



# (12)发明专利申请

(10)申请公布号 CN 107567501 A

(43)申请公布日 2018.01.09

(21)申请号 201680008727.1

(74)专利代理机构 北京市中咨律师事务所  
11247

(22)申请日 2016.02.10

代理人 凌立 黄革生

(30)优先权数据

1502152.0 2015.02.10 GB

(51)Int.Cl.

C12P 19/34(2006.01)

(85)PCT国际申请进入国家阶段日

C12Q 1/68(2006.01)

2017.08.03

C12N 9/12(2006.01)

(86)PCT国际申请的申请数据

PCT/GB2016/050301 2016.02.10

(87)PCT国际申请的公布数据

W02016/128731 EN 2016.08.18

(71)申请人 核酸有限公司

地址 英国赫特福德郡

(72)发明人 M·C·陈 R·A·拉扎尔 J·黄

G·R·麦金罗伊

权利要求书3页 说明书17页

序列表12页 附图6页

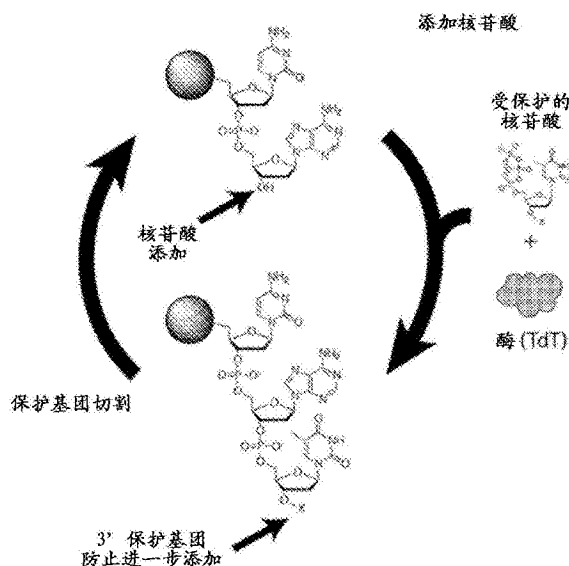
按照条约第19条修改的权利要求书3页

(54)发明名称

新用途

(57)摘要

本发明涉及特异性末端脱氧核苷酸转移酶(TdT)在核酸合成方法中的应用,合成核酸的方法,以及包含所述酶的试剂盒在核酸合成方法中的应用。本发明还涉及末端脱氧核苷酸转移酶和3'-封闭的核苷酸三磷酸在非模板依赖性核酸合成方法中的应用。



1. 包含选自以下的氨基酸序列的末端脱氧核苷酸转移酶 (TdT) 在核酸合成的方法中的用途: (a) SEQ ID NO:1至5和8中的任一个, 或与所述氨基酸序列具有至少20%序列同源性的其功能性等同物或其片段; 或 (b) SEQ ID NO:6的修饰的衍生物。

2. 如权利要求1所述的用途, 其中末端脱氧核苷酸转移酶 (TdT) 包含选自SEQ ID NO:1至5和8中任一项的氨基酸序列, 或与所述氨基酸序列具有至少20%序列同源性的其功能性等同物或片段。

3. 如权利要求1或2所述的用途, 其中末端脱氧核苷酸转移酶 (TdT) 包含选自SEQ ID NO:1、2或8的氨基酸序列, 或与所述氨基酸序列具有至少20%序列同源性的其功能性等同物或片段。

4. 一种核酸合成方法, 其包括以下步骤:

(a) 提供起始序列;

(b) 在权利要求1至3中任一项定义的末端脱氧核苷酸转移酶 (TdT) 存在下, 向所述起始序列中加入3'-封闭核苷酸三磷酸;

(c) 从起始序列中除去所有试剂;

(d) 在裂解剂存在下, 从3'-封闭的核苷酸三磷酸切割封闭基团;

(e) 去除裂解剂。

5. 如权利要求4所述的方法, 其中通过重复步骤 (b) 至 (e) 加入大于1个核苷酸。

6. 如权利要求4或权利要求5所述的方法, 其中3'-封闭的核苷酸三磷酸被3'-O-叠氮甲基、3'-氨基或3'-O-烯丙基封闭。

7. 如权利要求4至6中任一项所述的方法, 其中末端脱氧核苷酸转移酶 (TdT) 是在延伸溶液存在下加入, 所述延伸溶液包含一种或多种缓冲液如Tris或二甲基胍酸盐, 一种或多种盐以及无机焦磷酸酶如来自酿酒酵母的纯化的重组无机焦磷酸酶。

8. 如权利要求4至7中任一项所述的方法, 其中步骤 (b) 在5至10的pH范围内进行。

9. 如权利要求4至8中任一项所述的方法, 其中裂解剂是化学裂解剂。

10. 如权利要求4至9中任一项所述的方法, 其中裂解剂是酶裂解剂。

11. 如权利要求4至10中任一项所述的方法, 其中裂解剂选自: 三(2-羧乙基)膦 (TCEP)、钼络合物或亚硝酸钠。

12. 如权利要求4至11中任一项所述的方法, 其中裂解剂是在包含变性剂如尿素、氯化胍、甲酰胺或甜菜碱和一种或多种缓冲液的裂解溶液的存在下加入。

13. 如权利要求4至12中任一项所述的方法, 其中步骤 (d) 是在低于99°C的温度下进行。

14. 如权利要求4至13中任一项所述的方法, 其中步骤 (c) 和 (e) 是通过施加洗涤溶液进行。

15. 如权利要求4至14中任一项所述的方法, 其中方法是在诸如微流体或基于柱的流动装置的流动装置内进行。

16. 如权利要求4至14中任一项所述的方法, 其中方法是在平板或微阵列设置内进行。

17. 如权利要求4至16中任一项所述的方法, 其中起始序列具有5至50个核苷酸长度, 例如10至30个核苷酸长度, 特别是约20个核苷酸长度。

18. 如权利要求4至17中任一项所述的方法, 其中起始序列是单链或双链的。

19. 如权利要求4至18中任一项所述的方法, 其中起始序列固定在固体载体上。

20. 如权利要求19所述的方法,其中起始序列通过可逆相互作用部分固定。

21. 如权利要求20所述的方法,其另外包括通过除去所述可逆相互作用部分来提取所得核酸。

22. 如权利要求20所述的方法,其中可逆相互作用部分是化学可切割的接头,例如二硫键、烯丙基或叠氮化物掩蔽的半缩醛胺醚。

23. 如权利要求22所述的方法,其另外包括通过切割化学可切割接头来提取所得核酸,例如通过加入三(2-羧乙基)膦(TCEP)、二硫苏糖醇(DTT)或钼络合物。

24. 如权利要求4至19中任一项所述的方法,其中起始序列含有至少一个限制性位点。

25. 如权利要求24所述的方法,其另外包括通过使用限制酶提取所得核酸。

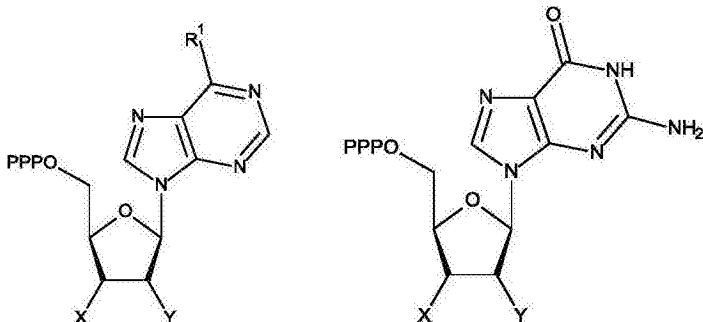
26. 如权利要求4至25中任一项所述的方法,其中起始序列含有至少一个尿苷。

27. 如权利要求4至26中任一项所述的方法,其另外包括例如通过PCR扩增得到的核酸。

28. 一种试剂盒在核酸合成方法中的用途,其中试剂盒包含如权利要求1至3中任一项定义的TdT,任选地与选自以下的一种或多种组分组合:起始序列,一种或多种3'-封闭的磷酸三核苷酸、无机焦磷酸酶,如来自酿酒酵母的纯化重组无机焦磷酸酶,以及裂解剂;进一步任选地按照权利要求4至27中任一项所述的方法使用所述试剂盒的说明书。

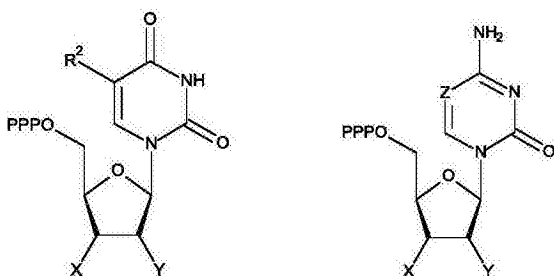
29. 如权利要求28所述的用途,其另外包含选自以下的一种或多种组分:延伸溶液、洗涤溶液和/或裂解溶液。

30. 3'-封闭的核苷酸三磷酸在非模板依赖性核酸合成方法中的用途,其中3'-封闭的核苷酸三磷酸是选自式(I)、(II)、(III)或(IV)的化合物:



(I)

(II)



(III)

(IV)

其中,

$R^1$ 表示 $NR^aR^b$ ,其中 $R^a$ 和 $R^b$ 独立地表示氢或 $C_{1-6}$ 烷基,

$R^2$ 表示氢、 $C_{1-6}$ 烷基、 $C_{1-6}$ 烷氧基、COH或COOH,

X表示C<sub>1-6</sub>烷基、NH<sub>2</sub>、N<sub>3</sub>或-OR<sup>3</sup>,

R<sup>3</sup>表示C<sub>1-6</sub>烷基、CH<sub>2</sub>N<sub>3</sub>、NH<sub>2</sub>或烯丙基,

Y表示氢,卤素或羟基,和

Z表示CR<sup>4</sup>或N,其中R<sup>4</sup>表示氢、C<sub>1-6</sub>烷基、C<sub>1-6</sub>烷氧基、COH或COOH。

31. 如权利要求30所述的用途,其中R<sup>a</sup>和R<sup>b</sup>均表示氢,或R<sup>a</sup>表示氢而R<sup>b</sup>表示甲基。

32. 如权利要求30或31所述的用途,其中R<sup>2</sup>表示氢、甲基或甲氧基。

33. 如权利要求30至32中任一项所述的用途,其中X表示-OR<sup>3</sup>,并且R<sup>3</sup>表示C<sub>1-6</sub>烷基, CH<sub>2</sub>N<sub>3</sub>、NH<sub>2</sub>或烯丙基。

34. 如权利要求30至33中任一项所述的用途,其中Y表示氢。

35. 如权利要求30至34中任一项所述的用途,其中Z表示CR<sub>4</sub>,并且其中R<sup>4</sup>表示甲氧基、COH或COOH。

36. 如权利要求30至35中任一项所述的用途,其中3'-封闭的核苷酸三磷酸选自E1至E11。

37. 无机焦磷酸酶在核酸合成方法中的应用。

38. 如权利要求37所述的用途,其中无机焦磷酸酶是来自酿酒酵母的纯化的重组无机焦磷酸酶。

## 新用途

### 技术领域

[0001] 本发明涉及特异性末端脱氧核苷酸转移酶(TdT)在核酸合成方法中的应用、合成核酸的方法以及包含所述酶的试剂盒在核酸合成方法中的应用。本发明还涉及末端脱氧核苷酸转移酶和3'-封闭的核苷酸三磷酸在非模板依赖性核酸合成方法中的应用。

### 背景技术

[0002] 核酸合成对现代生物技术至关重要。科学界人工合成DNA、RNA和蛋白质的能力使生物技术领域的快速发展成为可能。

[0003] 人工DNA合成——一个10亿英镑的、不断增长的市场——允许生物技术和制药公司开发一系列肽治疗剂,如用于治疗糖尿病的胰岛素。它使研究人员能够表征细胞蛋白质来开发新的小分子治疗方法,用于治疗当今老龄化人群所面临的疾病,如心脏病和癌症。它甚至为创建生命铺平了道路,正如Venter Institute在2010年将人工合成的基因组置于细菌细胞中时所展示的那样。

[0004] 然而,目前的DNA合成技术不能满足生物技术行业的需求。虽然DNA合成的好处众多,但是一个经常提及的问题阻碍了人工DNA合成工业的进一步增长,从而阻碍了生物技术领域的进一步发展。尽管是成熟的技术,但实际上不可能合成长度超过200个核苷酸的DNA链,并且大多数DNA合成公司只提供多达120个核苷酸。相比之下,平均的蛋白质编码基因的数量级为2000-3000个核苷酸,并且平均真核基因组数目是数十亿个核苷酸。因此,目前所有主要的基因合成公司都依赖于“合成和缝合”技术的变型,其中通过PCR合成重叠的40-60-mer片段并将其缝合在一起(参见Young,L.等人(2004)Nucleic Acid Res 32,e59)。基因合成工业提供的当前方法通常允许常规生产多达3kb长度。

[0005] DNA一次无法合成超过120-200个核苷酸的原因是由于目前的生成DNA的方法,它使用合成化学(即亚磷酸胺技术)将核苷酸一次偶联以制备DNA。由于每个核苷酸偶联步骤的效率为95.0-99.5%,因此在数学上不可能以可接受的产率合成长于200个核苷酸的DNA。Venter Institute花费4年和2000万美元来合成相对较小的细菌基因组(见Gibson,D.G.等人,(2010)Science 329,52-56),说明了这个艰巨的过程。

[0006] DNA测序的已知方法使用模板依赖性DNA聚合酶向生长的双链底物加入3'-可逆终止的核苷酸(参见Bentley,D.R.等人,(2008)Nature456,53-59)。在“按合成测序”过程中,每个加入的核苷酸含有染料,允许使用者识别模板链的确切序列。尽管双链DNA,但是这种技术能够产生500-1000bps长的链。然而,由于需要现有的核酸链作为模板,该技术不适用于从头核酸合成。

[0007] 已经对使用末端脱氧核苷酸转移酶用于受控的从头单链DNA合成进行了各种尝试(参见Ud-Dean,S.M.M.(2009)Syst Synth Boil 2,67-73,US 5,763,594和US 8,808,989)。不受控制的从头单链DNA合成,与受控制的相反,利用TdT在单链DNA上的脱氧核苷酸三磷酸(dNTP)3'拖尾性质产生例如用于下一代测序文库制备的均聚接头序列(参见Roychoudhury R.等人,(1976)Nucleic Acids Res 3,101-116和WO 2003/050242)。需要采用可逆的脱氧

核苷酸三磷酸终止技术来防止dNTP不受控地加入到正在生长的DNA链的3'末端。通过TdT开发受控的单链DNA合成方法对于基因组或杂交微阵列的原位DNA合成将是无价的,因为它消除了对无水环境的需要,并允许使用与有机溶剂不相容的各种聚合物(参见Blanchard, AP(1996) Biosens Bioelectron 11,687-690和US 7,534,561)。

