

[19] 中华人民共和国国家知识产权局

[51] Int. Cl<sup>7</sup>

C12N 15/12

C12Q 1/68



# [12] 发明专利说明书

[21] ZL 专利号 98105514.1

[45] 授权公告日 2004 年 7 月 7 日

[11] 授权公告号 CN 1156572C

[22] 申请日 1998.3.11 [21] 申请号 98105514.1

[30] 优先权

[32] 1997.3.12 [33] DE [31] 19710159.3

[71] 专利权人 赫彻斯特股份公司

地址 联邦德国法兰克福

[72] 发明人 B·基尔施巴姆 W·斯塔尔

I·温克勒

M·梅斯特里恩斯特

审查员 姜 涛

[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

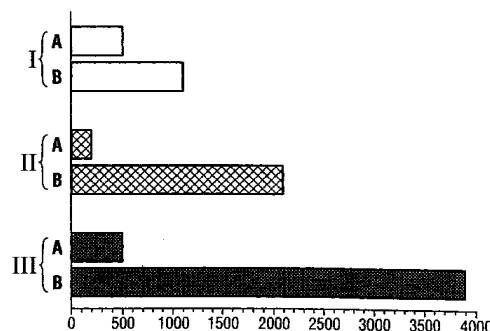
代理人 郭建新

权利要求书 5 页 说明书 28 页 附图 3 页

[54] 发明名称 筛选天然产物和其它化学物质的体外转录方法

[57] 摘要

本发明涉及 DNA 模板的无细胞体外转录方法，该 DNA 模板包括至少受一种基因调节元件控制的待转录的 DNA 序列，该方法包括：a) 提供含有要转录的 DNA 序列的 DNA 模板，其处于至少一种基因调节元件的控制下，b) 为了转录而应用浓缩的细胞核提取物，并加入至少一种标记的核苷酸，c) 加入一种具有潜在药理活性的物质，d) 将标记的转录物与固体载体结合，e) 除去过剩的标记核苷酸，f) 测定标记的转录物量，和 g) 将标记的转录物量和一种对比混合物进行对比。本发明还涉及该方法所用的 DNA 模板及其应用。



ISSN 1008-4274

- 1.一种用于鉴定药理活性物质的方法，包括步骤：
  - a) 提供含有要转录的 DNA 序列的 DNA 模板，其处于至少一种基因调节元件的控制下，
  - b) 为了转录而应用浓缩的细胞核提取物，并加入至少一种标记的核苷酸，
  - c) 加入一种具有潜在药理活性的物质，
  - d) 将标记的转录物与固体载体结合，
  - e) 除去过剩的标记核苷酸，
  - f) 测定标记的转录物量，和
  - g) 将标记的转录物量和一种对比混合物进行对比。
- 2.权利要求 1 的方法，其中步骤 b) 中应用的细胞核提取物是经过纯化的。
- 3.权利要求 1 的方法，其中步骤 b) 中应用的细胞核提取物被细胞因子、辅因子或者其组合所补充或置换。
- 4.权利要求 3 的方法，其中步骤 b) 中应用的细胞核提取物被细胞因子、辅因子或者其组合所部分或全部置换。
- 5.权利要求 1 的方法，其中还包括一个在转录完成后进行的分离反应混合物中的蛋白质的步骤。
- 6.权利要求 1 的方法，其中待转录的 DNA 序列缺乏至少一种选自鸟嘌呤、胞嘧啶和胸腺嘧啶的核碱基，即无 G、无 C 或无 T 序列。
- 7.权利要求 6 的方法，其中无 G、无 C 或无 T 序列的长度多于 400 个核苷酸。
- 8.权利要求 7 的方法，其中无 G、无 C 或无 T 序列的长度为 800 个核苷酸。
- 9.权利要求 1 的方法，其中的基因调节元件包括启动子。
- 10.权利要求 9 的方法，其中的启动子包括“TATA”盒和/或起始区。

11.权利要求 1 的方法，其中的基因调节元件包括增强子。

12.权利要求 1 的方法，其中的基因调节元件包括沉默子。

13.权利要求 1 的方法，其中的基因调节元件是模块启动子。

14.权利要求 13 的方法，其中的模块启动子包括“TATA”盒、起始区和其它蛋白质结合 DNA 序列。

15.权利要求 14 的方法，其中的模块启动子包括：人 T 细胞受体 V $\beta$ 8.1 的“TATA”盒、腺病毒主要晚期启动子的起始区和用于酵母 Gal 4 蛋白的 5 个结合位点和，介于“TATA”盒和 Gal 4 结合位点之间、用于限制内切酶的至少一个单一酶切位点。

16.权利要求 13 的方法，其中模块启动子的一部分或全部被至少一个待研究的基因调节元件置换和/或该模块启动子被至少一个待研究的基因调节元件补充。

17.权利要求 16 的方法，其中基因调节元件和/或待研究的基因调节元件选自下列物质的基因调节元件：粘附分子、生长因子、磷酸二酯酶、磷酸酶、激酶、ATP 酶、膜受体、第二信使受体、激素受体、类固醇受体、金属蛋白酶、NO-合成酶、5-脂肪氧化酶、细胞因子、白介素、T 细胞受体、B 细胞受体、核受体、病毒和逆转录病毒的基因或者上述基因的基因调节元件的一部分。

18.权利要求 1 的方法，其中的 DNA 模板是通用报道质粒。

19.权利要求 18 的方法，其中的 DNA 模板是通用报道质粒 pGS100。

20.权利要求 18 的方法，其中通用报道质粒具有 SEQ ID No.1 所示的序列。

21.权利要求 18 的方法，其中的通用报道质粒业已保藏，保藏号为 DSM 11450。

22.权利要求 1 的方法，其中得自细胞核的浓缩提取物即浓缩核提取物通过至少一个色谱步骤纯化。

23.权利要求 1 的方法，其中的浓缩核提取物在选自下列物质的核-蛋白质-结合材料上纯化：磷酸纤维素、DEAE 纤维素或肝素-琼脂糖 Sepharose 凝胶。

24.权利要求 1 的方法，其中的浓缩核提取物在 P11®柱上纯化。

25.权利要求 1 的方法，其中的浓缩核提取物由至少一种通用的和/或特异性的转录因子补充。

26.权利要求 25 的方法，其中的浓缩核提取物由至少一种选自下列组的通用转录因子补充：TFIIA、TFIIB、TFIID 或 TBP、TFIIE、TFIIF、TFIIH 和 TFII.

27.权利要求 25 的方法，其中的浓缩核提取物由选自下列组的至少一种特异性转录因子补充：原癌基因、激素受体、肿瘤抑制物、病毒病原体、STAT 蛋白、参与第二信使转导级联系统的蛋白质、核受体、基因特异性转录因子、细胞类型特异性转录因子、组织特异性转录因子、依赖于分化地表达的转录因子、发育特异性地表达的和细胞周期特异性地表达的转录因子和上述各种转录因子的部分。

28.权利要求 27 的方法，其中所述转录因子部分以融合蛋白的形式应用。

29.权利要求 1 的方法，其中的浓缩核提取物由至少一种选自下列的辅因子补充：通用辅因子、特异性辅因子、正辅因子、负辅因子、TAFs 和介体。

30.权利要求 1 的方法，其中转录反应混合物中的蛋白质在转录后通过选自下述组中的一个或多个处理步骤而得以消除和/或降解：化学处理，机械处理和酶促处理。

31.权利要求 30 的方法，其中的处理步骤指蛋白酶消化。

32.权利要求 31 的方法，其中将蛋白酶 K 用于蛋白酶消化。

33.权利要求 1 的方法，其中将标记的转录物与带电的固体基质结合。

34.权利要求 33 的方法，其中的基质指 DEAE - 纤维素膜。

35.权利要求 1 的方法，其中至少将一种核苷酸进行放射性标记。

36.权利要求 1 的方法，其中处理步骤 a) ~ d) 的全部或至少部分借助于自动系统进行。

37.权利要求 36 的方法，其中的自动系统指 Biomek 2000®。

- 38.权利要求 1 的方法，其中所述药理活性物质抑制转录。
- 39.权利要求 1 的方法，其中所述药理活性物质激活转录。
- 40.权利要求 1 的方法，其中
- a) 至少有两种转录反应在相同条件下平行进行，此时
  - b) 这些转录反应混合物不同之处仅仅在于它们包括不同量的一种或多种下述组分：待检验活性物质，至少一种转录因子，至少一种辅因子，和浓缩的核提取物，
  - c) 转录后对于每种反应混合物测定生成的标记转录物量，和
  - d) 通过生成的标记转录物量的差异测定待检验的活性物质、转录因子、辅因子和/或浓缩核提取物对基因调节元件的活性和/或特异性。
- 41.权利要求 1 的方法，其中得自细胞核的提取物被浓缩但未进一步纯化。
- 42.根据权利要求 1 的方法，其特征在于所述药理活性物质在规定的条件下抑制受基因调节元件控制的待转录 DNA 序列的转录。
- 43.根据权利要求 1 的方法，其特征在于所述药理活性物质在规定的条件下激活受基因调节元件控制的待转录 DNA 序列的转录。
- 44.根据权利要求 1 的方法，其特征在于所述 DNA 模板具有限制性内切酶 PstI、EcoRI、SacI、KpuI、SacII、BamHI、SwaI 的各个酶切位点、质粒 pUC19 的一部分、5 个酵母 Gal4 蛋白的结合位点，介于 SacII 和 BamHI 酶切位点之间的人类 T-细胞受体 V $\beta$ 8.1 的 TATA-盒、介于 BamHI 和 SwaI 酶切位点之间的腺病毒主要晚期启动子的起始区和 800 碱基对的无 G 序列。
- 45.应用于权利要求 1 所述鉴定药理活性物质的方法中的 DNA 模板，它包括待转录的 DNA 序列，并且处于至少一种基因调节元件的控制之下，其中待转录的 DNA 序列至少缺乏一种选自鸟嘌呤、胞嘧啶或胸腺嘧啶的核碱基，该无 G、无 T 或无 C 序列的长度为多于 400 个核苷酸。
- 46.权利要求 45 的 DNA 模板，其中待转录的无 G、无 T 或无 C 序列长度为 800 个核苷酸。

47.权利要求 45 的 DNA 模板, 其中的基因调节元件包括启动子。

48.权利要求 45 的 DNA 模板, 其中的基因调节元件还包括增强子和/或沉默子。

49.权利要求 45 的 DNA 模板, 其中该 DNA 模板是通用报道质粒 pGS100。

50.权利要求 45 的 DNA 模板, 其中的通用报道质粒具有如 SEQ ID No.1 所示的序列。

51.权利要求 45 的 DNA 模板, 其中的通用报道质粒业已保藏, 保藏号为 DSM 11450。

52.根据权利要求 45 的 DNA 模板, 其特征在于所述 DNA 模板具有限制性内切酶 PstI、EcoRI、SacI、KpuI、SacII、BamHI、SwaI 的各个酶切位点、质粒 pUC19 的一部分、5 个酵母 Gal4 蛋白的结合位点, 介于 SacII 和 BamHI 酶切位点之间的人类 T- 细胞受体 V $\beta$ 8.1 的 TATA- 盒、介于 BamHI 和 SwaI 酶切位点之间的腺病毒主要晚期启动子的起始区和 800 碱基对的无 G 序列。

