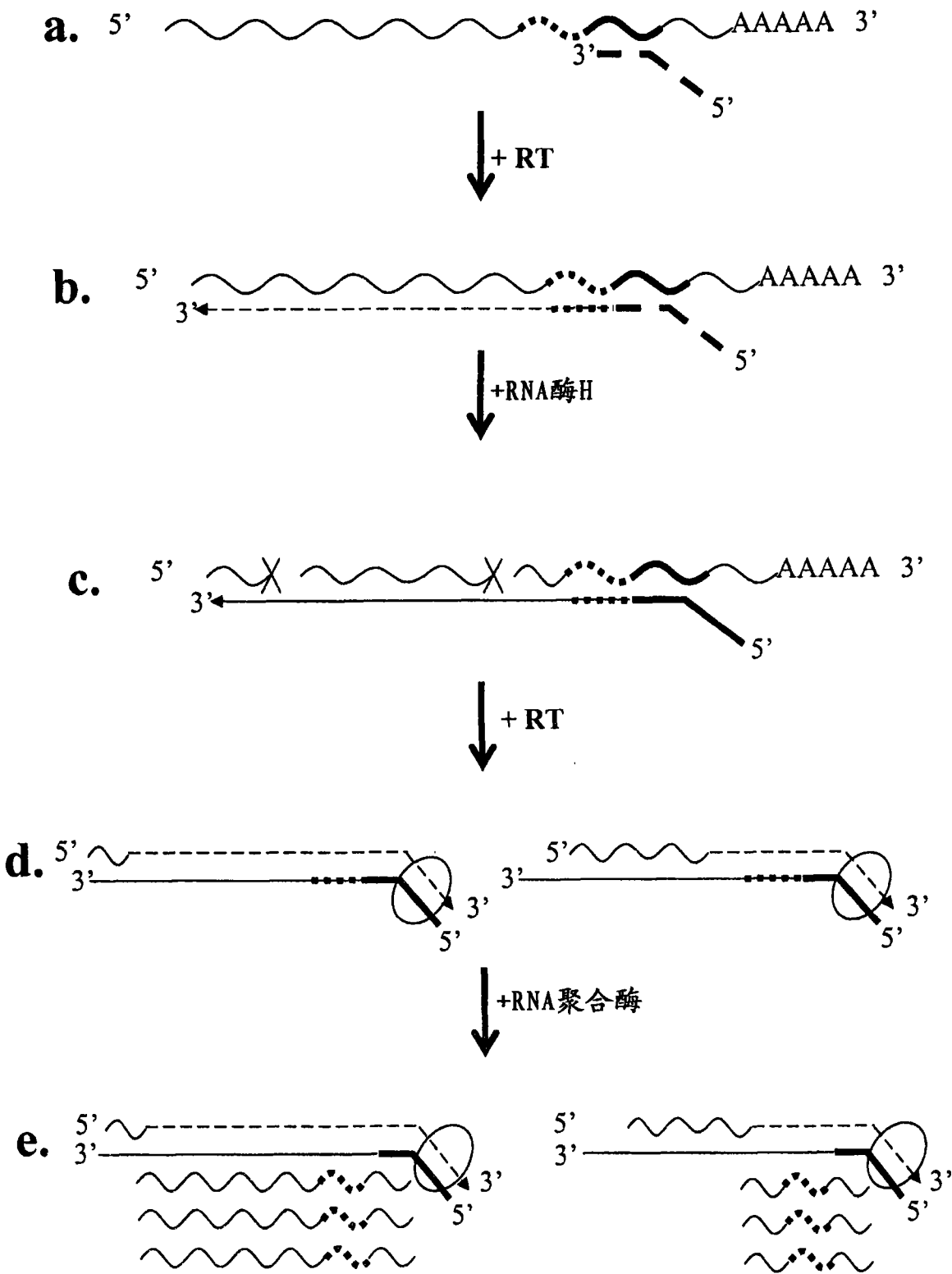


图 1