[0008] 然而,TdT尚未被证明可有效地加入含有3'-O可逆终止部分的核苷酸三磷酸,以建立从头合成循环所必需的新生单链DNA链。3'-O可逆终止部分将阻止如TdT的末端转移酶催化正在生长的DNA链的3'末端和进入的核苷酸三磷酸的5'-三磷酸之间的核苷酸转移酶反应。本文呈现了数据,其表明来自小牛胸腺的广泛市售的重组TdT不能以定量的方式加入3'-O-封闭的三磷酸三磷酸(参见图3)。在以前的报告中,特别提到的TdT是来自小牛胸腺的重组TdT(参见Ud-Dean, SMM(2009) Syst Synth Boil 2,67-73, US 5,763,594和US 8,808,989),或者是使用不在脱氧核糖部分的3'端上定位的不同的可逆终止机制(参见US 8,808,989)。

[0009] 大多数DNA和RNA聚合酶含有高度选择性的糖立体门,用于严格区分脱氧核糖和核糖三磷酸底物(参见Joyce C.M. (1997) Proc Natl Acad Sci 94,1619-22)。这种糖立体门的结果是为了生物技术原因而发现和/或改造聚合酶以接受糖变体的巨大挑战,例如按合成顺序(见Metzker ML (2010) Nat Rev Genet 11,31-46和US 8460910)。发现接受3'-O可逆终止核苷酸的聚合酶的挑战如此之大,已经做出各种努力来产生可逆终止核苷酸,其中聚合酶终止机制位于终止核苷酸的含氮碱基上(参见Gardner, A.F. (2012) Nucleic Acids Res 40,7404-15和US 8,889,860)。

[0010] 因此,需要鉴定出可以容易地掺入3'-O可逆终止的核苷酸的末端脱氧核苷酸转移酶以及修饰的所述末端脱氧核苷酸转移酶,以通过可用于生物技术和单链DNA合成过程的方式掺入3'-O可逆终止的核苷酸,从而提供能够克服与当前可用方法相关的问题的改进的核酸合成方法。

## 发明内容

[0011] 根据本发明的第一方面,提供了包含选自以下的氨基酸序列的末端脱氧核苷酸转移酶(TdT)在核酸合成的方法中的用途:(a) SEQ ID NO:1至5和8中的任一个或与所述氨基酸序列具有至少20%序列同源性的功能性等同物或片段;或(b) SEQ ID NO:6的修饰的衍生物。

[0012] 根据本发明的第二方面,提供了一种核酸合成方法,其包括以下步骤:

[0013] (a) 提供起始序列(initiator sequence);

[0014] (b) 在本发明第一方面定义的末端脱氧核苷酸转移酶(TdT)的存在下,向所述起始序列中加入3'-封闭的核苷酸三磷酸;

[0015] (c) 从起始序列中除去所有试剂;

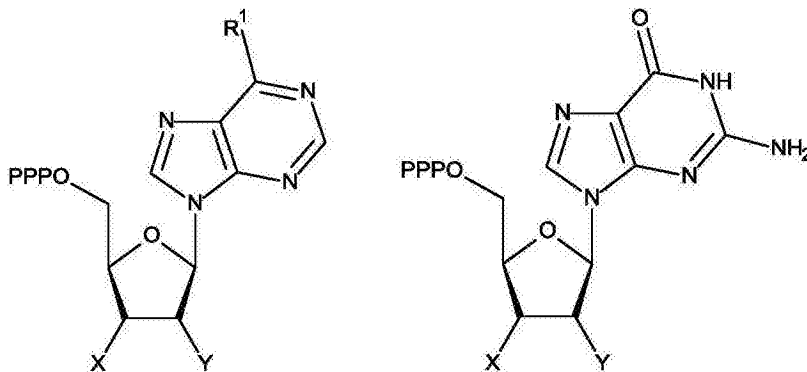
[0016] (d) 在裂解剂存在下,从3'-封闭的核苷酸三磷酸切割封闭基团;

[0017] (e) 去除裂解剂。

[0018] 根据本发明的另一方面,提供了试剂盒在核酸合成方法中的用途,其中该试剂盒包含如本发明的第一或第二方面中所定义的TdT,任选地与选自以下的一种或多种组分组合:起始序列,一个或多个3'-封闭的核苷酸三磷酸,无机焦磷酸酶,例如来自酿酒酵母的纯

化的重组无机焦磷酸酶,以及裂解剂;进一步任选包含根据本文定义的任何方法使用该试剂盒的说明书。

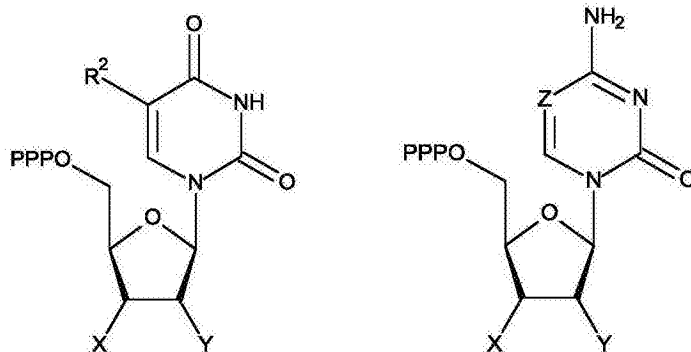
[0019] 根据本发明的另一方面,提供了3'-封闭的核苷酸三磷酸在非模板依赖性核酸合成的方法中的用途,其中3'-封闭的核苷酸三磷酸是选自式(I)、(II)、(III)或(IV)的化合物:



(I)

(II)

[0020]



(III)

(IV)

[0021] 其中,

[0022]  $R^1$ 表示 $NR^aR^b$ ,其中 $R^a$ 和 $R^b$ 独立地表示氢或 $C_{1-6}$ 烷基,[0023]  $R^2$ 表示氢、 $C_{1-6}$ 烷基、 $C_{1-6}$ 烷氧基、COH或COOH。[0024] X表示 $C_{1-6}$ 烷基、 $NH_2$ 、 $N_3$ 或 $-OR^3$ ,[0025]  $R^3$ 表示 $C_{1-6}$ 烷基、 $CH_2N_3$ 、 $NH_2$ 或烯丙基,

[0026] Y表示氢,卤素或羟基,和

[0027] Z表示 $CR^4$ 或N,其中 $R^4$ 表示氢、 $C_{1-6}$ 烷基、 $C_{1-6}$ 烷氧基、COH或COOH。

[0028] 根据本发明的另一方面,提供了无机焦磷酸酶在核酸合成方法中的用途。

## 附图说明

[0029] 图1:酶促DNA合成过程示意图。从图的顶部开始,将具有去保护的3'-末端的固定DNA链暴露于由TdT、碱基特异性3'-封闭核苷酸三磷酸盐、用于减少无机焦磷酸盐积累的无机焦磷酸酶以及用于最佳酶活性和稳定性的合适的缓冲液/盐组成的延伸混合物。该蛋白质将一个受保护的核苷酸加入到固定的DNA链上(图的底部)。然后用洗涤混合物除去延伸

混合物并任选地回收。然后用裂解混合物洗涤固定的(n+1) DNA链以切割3'-保护基团,使得能够在下一个循环中进行反应。在裂解混合物中,可能存在变性剂以破坏任何二级结构。在该步骤期间,可以升高温度以辅助二级结构的裂解和破坏。用洗涤混合物处理固定的DNA以除去残留的裂解混合物。可以用适当的核苷酸三磷酸重复步骤1-4,直到达到所需的寡核苷酸序列。

[0030] 图2:TdT向引发剂链加入3'-不可逆封闭的核苷酸三磷酸。(A)将单链起始DNA与15U欧洲普通牛(*Bos taurus*) TdT、所需盐(50mM乙酸钾,20mM Tris乙酸盐pH 7.9,1mM氯化钴)和3'-O-甲基dTTP在37°C的一起温育指定的时间量。3'-不可逆封闭的核苷酸三磷酸的浓度为1mM,并且起始DNA为200pM,以1:5000的比例进行,以促进核苷酸的加入。用EDTA(0.5M)终止反应。(B)与(A)类似,除了使用3'-叠氮基dTTP作为3'-封闭的核苷酸三磷酸。

[0031] 图3:各种TdT直向同源物将3'可逆受保护的(3'-O-叠氮甲基)核苷酸三磷酸加入至起始链,比欧洲普通牛TdT快3.8倍。(A)在上述缓冲液和反应条件下,将单链起始DNA与具有指定物种的TdT孵育60分钟。然后通过毛细管电泳分析反应,并出色谱图子集。(B)在1小时内转化成N+1种(加入3'-O-叠氮甲基dTTP)的起始DNA的部分。(C)三个TdT直向同源物与3'-O-叠氮甲基dTTP和3'-O-叠氮甲基dCTP反应20分钟。与欧洲普通牛TdT相比,斑点雀鳢(*Lepisosteus oculatus*)和袋獾(*Sarcophilus harrisii*) TdT持续表现更好。TdT直向同源物加入dTTP和dCTP核苷酸类似物两者的能力证明了DNA序列特异性的控制。

[0032] 图4:斑点雀鳢(*Lepisosteus oculatus*) TdT的改造变体显示出比野生型斑点雀鳢TdT更好的活性。将起始DNA与1mM 3'-可逆封闭的dNTP以及如前所述所需的盐在37°C下孵育20分钟。通过PAGE测量dNTP的加入,并且绘制相对于野生型TdT的转化率。

[0033] 图5:斑点雀鳢的改造变体显示出比野生型TdT(参见图3)和市售的欧洲普通牛TdT巨大改善的活性。TdT以与前述相似的方式与起始DNA以及3'-可逆封闭的dNTP一起温育。欧洲普通牛TdT的表现优于斑点雀鳢TdT的改造变体,如PAGE凝胶所示。

[0034] 图6:无机焦磷酸盐是TdT介导的从头序列特异性DNA合成所必需的。将核酸引发剂与酿酒酵母(*Saccharomyces cerevisiae*)无机焦磷酸酶和欧洲普通牛TdT或不与酿酒酵母无机焦磷酸酶和欧洲普通牛TdT一起在37°C,50mM乙酸钾,20mM Tris乙酸盐pH 7.9,1.5mM氯化钴中培养。以1mM引入3'-O-叠氮甲基dTTP(N3Me-dTTP),并以100μM引入双脱氧TTP(ddTTP)。通过PAGE分析反应。没有无机焦磷酸酶,链歧化占主导地位,并且发生序列特异性的灾难性损失。

[0035] 图7:DNA合成中使用的基于色谱柱的流式仪器的简化示意图。计算机(302)控制两个泵和溶液混合室(311)。泵1(304)选择性地将延伸溶液(301)、洗涤溶液(305)或裂解溶液(310)泵送到混合室中。泵2(306)选择性地将含有3'-封闭的A(腺嘌呤)TP(303),T(胸腺嘧啶)TP(307),G(鸟嘌呤)TP(308)的单个3'-封闭核苷酸三磷酸(TP)或C(胞嘧啶)TP(309)泵送进入室中。计算机控制的混合室然后将来自泵1和泵2的适当的溶液比率传递到基于柱(column)的DNA合成室(312)中。加热元件(313)确保DNA合成柱保持在用于合成过程发生的必要温度。在离开DNA合成室时,反应溶液或者进入供将来使用的循环容器(314)、或者废物容器(316)或者移动到用于扩增所得DNA的聚合酶链式反应(PCR)步骤(315)。PCR完成产生终产物(317)。



## 具体实施方式

[0036] 根据本发明的第一方面,提供了包含选自以下的氨基酸序列的末端脱氧核苷酸转移酶(TdT)在核酸合成的方法中的用途:(a) SEQ ID NO:1至5和8中的任一个,或与所述氨基酸序列具有至少20%序列同源性的其功能性等同物或片段;或(b) SEQ ID NO:6的修饰的衍生物。

[0037] 根据本发明的可以提及的一个特定方面,提供了包含选自以下的氨基酸序列的末端脱氧核苷酸转移酶(TdT)在核酸合成的方法中的用途:(a) SEQ ID NO:1至5中的任一个,或与所述氨基酸序列具有至少20%序列同源性的其功能性等同物或片段;或(b) SEQ ID NO:6的修饰的衍生物。

[0038] 本发明涉及从未研究过的末端脱氧核苷酸转移酶的鉴定,其出人意料地具有将大的3'-0可逆终止部分掺入脱氧核苷酸三磷酸的能力。

[0039] 由于来源于小牛胸腺的市售重组TdT不容易掺入3'-0可逆终止的核苷酸,所以本发明人已经定位了作为DNA聚合酶的末端脱氧核苷酸转移酶,从而可逆地接受3'-0终止核苷酸,例如用3'-0-叠氮基甲基修饰的dNTP。

[0040] 此外,本发明涉及改造的末端脱氧核苷酸转移酶,其实现了含有将用于受控的从头合成单链DNA的3'-0可逆终止部分的dNTP的掺入率的实质增加。

[0041] 由于受控的从头单链DNA合成是一种加入方法,偶联效率对于获得用于诸如基因组组装或杂交微阵列的应用的最终单链DNA产物的实际有用产率是非常重要的。因此,本发明涉及具有加入3'-0可逆终止的核苷酸的能力的TdT直向同源物的鉴定,以及TdT直向同源物的改造变体,其以实际上可用于单链DNA合成过程的定量方式加入3'-0可逆终止的核苷酸。

[0042] 本文所述的用途具有显著的优点,例如快速产生长的DNA长度的能力同时仍保持高产率并且不使用任何有毒的有机溶剂。

[0043] 在一个实施方案中,末端脱氧核苷酸转移酶(TdT)包含选自SEQ ID NO:1至5和8中任一项的氨基酸序列,或与所述氨基酸序列具有至少20%序列同源性的其功能性等同物或片段。

[0044] 在另一个实施方案中,末端脱氧核苷酸转移酶(TdT)包含选自SEQ ID NO:1至5中任一项的氨基酸序列,,或与所述氨基酸序列具有至少20%序列同源性的其功能性等同物或片段。

[0045] 在另一个实施方案中,末端脱氧核苷酸转移酶(TdT)包含选自SEQ ID NO:1的氨基酸序列。SEQ ID NO:1的氨基酸序列是来自袋獾(UniProt)的末端脱氧核苷酸转移酶(TdT) UniProt:G3VQ55)。袋獾(又名塔斯马尼亚恶魔)是袋鼬科(Dasyuridae)的食肉动物,现在只在澳大利亚岛屿塔斯马尼亚州的野外发现。

[0046] 在另一个实施方案中,末端脱氧核苷酸转移酶(TdT)包含选自SEQ ID NO:2的氨基酸序列。SEQ ID NO:2的氨基酸序列是来自斑点雀鳉的末端脱氧核苷酸转移酶(TdT) (UniProt:W5MK82)。斑点雀鳉(也称为斑点雀鳉(spotted gar))是雀鳉科(Lepisosteidae)科的原始淡水鱼,原产于北美,从伊利湖和南密歇根湖南部经密西西比河流域向南到海湾斜坡排水系统,从美国佛罗里达州的Apalachicola河下游到得克萨斯州的纽约河。