53.权利要求 45 的 DNA 模板在转录方法中的应用。

54.权利要求 53 中关于 DNA 模板的应用, 其中的转录方法指权利要求 1~44 中任何一项的方法。

55.根据权利要求 53 的应用, 其特征在于所述 DNA 模板具有限制性内切酶 PstI、EcoRI、SacI、KpuI、SacII、BamHI、SwaI 的各个酶切位点、质粒 pUC19 的一部分、5 个酵母 Gal4 蛋白的结合位点, 介于 SacII 和 BamHI 酶切位点之间的人类 T- 细胞受体 V $\beta$ 8.1 的 TATA- 盒、介于 BamHI 和 SwaI 酶切位点之间的腺病毒主要晚期启动子的起始区和 800 碱基对的无 G 序列。

56. 根据权利要求 53 的应用, 其特征在于所述药理活性物质在规定的条件下抑制受基因调节元件控制的待转录 DNA 序列的转录。

57.根据权利要求 53 的应用, 其特征在于所述药理活性物质在规定的条件下激活受基因调节元件控制的待转录 DNA 序列的转录。

## 筛选天然产物和其它化学物质的体外转录方法

### 技术领域

本发明涉及体外分析病毒基因和细胞基因的转录的方法，该方法可实现自动化并且适合于有效而经济的整体筛选以查明对于基因活性具有选择性效应的、特定的重要化学结构。

### 背景技术

当只是合理设计活性物质不能成功地寻找活性物质之后，有关筛选天然产物的生物活性组分的研究日益增多。因此，研究工作不仅集中于化学物质文库和组合文库，而且又集中于作为物质资源的天然产物传统提取物。这主要是因为这些提取物所含物质的多样性。经典分析方法证实，微生物发酵液的提取物中大约含有 500 种化合物，其结构大不相同。考虑到它们的多样性，它们远远超越于化学物质文库和组合物文库。

限制在药物上应用各种各样的、大体上未探明的潜在天然产物的一个因素在于，能测定候选活性物质的合适而有意义的方法数量有限。具体而言，要求所用方法能鉴定使用时引起最小副作用的高度特异性的药理活性物质。

下述方法基于这样的研究：其中测试物质在参与转换遗传信息的第一步即基因转录的调节中的可能性。这种方法旨在鉴定对于转录有直接的或间接的、正的或负的效应的物质。

基因的转录强度由该基因的基因调节元件来确定，具体而言，由启动子、由增强子或由沉默子确定。基因调节元件的作用由转录因子和辅因子介导和转换。这些转录因子对于基因的转录速度可产生负的或正的效应，因而对于转录强度有贡献。同时，业已鉴定，在很多细胞过程中有大量转录因子作为重要的“分子开关”，其中的过程包括信号转导、细胞周期控制、分化和控制的细胞死亡（编程性细胞死亡）。

被细胞接收的大多数信号（它们影响基因的转录强度）由跨膜蛋白质“寄存”，由信号转导链在细胞内递送并由转录因子转换。接收外部

信号的蛋白质的例子有 cAMP-结合蛋白、生长信号用的传感蛋白（例如血清效应因子、SRF）、参与细胞因子（cytokin）表达的激素受体或转录因子、所谓的 STAT 蛋白（信号转导蛋白及转录活化蛋白）。

同时，已知有许多物质对于基因的转录强度具有直接或间接效应。即使这类物质的作用通常不是特异性的，但这类物质特别是在药剂中用作药理活性物质。因而服用这类药剂通常会起不希望有的副作用。

例如，免疫疾病用含有环孢菌素和类固醇衍生物为活性物质的药剂治疗。环孢菌素 A 与亲环素形成复合体。后者抑制钙调磷酸酶，一种遍在的磷酸酶，它通过各种代谢途径使蛋白质去磷酸化。钙调磷酸酶例如调节转录因子 NFAT 的亚单位从胞质溶胶至细胞核的转运(Liu, J.(1993) *Immunology Today* 14, 290-295)。NFAT（活化 T 细胞的核因子）参与某些免疫相关基因的活化。环孢菌素 A（CsA）通过它对 NFAT（活化 T 细胞的核因子）的效应而间接地调节这些基因的表达。然而，由于环孢菌素 A 只是经由遍在的钙调磷酸酶间接地调节 NFAT 活性，所以环孢菌素 A 还通过其它代谢途径起血管收缩剂和肾毒素和神经毒素的作用。如果已知某药理活性物质能特异性地、可能直接地抑制 NFAT，那么含有该活性物质的药物就可能引起更少的副作用。

除了所需效应外还可能带来副作用的药理活性物质也包括糖皮质激素。糖皮质激素多年来被用于下列疾病的标准疗法中：变应性病、风湿病、炎症和由免疫系统反应过度引起的其它疾病。它们尤其能通过刺激形成细胞 NFkB 抑制剂，即 Ikb 蛋白质而抑制细胞类型特异性转录因子 Nfkb 的活化(Scheinmann, R.I., Cogswell, P.C., Lofquist, A.K. & Baldwin Jr., A.S. (1995) *Science* 270, 283-286; Auphan, N., DiDonato, J.A., Rosette, C., Helmborg, A & Karin M.(1995) *Science* 270, 286-290)。Ikb 本身又防止活性 NFkB 二聚体转移入细胞核，从而防止重要的免疫靶基因的活化。类似于有关 CsA 的所述那样，糖皮质激素对于基因表达的效应相对地为非特异性的，因为糖皮质激素不但对 NFkB 起作用，还对其它蛋白质起作用。

这些实例表明，很需要这样的药理活性物质，即它们的作用方式应尽可能是特异性的。为了探寻具有这种性能的新型重要化学结构，必须对于大量物质测试其特异的活性。

虽然具有相同的基因成分，但各个细胞通常只表达特定的蛋白质，这取决于细胞类型和/或一定的疾病或缺损和这些细胞发育和分化的各别的程度。据认为细胞的这种特性的基础是基因调节蛋白特定的所有组成成分，例如细胞类型特异性的和发育特异性的成分，它提供某些转录因子和辅因子（辅助蛋白），这些因子调节不同基因的协同、受控转录。

因此，在确定类型的细胞中，特定的药理活性物质应能选择性地活化或抑制病理相关基因的转录。为了鉴定这种活性物质，需要这样的转录方法：其中候选活性物质对于各个基因的转录的效应，即对于参与转录调节的蛋白质和对于基因调节元件的效应，可在规定的条件下直接测定。由于必须测定多种候选物质，所以其它前提是：该方法操作简便并且它能实现自动化。

Weil 等人首次描述了无细胞转录方法（Weil, P.A., Luse, D.S., Segall, J., Roeder, R.G.(1979) Cell 18, 469-484）。在该方法中，将细胞核的浓缩提取物（称为 S100 提取物）（Weil, P.A., Segall, J., Harris, B.Ng, S.Y., Roeder, R.G.(1979) J. Biol. Chem. 254, 6163-6173），和纯化的 RNA 聚合酶 II 用于体外转录。如果没有外源的 RNA 聚合酶 II，则这些浓缩但未进一步纯化的细胞核提取物不能转录（Weil, P.A., Luse, D.S., Segall, J., Roeder, R.G.(1979) Cell 18, 469-484; Dignam, J.D., Martin, P.L., Shastry, B.S., Roeder, R.G.(1983) Methods in Enzymology 101, 582-598）。

从这些细胞核提取物开始，接着开发出多种方法，通过这些方法利用数步纯化步骤可分离转录因子。这些方法尤其包括纯化步骤，其中通过色谱法在结合核蛋白质的材料例如磷酸纤维素柱上纯化核提取物。在包括数步的这些复杂方法范围内，Dignam 等首先描述了将可商购的 P11® 系统（Whatman, Maidstone, England）用于纯化步骤之一中（Dignam, J.D., Martin, P.L., Shastry, B.S., Roeder, R.G. (1983) Methods in Enzymology 101, 582-598）。

这些包括数步的纯化方法日臻完善以至现在有可能通过复杂的方法从细胞核提取物中分离各种转录因子。此外，各种因子或者它们的亚单位现在还可以重组体形式获得，例如：TFIIA，TFIIB，TFII $\alpha$ ，

TFIIE $\beta$  和 TFIIF ( Zawel, L.和 Reinberg, D. (1995) *Annu Rev. Biochem.* 64, 533-561 ) .

因而, 现在已存在由重组体和天然纯化的因子的混合物构成的转录系统。然而, 这类转录系统从技术角度看过于复杂且对于样品流通量大的筛选方法来说太昂贵。反之, 在其它转录系统中, 例如那些利用得自细胞核的提取物来代替重组体或纯化的因子的系统, 可发现大量副反应。在未充分地纯化或未纯化的核提取物(粗提取物)中, 主要是核酸和 DNA - 结合蛋白, 例如阻抑蛋白如组蛋白, 对于体外转录产生不利效应。在粗提取物中发现的核酸中, 尤其是编码 t-RNAs 的 DNA 序列具有不利效应。由于 t-RNAs 的基因转录比 mRNAs 的强大约 100 倍, 所以这些 t-RNA-编码序列导致非特异性转录物过量。于是在特异性转录物可被检测之前, 必须通过复杂的纯化步骤除去非特异性转录物。

为了定量分析体外转录的结果, 开发了这样的载体: 它们待转录的 DNA 序列缺乏鸟嘌呤碱基(称为无 G 序列或无 G 盒), 适当时候可以对无 G 序列接上含有大量鸟嘌呤的序列区段。应用这类载体可使转录作用在不存在 GTP 下进行。于是, 只有无 G 序列而没有其它含 G 序列被转录。这就产生特异性转录物, 而且它们的长度(基本上)一致。Sawadogo 和 Roeder 最早描述了将载体用于这种转录作用即, 此时长度为 400 个核苷酸的序列受 ML (腺病毒主要晚期) 启动子的控制。该载体产生长度约为 400 个核苷酸的转录物 (Sawadogo, M.和 Roeder, R.G. (1985) *Proc. Natl. Acad. Sci. USA* 82, 4394-4398)。借助于这些载体得到显著地更少量的非特异性转录物, 这就是为什么在转录反应中应用这些载体已被描述过许多次。(Goppelt, A., Stelzer, G., Lottspeich, F., Meisterernst, M.(1996) *EMBO J.* 15, 3105-3115; Kretzschmar, M., Kaiser, K., Lottspeich, F., Meisterernst, M. (1994) *Cell* 78, 525-534; Meisterernst, M., Roy, A.L., Lieu, H.M.和 Roeder, R.G. (1991) *Cell* 66, 981-993)。然而, 迄今应用的载体都是这样的, 即无 G 序列长度不多于 400 个核苷酸。

为了定量和定性分析前述转录方法的结果, 在有放射性标记的核苷酸存在时进行转录作用, 先使放射性标记的转录物酚化并沉淀, 然后在凝胶上分离。这不但从特异性转录物中除去错误地引发或错误地终止的

转录物和非特异性地标记的核酸（例如由质粒引起的转录物或 tRNAs），还从特异性转录物中除去过多的核苷酸。过量的放射性标记的核苷酸与放射性标记的转录物的活性比在不适宜条件下约为 10,000:1，于是标记的转录物必须浓缩大约 10,000 倍。特异性转录物的浓缩通过沉淀步骤和电泳分离达到。然而，这些浓缩步骤不适于自动化整体筛选，这就是为什么必须开发其它方法以这样的程度除去标记的核苷酸，即仍有可能定量分析转录结果。