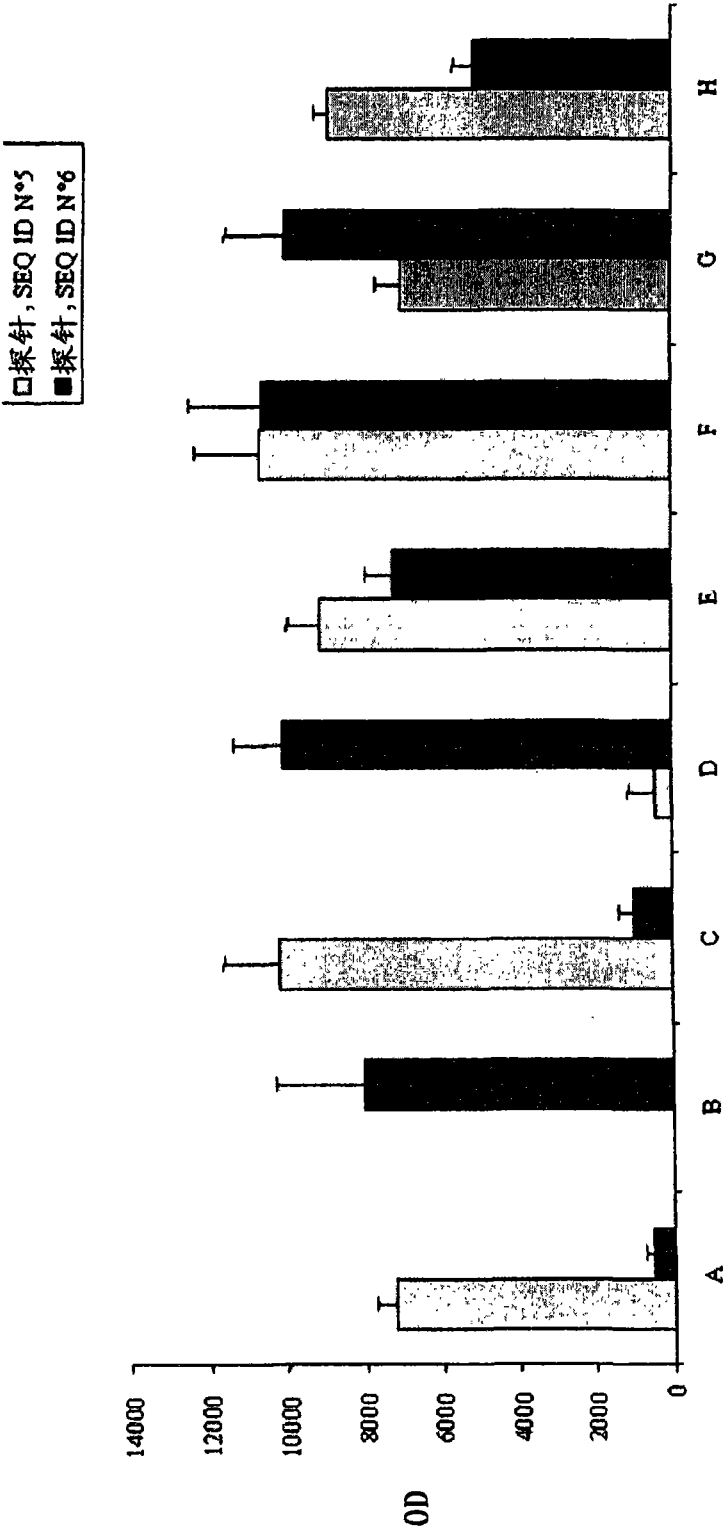


图2

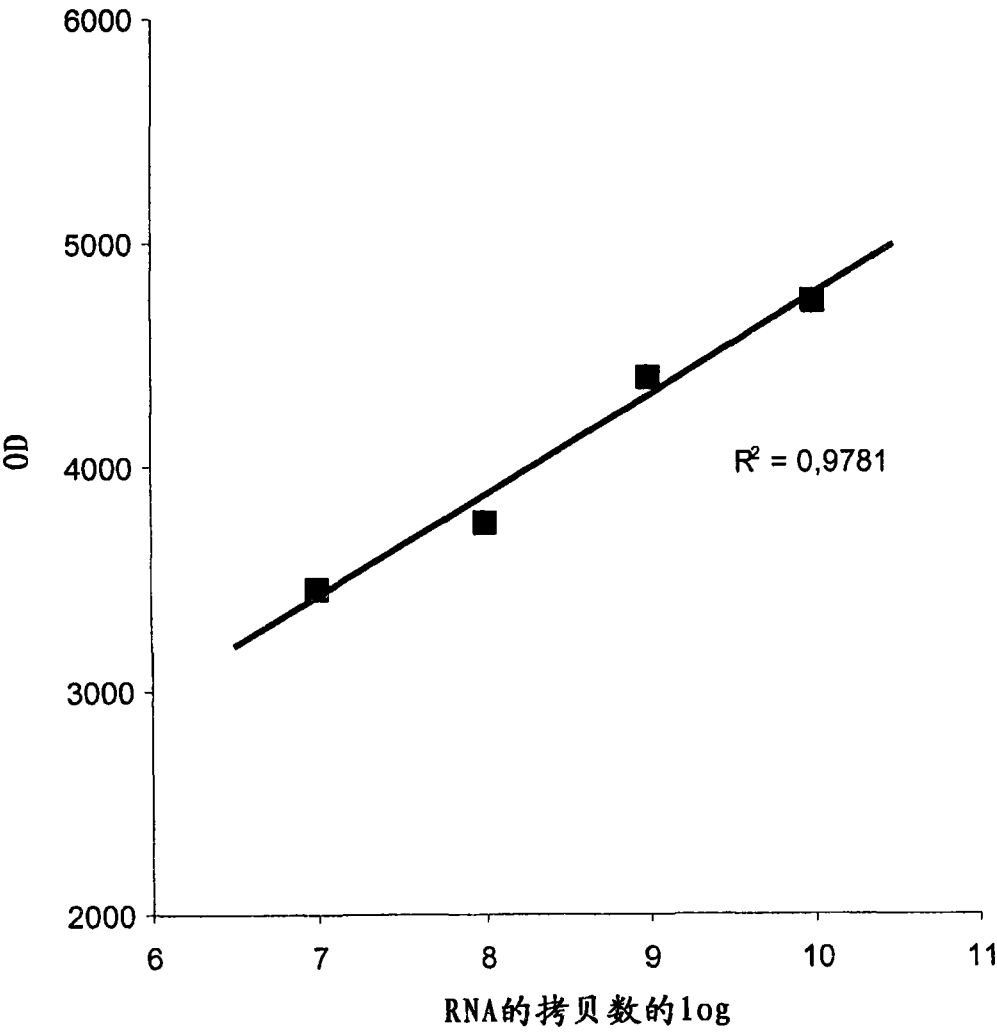