[0047] 在另一个实施方案中,末端脱氧核苷酸转移酶(TdT)包含选自SEQ ID NO:3的氨基

酸序列。SEQ ID NO:3的氨基酸序列是来自长尾龙猫 (*Chinchilla lanigera*) 的末端脱氧核苷酸转移酶 (TdT) 序列 (NCBI参考序列:XP\_005407631.1;<http://www.ncbi.nlm.nih.gov/protein/533189443>)。长尾龙猫 (又名长尾栗鼠,智利,沿海,常见的栗鼠或较小的栗鼠) 是栗鼠属的两种啮齿类动物之一,另一种是龙猫 (*Chinchilla chinchilla*)。

[0048] 在另一个实施方案中,末端脱氧核苷酸转移酶 (TdT) 包含选自SEQ ID NO:4的氨基酸序列。SEQ ID NO:4的氨基酸序列是来自小耳大婴猴 (*Otolemur garnettii*) 的末端脱氧核苷酸转移酶 (TdT) (UniProt:A4PCE6)。小耳大婴猴 (也被称为北部大型婴猴属,加内特的大型婴猴属或小耳大型婴猴属) 是一种夜行的非洲野生灵长类动物。

[0049] 在另一个实施方案中,末端脱氧核苷酸转移酶 (TdT) 包含选自SEQ ID NO:5的氨基酸序列。SEQ ID NO:5的氨基酸序列是来自野猪 (*Sus scrofa*) 的末端脱氧核苷酸转移酶 (TdT) 序列 (UniProt:F1SBG2)。野猪 (也被称为野猪 (boar), 野猪 (swine) 或欧亚野猪 (pig)) 是欧亚大陆、北非和大巽他群岛的suid土著的。

[0050] 在另一个实施方案中,末端脱氧核苷酸转移酶 (TdT) 包含选自SEQ ID NO:8的氨基酸序列。SEQ ID NO:8的氨基酸序列是SEQ ID NO:2的变体,其已通过改变氨基酸序列被改造用于改善活性。在实施例3和图4中提供了数据,其表明了改造变体如SEQ ID NO:8相比于野生型SEQ ID NO:2的益处。

[0051] 在另一个实施方案中,末端脱氧核苷酸转移酶 (TdT) 包含选自SEQ ID NO:1、2或8的氨基酸序列。

[0052] 在另一个实施方案中,末端脱氧核苷酸转移酶 (TdT) 包含选自SEQ ID NO:1或2的氨基酸序列。实施例2和图3提供了数据,其表明了相对于来自欧洲普通牛的天然重组TdT酶的有益结果。

[0053] 在替代实施方案中,末端脱氧核苷酸转移酶 (TdT) 包含选自SEQ ID NO:6的修饰的衍生物 (即SEQ ID NO:6的非天然的、突变的衍生物) 的氨基酸序列。SEQ ID NO:6的氨基酸序列是来自欧洲普通牛的末端脱氧核苷酸转移酶 (TdT) 序列 (UniProt:P06526)。欧洲普通牛 (也称为牛 (cattle) 或口语牛 (cow)) 是最常见的大型驯养有蹄类动物。它们是牛 (*Bovinae*) 亚科的突出的现代成员,是牛 (*Bos*) 属最广泛的种类。

[0054] 本文中对“TdT”的参考是指末端脱氧核苷酸转移酶 (TdT), 并且包括对所述酶的纯化和重组形式的参考。TdT也通常称为DNIT (DNA核苷酸转移酶), 任何这样的术语都应该互换使用。

[0055] 本文中提及的“核酸合成方法”包括合成DNA (脱氧核糖核酸) 或RNA (核糖核酸) 长度的方法,其中通过加入另外的核苷酸 (n+1) 延伸核酸 (n) 的链。在一个实施方案中,核酸是DNA。在替代实施方案中,核酸是RNA。

[0056] 本文中提及的“DNA合成方法”是指通过加入另外的核苷酸 (n+1) 延伸DNA链 (n) 的DNA链合成方法。本文描述的方法提供了本发明的末端脱氧核苷酸转移酶和3' -可逆封闭的核苷酸三磷酸在从头DNA链合成中依次加入核苷酸的新用途,其具有几个优于本领域当前已知的DNA合成方法的优点。

[0057] 目前用于偶联核苷酸以形成序列特异性DNA的合成方法已经达到渐近长度极限,因此需要从头DNA合成的新方法。合成的DNA的合成方法也具有使用对环境有害的有毒有机溶剂和添加剂 (例如乙腈、乙酸酐、三氯乙酸、吡啶等) 的缺点。

[0058] 备选地,核酸合成的酶促方法是合乎需要的。天然酶如DNA聚合酶能够在分离前加入50,000个核苷酸。然而,DNA聚合酶需要模板链,从而破坏从头链合成的目的。

[0059] 然而,在脊椎动物中发现了能够进行非模板依赖性DNA合成的DNA聚合酶,称为TdT。鉴于游离的3'末端和核苷酸三磷酸,来自欧洲普通牛和小鼠(*Mus musculus*)的重组TdT显示出在DNA链的3'末端加入10到数百个核苷酸。如Basu, M等人的一篇文章所示(*Biochem. Biophys. Res. Commun.* (1983) 111, 1105-1112) TdT将不可控地将核苷酸三磷酸加入到DNA链的3'末端。然而,这种不受控制的加入不适用于需要序列特异性寡核苷酸的受控从头链合成。因此,市售重组TdT主要是分子生物学家采用有用的化学标签标记DNA的工具。

[0060] 本发明人已经发现了几种欧洲普通牛TdT的直系同源物,其与3'-可逆保护的核苷酸三磷酸相结合,能够以受控的方式合成DNA。欧洲普通牛TdT在将3'-保护基掺入核苷酸三磷酸中并不有效,这可能是由于TdT活性位点的空间问题所致。本文在实施例2和图3中给出了数据,其示出了欧洲普通牛TdT的直向同源物(例如本发明第一方面的纯化重组TdT(特别是SEQ ID NO:1和2))在掺入3'-OH封闭的核苷酸三磷酸,从而使非模板依赖性的核酸链的序列特异性合成更有效。

[0061] 这种酶促方法意味着该方法具有能够产生超过目前的合成DNA合成方法的120-200核苷酸极限的DNA链的特别优点。此外,这种酶促法避免了使用可能对环境有害的强有机溶剂的需要。

[0062] 应当理解,术语“功能性等同物”是指与本发明第一方面的TdT的确切序列不同,但是可以发挥相同功能的多肽,即以模板依赖的方式催化将核苷酸三磷酸加入到DNA链的3'末端。

[0063] 在一个实施方案中,末端脱氧核苷酸转移酶(TdT)是TdT的非天然衍生物,例如本发明第一方面的TdT的功能片段或同系物。

[0064] 本文中对“片段”的参考包括例如具有C-末端截短或具有N-末端截短的功能片段。片段长度上适当地大于10个氨基酸,例如大于15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、100、110、120、130、140、150、160、170、180、190、200、210、220、230、240、250、260、270、280、290、300、310、320、330、340、350、360、370、380、390、400、410、420、430、440、450、460、470、480、490或500个氨基酸长度。

[0065] 应当理解,本文中“同源性”的参考应理解为表示两个蛋白质序列之间的百分比同一性,例如:SEQ ID NO:X和SEQ ID NO:Y,其是在比对序列SEQ ID NO:X和SEQ ID NO:Y之间的共同氨基酸之和除以SEQ ID NO:X或SEQ ID NO:Y中较短长度,以百分比表示。

[0066] 在一个实施方案中,末端脱氧核苷酸转移酶(TdT)与本发明第一方面的TdT具有至少25%的同源性,例如至少30%、40%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、96%、97%、98%或99%的同源性。

[0067] 根据本发明的第二方面,提供了一种核酸合成方法,其包括以下步骤:

[0068] (a) 提供起始序列;

[0069] (b) 在本发明第一方面定义的末端脱氧核苷酸转移酶(TdT)的存在下,向所述起始序列中加入3'-封闭的核苷酸三磷酸;

[0070] (c) 从起始序列中除去所有试剂;

[0071] (d) 在裂解剂存在下,从3'-封闭的核苷酸三磷酸切割封闭基团;

[0072] (e) 去除裂解剂。

[0073] 在一个实施方案中,步骤(c)包括去除脱氧核苷酸三磷酸和TdT,例如TdT。因此,根据本发明的一个特定方面,提供了一种核酸合成方法,其包括以下步骤:

[0074] (a) 提供起始序列;

[0075] (b) 在本发明第一方面定义的末端脱氧核苷酸转移酶(TdT)的存在下,向所述起始序列中加入3'-封闭的核苷酸三磷酸;

[0076] (c) 去除TdT;

[0077] (d) 在裂解剂存在下,从3'-封闭的核苷酸三磷酸切割封闭基团;

[0078] (e) 去除裂解剂。

[0079] 应当理解,可以多次重复该方法的步骤(b)至(e),以产生所需长度的DNA或RNA链。因此,在一个实施方案中,通过重复步骤(b)至(e)将大于1个核苷酸加入到起始序列中,例如大于5、10、20、30、40、50、60、70、80、90、100、110或120个核苷酸加入到起始序列。在另一个实施方案中,加入大于200个核苷酸,例如大于300、400、500、600、700、800、900、1000、1250、1500、1750、2000、2250、2500、2750、3000、4000、5000、6000、7000、8000、9000或10000个核苷酸。

[0080] 3'-封闭的核苷酸三磷酸

[0081] 本文中提及的“核苷酸三磷酸”是指含有与三个磷酸酯基团结合的核苷(即与脱氧核糖或核糖分子连接的碱基)的分子。含有脱氧核糖的核苷酸三磷酸的实例是:脱氧腺苷三磷酸(dATP)、脱氧鸟苷三磷酸(dGTP)、脱氧胞苷三磷酸(dCTP)或脱氧胸苷三磷酸(dTTP)。含有核糖的核苷酸三磷酸的实例是:腺苷三磷酸(ATP)、鸟苷三磷酸(GTP)、胞苷三磷酸(CTP)或尿苷三磷酸(UTP)。其他类型的核苷可以与三个磷酸酯基团结合以形成核苷酸三磷酸,例如天然存在的修饰核苷和人工核苷。

[0082] 因此,本文中对“3'-封闭的核苷酸三磷酸”的参考是指在3'末端具有额外的基团的核苷酸三磷酸(例如dATP、dGTP、dCTP或dTTP),其防止进一步加入核苷酸,即通过用保护基替代3'-OH基。

[0083] 应当理解,本文中对“3'封闭”,“3'-封闭基团”或“3'保护基团”的参考是指连接于核苷酸三磷酸3'末端的基团,其防止进一步的核苷酸加入。本方法使用可逆的3'-封闭基团,其可通过切割而被除去,以允许加入额外的核苷酸。相比之下,不可逆的3'-封闭基团是指其中3'-OH基团既不能暴露也不会被切割解包覆的dNTP。

[0084] 存在几种已证实的可逆保护基团,例如叠氮基甲基、氨基和烯丙基,其可以应用于本文所述的方法。Greene's Protective Groups in Organic Synthesis, (Wuts, P.G.M.&Greene, T.W. (2012) 第4版, John Wiley&Sons)中描述了合适的保护基团的实例。

[0085] 在一个实施方案中,3'-封闭的核苷酸三磷酸被可逆保护基封闭。在替代的实施方案中,3'-封闭的核苷酸三磷酸被不可逆保护基团封闭。

[0086] 因此,在一个实施方案中,3'-封闭的核苷酸三磷酸被3'-O-甲基、3'-叠氮基、3'-O-叠氮基甲基、3'-氨基或3'-O-烯丙基基团封闭。在另一个实施方案中,3'-封闭的核苷酸三磷酸被3'-O-叠氮基甲基、3'-氨基或3'-O-烯丙基基团封闭。在另一个实施方案中,3'-封闭的核苷酸三磷酸被3'-O-甲基或3'-叠氮基封闭。

[0087] 裂解剂

[0088] 本文所指的“裂解剂”是指能够从3'-封闭的核苷酸三磷酸切割3'-封闭基团的物质。

[0089] 本文所述的3'-封闭基团可以全部用已记录的可用作裂解剂的化合物在水溶液中定量除去(例如,参见:Wuts, PGM&Greene, TW (2012) 第四版, John Wiley&Sons; Hutter, D. 等人(2010) Nucleosides Nucleotides Nucleic Acids 29, 879-895; EP 1560838和US 7, 795, 424)。

[0090] 在一个实施方案中,裂解剂是化学裂解剂。在替代的实施方案中,裂解剂是酶裂解剂。

[0091] 本领域技术人员将理解,裂解剂的选择取决于所用的3'-核苷酸封闭基团的类型。例如,三(2-羧乙基)膦(TCEP)可用于切割3'-O-叠氮基甲基,钯络合物可用于切割3'-O-烯丙基,或亚硝酸钠可用于切割3'-氨基。因此,在一个实施方案中,裂解剂选自:三(2-羧乙基)膦(TCEP)、钯络合物或亚硝酸钠。

[0092] 在一个实施方案中,在包含变性剂如尿素、氯化胍(guanidinium chloride)、甲酰胺或甜菜碱的裂解溶液的存在下加入裂解剂。加入变性剂具有能够破坏DNA中任何不希望的二级结构的优点。在另一个实施方案中,裂解溶液包含一种或多种缓冲液。本领域技术人员将理解,缓冲液的选择取决于所需的确切的裂解化学品和裂解剂。

[0093] 起始序列

[0094] 本文所指的“起始序列”是指具有游离的3'末端、可以连接3'-封闭的核苷酸三磷酸的短寡核苷酸。在一个实施方案中,起始序列是DNA启动子序列。在替代的实施方案中,起始序列是RNA启动子序列。

[0095] 本文所指的“DNA起始序列”是指可以连接3'-封闭的核苷酸三磷酸的DNA小序列,即将从DNA起始序列末端合成DNA。

[0096] 在一个实施方案中,起始序列为5至50个核苷酸长,例如5至30个核苷酸长(即10至30之间),特别是5至20个核苷酸长(即约20个核苷酸长),更特别是5至15个核苷酸长,例如10至15个核苷酸长,特别是12个核苷酸长。

[0097] 在一个实施方案中,起始序列具有以下序列:5' CGTTAACATATT-3' (SEQ ID NO:7)。

[0098] 在一个实施方案中,起始序列是单链的。在替代的实施方案中,起始序列是双链的。本领域技术人员将理解,3'突出端(即,游离的3'端)允许有效的加入。

[0099] 在一个实施方案中,起始序列被固定在固体载体上。这允许在不洗去合成的核酸的情况下除去TdT和裂解剂(分别在步骤(c)和(e)中)。起始序列可以连接到在水性条件下稳定的固体载体上,使得该方法可以通过流动设置容易地执行。