也可通过对于膜例如 DEAE-纤维素膜应用反应溶液来监测转录作用。可在该膜上直接检测放射性标记的转录物。然而迄今只是成功地利用膜来检测得自体外转录的转录物，这里所说的体外转录是在纯化的 RNA 聚合酶 II 的存在下（Roeder, R.G.(1974) J. Biol. Chem. 249, 241-248）或者在纯化的基础转录因子的存在下（Ohkuma, Y., Sumimoto, H., Horikoshi, M., Roeder, R.G. (1990) Proc. Natl. Acad. Sci. USA 87, 9163-9167）进行的。没有任何迹象表明，可以按这种方法检测借助于从细胞核浓缩的和，合适的话，预先纯化的提取物而获得的转录物。

在 WO96/26959 中业已描述了利用基因转录来筛选活性物质。该文献公开了人 NFATs (hNFAT) 的序列及其在转录分析中的潜在应用，这些分析方法转过来又可用于有可能自动化的整体筛选天然产物。与下述转录方法相反，该转录分析纯粹是一种结合分析法，其中未进行转录反应。

US 5,563,036 描述了另一种结合分析法，它可用于筛选能抑制转录因子与核酸的结合的物质。该分析法中也未进行转录作用。

US 5,563,039 描述了结合分析法的又一个实例，该方法也旨在用于寻找能抑制该场合中的蛋白质与一定的 DNA 序列结合的物质，该蛋白质与肿瘤坏死因子受体 (TRADD) 相关。

#### 发明内容

本发明的一个目的是提供一种分析方法，用于分析例如细胞基因和病毒基因的基因在规定的反应条件下的转录作用，该方法操作简便，可重复且尤其可广泛用于整体筛选。

#### 具体实施方式

本发明提供了一种用于鉴定药理活性物质的方法，包括步骤：

- i) 提供含有要转录的 DNA 序列的 DNA 模板，其处于至少一种基因调节元件的控制下，
- ii) 为了转录而应用浓缩的细胞核提取物，并加入至少一种标记的核苷酸，
- iii) 加入一种具有潜在药理活性的物质，
- iv) 将标记的转录物与固体载体结合，

- v) 除去过剩的标记核苷酸,
- vi) 测定标记的转录物量, 和
- vii) 将标记的转录物量和一种对比混合物进行对比.

本发明涉及 DNA 模板的无细胞体外转录方法, 该 DNA 模板含有待转录的 DNA 序列, 该序列受一个或多个基因调节元件的控制, 且此时

a) 为了转录而应用一种浓缩的和合适的话纯化的细胞核提取物, 合适的话该提取物可被转录因子和/或辅因子补充或者部分或全部置换, 和至少一种标记的核苷酸,

- b) 合适的话, 在转录之后分离和/或降解反应混合物中的蛋白质,
- c) 将标记的转录物与固体基质结合,
- d) 除去过剩的标记核苷酸和
- e) 测定标记的转录物量.

本方法包括: 实际的转录反应 (a), 分离特异性转录物 (b,c,d) 和检测特异性转录物 (e)。该方法包括分离特异性转录物的具体实施方案, 上述 b、c 和 d 的顺序只是一种可能性。适当时, 分离特异性转录物的顺序也可以是 c、d、b 或者 d、b、c。此外, 本发明的具体实施方案可以省去个别分离步骤。例如, 本转录方法可以只包括步骤 a、c、d 和 e 或者只有步骤 a、c 和 e 或者只有步骤 a、b、c 和 e 或者只有步骤 a、b、d 和 e 或者只有 a、b 和 e 或者只有 a、d 和 e。

本方法的一个独特特征在于, 所有操作步骤、即实际的转录反应 (转录作用) 以及特异性转录物的分离和检测, 可实现自动化, 使得简便而可靠地测定在特定的反应条件下获得的特异性转录物的量, 因而能测定转录速率。

转录速率表示: 单位时间某特定基因被转录的频率, 或者在所述转录方法中, 单位时间待转录的 DNA 序列被转录的频率。若要测定转录速率, 就要测定在规定的单位时间之后所得放射性标记的转录物量。

本方法的一个方面在于, 该转录作用是在激活剂和/或抑制剂存在下进行的, 即在对于转录作用产生正的或负的效应的组分存在下进行。例如, 能被转录的细胞核提取物可用于基础转录。该基础转录系统可由激活剂和/或抑制剂补充。与基础转录相比, 转录抑制导致转录速率的降低, 于是单位时间特异性转录物的量更小; 而转录激活导致转录速率的升高, 于是单位时间特异性转录物的量更大。

基因的转录可分成数步 - 前起始复合体 (PIC) 的形成、PIC 激活、引发、启动子清除、增长和终止。在真核生物中, 引发转录作用需要 RNA 聚合酶 (至于蛋白质 - 编码基因的转录, 应用 RNA 聚合酶 II)

和 DNA - 结合蛋白, 该蛋白可使 RNA 聚合酶 II 与 DNA 进行特异性的相互作用。这些 DNA - 结合蛋白被称为转录因子, 通用转录因子基本上参与同启动子的相互作用, 而特异性的转录因子介导位于启动子下游或上游的基因调节元件的作用。

通用转录因子 TFIIA、TFIIB、TFIID、TFIIE、TFIIF 和 TFIIH 在真核基因的转录中起作用。视启动子而定, 负责基因较低而基本的活性的、可被转录的蛋白质级分含有全部或者大多数这些通用转录因子和 RNA 聚合酶 II (RNA Pol II)。ML 启动子的基本活性例如可通过 TBP (TFIID 的 TATA - 结合亚单位)、TFIIB、TFIIE、TFIIF、TFIIH 和 RNA Pol II 达到。带来这一基本活性的蛋白质级分被称为基础转录系统。为了本发明的目的, 该术语也用于浓缩的和适当的话, 纯化的细胞核提取物。借助于基础转录系统进行的转录作用被称为基础转录。在体外进行无细胞转录作用需要至少一种基础转录系统, 核苷酸和待转录的 DNA 模板。

除了通用转录因子, 除了基础转录系统之外, 活化转录还需要特异性转录因子和辅因子(辅助蛋白)(Kaiser, K., Stelzer, G 和 Meisterernst, M. (1995)EMBO J.14, 3520-2527)。特异性转录因子能成倍地增加特异性基因的很低的基础转录强度, 能调控转录引发的频率。所以, DNA - 结合蛋白高度决定基因被转录的频率(转录速率)。不直接与 DNA 结合但能通过蛋白质 - 蛋白质相互作用影响转录因子或 RNA 聚合酶 II 的活性的其它蛋白质, 例如辅因子, 也在该调节过程中起使用。

本方法的一方面在于, 将得自细胞核的浓缩提取物(得自细胞的核提取物, 核提取物)用于基础转录。考虑到这个特性, 本方法可以通用; 它适合所用的细胞, 也适用于所用的真核种类。

例如, 可以从得自人或动物细胞的细胞系获取浓缩的核提取物。能在发酵罐内大规模地生长和繁殖的那些细胞是特别合适的, 例如 HeLa 细胞。此外, 还可应用得自选择的细胞种类的细胞核提取物, 特别是那些例如可通过下列特异性补充区别的细胞: 用转录因子和/或辅因子对它们进行的细胞种类特异性、细胞周期特异性、发育特异性、分化特异性或疾病特异性补充。特别是可应用对于疾病起源产生关键性作用的

细胞种类，例如，免疫系统的细胞（如 B 细胞和 T 细胞）。

本方法的独特优点在于，还可以分离和应用来自组织或肿瘤细胞的核提取物。这特别适用于得不到合适的细胞系时的场合。具体而言，可从容易得到的组织分离核提取物并用于本方法中，其中的组织例如有：动物或人的脐带、动物或人的移植物废品、动物或人的活检材料或者动物或人的肿瘤组织（例如在外科手术期间取出的组织）或者动物或人的胎盘。

本方法的一方面在于，为了从新鲜的或冷冻的细胞的细胞核或者从新鲜的或冷冻的细胞核浓缩蛋白质而通过已知方法制备浓缩的核提取物，例如采用 Dignam 等人描述的方法（Dignam, J.D., Martin, P.L., Shastry, B.S., Roeder, R.G. (1983) *Methods in Enzymology* 101, 582-598; Dignam, J.D., Lebovitz, R.M., Roeder, R.G. (1983) *Nucleic Acid Res.* 11, 1475-1489）。本方法一个特定的实施方案在于，为了制备浓缩的核提取物，所用方法包括细胞核的匀浆，接着透析该匀浆物。

本方法一个重要实施方案的一方面在于，通过一个或多个纯化步骤特别是简便的纯化步骤来纯化得自细胞核的已浓缩提取物，纯化至可被转录的程度，即当实施本方法时可得到特异性转录物。例如，提取物可由色谱法纯化。纯化操作例如可在核-蛋白质结合材料如磷酸纤维素、DEAE-纤维素或肝素-Sephrose 上进行。也可将阳离子型和/或阴离子型交换柱或特异性亲和柱例如那些抗体或寡核苷酸与柱材料结合的柱用于纯化。

本方法一个特定实施方案的一方面在于，浓缩的核提取物在磷酸纤维素柱、特别是在 P11® 柱（P11® 系统，Whatman, Maidstone, England）上纯化。本方法又一特定实施方案的一方面在于，浓缩的核提取物只通过一步纯化，例如在单一的 P11® 柱上，或者在用含有 DEAE-纤维素或肝素-Sephrose 的柱材料的单一柱上。

在 P11® 纯化的特定实施方案中，在除了其它组分外还包括 0.05 ~ 0.15M 优选为 0.1M KCl 的缓冲液的存在下，先将核提取物与磷酸纤维素结合。用适当的缓冲液，优选包括 0.05 ~ 0.15M KCl、优选为 0.1M KCl 的缓冲液洗涤荷载的柱，将非特异性和干扰组分从柱中洗出。能转录的

组分优选呈 2 个级分从柱上洗脱，首先应用例如包含 0.4 ~ 0.6M KCl、优选为 0.5M KCl 的缓冲剂，接着应用例如含有 0.7 ~ 1M KCl、优选为 0.85M KCl 的缓冲剂。

本方法的一个特定实施方案在于，应用浓缩的、适当的话纯化的核提取物在外源 RNA 聚合酶 II 的存在下进行转录。所用的 RNA 聚合酶优选是真核 II 类 RNA 聚合酶（RNA 聚合酶 II），特别是动物或人 RNA 聚合酶 II。

本方法另一个实施方案的一方面在于，浓缩的和适当的话纯化的核提取物通过加入蛋白质例如转录因子和/或辅因子（辅助蛋白）而被补充或者全部地或部分地置换。例如，这些蛋白质可从细胞核分离出，或者它们可由重组技术制备。

本方法的一个特定实施方案在于，核提取物只由转录因子和/或辅因子补充。在其它极个别情况下，本方法的一个方面在于，能转录的蛋白质级分（基础转录系统）只包括业已分离的或者业已由重组技术制备的转录因子和/或辅因子，和 RNA 聚合酶。