图 3

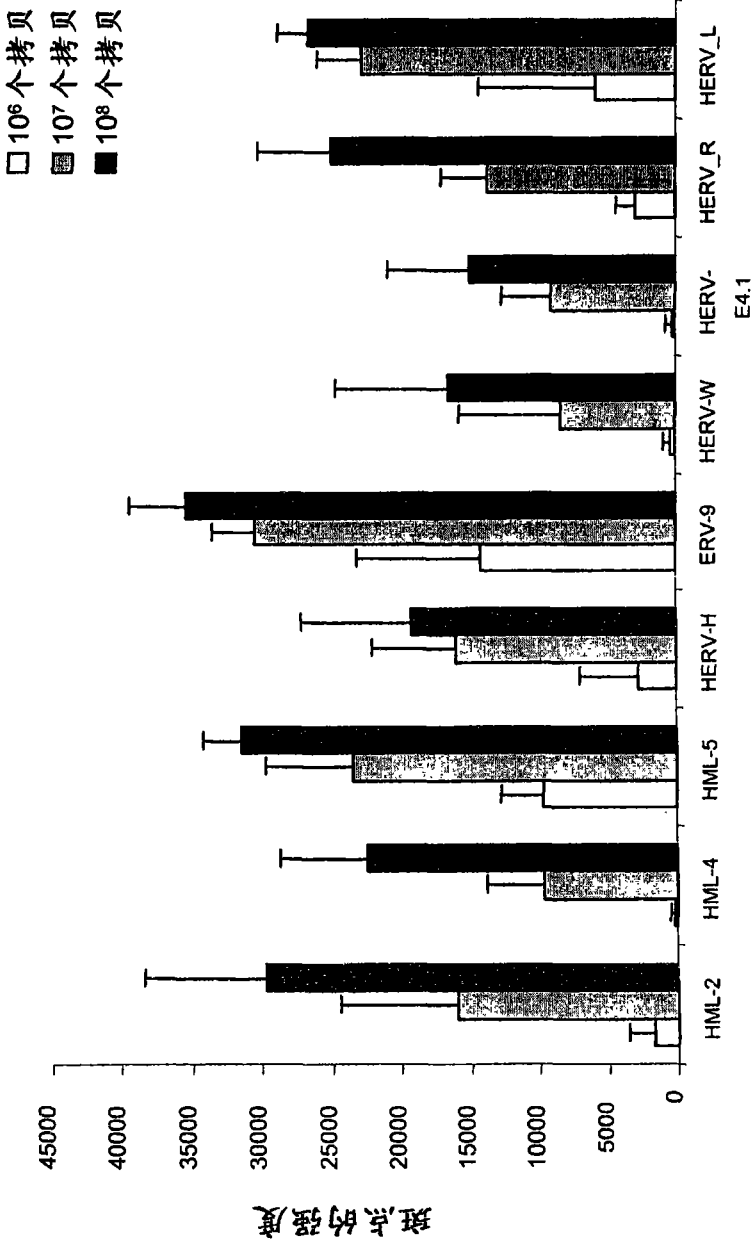


图 4