[0100] 在一个实施方案中,起始序列通过可逆相互作用部分例如化学切割的接头,抗体/免疫原性表位,生物素/生物素结合蛋白(例如抗生物素蛋白或链霉抗生物素蛋白)或谷胱甘肽-GST标签固定到固体载体上。因此,在另一个实施方案中,该方法另外包括通过除去起始序列中的可逆相互作用部分,例如通过用蛋白酶K孵育来提取所得到的核酸。

[0101] 在另一个实施方案中,起始序列经由化学可切割的接头如二硫键、烯丙基或叠氮掩蔽的半缩醛胺醚接头固定在固体载体上。因此,在一个实施方案中,该方法另外包括通过加入用于二硫链接头的三(2-羧乙基)膦(TCEP)或二硫苏糖醇(DTT)裂解化学接头来提取所得核酸;烯丙基接头的钯络合物;或用于叠氮掩蔽的半缩醛胺醚接头的TCEP。

[0102] 在一个实施方案中,使用与固体载体结合的核酸作为模板,通过聚合酶链反应提取所得核酸并进行扩增。因此,起始序列可以包含合适的正向引物序列,并且可以合成适当的反向引物。

[0103] 在替代的实施方案中,固定的起始序列含有至少一个限制性位点。因此,在另一个实施方案中,该方法另外包括通过使用限制酶提取所得核酸。

[0104] 使用限制酶和限制性在特定位点切割核酸是本领域公知的。限制性位点和酶的选择可以取决于所需的性质,例如是否需要“平”或“粘”末端。限制酶的实例包括:AluI、BamHI、EcoRI、EcoRII、EcoRV、HaeII、HgaI、HindIII、HinfI、NotI、PstI、PvuII、SalI、Sau3A、ScaI、SmaI、TaqI和XbaI。

[0105] 在替代实施方案中,起始序列含有至少一个尿苷。用尿嘧啶-DNA糖基化酶(UDG)处理产生无碱基位点。用无嘌呤/无嘧啶(AP)位点核酸内切酶处理适当的底物将提取核酸链。

[0106] 核酸合成方法

[0107] 在一个实施方案中,本发明的末端脱氧核苷酸转移酶(TdT)在包含一种或多种缓冲液(例如,Tris或二甲基胍酸盐)、一种或多种盐(例如 $\text{Na}^+$ 、 $\text{K}^+$ 、 $\text{Mg}^{2+}$ 、 $\text{Mn}^{2+}$ 、 $\text{Cu}^{2+}$ 、 $\text{Zn}^{2+}$ 、 $\text{Co}^{2+}$ 等,均具有适当的平衡离子,如 $\text{Cl}^-$ )和无机焦磷酸酶(例如,酿酒酵母同源物)的延伸溶液存在下加入。应当理解,缓冲液和盐的选择取决于最佳的酶活性和稳定性。

[0108] 无机焦磷酸酶的使用有助于减少由于TdT的核苷酸三磷酸水解引起的焦磷酸盐积聚。因此,使用无机焦磷酸酶具有降低(1)反向反应速率和(2)TdT链歧化(dismutation)速率的优点。因此,根据本发明的另一方面,提供了无机焦磷酸酶在核酸合成方法中的用途。在实施例5和图6中给出了数据,其显示出在核酸合成期间使用无机焦磷酸酶的益处。在一个实施方案中,无机焦磷酸酶包含来自酿酒酵母的纯化的重组无机焦磷酸酶。

[0109] 在一个实施方案中,步骤(b)是在5和10之间的pH范围内进行。因此,可以理解,可以使用具有pH 5-10的缓冲范围的任何缓冲液,例如二甲基胍酸盐、Tris、HEPES或Tricine,特别是二甲基胍酸盐或Tris。

[0110] 在一个实施方案中,步骤(d)是在低于99°C的温度下进行,例如低于95°C、90°C、85°C、80°C、75°C、70°C、65°C、60°C、55°C、50°C、45°C、40°C、35°C或30°C。应当理解,最佳温度将取决于所使用的裂解剂。所使用的温度有助于帮助切割并破坏在核苷酸加入期间形成的任何二级结构。

[0111] 在一个实施方案中,步骤(c)和(e)是通过施加洗涤溶液进行。在一个实施方案中,洗涤溶液包含与本文所述的延伸溶液中使用的相同的缓冲液和盐。这具有当重复方法步骤时允许在步骤(c)后收集洗涤溶液并且作为步骤(b)中的延伸溶液再循环的优点。

[0112] 在一个实施例中,该方法在如图7所示的流动(flow)仪器内执行,例如微流体或基于柱的流动装置。在使得该方法简单使用的流程(flow)设置中,本文描述的方法可以容易地进行。应当理解,可以对所需的反应条件优化市售DNA合成器(例如来自BioAutomation的MerMade 192E或来自K&A的H-8 SE)的实例,并用于执行本文所述的方法。

[0113] 在一个实施方案中,该方法在平板或微阵列设置上进行。例如,可以使用任何适用的喷射技术(包括压电和热射流)通过一系列微分散喷嘴单独地处理核苷酸。这种高度平行的方法可用于产生杂交微阵列,并且还适于通过标准分子生物学技术进行DNA片段组装。

[0114] 在一个实施方案中,该方法另外包括扩增所得到的核酸。DNA/RNA扩增的方法是本

领域公知的。例如,在另一个实施方案中,通过聚合酶链反应(PCR)进行扩增。该步骤的优点是能够在一步骤中提取和扩增所得到的核酸。

[0115] 本文描述的非模板依赖性核酸合成方法具有将限定组成和长度的核酸序列加入到起始序列的能力。因此,本领域技术人员将理解,本文所述的方法可以用作将接头序列引入核酸文库的新方法。

[0116] 如果起始序列不是一个限定的序列,而是核酸片段的文库(例如通过基因组DNA的超声处理产生,或者例如信使RNA),则该方法能够在每个片段上从头合成“接头序列”。接头序列的安装是下一代文库核酸测序方法的水库制备的组成部分,因为它们含有允许与流动池/固相载体杂交以及测序引物杂交的序列信息。

[0117] 目前使用的方法包括单链连接,但是这种技术是受限的,因为连接效率随着片段长度的增加而强烈降低。因此,目前的方法不能结合长度超过100个核苷酸的序列。因此,本文描述的方法允许以目前可能的方式改进文库制备。

[0118] 因此,在一个实施方案中,将接头序列加入到起始序列。在另一个实施方案中,起始序列可以是来自核酸片段文库的核酸。

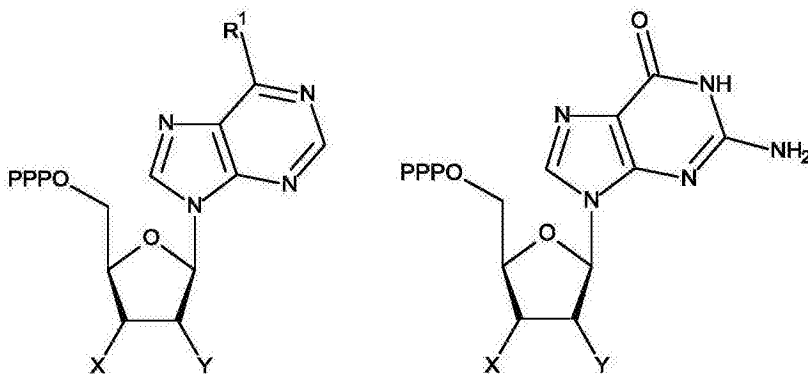
[0119] 试剂盒

[0120] 根据本发明的另一方面,提供了试剂盒在核酸合成方法中的用途,其中试剂盒包含如本发明的第一或第二方面中所定义的TdT,任选地与选自以下的一种或多种组分组合:起始序列、一个或多个3'-封闭的核苷酸三磷酸、无机焦磷酸酶,例如来自酿酒酵母的纯化重组无机焦磷酸酶,以及裂解剂;进一步任选包含根据本文定义的任何方法使用试剂盒的说明书。

[0121] 适当地,根据本发明的试剂盒还可以含有一种或多种选自以下的组分:如本文所定义的延伸溶液、洗涤溶液和/或裂解溶液;任选包含根据本文定义的任何方法使用试剂盒的说明书。

[0122] 3'-封闭核苷酸三磷酸的应用

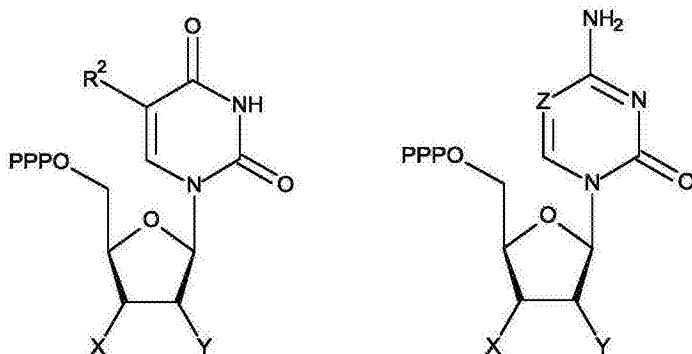
[0123] 根据本发明的另一方面,提供3'-封闭的核苷酸三磷酸在非模板依赖性核酸合成的方法中的应用,其中3'-封闭的核苷酸三磷酸是选自式(I)、(II)、(III)或(IV)的化合物:



(I)

(II)

[0124]



(III)

(IV)

[0125] 其中,

[0126]  $R^1$ 表示 $NR^aR^b$ ,其中 $R^a$ 和 $R^b$ 独立地表示氢或 $C_{1-6}$ 烷基,[0127]  $R^2$ 表示氢、 $C_{1-6}$ 烷基、 $C_{1-6}$ 烷氧基、COH或COOH。[0128] X表示 $C_{1-6}$ 烷基、 $NH_2$ 、 $N_3$ 或 $-OR^3$ ,[0129]  $R^3$ 表示 $C_{1-6}$ 烷基、 $CH_2N_3$ 、 $NH_2$ 或烯丙基,

[0130] Y表示氢,卤素或羟基,和

[0131] Z表示 $CR^4$ 或N,其中 $R^4$ 表示氢、 $C_{1-6}$ 烷基、 $C_{1-6}$ 烷氧基、COH或COOH。

[0132] 本文所指的“非模板依赖性核酸合成方法”是指不需要模板DNA/RNA链的核酸合成方法,即核酸可以从头合成。

[0133] 在一个实施方案中,核酸是DNA。本文所指的“非模板依赖性DNA合成方法”是指不需要模板DNA链的DNA合成方法,即可以从头合成DNA。在替代实施方案中,核酸是RNA。

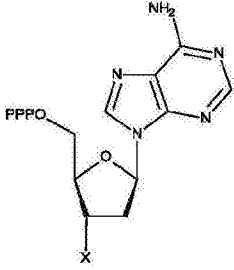
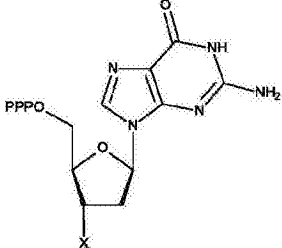
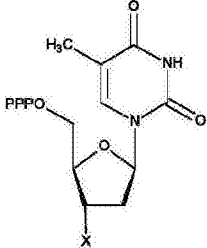
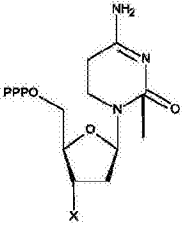
[0134] 应当理解,本文所示结构中的“PPP”表示三磷酸基团。

[0135] 本文中用作基团或部分基团的术语“ $C_{1-6}$ 烷基”是指含有1至6个碳原子的直链或支链饱和烃基。这些基团的实例包括甲基、乙基、丁基、正丙基、异丙基等。[0136] 本文所用的术语“ $C_{1-6}$ 烷氧基”的术语是指通过单键(即R-O)键合至氧的烷基。这些参考包括具有1至6个碳原子的直链和支链烷基链的那些,例如甲氧基(或甲基氧基)、乙氧基、正丙氧基、异丙氧基、正丁氧基和2-甲基丙氧基。[0137] 本文所用的术语“烯丙基”是指具有结构式 $RCH_2-CH=CH_2$ 的取代基,其中R是分子的其余部分。它由连接到乙烯基( $-CH=CH_2$ )的甲基( $-CH_2-$ )组成。

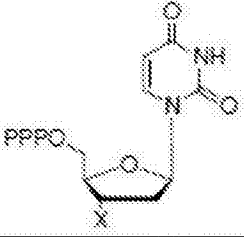
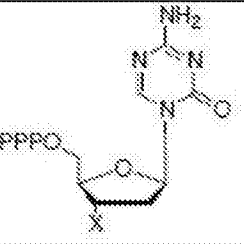
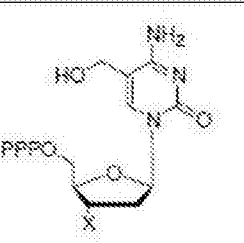
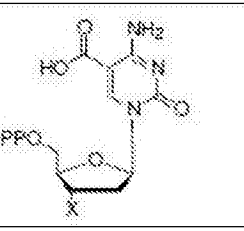
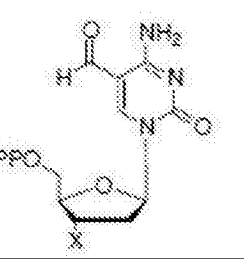
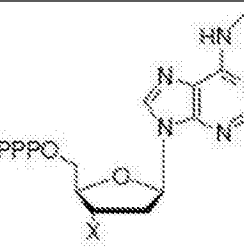
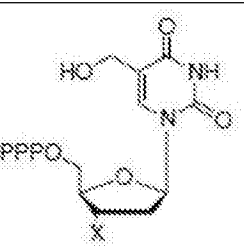