可应用的转录因子例如有它的部分的通用和/或特异性的转录因子，合适的话呈融合蛋白的形式。

可应用的通用转录因子例如有：TFIIA、TFIIB、TFIID、TFIIE、TFIIF、TFIIH、TFIIJ 和 TBP（TATA - 结合蛋白）。

可应用的特异性转录因子例如有：NFkB、AP1、NFAT、GATA3、TCF/Lef、CBF、Tat、fos/jun 族系的因子、Oct 族系（Oct-1、Oct-2）的因子和与它们相互作用的因子例如 Ets 族系的 Bob1、OCA - B 或 OBF，ATF/CREB 蛋白质族系的激活蛋白，核受体例如 PPAR $\alpha$  或者相应的细胞类型特异性的转录因子的同种型（Kel, O.V., Romaschenko, A.G., Kel, A.E., Wingender, E., Kolachenov, N.A., (1995) Nucl. Acids. Res. 20,3-16）。

特异性转录因子的其它实例有：

1. 原癌基因，例如 jun、fos、ets、myc、bcl-同种型和 erb
2. 激素受体，例如（erb）、糖皮质激素受体、雌激素受体、视黄酸受体、维生素 D 受体或者

3.肿瘤抑制剂, 例如 p53、NF1、WT1、RB

4.病毒病原体, 例如: 单纯疱疹病毒的蛋白质如 VP16 或 ICP4, 乳头瘤病毒的蛋白质如 E1、E2、E6 或 E7, 腺病毒的蛋白质如 E1A 或 E2A, 巨细胞病毒的蛋白质如 IE86, 乙肝病毒的蛋白质如 pX, HIV 病毒的蛋白质如 Tat 或 Rev

5.细胞类型特异性因子和/或组织特异性因子例如肌浆蛋白因子 (myogenic factor), Pit-1, Oct-2, Pu-1, OCA - B 或 HNFs 或者 T 细胞特异性因子如 Ets-1、GATA3、TCF/Lef、CBF

6.STAT 蛋白质 (信号转导蛋白和转录的激活蛋白), 例如细胞因子激活的转录因子如: IL - 1 Stat, IL-2 Stat, IL-3 Stat, IL-4, IL-5 Stat, IL-6 Stat, IL-7 Stat, IL-8 Stat, IL-9 Stat, IL-10 Stat, IL-11 Stat, IL-12 Stat ( " Stat " 表示介导该作用的蛋白质) 或者

7.参与第二信使转导级联系统的蛋白质例如 CREB 或 abl,

8.核受体, 例如第二信使受体 (如 cAMP 或 IP<sub>3</sub> 受体、Ca<sup>2+</sup>依赖性受体), 视黄酸受体, 糖皮质激素受体或类固醇受体,

9.基因特异性激活蛋白或抑制蛋白, 例如 IL-2 基因的特异性激活蛋白如 NFkB、AP1 或 NFAT

10.发育特异性地、细胞周期特异性地和分化依赖性表达的转录因子。

辅因子例如通过蛋白质 - 蛋白质相互作用和/或蛋白质 - DNA 相互作用而在转录中起直接的或间接的作用。有些辅因子已经存在于基础转录系统中, 而其它辅因子则只存在于激活的转录系统中。辅因子可对转录速率产生正的或负的效应。可应用的辅因子例如有:

- TBP 相关性因子 (TAFs), 例如 TAF<sub>30</sub>, TAF<sub>40</sub>, TAF<sub>55</sub>, TAF<sub>60</sub>, TAF<sub>110</sub>, TAF<sub>150</sub>, TAF<sub>250</sub>, (Verrijzer, C.P. 和 Tjian, R.(1996) Trends Biochem. Sci, 21, 338-342; TAFs together with TBP form the TFIID complex, it being possible for the composition of the TAFs in the TFIID to vary considerably);

- 中介体, 即与 RNA 聚合酶 II 相关的辅因子例如 CTD (羧 - 端 - 区) 活性的蛋白质和/或阻抑蛋白和/或 RNA 聚合酶 II 的激活蛋白, 特别

是 RAP30、RAP74、RAP38、SR7 (RNA 聚合酶 B 的抑制因子, SRB), 细胞周期蛋白或激酶 (例如 CKII);

- 通用辅因子;

- 含于 USA (上游刺激活性) 级分中的辅因子 (Kaiser, K. 和 Meisterernst, M. (1996) Trends Biochem. Sci. 342-345);

- 正辅因子, 例如 PC1、PC2、PC4 (p15)、PC5、PC6、Dr2 (D 阻抑蛋白 2)/PC3、ACF (激活辅因子) CofA (辅因子 A)、HMG - 蛋白 (染色质相关性高迁移率族蛋白质);

- 负辅因子, 例如 NC1、NC2 和/或

- 特异性辅因子。

本方法的一方面在于, 用于转录的 DNA 模板包括一个或多个基因调节元件和一个待转录的 DNA 序列。

本发明的主题是 DNA 模板, 它可在上述用于无细胞体外转录方法中。该 DNA 模板包括一个或多个基因调节元件和一个待转录的 DNA 序列。该 DNA 模板另外还可包括其它序列区段。

基因调节元件可以是已知的基因调节元件或待研究的基因调节元件, 或是一个或多个已知基因调节元件和一个或多个待研究基因调节元件的构建物。基因调节元件可包括参与基因调节的任何 DNA 序列或其区段。考虑到该基因调节元件、DNA 模板或本方法的通用性, 该基因调节元件优选得自真核基因, 或相应于后者。该基因调节元件可以是细胞的或病毒的基因调节元件或合成的基因调节元件。基因调节元件优选包括特别是 DNA 序列, 这些序列代表 DNA 结合蛋白的结合位点 (蛋白质结合 DNA 序列, 例如转录因子或融合蛋白的结合位点)。基因调节元件可包括一个启动子 (启动子序列) 和/或一个或多个增强子 (增强子序列) 和/或一个或多个沉默子 (沉默子序列)。该基因调节元件可优选包括自然排列和/或人工排列的启动子、增强子和/或沉默子序列或其部分序列。

一个启动子可包括一个 "TATA" 盒和/或一个起始区 (INR) (引发转录)。该启动子可包括一个 "GC" 盒和/或 "GAAT" 盒。

在该 DNA 模板的一个特定实施方案中, 该基因调节元件是一个模型启动子。

一个模型启动子包括：一个启动子和另外的蛋白质结合 DNA 序列和，适当的话，其它基因调节元件。该模型启动子优选包括一个“TATA”盒和一个起始区。在一个具体实施方案中，该模型启动子包括人 T 细胞受体 V $\beta$ 8.1 的“TATA”盒和 ML 启动子的起始区。这两种基础启动子元件允许进行基础体外转录。此外，该特异性模型启动子具有用于酵母 Gal 4 蛋白的 5 个结合位点。该模型启动子可根据需要而被改变，例如采用分子生物学的方法，例如由待研究的基因调节元件补充该模型启动子，和/或由其它基因调节元件例如待研究的基因调节元件置换该模型启动子的个别部分。

为了可操纵该模型启动子，它优选含有一个或多个用于限制酶的单一切割位点。在该模型启动子的一个具体实施方案中，至少有一个用于限制酶的单一切割位点处于“TATA”盒和 Gal 4 结合位点之间。

除了已存在的蛋白质结合序列（例如除了 Gal 4 序列）之外，也可将其它蛋白质结合序列整合入该模型启动子。例如，如果除了业已研究的转录系统之外加入其它待研究转录因子如特异性转录因子，这就尤为有趣。

本方法的一方面在于，融合蛋白与该模型启动子组合也可用作转录激活物和/或抑制物。这种融合蛋白例如可包括：DNA 结合域如酵母 Gal 4 蛋白的 DNA 结合域，和特异性激活域如 HSV 激活物 VP16 的激活域。融合蛋白可使定义的激活物和/或抑制物或其部分例如它们的激活域或抑制域被分别分析，而没有待研究的基因调节元件的背景。例如，如果该 DNA 结合域得自所应用的转录系统（如浓缩的核提取物）缺乏的 DNA 结合蛋白，则这是可能的。例如，如果将得自哺乳动物细胞的浓缩核提取物用于本方法中，就可以无背景地分析以与酵母 Gal 4 结合域的融合蛋白形式而存在的转录因子的激活物（域），因为得自哺乳动物细胞的核提取物不含 Gal 4。

可应用的基因调节元件和待研究基因调节元件规定为人的和/或动物和/或病毒的基因调节元件，特别是病理学中有意义基因的基因调节元件。实例有，下列物质的基因的基因调节元件：粘附分子，生长因子，磷酸二酯酶，磷酸酶，激酶，ATP 酶，膜受体，第二信使受体，激素受

体如类固醇受体, 金属蛋白酶, 亲免疫素, NO 合酶, 5 - 脂肪氧合酶; 或者下列免疫靶例如: 细胞因子如白介素的基因调节元件, T - 或 B - 细胞特异性地表达的基因如 CD4 受体、TCR 或 BCR ( T - 或 B - 细胞受体) 的启动子, 淋巴特异性基因如 TNF (肿瘤坏死因子) 的启动子, 或者 T 细胞特异性逆转录病毒如 HTLV - 1 或 HIV - 1 的基因调节元件。

本方法、该模型启动子的一个独特优点在于, 可应用的该基因调节元件不但是含有“TATA”盒的那些启动子如编码白介素 - 2 ( IL - 2 ) 的基因的启动子, 还有缺乏“TATA”盒的启动子如编码 T 细胞受体 $\beta$ 链的基因的启动子。例如, 缺乏“TATA”盒的启动子可被整合入该模型启动子, 或整合入通用报道质粒 pGS100, 并应用于转录。

待转录的 DNA 序列的一个特征在于, 在该序列中它缺乏一个或多个核碱基。待转录的 DNA 序列优选包括: 任选无鸟嘌呤或者无胞嘧啶或者无胸腺嘧啶, 即该序列无 G, 无 C 或者无 T。此外, 该序列也可缺乏一个以上核碱基, 例如鸟嘌呤和胸腺嘧啶或者鸟嘌呤和胞嘧啶或者胞嘧啶和胸腺嘧啶。

应用的无 G、无 T 或无 C 序列尤其是指那些长度多于 400 个核苷酸, 优选在 400 ~ 2000 个核苷酸之间或者更长的无 G、无 T 或无 C 序列。具体而言, 所用序列的长度约为 500、600、700、800、900、1000、1100、1200、1300、1400、1500、1600、1700、1800、1900、2000、5000 或更多个核苷酸。缺乏一种特定的核碱基的序列特别优选的长度是约为 800、1200、1600 或 2200 个核苷酸的长度。

该 DNA 模板可以是线型或环状 DNA 序列, 例如该 DNA 模板可以由聚合酶链反应产生的线型序列或是质粒。

在一个实施方案中, 该 DNA 模板是质粒, 并且由整个或部分载体、模型启动子和例如无 G、无 T 或无 C 的待转录 DNA 序列构建而成。

本发明的目的在于可广泛利用的报道质粒 (通用报道质粒)。该通用报道质粒含有用于限制酶 PstI、EcoRI、SacI、KpnI、SacII、BamHI、SwaI 的单一切割位点, 质粒 pUC19 的部分, 用于酵母 Gal 4 蛋白的 5 个结合位点, 介于 SacII 和 BamHI 限制酶切位点之间的人 T 细

胞受体 V $\beta$ 8.1 的“TATA”盒，介于 BAMHI 和 SwaI 限制酶切位点之间的 ML 启动子（腺病毒主要晚期启动子）的 INR（起始）区，和长度约为 800 个核苷酸（碱基对）的无 G 序列。通用报道质粒中的该模型启动子是一种合成启动子，它含有 5 个酵母 Gal 4 结合位点，用于 PstI、EcoRI、SacI、KpnI、SacII、BamHI 和 SwaI 的单一切割位点，人 T 细胞受体 V $\beta$ 8.1 的 TATA 盒和 ML 启动子的 INR。待研究的任何基因调节元件可被整合入该启动子区。该模型启动子的部分或整个模型启动子可以除去并由待研究的基因调节元件置换。

通用报道质粒的一个具体实例称为 pGS100（图 2）。

在通用报道质粒 pGS100 的一个具体实例中，合成启动子在序列区中处于核苷酸位置 2168 和 2337 之间。腺病毒主要晚期（ML）启动子的起始区处于核苷酸位置 2322 ~ 2337 上，而含有人 T 细胞受体 V $\beta$ 8.1 启动子的 TATA 盒的区在核苷酸位置 2289 ~ 2316 上。酵母 Gal 4 蛋白的 5 个结合位点处于核苷酸位置 2168 ~ 2260 之间。

通用报道质粒 pGS100 进一步的具体实例示于序列表中（SEQ ID No.1）。

通用报道质粒 pGS100 进一步的具体实例保藏在 the DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Mascheroder Weg 1b, D-38124 Braunschweig, 这遵照有关国际承认的用于专利程序的微生物保藏的布达佩斯条约；DSM 保藏号：11450。

本方法的一方面在于，直接或间接地干扰特异性转录物的检测致使不再能明确地检测特异性转录物的蛋白质，在转录后被消除和/或降解。消除或降解尤其包括这样的处理步骤：它们操作简便，例如其中的反应可通过化学的和/或机械的和/或酶促的处理步骤猝灭的方法和，合适的话，与此同时使转录物不含干扰蛋白，即在该处理步骤之后，过剩的标记核苷酸例如可通过洗涤步骤从特异性转录物中除去。

本方法的该实施方案一方面在于，例如在转录反应之后可借助于蛋白酶使蛋白质降解。可应用的蛋白酶例如有：含锌蛋白酶、丝氨酸蛋白酶、巯基蛋白酶和羧基蛋白酶。蛋白酶 K、胰蛋白酶、糜蛋白酶、羧肽酶 A、木瓜蛋白酶和胃蛋白酶优选可应用。本方法一个特定的实施方案

在于，转录之后用蛋白酶 K 进行消化。

为了测定转录程度即转录速率，就必须测定特异性转录物的量。

本方法的一方面在于，转录作用在标记的核苷酸存在下进行，或者在于，标记特异性转录物。例如，如果转录时应用适当的非放射性标记核苷酸，则该转录物可以是非放射性标记的。可应用的核苷酸标记基例如有：荧光基如丹酰（= N - 二甲基 - 1 - 氨基萘 - 5 - 磺酰）衍生物，荧光素衍生物或香豆素衍生物，或者化学发光基如吡啶衍生物。上述标记基可使得直接检测特异性转录物。此外，也可以应用适于间接检测该转录物的标记基。实例有：异羟洋地黄毒甙元（digoxigenin），它可用特异性抗异羟洋地黄毒甙元的抗体例如在 ELISA 中检测；生物素，它可由生物素/阿维丁系统检测；和含有功能基的连接臂，它可随后用可检测的报道基进行衍生化作用。最后提及的可能性实例如氨基烷基连接基，在转录后它可在化学发光试验中与吡啶鎓（acridinium）活化的酯反应并检测。

本方法的一个具体实施方案是在放射性标记的核苷酸存在下进行转录。该核苷酸例如可用磷（ $^{32}\text{P}$  或  $^{33}\text{P}$ ）、硫（ $^{35}\text{S}$ ）或氘（ $^3\text{H}$ ）进行放射性标记。

本方法的一个特定实施方案在于，特异性转录物通过结合到固相（固体基质）上而从反应混合物中分离，例如通过结合到微量滴定板上或通过特将特异性转录物与特定的滤器、膜或其它固相结合，尤其是通过结合到带电的膜或带电的滤器上，它们优选由尼龙或硝酸纤维素构成，特别优选的是通过结合到含有带电基例如二乙氨基乙基的膜上，例如 DEAE - 纤维素膜上。该实施方案的一方面在于，结合在固体基质上的特异性转录物可通过洗涤步骤与过剩的标记核苷酸分离，其中的洗涤可进行到如此程度以致可以清晰地检测该特异性转录物。

对比通常的分离和检测方法（包括酚化处理、沉淀和随后在变性凝胶上分离该转录物），本方法惊人地产生可同样清晰地检测的特殊信号（参考实施例 7 和图 3）。

本方法适于产生例如由特异性转录引起的特殊信号，其中的转录例如在激活物（激活转录）或抑制物（抑制转录）的存在下进行，这些信

号不同于由基础转录例如当不存在激活物或抑制物时引起的基本信号强度；不同之处至少有7倍优选有8倍，在特殊情况下甚至为9或10倍（参考实施例7和图3）。

本方法一个重要的实施方案在于，它可用于筛选有药理活性的物质。为了测试候选药理活性物质的活性（例如活化或抑制性能），在该待测活性物质存在下进行转录。适当时，可将各组分如浓缩的核提取物与待测活性物质预保温。此外，待测试活性物质的特异性的表征可这样进行：在该待测活性物质的存在下于多种转录反应混合物中平行进行转录作用，其中每种反应混合物包含不同组分，然后测定和比较转录速率。

待测试活性物质的实例可以是天然产物和/或得自化学的和组合的物质文库的物质。天然产物例如可分离自植物、动物、植物分泌物、动物分泌物和，特别是得自微生物如得自真菌、酵母、细菌或藻类。

本方法的另一具体实施方案在于，在待测活性物质、转录因子、辅因子或细胞类型特异性核提取物的存在下进行转录，并不存在待测活性物质、转录因子、辅因子或细胞类型特异性核提取物时平行进行反应混合物的反应，但反应条件完全相同；然后通过标记转录物量的差异测定待测活性物质、转录因子、辅因子或细胞类型特异性核提取物相对于基因调节元件（或基因）和/或转录因子和/或辅因子的活性（例如抑制性，激活性）或效应。

本方法的一方面在于，借助于在相同条件下平行进行的转录作用测定参与基因调节的、待研究的蛋白质的效应，待研究的蛋白质只存在于两种反应混合物之一中。

本方法一个实施方案的一方面在于：

- a) 在相同条件下平行进行至少两种转录作用，此时
- b) 转录反应混合物的不同点只是在于，它们包括不同量的待测活性物质和/或至少一种转录因子和/或至少一种辅因子和/或一种浓缩的核提取物，
- c) 转录后测定每种反应混合物生成的标记转录物量，和
- d) 通过生成的标记转录物量的差异确定待测活性物质、转录因子、辅因子和/或浓缩的核提取物相对于基因调节元件的活性和/或特异性。

本方法能靶特异性地测试转录反应混合物所含的各组分（例如对于核提取物（即对于特异性的细胞类型）或对于特异的转录因子）的效应。例如可分析待研究的活性物质对于规定的基因调节元件的效应。此处的决定因素在于转录的反应条件可以准确限定。

本方法以靶特异性的方式提供影响各因子对病理基因表达的作用的方法，因为可通过选择基因调节元件和转录因子和/或辅因子和/或细胞类型特异性核提取物而调节或准确设定合适的反应条件。

本方法的一个独特优点在于，它能鉴定能对规定条件下的转录作用产生正的或负的效应的药理活性物质，特别是那些激活或抑制规定的（靶）基因的转录作用的活性物质，这些活性物质对规定的基因调节元件和/或规定的转录因子、辅因子和/或细胞类型特异性核提取物（或细胞）具有特异性效应。

本发明涉及本方法在鉴定特异性活性物质中的应用。本方法例如可用于表征待测物质。本方法例如可用于表征待测物质在规定条件下的特异性。由本方法鉴定的药物活性物质例如应能抑制或激活受基因调节元件控制的 DNA 序列的转录作用。

本方法又一个特别优异的特征在于，所有步骤可以简便方式实现自动化。例如可应用与自动供料微型组件联接的自动移液装置 Biomek 2000®（Beckman, Munich）。然后可以由手工或自动方式例如借助于传输带洗涤与固相结合转录物。

本发明一方面提供自动移液装置，例如 Biomek 2000®，它装备有各反应组分例如蛋白质级分（如核提取物和其它蛋白质）、DNA 模板、转录缓冲液、在转录作用中待研究的物质、移液管滴头、微量滴定板和膜。各个反应由这些组分构成例如微量滴定孔中的组分，自动地进行所有移液步骤。这样，可以同时处理每块板上的 96 个或更多的不同转录反应混合物，结合少量样品体积，特别是小于 100 $\mu$ l、优选为 10 ~ 50 $\mu$ l、尤其是 20 $\mu$ l 的体积。每次转录需大约 1 ~ 1.5 小时，于是当所用保温时间呈最佳可能方式时，每天以这种方式可进行多达 1000 或更多次转录。

转录作用例如可在 20 ~ 50 $^{\circ}$ C 的温度下进行。在大约 30 $^{\circ}$ C 下进行转录是特别优选的。

由于特别是待测活性物质、转录因子、辅因子和适当时还有待测浓缩的核提取物可获得量较少，但这些物质应在大量转录作用中很多反应条件下测试，所以本方法必须满足这一要求，即转录反应所需样品体积应尽可能小。因此特别重要的是，本方法也适于在毫微升规模进行，即反应体积约为 50 ~ 500nl。

本转录方法可以通用。它可用于鉴定和表征在基因转录的调节中起直接或间接作用的转录因子和/或辅因子和/或其它蛋白质（即作为靶的特异性蛋白质）和/或浓缩的核提取物（即作为靶的特异性核提取物或特种细胞）的基因调节元件（即作为靶的特异性基因）。本转录方法尤其可用于鉴定病理学中有意义的新型基因调节元件和用于排布基因调节蛋白，它们介导这些元件的效应至相应基因调节元件。

与细胞分析法相比，上述转录方法具有更高的靶特异性。与细胞分析法相反，细胞对各组分的摄取对于转录效率无影响。此外，上述方法可简便而快速地进行（例如，可制备各组分并冷冻贮存）。本方法易于标准化并可通用，因为它基本上可用于任何类型细胞和任意基因。

本方法可鉴定能用于制备药剂的药理活性物质。与已知活性物质相比，由这些转录方法鉴定的活性物质应引起相当少的副作用。例如，可能测试或鉴定可用于制备治疗下述疾病的药剂的活性物质：（自身）免疫病、代谢病、癌、心血管病、传染病、风湿病、糖尿病、变性病和精神病，尤其用于制备治疗下述疾病的药剂：类风湿性关节炎、多发性硬化、糖尿病、变态反应、气喘、过敏症、特应性皮炎、老年性痴呆、帕金森氏病、AIDS、克雅氏病、癫痫、精神分裂症、动脉硬化和结核病。