- [0138] 术语“COOH”或“CO<sub>2</sub>H”是指由羰基(C=O)和羟基(O-H)基团组成的羧基基团(或羧基)。术语“COH”是指由与氢键合的羰基(C=O)基团组成的甲酰基。
- [0139] 术语“N<sub>3</sub>”(结构上由-N=N<sup>+</sup>=N)是指叠氮基。
- [0140] 在一个实施方案中,R<sup>a</sup>和R<sup>b</sup>都表示氢(即,R<sup>1</sup>表示NH<sub>2</sub>)。
- [0141] 在替代实施方案中,R<sup>a</sup>表示氢,R<sup>b</sup>表示甲基(即,R<sup>1</sup>表示NHCH<sub>3</sub>)。
- [0142] 在一个实施方案中,R<sup>2</sup>表示氢、甲基或甲氧基。在另一个实施方案中,R<sup>2</sup>表示氢。在替代实施方案中,R<sup>2</sup>表示甲基。在另一个替代实施方案中,R<sup>2</sup>表示甲氧基。
- [0143] 在一个实施方案中,X表示-OR<sup>3</sup>,且R<sup>3</sup>表示C<sub>1-6</sub>烷基、CH<sub>2</sub>N<sub>3</sub>、NH<sub>2</sub>或烯丙基。
- [0144] 在替代实施方案中,X表示C<sub>1-6</sub>烷基(例如甲基)或N<sub>3</sub>。
- [0145] 在一个实施方案中,Y表示氢或羟基。
- [0146] 在一个实施方案中,Y表示氢。
- [0147] 在替代实施方案中,Y表示羟基。
- [0148] 在替代实施方案中,Y表示卤素,例如氟。
- [0149] 在一个实施方案中,Z表示N。
- [0150] 在替代实施方案中,Z表示CR<sup>4</sup>。
- [0151] 在一个实施方案中,R<sup>4</sup>表示C<sub>1-6</sub>烷基、C<sub>1-6</sub>烷氧基、COH或COOH。
- [0152] 在另一个实施方案中,R<sup>4</sup>表示甲氧基、COOH或COH。在另一个实施方案中,R<sup>4</sup>表示甲氧基。在替代实施方案中,R<sup>4</sup>表示COOH。在又一替代实施方案中,R<sup>4</sup>表示COH。
- [0153] 在一个实施方案中,3'-封闭的核苷酸三磷酸选自:

[0154]

结构	名称	实施例号
	脱氧腺苷三磷酸	E1
	脱氧鸟苷三磷酸	E2
	脱氧胸苷三磷酸	E3
	脱氧胞苷三磷酸	E4

[0155]

	2'-脱氧尿苷三磷酸	E5
	5-氨基-2'-脱氧胞苷三磷酸	E6
	5-羟甲基-脱氧胞苷三磷酸	E7
	5-羧基-脱氧胞苷三磷酸	E8
	5-甲酰基脱氧胞苷三磷酸	E9
	N6-甲基腺苷三磷酸	E10
	5-羟甲基-脱氧尿苷三磷酸	E11

[0156] 其中“X”如前文所定义。

[0157] 以下研究和方案说明了本文所述的方法的实施方案：

[0158] 比较实施例1：使用欧洲普通牛TdT将3'-不可逆封闭的核苷酸三磷酸加入到起始DNA

[0159] 将单链起始DNA (SEQ ID NO:7) 与15U欧洲普通牛TdT (Thermo Scientific), 所需盐 (50mM乙酸钾, 20mM Tris乙酸盐pH 7.9, 1mM氯化钴) 和3'-O-甲基dTTP (TriLink) 在37°C下孵育长达1小时。3'-不可逆封闭的核苷酸三磷酸的浓度为1mM, 起始DNA为200pM, 以1:5000的比例进行, 以促进核苷酸的加入。用EDTA (0.5M) 以不同的间隔停止反应, 结果示于图2 (A)。

[0160] 重复实验, 不同之处在于使用3'-叠氮基dTTP代替3'-O-甲基dTTP作为3'-不可逆封闭的核苷酸三磷酸。用EDTA (0.5M) 以不同的间隔停止反应, 结果示于图2 (B)。

[0161] 这些研究表明, 市售的欧洲普通牛TdT将不可逆的封闭核苷酸, 特别是3'-O-甲基dTTP和3'-叠氮基dTTP加入到起始DNA链上时将不可逆的封闭核苷酸 (图2)。此外, 完成向单链起始DNA加入3'-封闭的核苷酸三磷酸, 使得n链向n+1链的转化率大于90%。

[0162] 实施例2：使用除了欧洲普通牛TdT之外的TdT直向同源物用于受控的非模板依赖性DNA合成

[0163] 将单链起始DNA (SEQ ID NO:7) 与纯化的重组TdT直向同源物, 所需的盐 (50mM乙酸钾, 20mM三醋酸盐pH7.9, 1.5mM氯化钴, 酿酒酵母无机焦磷酸酶) 和3'-O-叠氮甲基dTTP在37°C下孵育60分钟。3'-封闭的核苷酸三磷酸盐的浓度为1mM, 起始DNA为200nM, 并通过毛细管电泳分析, 如图3 (A) 所示。反应在60分钟后进行定量, 并且结果如图3 (B) 所示。

[0164] 重复实验, 不同之处在于使用3'-O-叠氮甲基dTTP和3'-O-叠氮甲基dCTP作为3' 封闭的核苷酸三磷酸。20分钟后停止反应并用毛细管电泳分析。定量反应如图3 (C) 所示。

[0165] 欧洲普通牛TdT仅有效地加入不可逆的封闭核苷酸, 这对于受控的酶促单链DNA合成无效。这个例子表明, 天然存在的TdT直向同源物而非欧洲普通牛TdT, 在加入3'-可逆封闭的核苷酸三磷酸时表现明显更好。通过N+1加入速率 (图3) 判断的更好的性能导致更长的可实现长度和更好的核酸序列特异性控制。

[0166] 实施例3：改造的TdT直向同源物显示出比野生型蛋白质更好的功能

[0167] 将单链起始DNA (SEQ ID NO:7) 与纯化的野生型斑点雀鳢TdT或纯化的重组改造的斑点雀鳢TdT (SEQ ID NO:8), 所需的盐, 氯化钴, 酿酒酵母无机焦磷酸酶, 3'-O-叠氮甲基dTTP在37°C孵育20分钟。

[0168] 当提供3'-可逆封闭的dCTP、dGTP或dTTP时, 野生型斑点雀鳢TdT的表现优于改造变体 (SEQ ID NO:8), 如通过将长度为n的引发剂链转化为长度为n+1的链的改进能力所证明。

[0169] 实施例4：使用改造的TdT直向同源物进行非模板依赖性和序列特异性DNA合成

[0170] 将单链起始DNA (SEQ ID NO:7) 与纯化的重组工程形式的斑点雀鳢TdT, 所需的盐, 氯化钴, 酿酒酵母无机焦磷酸酶和3'-O-叠氮甲基dTTP在37°C下孵育10分钟。

[0171] 欧洲普通牛TdT的表现优于斑点雀鳢TdT的改造变体, 如图5中变性PAGE凝胶所证明。该实施例表明了改造形式的斑点雀鳢TdT掺入3'-可逆封闭的核苷酸 (1) 比野生型斑点雀鳢TdT更好, 和 (2) 比欧洲普通牛TdT好得多 (参见图2-5)。

[0172] 实施例5：使用无机焦磷酸酶来控制核酸序列特异性

[0173] 将单链起始DNA (SEQ ID NO:7) 与欧洲普通牛TdT在37°C下在上述反应缓冲液下孵育60分钟,不同之处在于酿酒酵母无机焦磷酸酶的浓度改变。当使用二脱氧TTP (ddTTP)时, NTP的浓度为0.1mM。反应通过PAGE分析,如图6所示。

[0174] 用TdT研究的无机焦磷酸酶表明,TdT介导的DNA合成必须利用无机焦磷酸酶来控制核酸序列特异性。

[0175] 实施例6:使用TdT的DNA合成方法实例

[0176] 1.将固定化的单链起始DNA暴露于由TdT;碱基特异性3'-封闭核苷酸三磷酸;无机焦磷酸酶(例如酿酒酵母同系物);和任何需要的缓冲液(例如,三(羟甲基)氨基甲烷(Tris),二甲基胍酸盐或缓冲范围在pH 5-10之间的任何缓冲液)以及盐(例如Na<sup>+</sup>、K<sup>+</sup>、Mg<sup>2+</sup>、Mn<sup>2+</sup>、Cu<sup>2+</sup>、Zn<sup>2+</sup>、Co<sup>2+</sup>等等,都具有适当的平衡离子,如Cl<sup>-</sup>)组成的延伸溶液中,并在优化的浓度、时间和温度下反应。3'-封闭的核苷酸三磷酸将含有含氮碱基腺嘌呤、鸟嘌呤、胞嘧啶或胸腺嘧啶中的一种。

[0177] 2.然后用洗涤混合物除去延伸混合物并回收。洗涤混合物是不含TdT和3'-封闭核苷酸三磷酸的延伸混合物。

[0178] 3.固定化的(n+1)DNA链用由适当的缓冲液、变性剂(例如尿素、氯化胍、甲酰胺、甜菜碱等)和裂解剂组成的裂解混合物处理(例如,三(2-羧乙基)膦(TCEP)解封闭3'-O-叠氮基甲基;钯络合物解封闭3'-O-烯丙基;或亚硝酸钠解封闭3'-氨基)。温度可升至99°C,有助于二级结构的裂解和破坏。最佳温度取决于所用的裂解剂。

[0179] 4.用洗涤混合物处理固定化的解封(n+1)DNA链以除去裂解混合物。

[0180] 5.用碱基特异性3'-封闭的核苷酸三磷酸重复循环1-4,直到达到所需的寡核苷酸序列。

[0181] 6.一旦达到所需的序列,使用具有DNA产物特异性引物的聚合酶链反应直接“提取”并扩增产物。



Val Leu Asn Asp Glu Arg Tyr Gln Ala Phe Lys Leu Phe Thr Ser Val  
 245 250 255  
 Phe Gly Val Gly Leu Lys Thr Ala Glu Lys Trp Tyr Arg Met Gly Phe  
 260 265 270  
 Arg Thr Leu Ser Lys Ile Gln Ser Asp Lys Ser Leu Lys Phe Thr Lys  
 275 280 285  
 Met Gln Lys Ala Gly Phe Leu Tyr Tyr Glu Asp Leu Ile Ser Cys Val  
 290 295 300  
 Ser Lys Ala Glu Ala Asp Ala Val Ser Leu Ile Val Lys Glu Ala Val  
 305 310 315 320  
 Trp Thr Phe Leu Pro Asp Ala Leu Ile Thr Ile Thr Gly Gly Phe Arg  
 325 330 335  
 Arg Gly Lys Glu Phe Gly His Asp Val Asp Phe Leu Ile Thr Ser Pro  
 340 345 350  
 Gly Gly Glu Lys Glu Gln Val Asp Gln Leu Leu Gln Lys Val Thr Asn  
 355 360 365  
 Leu Trp Glu Lys Gln Gly Leu Leu Leu Tyr Tyr Asp Leu Met Glu Ser  
 370 375 380  
 Thr Phe Glu Asp Leu Lys Leu Pro Ser Arg Lys Val Asp Ala Leu Asp  
 385 390 395 400  
 His Phe Gln Lys Cys Phe Leu Ile Leu Lys Leu Tyr Cys Gln Arg Gly  
 405 410 415  
 [0002] Asp Arg Ser Lys Trp Glu Gly Pro Glu Gly Ser Asn Gly Leu Gln Thr  
 420 425 430  
 Lys Asn Trp Lys Ala Ile Arg Val Asp Leu Val Val Cys Pro Tyr Asp  
 435 440 445  
 Arg Tyr Ala Tyr Ala Leu Leu Gly Trp Ser Gly Ser Arg Gln Phe Glu  
 450 455 460  
 Arg Asp Leu Arg Arg Tyr Ala Thr His Glu Lys Lys Met Met Leu Asp  
 465 470 475 480  
 Asn His Ala Leu Tyr Asp Lys Thr Lys Arg Thr Phe Leu Lys Ala Glu  
 485 490 495  
 Ser Glu Glu Glu Ile Phe Ser His Leu Gly Leu Glu Tyr Ile Glu Pro  
 500 505 510  
 Trp Glu Arg Asn Ala  
 515  
 <210> 2  
 <211> 494  
 <212> PRT  
 <213> 斑点雀鳝 (*Lepisosteus oculatus*)  
 <400> 2  
 Met Leu His Ile Pro Ile Phe Pro Pro Ile Lys Lys Arg Gln Lys Leu  
 1 5 10 15

Pro Glu Ser Arg Asn Ser Cys Lys Tyr Glu Val Lys Phe Ser Glu Val  
 20 25 30  
 Ala Ile Phe Leu Val Glu Arg Lys Met Gly Ser Ser Arg Arg Lys Phe  
 35 40 45  
 Leu Thr Asn Leu Ala Arg Ser Lys Gly Phe Arg Ile Glu Asp Val Leu  
 50 55 60  
 Ser Asp Ala Val Thr His Val Val Ala Glu Asp Asn Ser Ala Asp Glu  
 65 70 75 80  
 Leu Trp Gln Trp Leu Gln Asn Ser Ser Leu Gly Asp Leu Ser Lys Ile  
 85 90 95  
 Glu Val Leu Asp Ile Ser Trp Phe Thr Glu Cys Met Gly Ala Gly Lys  
 100 105 110  
 Pro Val Gln Val Glu Ala Arg His Cys Leu Val Lys Ser Cys Pro Val  
 115 120 125  
 Ile Asp Gln Tyr Leu Glu Pro Ser Thr Val Glu Thr Val Ser Gln Tyr  
 130 135 140  
 Ala Cys Gln Arg Arg Thr Thr Met Glu Asn His Asn Gln Ile Phe Thr  
 145 150 155 160  
 Asp Ala Phe Ala Ile Leu Ala Glu Asn Ala Glu Phe Asn Glu Ser Glu  
 165 170 175  
 Gly Pro Cys Leu Ala Phe Met Arg Ala Ala Ser Leu Leu Lys Ser Leu  
 180 185 190  
 [0003] Pro His Ala Ile Ser Ser Ser Lys Asp Leu Glu Gly Leu Pro Cys Leu  
 195 200 205  
 Gly Asp Gln Thr Lys Ala Val Ile Glu Asp Ile Leu Glu Tyr Gly Gln  
 210 215 220  
 Cys Ser Lys Val Gln Asp Val Leu Cys Asp Asp Arg Tyr Gln Thr Ile  
 225 230 235 240  
 Lys Leu Phe Thr Ser Val Phe Gly Val Gly Leu Lys Thr Ala Glu Lys  
 245 250 255  
 Trp Tyr Arg Lys Gly Phe His Ser Leu Glu Glu Val Gln Ala Asp Asn  
 260 265 270  
 Ala Ile His Phe Thr Lys Met Gln Lys Ala Gly Phe Leu Tyr Tyr Asp  
 275 280 285  
 Asp Ile Ser Ala Ala Val Cys Lys Ala Glu Ala Gln Ala Ile Gly Gln  
 290 295 300  
 Ile Val Glu Glu Thr Val Arg Leu Ile Ala Pro Asp Ala Ile Val Thr  
 305 310 315 320  
 Leu Thr Gly Gly Phe Arg Arg Gly Lys Glu Cys Gly His Asp Val Asp  
 325 330 335  
 Phe Leu Ile Thr Thr Pro Glu Met Gly Lys Glu Val Trp Leu Leu Asn  
 340 345 350  
 Arg Leu Ile Asn Arg Leu Gln Asn Gln Gly Ile Leu Leu Tyr Tyr Asp  
 355 360 365  
 Ile Val Glu Ser Thr Phe Asp Lys Thr Arg Leu Pro Cys Arg Lys Phe