此外，本通用方法提供大量的其它可能用途。例如，假如使用下列生物体的有关基础转录系统以及特异性基因调节元件、转录因子和/或辅因子和/或直接或间接地参与转录反应的其它蛋白质，则本方法类似地可用于动物病理学和动物饲养，用于作物保护或植物栽培以寻找特异性的药理活性物质。

原则上，如果相应地使用微生物的或昆虫的基因调节元件、能转录的系统，转录因子和/或辅因子，则该体外转录方法也可类似地用来鉴定可用于保藏原料和食料的物质，所述微生物例如有酵母、真菌或细菌。

### 附图说明

图 1: 读出转录反应的放射活性, 其中应用了具有不同长度无 G 序列的各种基因调节元件和报道质粒。

示出了两种标准启动子与通用报道质粒 pGS100 (图 2) 的比较。该转录反应按实施例 6 中所述方法进行。

A) 基础转录: 未活化启动子而得到的放射活性转录物的量〔以相对单位表示〕(基础信号强度);

B) 活化转录: 用 Gal 4-聚谷氨酰胺活化启动子而得到的放射活性转录物的量〔以相对单位表示〕。

I) 所用报道质粒为 pMRG5 (Kretschmar, M., Kaiser, K., Lottspeich, F., Meisterernst, M. (1994) Cell 78, 525-534)。pMRG5 中的合成启动子含有 HIV 启动子的 TATA 盒, ML 启动子的起始子以及 pUC19 中长度约为 400 个核苷酸的无 G 序列。

II) 所用报道质粒为 pV $\beta$ ML。报道质粒 pV $\beta$ ML 的构建方法与报道质粒 pGS100 的相同, 但它所含的无 G 序列长度只有 400 个核苷酸而不是 800 个核苷酸。

III) 所用报道质粒为 pGS100, 它包含约有 800 个核苷酸的无 G 序列。

每种情况下比较的是基础转录 (A) 和活化转录 (B)。采用的激活物是一种融合蛋白, 它包含 Gal 4 结合域 (94 个氨基端氨基酸) 和聚谷氨酰胺活化域 (有 11 个谷氨酸单元的合成肽)。标绘的数据表示扣除背景后的绝对值 (以相对单位表示, 是由磷酸图象仪 (phosphoimager) 测得的)。在转录反应中, 其中用 pGS100 为报道质粒, 测得更高的绝对值和由于合成启动子的不同构建而导致比 mRG5 更优良的可激活性。此外, 长度在 400 个核苷酸以上的无 G 序列的正效应是显而易见的。

图 2: 通用报道质粒 pGS100。

通用报道质粒 pGS100 在 pUC19 中在长度约为 800 个碱基对的无 G 区之前包含一个合成启动子区作为模型启动子。在 BamHI 和 SmaI 限制酶切位点内定位有腺病毒主要晚期 (ML) 启动子的起始区。在 SacII 和 BamHI 限制酶切位点之间, 定位有含人 T 细胞受体 V $\beta$ 8.1 启动子的

TATA 盒的区。这两个基本启动子元件（起始子和 TATA 盒）使得可在体外进行基础转录作用。由于这些候选靶基因可能对于筛选尤为重要，所以有可能将它们个别地换成待研究基因的相应区（相应的基因调节元件）。靶基因的任何调节区都可引入基础启动子的多接头下游（特异性激活）。也可以交换整个启动子区。多接头上游（SacII 至 PstI），pGS100 含有用于酵母 Gal4 蛋白的 5 个结合位点。这也使得可分析合成转录激活物例如融合蛋白，它们包括任意所需的活化域（例如单纯疱疹反式激活蛋白 VP16）和包括 Gal 4 DNA 结合域。该无 G 序列是 pGS100 的基本单元。

图 3：常规的标准转录方法与上述无细胞体外转录方法功效的比较。

图 3 示出了在不同反应条件下测得的信号强度（作为转录物量的量度，因而可反映转录强度）的比较。转录反应按实施例 7 中所述方法进行。

A) 常规的标准转录方法，其中通过苯酚处理、用乙醇沉淀和随后的变性凝胶电泳处理将特异性转录物与蛋白质和过剩的核苷酸分离。

B) 上述体外转录方法，其中蛋白质用蛋白酶 K 消化，而过剩的核苷酸则通过洗涤与 DEAE 滤器结合的特异性转录物而得以除去；

I) 用于转录的原料是得自 HeLa 细胞核的提取物，它能转录且已经在 P11® 柱上纯化过，没有相应的 DNA 模板（对照试验）；

II) 基础转录：将与 I 的反应混合物相当的反应混合物用作为 DNA 模板的报道质粒 pMRG5 处理，并进行转录反应；以该方式测定基本信号强度；

III) 活化转录：将与反应混合物 II 相当的反应混合物另外用转录激活物处理，后者是一种融合蛋白，它包括 Gal 4 的 DNA 结合域和 VP16 的活化域（Gal 4-VP16），并进行转录作用；这样，可得活化信号强度；

IV) 将与 III 的转录反应混合物相当的反应混合物用作为 RNA 聚合酶 II 抑制物的  $\alpha$ -鹅膏蕈碱处理（对照试验）。

比较在不同反应条件下由两种不同方法得到的信号强度，表明利用本文所述体外转录方法不但可测定基础转录（II），而且可测定活化转录（III），信号强度类似于利用常规的转录方法获得的那些。甚至基本

信号强度也可测定这一现象极为重要，这样，上述转录方法也可用于筛选转录抑制物。借助于特异性 RNA 聚合酶 II 抑制剂  $\alpha$ -鹅膏蕈碱可使信号大约减小到 1/8（比较 III 和 IV）

实施例：

#### 实施例 1：HeLa 核提取物的制备

用 HeLa 细胞核为起始原料制备核提取物。HeLa 细胞核可从许多公司（例如“4℃”，Mons; Sigma, Munich, Santa Cruz Biotechnology, Heidelberg）商购。下文中所述对该细胞核的处理是在 4℃（冷室）下进行的。缓冲液都是在室温（RT）下用 Tris pH6.8 调节的，它相当于 4℃下的 pH7.3。使用前，缓冲液用 DTT（贮存液 1M，于水中）处理至最终浓度为 5mM，并用 PMSF（贮存液 200mM，于 DMSO 中）处理至最终浓度为 1mM。

细胞核的处理包括下列步骤：

1. 在冰上解冻细胞核并测定 NPV 体积（核沉淀体积）。
2. 在两个不同的玻璃烧杯中分别加入 1/2 NPV 体积的 0.02M KCl 缓冲液（含盐量低的缓冲液：20ml 1M Tris pH6.8RT, 250ml 100%甘油, 6.67ml 3M KCl, 1.5ml 1M MgCl<sub>2</sub>, 0.4ml 0.5M EDTA, 加 H<sub>2</sub>O 至 1000ml）和 1.2M KCl 缓冲液（含盐量高的缓冲液：20ml 1M Tris pH6.8 RT, 250ml 100%甘油, 400ml 3M KCl, 1.5ml 1M MgCl<sub>2</sub>, 0.4ml 0.5M EDTA, 加 H<sub>2</sub>O 至 1000ml）。

每种缓冲液用  $0.0007 \times \text{NPV}/2 \beta$ -巯基乙醇和  $0.001 \times \text{NPV}/2$  0.2M PMSF 处理。

该沉淀重悬浮于 1/2 体积 0.02M KCl 缓冲液并用研杵（6X）轻轻地混匀。

3. 将匀浆转入玻璃烧杯，在 30 分钟内和不断搅拌下滴加 1.2M KCl 缓冲液进行处理。又搅拌 30 分钟后，提取完毕。离心分离（Beckman 离心机，SS34 转子，在 14,000rpm 下，30min, 4℃），进一步分别处理沉淀和上清液。

4. 在缓冲液 1（40ml 1M Tris pH6.8RT, 400ml 甘油, 0.8ml 0.5M EDTA, 加 H<sub>2</sub>O 至 2000ml）中透析上清液，直至达到缓冲液 2（40ml

1M Tris pH6.8RT, 400ml 甘油, 0.8ml 0.5M EDTA, 66.7ml 3M KCl, 加 H<sub>2</sub>O 至 2000ml) 的电导率 (45 ~ 55min)。

5. 离心透析后的上清液 (Beckman, SS34 转子, 18,000rpm, 20min, 4 °C)。HeLa 核提取物处于该上清液中 (HeLa 核提取物 = HeLa NE)。在液 N<sub>2</sub> 中冷冻该提取物等分样。将沉淀物从核提取液转入匀浆器, 用 10ml TGME/5 mM DTT 处理 (TGME 达 1L: 250ml 100%甘油, 50ml Tris pH7.3RT, 5ml 1M MgCl<sub>2</sub>, 0.2ml 500mM EDTA pH8.0), 用研杵 (强烈地) 搅拌 (20X), 再在液 N<sub>2</sub> 中冷冻 (Hela 核沉淀物)。

实施例 2: 制备用于分离核提取物的磷酸纤维素柱。

1. 在水中反复洗涤 P11®柱材料。测定膨胀材料的体积。
2. 加入 5 份体积的 0.5N NaOH, 放置 5 分钟, 然后立即用折叠滤纸吸滤。
3. 水洗至 pH11。
4. 加入 25 份体积的 0.5N HCl。放置 5 分钟, 然后立即吸滤。
5. 水洗至 pH3。
6. 用 1M Tris pH7 洗涤至 pH 恒定为 7。在 4 °C 下贮存柱材料。最好使之平衡一夜。

实施例 3: 用磷酸纤维素色谱法纯化核提取物:

P11®材料的容量为 10mg 蛋白质/1ml 材料。色谱法进行如下:

1. 用 P11®材料装填柱并在缓冲液 2 中平衡 (加入新鲜的 DTT 和 PMSF, 见实施例 1)。将柱与泵连接。
2. 装载 (以 1 柱体积 (CV) /h) 核提取物 (于缓冲液 2 中)。
3. 用 5CV 缓冲液 2 洗涤柱 (5 ~ 10CV/h)。
4. 用缓冲液 3 (40ml 1M Tris pH6.8RT, 400ml 甘油, 0.8ml 0.5M EDTA, 200ml 3M KCl, 加 H<sub>2</sub>O 至 2000ml; 2CV/h) 洗脱。当不再能洗脱出蛋白质后, 又通入 2CV 缓冲液 3。
5. 用缓冲液 4 (40ml 1M Tris pH6.8RT, 400ml 甘油, 0.8ml 0.5M EDTA, 333ml 3M KCl, 加 H<sub>2</sub>O 至 2000ml; 2CV/h) 洗脱并收集峰级分。
6. 用缓冲液 5 (40ml 1M Tris pH6.8RT, 400ml 甘油, 0.8ml 0.5M

EDTA, 567ml 3M KCl, 加 H<sub>2</sub>O 至 2000ml; 2CV/h) 洗脱并收集峰级分。

7.用缓冲液 2 透析这 2 种洗脱液中每一种直至电导率保持恒定, 然后在液 N<sub>2</sub> 中冷冻等分样。

实施例 4: 利用聚丙烯酰胺凝胶分离转录物的标准转录方法。

转录反应混合物(终体积 20 $\mu$ l)可包括下列组分:

- a) 呈预纯化的核提取物形式或由重组技术生产的通用转录因子, 适当由其它特异性转录因子、激活物、抑制物或融合蛋白补充;
- b) DNA 模板(例如带有待研究的基因调节元件的载体 pGS100);
- c) 转录缓冲液, 例如包括待研究的活性物质;
- d) 标记的和未标记的核苷酸;

其中 a) (实施例 3 的) 缓冲液 4 洗脱物和缓冲液 5 洗脱物都含于能转录的核提取物中(所有通用转录因子)。每次单独制备都要确定缓冲液 4 洗脱物和缓冲液 5 洗脱物的最适量(每种反应混合物大约 3 $\mu$ l 缓冲液 4 洗脱物和 2 $\mu$ l 缓冲液 5 洗脱物);

其中 b) 100ng DNA 模板(报道质粒例如 pGS100 中的基因调节元件)一般被用于每种反应混合物;

其中 c) 每种反应混合物中含转录缓冲液(5mM MgCl<sub>2</sub>, 25mM HEPES KOH pH8.2, 0.5 $\mu$ l BSA 乙酰化(贮存物 20mg/ml), 大约 10% 甘油, 大约 70mM KCl, 0.2mM PMSF, 10mM DTT)加上每种情况下 20U RNA 酶抑制剂;

其中 d) NTPs (ATP, UTP 在各情况下最终浓度为 100 $\mu$ M, CTP 最终浓度为 5 $\mu$ M, o-m-GTP 最终浓度为 20 $\mu$ M,  $\alpha$ -<sup>32</sup>P-CTP 最终浓度约为 0.12 $\mu$ M, 3000Ci/mmol, 10mCi/ml)。

- 1.引入转录缓冲液和 dNTPs 并加入 DNA 模板。
- 2.加入转录激活物和缓冲液 4 洗脱物与缓冲液 5 洗脱物。
- 3.在 30 °C 下保温 1h 以进行转录反应。
- 4.加入 400 $\mu$ l 终止混合物 (stopmix) (7M 脲, 10mM Tris HCl pH7.8, 10mM EDTA/NaOH pH7.8, 0.5% SDS, 100mM LiCl, 100 $\mu$ g/ml tRNA, 300mM 乙酸钠)和 400 $\mu$ l 苯酚/氯仿/异戊醇 (25/24/1)。混合后

离心分离 ( SS34 转子, Beckman 离心机, 5min, 14,000rpm, RT ) 。吸出上清液, 加入 400 $\mu$ l 异丙醇并混合, 然后在 - 20  $^{\circ}$ C 下保温 1h 。

5. 离心分离 ( SS34 rotor, Beckman 离心机, 14,000rpm, 30min, 4  $^{\circ}$ C ) , 用 70 % 乙醇洗涤沉淀物, 再在 Speedvac 中干燥。

6. 在 10 $\mu$ l 荷载缓冲液 ( 955 $\mu$ l 100% 甲酰胺(去离子化), 10 $\mu$ l 0.5M EDTA, 20 $\mu$ l 1M Tris pH7, 各情况下 0.003 % 溴酚蓝/二甲苯青 ( xylene cyanole ) , 加 H<sub>2</sub>O 至 1ml ) 中溶解沉淀物, 并在 50  $^{\circ}$ C 下保温 15 分钟。

7. 在 5 % 加强变性 ( strength denaturing ) 聚丙烯酰胺凝胶内用 1  $\times$  TBE 作流动缓冲液分离转录物。在 60mA 时预走 20min 。装载凝胶。在 60mA 时走凝胶约 1 小时。

8. 在 10 % 乙酸中固定凝胶, 干燥, 再在 Phosphoimager<sup>®</sup> 中暴露于 X 光片。

实施例 5: 可实现自动化的、利用滤膜结合转录物的上述体外转录方法的操作流程。

1. 引入转录缓冲液和 dNTRs , 并加入 DNA 模板。

2. 加入缓冲液 4 洗脱物和缓冲液 5 洗脱物和适当的话, 转录激活物或转录抑制物。

3. 在 30  $^{\circ}$ C 下将混合物保温 1h 而进行转录反应。

4. 加入 5 $\mu$ l 蛋白酶 K 混合物 ( 1 $\mu$ l 蛋白酶 K 溶液 [20mg/ml], 1 $\mu$ l 10 % SDS , 0.5 $\mu$ l 0.5M EDTA pH8, 1 $\mu$ l 50mM Tris pH7.8 , 加水至 5 $\mu$ l ) 。

5. 在 30  $^{\circ}$ C 下保温 15min 。

6. 将 DEAE 膜 NA45<sup>®</sup> ( Schleicher und Schuell ) 置于洗膜缓冲液 ( 100mM 磷酸钠缓冲液 pH7.5, 250mM NaCl, 2% 焦磷酸钠 ) 中短暂地洗涤。

7. 从该反应混合物中吸移 5 $\mu$ l 于膜上, 让它干燥 5 分钟。

8. 将该膜在大约 100ml 洗膜缓冲液 ( + 1 % Triton ) 中轻轻振动 4  $\times$  15 分钟。

9. 将膜转至 Whatmann 3MM<sup>®</sup> 纸 ( Whatman, Maidstone, England ) 上, 风干, 用胶片覆盖后利用 Phosphoimager<sup>®</sup> 使滤膜暴露于 X 光片。

**实施例 6:** 按照上述应用滤膜的体外转录方法, 用 3 种不同的报道质粒平行进行转录反应。

程序同实施例 1 - 3 和 5。应用的报道质粒是 pMRG5 ( HIV 启动子的 TATA 盒, ML 启动子的起始区和 pUC19 中长度为大约 400 个核苷酸的无 G 盒 ( Kretschmar, M., Kaiser, K., Lottspeich, F., Meisterernst, M. (1994) Cell. 78, 525-534 ) ), pGS100 ( 图 2 ) 和 pV $\beta$ ML ( 象 pGS100 那样构建的, 但只包含 400bp 无 G 盒而不是 800bp 无 G 盒)。每种情况下将 100ng DNA 模板和 3 $\mu$ l P11 $\text{\textcircled{R}}$ 级分( 缓冲液 4 洗脱物和缓冲液 5 洗脱物) 用于转录反应。这些反应每一种都是在不含和含有激活物 ( 由 Gal 4 结合域和聚谷氨酰胺活化域构成的融合蛋白) 时平行进行的。结果如图 1 所示和阐明。

**实施例 7:** 比较应用标准转录方法获得的和用上述体外转录方法获得的放射性标记转录物的信号强度 ( 作为转录物量的量度, 因而反映转录强度) 。

按实施例 1 ~ 5 中所述方法进行转录反应, 并利用滤膜分析法 ( 实施例 5 ) 或凝胶分析法 ( 实施例 4 ) 分析平行反应混合物。在每种情况下将 200ng pMRG5 用作反应的 DNA 模板。某些转录反应是在 30ng 用作激活物的 Gal 4-VP16 存在下进行的。结果示于和描述于图 3 中。