370                      375                      380  
 Glu Ala Met Asp His Phe Gln Lys Cys Phe Ala Ile Ile Lys Leu Lys  
 385                      390                      395                      400  
 Lys Glu Leu Ala Ala Gly Arg Val Gln Lys Asp Trp Lys Ala Ile Arg  
                          405                      410                      415  
 Val Asp Phe Val Ala Pro Pro Val Asp Asn Phe Ala Phe Ala Leu Leu  
                          420                      425                      430  
 Gly Trp Thr Gly Ser Arg Gln Phe Glu Arg Asp Leu Arg Arg Phe Ala  
                          435                      440                      445  
 Arg His Glu Arg Lys Met Leu Leu Asp Asn His Ala Leu Tyr Asp Lys  
                          450                      455                      460  
 Thr Lys Lys Tyr Leu Lys Lys Lys Thr Thr Asn Asn Tyr Leu Ala Leu  
 465                      470                      475                      480  
 Asn Asp Val Cys Ser Asp Leu Ser Glu Trp His Tyr Lys Gly  
                          485                      490

<210> 3

<211> 510

<212> PRT

<213> 长尾龙猫 (Chinchilla lanigera)

<400> 3

[0004]

Met Asp Pro Leu Gln Ala Ala His Ser Gly Pro Arg Lys Lys Arg Pro  
 1                      5                      10                      15  
 Arg Gln Thr Gly Thr Leu Met Val Ser Ser Pro His Asp Val Arg Phe  
                          20                      25                      30  
 Gly Asp Leu Val Leu Phe Ile Leu Glu Lys Lys Met Gly Thr Thr Arg  
                          35                      40                      45  
 Arg Ala Phe Leu Met Glu Leu Ala Arg Arg Lys Gly Phe Arg Val Glu  
                          50                      55                      60  
 Asn Glu Leu Ser Asp Ser Val Thr His Ile Val Ala Glu Asn Asn Ser  
 65                      70                      75                      80  
 Gly Asn Asp Val Leu Glu Trp Leu Gln Val Gln Asn Ile Gln Ala Ser  
                          85                      90                      95  
 Ser Arg Leu Glu Leu Leu Asp Val Ser Trp Leu Ile Glu Cys Met Gly  
                          100                      105                      110  
 Ala Gly Lys Pro Val Glu Met Thr Gly Lys His Gln Leu Leu Val Arg  
                          115                      120                      125  
 Arg Asp Tyr Pro Ala Ser Pro Lys Pro Gly Pro Gln Lys Thr Pro Ser  
                          130                      135                      140  
 Leu Ala Val Gln Lys Ile Ser Glu Tyr Ala Cys Gln Arg Arg Thr Thr  
 145                      150                      155                      160  
 Leu Asn Asn Cys Asn Cys Ile Phe Thr Asn Ala Phe Glu Ile Leu Ala  
                          165                      170                      175  
 Glu Asn Cys Glu Phe Arg Glu Asn Glu Ser Ser Tyr Val Thr Tyr Met  
                          180                      185                      190

Arg Ala Ala Ser Val Leu Lys Ser Leu Pro Phe Thr Ile Ile Ser Met  
 195 200 205  
 Lys Asp Thr Glu Gly Ile Pro Cys Leu Gly Glu Lys Val Lys Cys Ile  
 210 215 220  
 Ile Glu Glu Ile Ile Glu Asp Gly Glu Ser Ser Glu Val Asn Ala Val  
 225 230 235 240  
 Leu Asn Asp Glu Arg Tyr Gln Ser Phe Lys Leu Phe Thr Ser Val Phe  
 245 250 255  
 Gly Val Gly Leu Lys Thr Ser Glu Lys Trp Phe Arg Met Gly Phe Arg  
 260 265 270  
 Ser Leu Asn Lys Ile Lys Ser Asp Lys Ser Leu Lys Phe Thr Arg Met  
 275 280 285  
 Gln Lys Ala Gly Phe Leu Tyr Tyr Glu Asp Leu Val Ser Cys Val Thr  
 290 295 300  
 Arg Ala Glu Ala Glu Ala Val Ser Met Leu Val Lys Glu Ala Val Trp  
 305 310 315 320  
 Ala Phe Leu Pro Gly Ala Phe Ile Ser Met Thr Gly Gly Phe Arg Arg  
 325 330 335  
 Gly Lys Glu Ile Gly His Asp Val Asp Phe Leu Ile Thr Ser Pro Glu  
 340 345 350  
 Val Thr Glu Asp Glu Lys Gln Gln Leu Leu His Lys Val Ile Ser Leu  
 355 360 365  
 [0005] Trp Glu Lys Lys Gly Leu Leu Leu Tyr Ser Asp Leu Val Glu Ser Thr  
 370 375 380  
 Phe Glu Lys Leu Lys Leu Pro Ser Arg Lys Val Asp Ala Leu Asp His  
 385 390 395 400  
 Phe Gln Lys Cys Phe Leu Ile Leu Lys Leu His His Gln Arg Val Asp  
 405 410 415  
 Ser Asp Lys Ser Pro Gln Gln Gly Gly Lys Thr Trp Lys Ala Ile Arg  
 420 425 430  
 Val Asp Leu Val Val Cys Pro Tyr Glu Arg Arg Ala Phe Ala Leu Leu  
 435 440 445  
 Gly Trp Thr Gly Ser Arg Gln Phe Glu Arg Asp Leu Arg Arg Tyr Ala  
 450 455 460  
 Thr His Glu Arg Lys Met Met Leu Asp Asn His Ala Leu Tyr Asp Lys  
 465 470 475 480  
 Thr Lys Arg Ile Phe Leu Lys Ala Glu Ser Glu Glu Glu Ile Phe Ala  
 485 490 495  
 His Leu Gly Leu Asp Tyr Ile Glu Pro Trp Glu Arg Asn Ala  
 500 505 510

<210> 4

<211> 511

<212> PRT

<213> 小耳大婴猴 (*Otolemur garnettii*)

<400> 4  
 Met Asp Pro Leu His Met Ala His Ser Gly Pro Arg Lys Lys Arg Pro  
 1 5 10 15  
 Arg Gln Thr Ala Ala Ser Met Val Ser Thr Pro Gln Asp Ile Lys Phe  
 20 25 30  
 Arg Asp Leu Val Leu Phe Ile Leu Glu Lys Lys Met Gly Thr Thr Arg  
 35 40 45  
 Arg Thr Phe Leu Met Glu Leu Ala Arg Thr Lys Gly Phe Arg Val Glu  
 50 55 60  
 Asn Glu Phe Ser Asp Ser Val Thr His Ile Ile Ala Glu Asn Asn Ser  
 65 70 75 80  
 Gly Ser Asp Val Leu Glu Trp Ile Gln Val Gln Lys Ile Lys Ala Gly  
 85 90 95  
 Ser Gln Met Glu Val Leu Asp Val Ser Trp Leu Ile Glu Cys Met Arg  
 100 105 110  
 Ala Gly Lys Pro Val Glu Met Thr Gly Lys His Gln Leu Val Val Arg  
 115 120 125  
 Gly Asp Tyr Ser Pro Ser Pro Asn Pro Ala Pro Gln Lys Thr Pro Pro  
 130 135 140  
 Leu Ala Val Gln Lys Ile Ser Gln Tyr Ala Cys Gln Arg Arg Thr Thr  
 145 150 155 160  
 Leu Asn Asn Cys Asn His Ile Phe Thr Asp Ala Phe Glu Ile Met Ala  
 [0006] 165 170 175  
 Glu Asn Tyr Glu Phe Arg Glu Asn Glu Gly Tyr Ser Ala Ala Phe Met  
 180 185 190  
 Arg Ala Ala Ser Val Leu Lys Ser Leu Pro Phe Thr Ile Ile Ser Met  
 195 200 205  
 Lys Asp Thr Glu Gly Val Pro Cys Leu Gly Asp Asn Val Lys Cys Ile  
 210 215 220  
 Ile Glu Glu Ile Ile Glu Glu Gly Glu Ser Ser Glu Val Lys Ala Val  
 225 230 235 240  
 Leu Asn Asp Glu Arg Tyr Gln Ser Phe Lys Leu Phe Thr Ser Val Phe  
 245 250 255  
 Gly Val Gly Leu Lys Thr Ser Glu Lys Trp Phe Arg Met Gly Phe Arg  
 260 265 270  
 Thr Leu Ser Lys Ile Arg Ser Asp Lys Ser Leu Arg Phe Thr Arg Met  
 275 280 285  
 Gln Gln Ala Gly Phe Leu Tyr Tyr Glu Asp Leu Val Ser Cys Val Thr  
 290 295 300  
 Arg Ala Glu Ala Glu Ala Val Gly Val Leu Val Lys Glu Ala Val Arg  
 305 310 315 320  
 Ala Phe Leu Pro Asp Ala Phe Val Thr Met Thr Gly Gly Phe Arg Arg  
 325 330 335  
 Gly Lys Asn Ile Gly His Asp Val Asp Phe Leu Ile Thr Ser Pro Gly  
 340 345 350

Ser Thr Glu Glu Glu Glu Gln Gln Leu Leu His Lys Ile Met Asp Leu  
 355 360 365  
 Trp Glu Lys Lys Gly Leu Leu Leu Tyr Cys Asp Leu Val Glu Ser Thr  
 370 375 380  
 Phe Glu Lys Leu Lys Leu Pro Ser Arg Lys Val Asp Ala Leu Asp His  
 385 390 395 400  
 Phe Gln Lys Cys Phe Leu Ile Phe Lys Leu His His Gln Arg Val Val  
 405 410 415  
 Asp Ser Glu Gln Ser Asn Gln Gln Glu Gly Lys Thr Trp Lys Ala Ile  
 420 425 430  
 Arg Val Asp Leu Val Met Cys Pro Tyr Glu Arg Arg Ala Tyr Ala Leu  
 435 440 445  
 Leu Gly Trp Thr Gly Ser Arg Gln Phe Glu Arg Asp Leu Arg Arg Tyr  
 450 455 460  
 Ala Thr His Glu Arg Lys Met Ile Leu Asp Asn His Gly Leu Trp Asp  
 465 470 475 480  
 Lys Thr Lys Arg Ile Phe Leu Lys Ala Glu Ser Glu Glu Glu Ile Phe  
 485 490 495  
 Ala His Leu Gly Leu Asp Tyr Ile Glu Pro Trp Glu Arg Asn Ala  
 500 505 510

<210> 5

[0007] <211> 510

<212> PRT

<213> 野猪 (*Sus scrofa*)

<400> 5

Met Asp Pro Pro Gln Thr Val Pro Ser Ser Pro Arg Lys Lys Arg Pro  
 1 5 10 15  
 Arg Gln Val Gly Ala Ser Met Ala Ser Pro Ala His Asn Ile Lys Phe  
 20 25 30  
 Arg Glu Leu Val Leu Phe Ile Leu Glu Lys Lys Met Gly Thr Thr Arg  
 35 40 45  
 Arg Thr Phe Leu Met Glu Leu Ala Arg Arg Lys Gly Phe Arg Val Glu  
 50 55 60  
 Asn Glu Leu Ser Asp Ser Val Thr His Ile Val Ala Glu Asn Asn Ser  
 65 70 75 80  
 Gly Ser Glu Val Leu Glu Trp Leu Gln Ala Gln Lys Ile Arg Ala Ser  
 85 90 95  
 Ser Gln Leu Thr Leu Leu Asp Val Ser Trp Leu Ile Glu Ser Met Gly  
 100 105 110  
 Ala Gly Lys Pro Val Glu Met Thr Gly Lys His Gln Leu Val Val Arg  
 115 120 125  
 Thr Asp Cys Ser Ala Ser Pro Ser Pro Gly Ser Gln Asn Thr Leu Pro  
 130 135 140  
 Pro Ala Val Lys Lys Ile Ser Gln Tyr Ala Cys Gln Arg Arg Thr Thr



<210> 6  
 <211> 509  
 <212> PRT  
 <213> 欧洲普通牛(Bos taurus)  
 <400> 6  
 Met Asp Pro Leu Cys Thr Ala Ser Ser Gly Pro Arg Lys Lys Arg Pro  
 1 5 10 15  
 Arg Gln Val Gly Ala Ser Met Ala Ser Pro Pro His Asp Ile Lys Phe  
 20 25 30  
 Gln Asn Leu Val Leu Phe Ile Leu Glu Lys Lys Met Gly Thr Thr Arg  
 35 40 45  
 Arg Asn Phe Leu Met Glu Leu Ala Arg Arg Lys Gly Phe Arg Val Glu  
 50 55 60  
 Asn Glu Leu Ser Asp Ser Val Thr His Ile Val Ala Glu Asn Asn Ser  
 65 70 75 80  
 Gly Ser Glu Val Leu Glu Trp Leu Gln Val Gln Asn Ile Arg Ala Ser  
 85 90 95  
 Ser Gln Leu Glu Leu Leu Asp Val Ser Trp Leu Ile Glu Ser Met Gly  
 100 105 110  
 Ala Gly Lys Pro Val Glu Ile Thr Gly Lys His Gln Leu Val Val Arg  
 115 120 125  
 Thr Asp Tyr Ser Ala Thr Pro Asn Pro Gly Phe Gln Lys Thr Pro Pro  
 [0009] 130 135 140  
 Leu Ala Val Lys Lys Ile Ser Gln Tyr Ala Cys Gln Arg Lys Thr Thr  
 145 150 155 160  
 Leu Asn Asn Tyr Asn His Ile Phe Thr Asp Ala Phe Glu Ile Leu Ala  
 165 170 175  
 Glu Asn Ser Glu Phe Lys Glu Asn Glu Val Ser Tyr Val Thr Phe Met  
 180 185 190  
 Arg Ala Ala Ser Val Leu Lys Ser Leu Pro Phe Thr Ile Ile Ser Met  
 195 200 205  
 Lys Asp Thr Glu Gly Ile Pro Cys Leu Gly Asp Lys Val Lys Cys Ile  
 210 215 220  
 Ile Glu Glu Ile Ile Glu Asp Gly Glu Ser Ser Glu Val Lys Ala Val  
 225 230 235 240  
 Leu Asn Asp Glu Arg Tyr Gln Ser Phe Lys Leu Phe Thr Ser Val Phe  
 245 250 255  
 Gly Val Gly Leu Lys Thr Ser Glu Lys Trp Phe Arg Met Gly Phe Arg  
 260 265 270  
 Ser Leu Ser Lys Ile Met Ser Asp Lys Thr Leu Lys Phe Thr Lys Met  
 275 280 285  
 Gln Lys Ala Gly Phe Leu Tyr Tyr Glu Asp Leu Val Ser Cys Val Thr  
 290 295 300  
 Arg Ala Glu Ala Glu Ala Val Gly Val Leu Val Lys Glu Ala Val Trp