## 序列表

### (1) 一般信息:

#### (i) 申请者:

(A) 姓名: **Hoechst Aktiengesellschaft**

(B) 街名: -

(C) 城市: **Germany**

(D) 州名: -

(E) 国别: **Deutschland**

(F) 邮编: **65926**

(G) 电话: **069 - 305 - 7072**

(H) 电传: **069 - 35 - 7175**

(I) 电报挂号: -

(ii) 发明名称: **筛选天然产物和其它化学物质的体外转录方法。**

(iii) 序列数: **1**

#### (iv) 计算机可读形式:

(A) 介质类型: **软盘**

(B) 计算机: **IBM PC 兼容机**

(C) 操作系统: **PC - DOS/MS - DOS**

(D) 软件: **Patent In Release #1.0, 版本 # 1.25 (EPO)**

### (2) SEQ ID No.: 1 的信息:

#### (i) 序列特征:

(A) 长度: **3130 个碱基对**

(B) 类型: **核酸**

(C) 链型: **双链**

(D) 拓扑学: **环状**

(ii) 分子类型: **DNA (基因组的)**

#### (ix) 特征:

(A) 名称/键: **外显子**

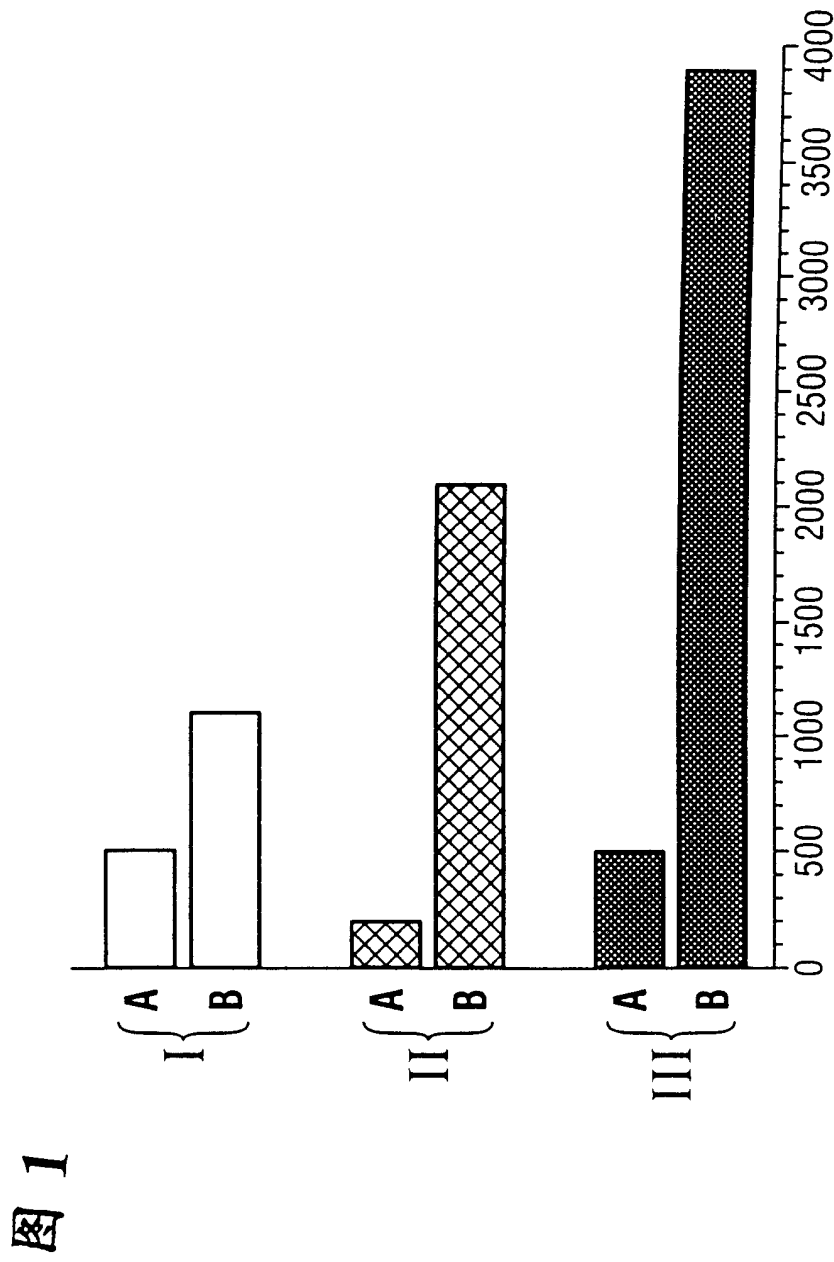
(B) 位置: **1..3130**

(xi) 序列描述: **SEQ ID No.: 1:**

TTCCTGTGT GAAATTGTTA TCCGCTCACA ATTCCACACA ACATACGAGC CGGAAGCATA	60
AAGTGTAAG CCTGGGGTGC CTAATGAGTG AGCTAACTCA CATTAAATGC GTTGCGCTCA	120
CTGCCCCGCTT TCCAGTCGGG AAACCTGTGC TGCCAGCTGC ATTAATGAAT CGGCCAACGC	180
GCGGGGAGAG GCGGTTTGC G TATTGGGCGC TCTTCCGCTT CCTCGCTCAC TGA CTGCTG	240
CGCTCGGTGC TTCGGCTGCG GCGAGCGGTA TCAGCTCACT CAAAGGCGGT AATACGGTTA	300
TCCACAGAAT CAGGGGATAA CGCAGGAAAG AACATGTGAG CAAAAGGCCA GCAAAAGGCC	360
AGGAACCGTA AAAAGGCCGC GTTGCTGGCG TTTTCCATA GGCTCCGCC CCCTGACGAG	420
CATCACAAA ATCGACGCTC AAGTCAGAGG TGGCGAAACC CGACAGGACT ATAAAGATAC	480
CAGGCGTTTC CCCCTGGAAG CTCCCTCGTG CGCTCTCCTG TTCCGACCCT GCCGTTACC	540
GGATACCTGT CCGCCTTTCT CCCTTCGGGA AGCGTGGCGC TTTCTCAATG CTCACGCTGT	600
AGGTATCTCA GTTCGGTGTA GGTGCTTCGC TCCAAGCTGG GCTGTGTGCA CGAACCCCC	660
GTTCAGCCCG ACCGCTGCGC CTTATCCGGT AACTATCGTC TTGAGTCCAA CCCGGTAAGA	720
CACGACTTAT CGCCACTGGC AGCAGCCACT GGTAACAGGA TTAGCAGAGC GAGGTATGTA	780
GGCGGTGCTA CAGAGTTCTT GAAGTGGTGG CCTAACTACG GCTACACTAG AAGGACAGTA	840
TTTGGTATCT GCGCTCTGCT GAAGCCAGTT ACCTTCGGAA AAAGAGTTGG TAGCTCTTGA	900
TCCGGCAAAC AAACCACCGC TGGTAGCGGT GGTTTTTTTG TTTGCAAGCA GCAGATTACG	960
CGCAGAAAA AAGGATCTCA AGAAGATCCT TTGATCTTTT CTACGGGGTC TGACGCTCAG	1020
TGGAACGAAA ACTCACGTTA AGGGATTTTG GTCATGAGAT TATCAAAAAG GATCTTCACC	1080
TAGATCCTTT TAAATTA AAA ATGAAGTTTT AAATCAATCT AAAGTATATA TGAGTAACT	1140
TGGTCTGACA GTTACCAATG CTTAATCAGT GAGGCACCTA TCTCAGCGAT CTGTCTATTT	1200
CGTTCATCCA TAGTTGCCTG ACTCCCCGTC GTGTAGATAA CTACGATACG GGAGGGCTTA	1260
CCATCTGGCC CCAGTGCTGC AATGATACCG CGAGACCCAC GCTCACCGGC TCCAGATTTA	1320
TCAGCAATA ACCAGCCAGC CGGAAGGGCC GAGCGCAGAA GTGGTCCTGC AACTTTATCC	1380
GCCTCCATCC AGTCTATTAA TTGTTGCCGG GAAGCTAGAG TAAGTAGTTC GCCAGTTAAT	1440
AGTTTGCGCA ACGTTGTTGC CATTGCTACA GGCATCGTGG TGTCACGCTC GTCGTTTGGT	1500
ATGGCTTCAT TCAGCTCCGG TTCCCAACGA TCAAGGCGAG TTACATGATC CCCCATGTTG	1560
TGCAAAAAG CGGTTAGCTC CTTCGGTCCT CCGATCGTTG TCAGAAGTAA GTTGGCCGCA	1620
GTGTTATCAC TCATGGTTAT GGCAGCACTG CATAATTCTC TTA CTGTCAT GCCATCCGTA	1680
AGATGCTTTT CTGTGACTGG TGAGTACTCA ACCAAGTCAT TCTGAGAATA GTGTATGCGG	1740
CGACCGAGTT GCTCTTGCCC GCGGTCAATA CGGGATAATA CCGCGCCACA TAGCAGA ACT	1800
TTAAAAGTGC TCATCATGG AAAACGTTCT TCGGGGCGAA AACTCTCAAG GATCTTACCG	1860
CTGTTGAGAT CCAGTTCGAT GTAACCCACT CGTGCACCCA ACTGATCTTC AGCATCTTTT	1920
ACTTTCACCA GCGTTTCTGG GTGAGCAAAA ACAGGAAGGC AAAATGCCGC AAAAAAGGGA	1980
ATAAGGGCGA CACGGAAATG TTGAATACTC ATACTCTTCC TTTTTC AATA TTATTGAAGC	2040

---

ATTTATCAGG	GTTATTGTCT	CATGAGCGGA	TACATATTTG	AATGTATTTA	GAAAAATAAA	2100
CAAATAGGGG	TCCGCGCAC	ATTTCCCCGA	AAAGTGCCAC	CTGGGGGACT	AGAGTCTCCG	2160
CTCGGAGGAC	AGTACTCCGC	TCGGAGGACA	GTACTCCGCT	CGGAGGACAG	TACTCCGCTC	2220
GGAGGACAGT	ACTCCGCTCG	GAGGACAGTA	CTCCGACCTG	CAGGAATTCCG	AGCTCGGTAC	2280
CCGCGGGGAT	AAAATGTCAC	AAAATTCATT	TGGATCCTCA	CTCTCTTCAT	TTAAATATCC	2340
CATACCCTTC	CTCCATCTAT	ACCACCCTAC	TCTCCTTTCC	TCATTATTCC	TCCTATTATC	2400
TTCTCCTCTT	CTCTCCTTCT	TCTATATTTT	CCAAATCTAT	CATCATTAC	TCTCATCCCC	2460
TCTTCCTTCA	CTCCATTCT	ATTCTACTCC	TTTCCCTTTC	CATATCCCCT	CCACCCCCCT	2520
TCCTCCCCTC	TTTCAATCTT	ATCCCCAATC	ATAAAATTAT	CTCAATTATA	TTCTCCTTCC	2580
ATACCCCCTA	TCATCCTCAT	CCCTATCACC	CCCTACTCAC	CCAATACTCC	CTACTCATCT	2640
CATATATCCT	TATCCTCTCC	TCACCTCTCC	CTCCTCTATC	TCCCCCCTC	ACACTCATTT	2700
CTCATTCCAC	TCCCAAATAT	CCCATACCCT	TCCTCCATCT	ATACCACCCT	ACTCTCCTTT	2760
CCTCATTATT	CCTCCTATTA	TCTTCTCCTC	TTCTCTCCTT	CTTCTATATT	TCCCAAATCT	2820
ATCATCATTC	ACTCTCATCC	CCTCTTCCTT	CACTCCCATT	CTATTCTACT	CCTTTCCCTT	2880
TCCATATCCC	CTCCACCCCC	CTTCCTCCCC	TCTTTCAATC	TTATCCCCAA	TCATAAAATT	2940
ATCTCAATTA	TATTCTCCTT	CCATACCCCC	TATCATCCTC	ATCCCTATCA	CCCCCTACTC	3000
ACCCAATACT	CCCTACTCAT	CTCATATATC	CTTATCCTCT	CCTCACCTCT	CCCTCCTCTA	3060
TCTCCCCCCC	TCACACTCAT	TTCTCATTCC	ACTCCCAGGG	ATCAGCTTGG	CGTAATCATG	3120
GTCATAGCTG						3130



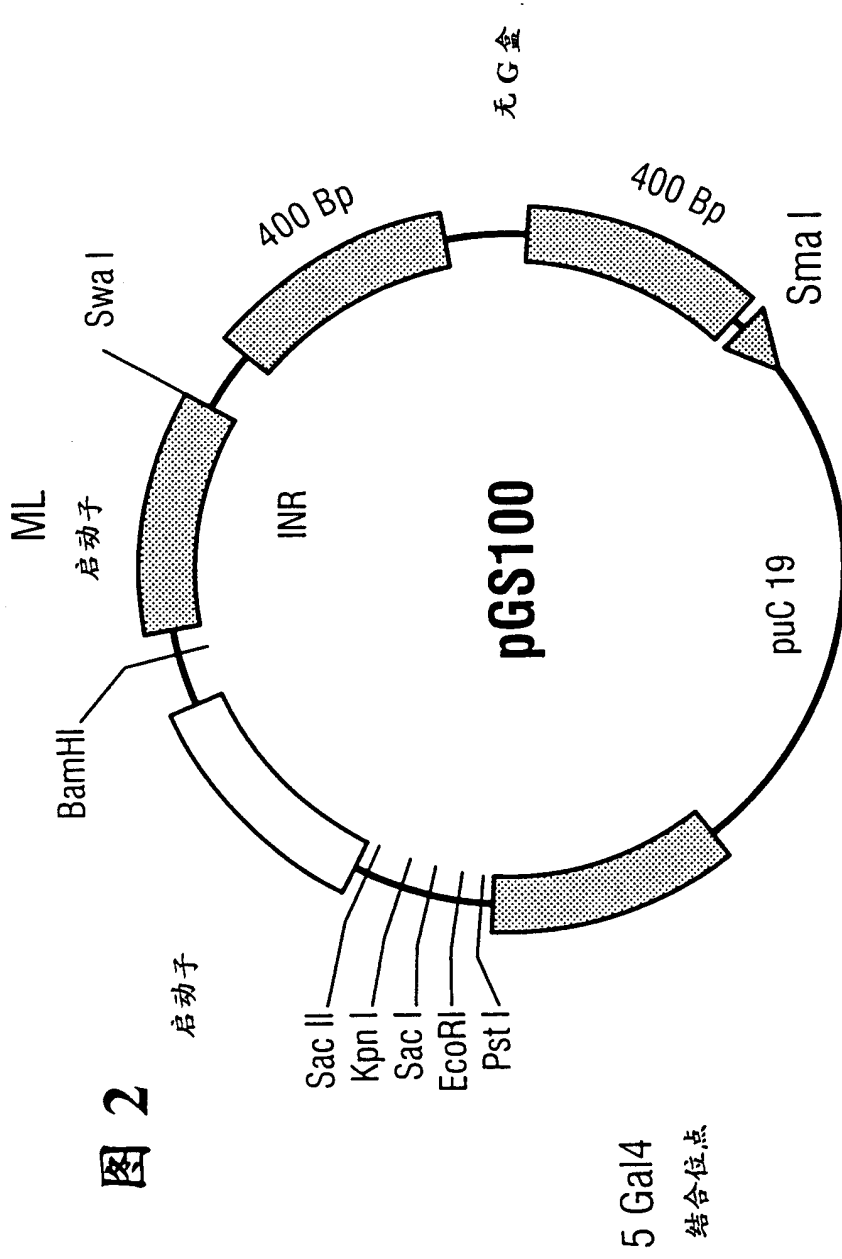


图 2

图 3