Ala Ile Phe Leu Val Glu Arg Lys Met Gly Ser Ser Arg Arg Lys Phe  
35 40 45  
Leu Thr Asn Leu Ala Arg Ser Lys Gly Phe Arg Ile Glu Asp Val Leu  
50 55 60  
Ser Asp Ala Val Thr His Val Val Ala Glu Asn Asn Ser Ala Asp Glu  
65 70 75 80  
Leu Leu Gln Trp Leu Gln Asn Ser Ser Leu Gly Asp Leu Ser Lys Ile  
85 90 95  
Glu Val Leu Asp Ile Ser Trp Phe Thr Glu Cys Met Gly Ala Gly Lys  
100 105 110  
Pro Val Gln Val Glu Ala Arg His Cys Leu Val Lys Ser Cys Pro Val  
115 120 125  
Ile Asp Gln Tyr Leu Glu Pro Ser Thr Val Glu Thr Val Ser Gln Tyr  
130 135 140  
Ala Cys Gln Arg Arg Thr Thr Met Glu Asn His Asn Gln Ile Phe Thr  
145 150 155 160  
Asp Ala Phe Ala Ile Leu Ala Glu Asn Ala Glu Phe Asn Glu Ser Glu  
165 170 175  
Gly Pro Cys Leu Ala Phe Met Arg Ala Ala Ser Leu Leu Lys Ser Leu  
180 185 190  
Pro His Ala Ile Ser Ser Ser Lys Asp Leu Glu Gly Leu Pro Cys Leu  
195 200 205  
[0011] Gly Asp Gln Thr Lys Ala Val Ile Glu Asp Ile Leu Glu Tyr Gly Gln  
210 215 220  
Cys Ser Lys Val Gln Asp Val Leu Cys Asp Asp Arg Tyr Gln Thr Ile  
225 230 235 240  
Lys Leu Phe Thr Ser Val Phe Gly Val Gly Leu Arg Thr Ala Glu Lys  
245 250 255  
Trp Tyr Arg Lys Gly Phe His Ser Leu Glu Glu Val Gln Ala Asp Asn  
260 265 270  
Ala Ile His Phe Thr Lys Met Gln Lys Ala Gly Phe Leu Tyr Tyr Asp  
275 280 285  
Asp Ile Ser Ala Ala Val Cys Lys Ala Glu Ala Gln Ala Ile Gly Gln  
290 295 300  
Ile Val Glu Glu Thr Val Arg Leu Ile Ala Pro Asp Ala Ile Val Thr  
305 310 315 320  
Leu Thr Gly Gly Phe Arg Arg Gly Lys Glu Cys Gly His Asp Val Asp  
325 330 335  
Phe Leu Ile Thr Thr Pro Glu Met Gly Lys Glu Val Trp Leu Leu Asn  
340 345 350  
Arg Leu Ile Asn Arg Leu Gln Asn Gln Gly Ile Leu Leu Tyr Tyr Asp  
355 360 365  
Ile Val Glu Ser Thr Phe Asp Lys Thr Arg Leu Pro Cys Arg Lys Phe  
370 375 380  
Glu Ala Met Asp His Phe Gln Lys Cys Phe Ala Ile Ile Lys Leu Lys





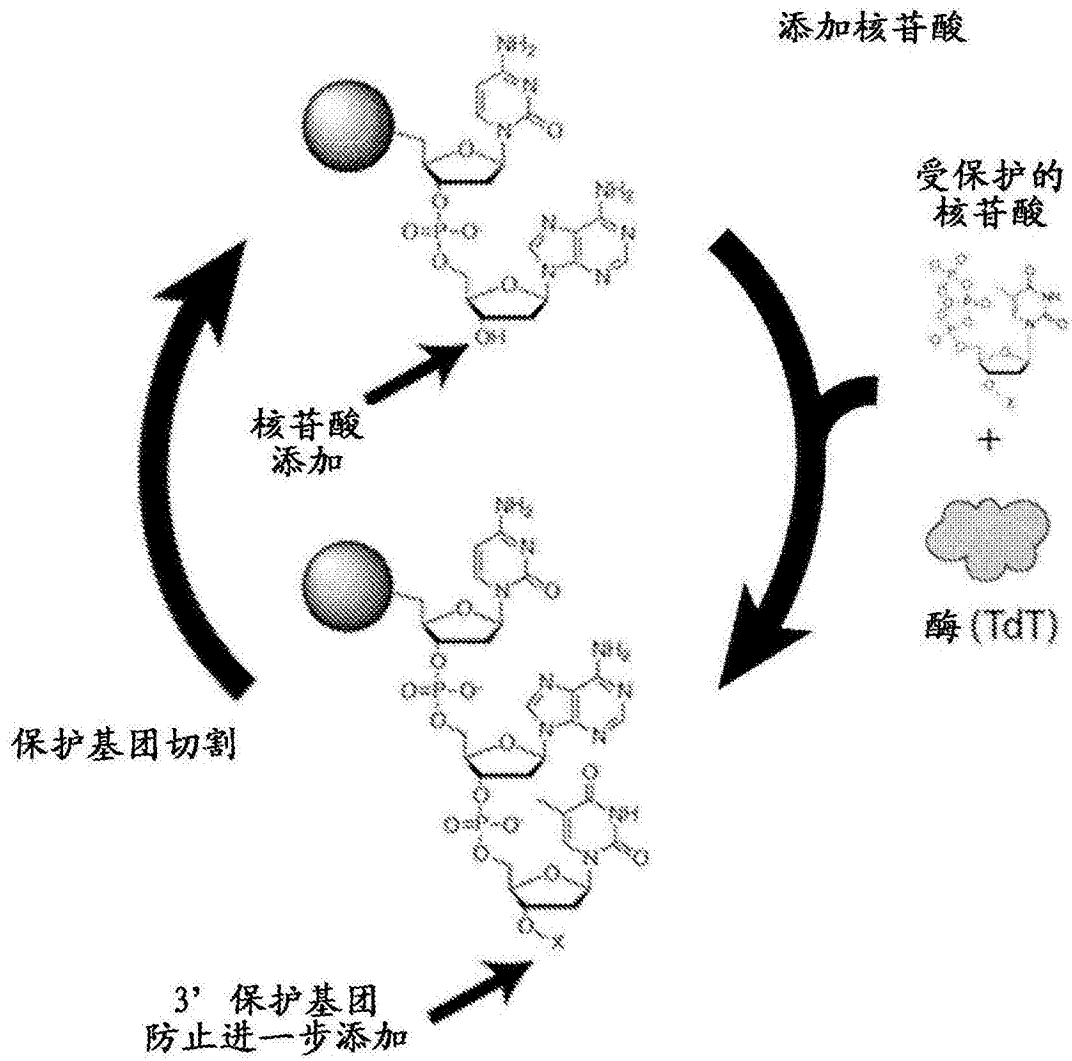


图1

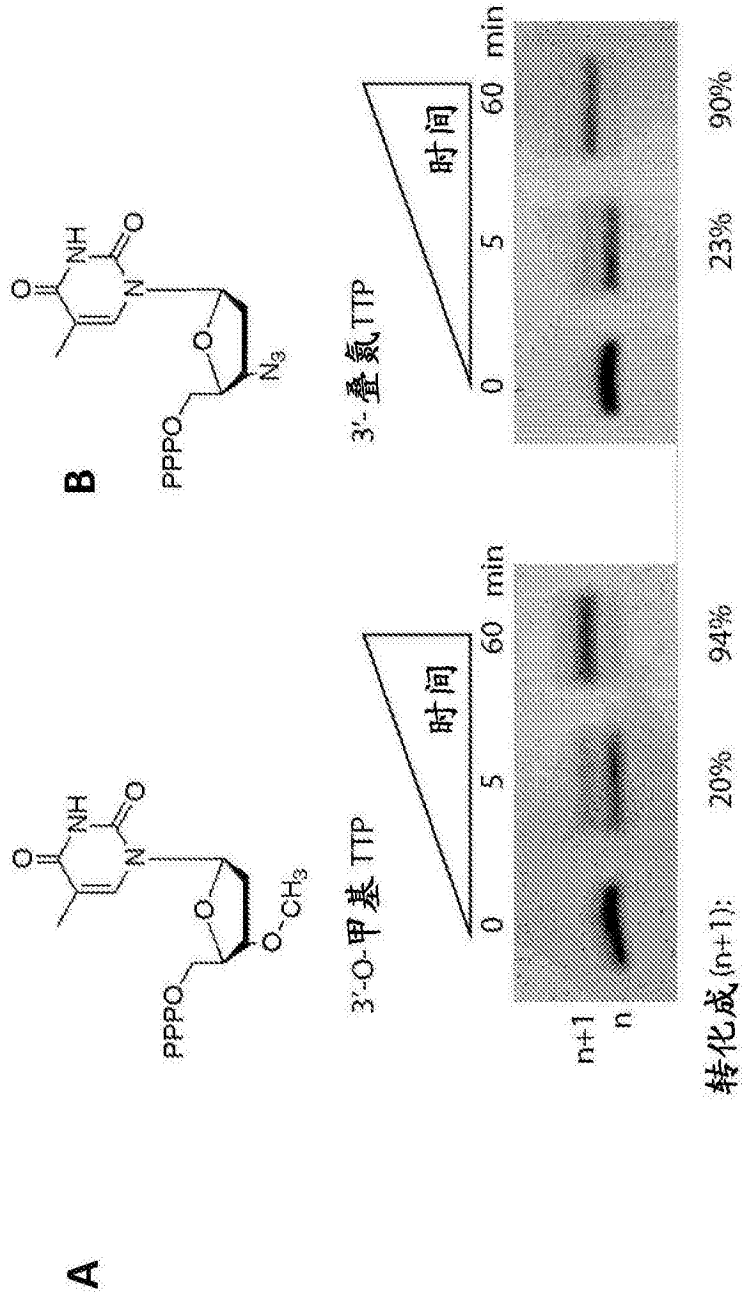


图2

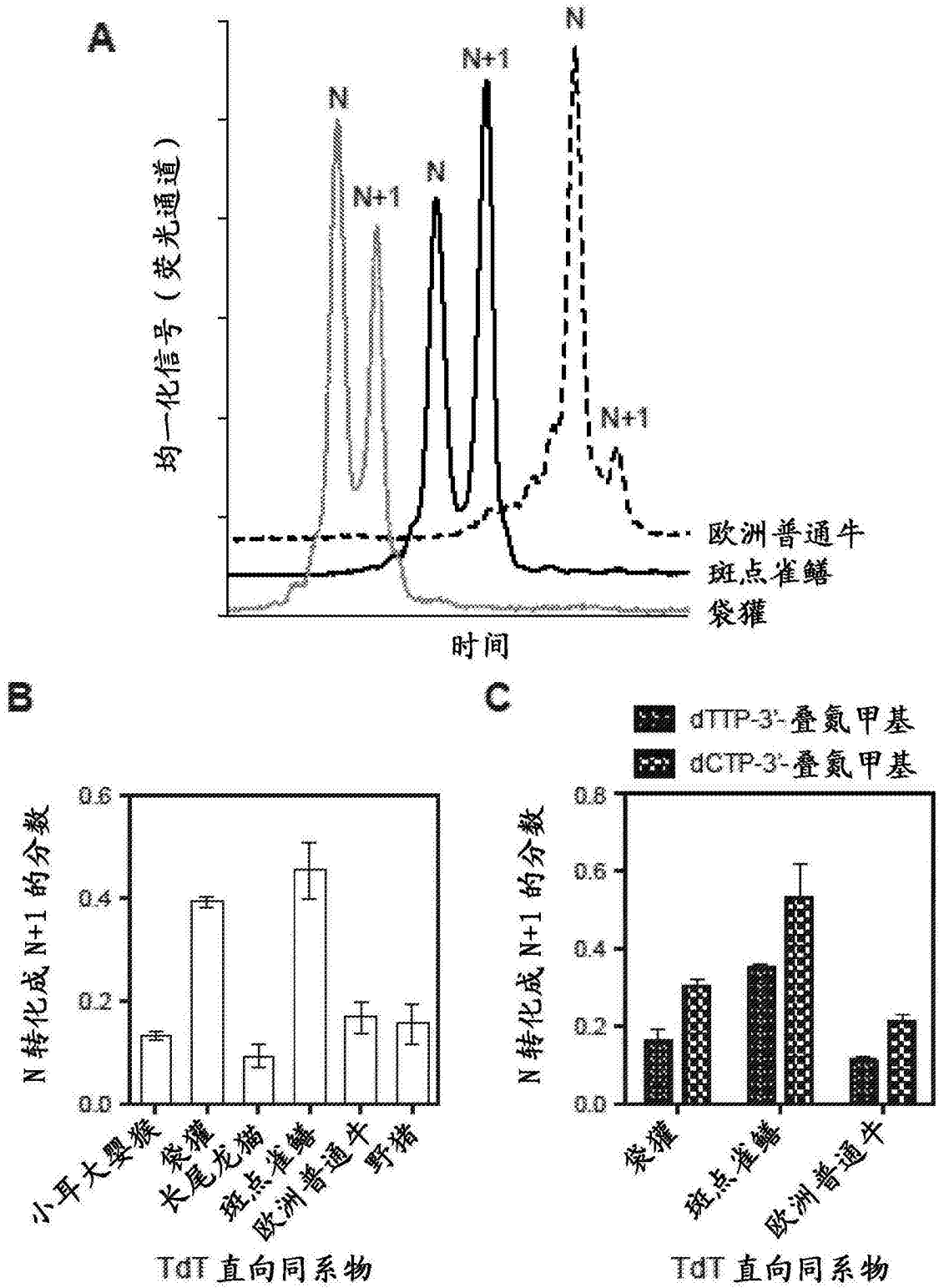


图3

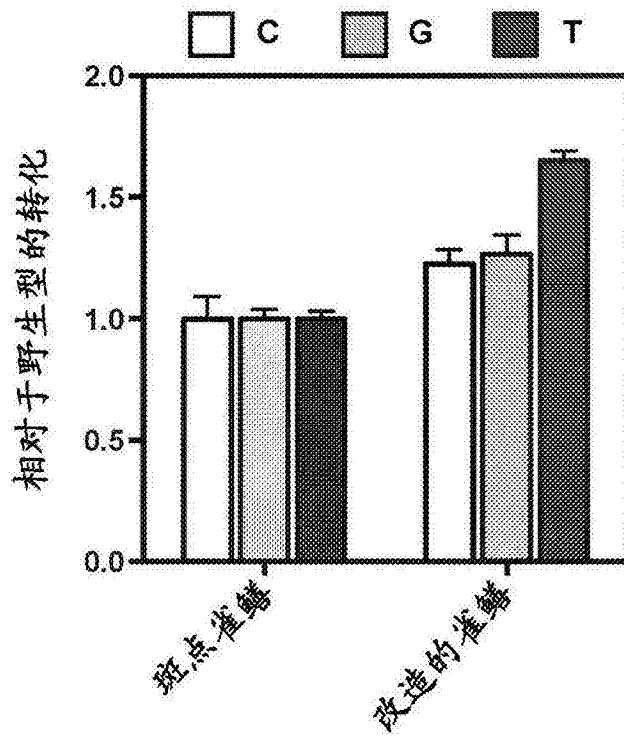


图4



阴性对照  
欧洲普通牛 T<sub>0</sub>T  
斑点雀鳝  
工程化变体

图5

ddTTP	+	+	-	-	-	-	+	+	-	-	-	-
N3Me-dTTP	-	-	+	+	+	+	-	-	+	+	+	+
YIPP (mU)	0	100	0	20	40	100	0	100	0	20	40	100
时间 (min)	5	5	5	5	5	5	30	30	30	30	30	30

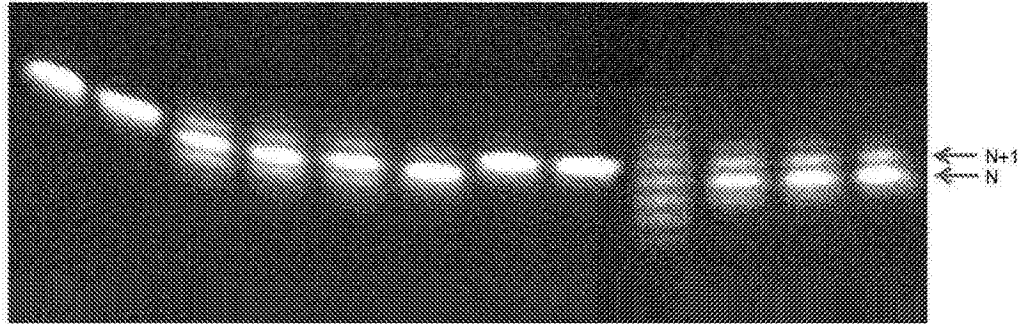


图6

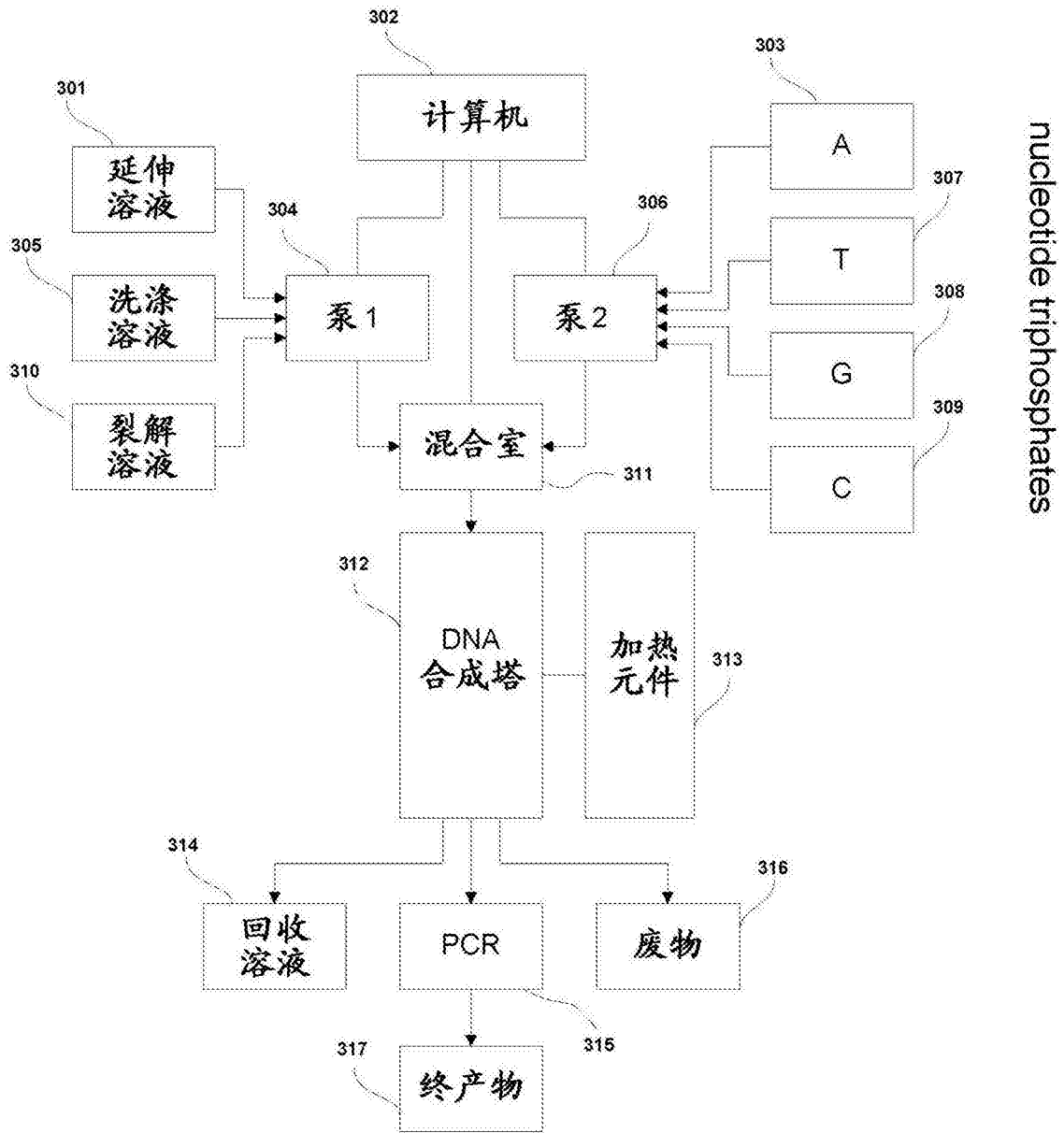


图7

1. 包含与SEQ ID NO:1至5和8中任一个具有至少95%序列同一性的氨基酸序列的末端脱氧核苷酸转移酶(TdT)在核酸合成方法中的用途。

2. 如权利要求1所述的用途,其中末端脱氧核苷酸转移酶(TdT)包含与SEQ ID NO:1、2或8中任一个具有至少95%序列同一性的氨基酸序列。

3. 一种核酸合成方法,其包括以下步骤:

(a) 提供起始序列;

(b) 在权利要求1或权利要求2定义的末端脱氧核苷酸转移酶(TdT)存在下,向所述起始序列加入3'-封闭核苷酸三磷酸;

(c) 从起始序列除去所有试剂;

(d) 在裂解剂存在下,从3'-封闭的核苷酸三磷酸切割封闭基团;

(e) 去除裂解剂。

4. 如权利要求3所述的方法,其中通过重复步骤(b)至(e)加入大于1个核苷酸。

5. 如权利要求3或权利要求4所述的方法,其中3'-封闭的核苷酸三磷酸被3'-O-叠氮甲基、3'-氨基或3'-O-烯丙基封闭。

6. 如权利要求3至5中任一项所述的方法,其中末端脱氧核苷酸转移酶(TdT)在延伸溶液存在下加入,所述延伸溶液包含一种或多种缓冲液,如Tris或二甲基胍酸盐、一种或多种盐,以及无机焦磷酸酶,如来自酿酒酵母的纯化重组无机焦磷酸酶。

7. 如权利要求3至6中任一项所述的方法,其中步骤(b)在5至10的pH范围内进行。

8. 如权利要求3至7中任一项所述的方法,其中裂解剂是化学裂解剂。

9. 如权利要求3至8中任一项所述的方法,其中裂解剂是酶裂解剂。

10. 如权利要求3至9中任一项所述的方法,其中裂解剂选自:三(2-羧乙基)膦(TCEP)、钯络合物或亚硝酸钠。

11. 如权利要求3至10中任一项所述的方法,其中裂解剂在包含变性剂如尿素、氯化胍、甲酰胺或甜菜碱和一种或多种缓冲液的裂解溶液的存在下加入。

12. 如权利要求3至11中任一项所述的方法,其中步骤(d)在低于99°C的温度下进行。

13. 如权利要求3至12中任一项所述的方法,其中步骤(c)和(e)是通过加入洗涤溶液进行。

14. 如权利要求3至13中任一项所述的方法,其中方法在诸如微流体或基于柱的流动装置的流动装置内进行。

15. 如权利要求3至13中任一项所述的方法,其中方法在平板或微阵列设置内进行。

16. 如权利要求3至15中任一项所述的方法,其中起始序列具有5至50个核苷酸长度,例如10至30个核苷酸长度,特别是约20个核苷酸长度。

17. 如权利要求3至16中任一项所述的方法,其中起始序列是单链序列或双链序列。

18. 如权利要求3至17中任一项所述的方法,其中起始序列固定在固体载体上。

19. 如权利要求18所述的方法,其中起始序列通过可逆相互作用部分固定。

20. 如权利要求19所述的方法,其另外包括通过除去所述可逆相互作用部分来提取得到的核酸。

21. 如权利要求19所述的方法,其中可逆相互作用部分是化学可切割的接头,例如二硫键、烯丙基或叠氮化物掩蔽的半缩醛胺醚。



22. 如权利要求21所述的方法,其另外包括通过切割化学可切割接头来提取得到的核酸,例如通过加入三(2-羧乙基)膦(TCEP)、二硫苏糖醇(DTT)或钨络合物。

23. 如权利要求3至18中任一项所述的方法,其中起始序列含有至少一个限制性位点。

24. 如权利要求23所述的方法,其另外包括通过使用限制酶提取所得核酸。

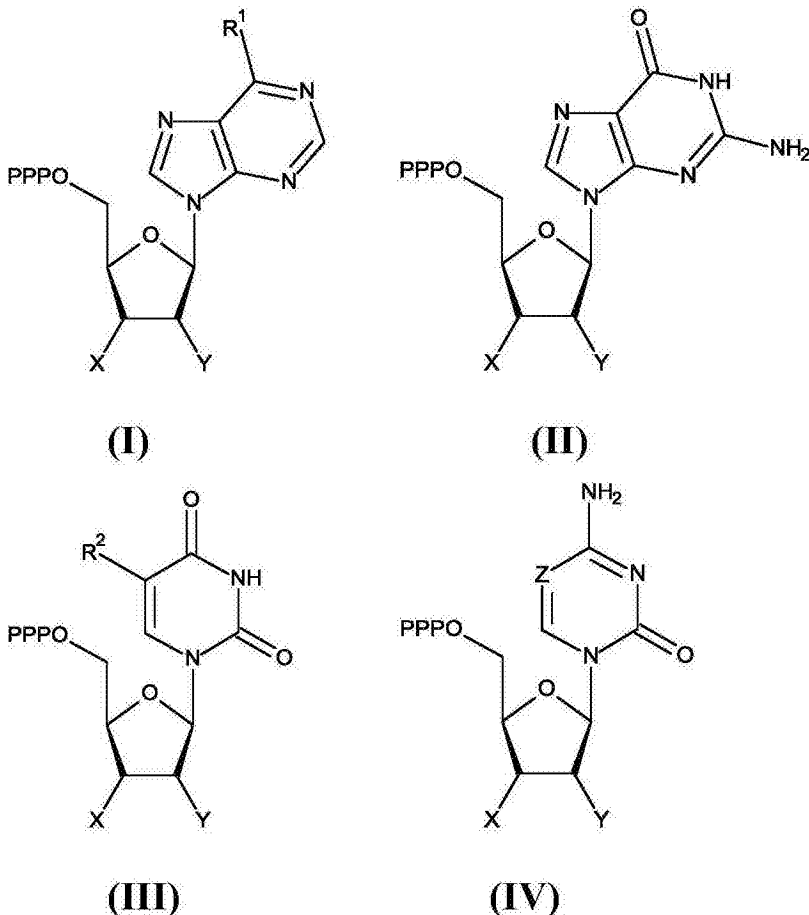
25. 如权利要求3至24中任一项所述的方法,其中起始序列含有至少一个尿苷。

26. 如权利要求3至25中任一项所述的方法,其另外包括例如通过PCR扩增得到的核酸。

27. 一种试剂盒在核酸合成方法中的用途,其中所述试剂盒包含如权利要求1或权利要求2定义的TdT,任选地与选自以下的一种或多种组分组合:起始序列、一种或多种3'-封闭的磷酸三核苷酸、无机焦磷酸酶,如来自酿酒酵母的纯化重组无机焦磷酸酶,以及裂解剂;进一步任选包含按照权利要求3至26中任一项所述的方法使用试剂盒的说明书。

28. 如权利要求27所述的用途,其另外包含选自以下的一种或多种组分:延伸溶液、洗涤溶液和/或裂解溶液。

29. 3'-封闭的核苷酸三磷酸在非模板依赖性核酸合成方法中的用途,其中3'-封闭的核苷酸三磷酸是选自式(I)、(II)、(III)或(IV)的化合物:



其中,

$R^1$ 表示 $NR^aR^b$ ,其中 $R^a$ 和 $R^b$ 独立地表示氢或 $C_{1-6}$ 烷基,

$R^2$ 表示氢、 $C_{1-6}$ 烷基、 $C_{1-6}$ 烷氧基、COH或COOH,

X表示 $C_{1-6}$ 烷基、 $NH_2$ 、 $N_3$ 或 $-OR^3$ ,

$R^3$ 表示 $C_{1-6}$ 烷基、 $CH_2N_3$ 、 $NH_2$ 或烯丙基,

Y表示氢,卤素或羟基,和

Z表示 $CR^4$ 或N,其中 $R^4$ 表示氢、 $C_{1-6}$ 烷基、 $C_{1-6}$ 烷氧基、COH或COOH。

30. 如权利要求29所述的用途,其中 $R^a$ 和 $R^b$ 均表示氢,或 $R^a$ 表示氢且 $R^b$ 表示甲基。

31. 如权利要求29或30所述的用途,其中 $R^2$ 表示氢、甲基或甲氧基。

32. 如权利要求29至31中任一项所述的用途,其中X表示 $-OR^3$ ,并且 $R^3$ 表示 $C_{1-6}$ 烷基, $CH_2N_3$ 、 $NH_2$ 或烯丙基。

33. 如权利要求29至32中任一项所述的用途,其中Y表示氢。

34. 如权利要求29至33中任一项所述的用途,其中Z表示 $CR^4$ ,并且其中 $R^4$ 表示甲氧基、COH或COOH。

35. 如权利要求29至34中任一项所述的用途,其中3'-封闭的核苷酸三磷酸选自E1至E11。

36. 无机焦磷酸酶在核酸合成方法中的应用。

37. 如权利要求36所述的用途,其中无机焦磷酸酶是来自酿酒酵母的纯化重组无机焦磷酸酶。