

(19)中华人民共和国国家知识产权局



(12)发明专利申请

(10)申请公布号 CN 109477080 A

(43)申请公布日 2019.03.15

(21)申请号 201780037427.0

(74)专利代理机构 北京安信方达知识产权代理

(22)申请日 2017.06.13

有限公司 11262

(30)优先权数据

代理人 王玮玮 郑霞

1655475 2016.06.14 FR

(51)Int.Cl.

(85)PCT国际申请进入国家阶段日

C12N 9/12(2006.01)

2018.12.14

C12P 19/34(2006.01)

(86)PCT国际申请的申请数据

PCT/FR2017/051519 2017.06.13

(87)PCT国际申请的公布数据

W02017/216472 FR 2017.12.21

(71)申请人 DNA斯克瑞普特公司

权利要求书3页 说明书22页

地址 法国巴黎

序列表18页 附图3页

申请人 巴斯德研究所 国家科学研究中心

(72)发明人 托马斯·伊贝尔特 马克·德拉吕

(54)发明名称

polX家族的DNA聚合酶的变体

(57)摘要

本发明涉及能够在没有模板链的情况下合成核酸分子的polX家族的DNA聚合酶的变体或该聚合酶的功能片段的变体以及所述变体的用途，特别是用于合成包含3'-OH-修饰的核苷酸的核酸分子的用途，所述变体包含在至少一个特定位置处的残基的至少一个突变。

1. 一种polX家族的DNA聚合酶的变体或该聚合酶的功能片段的变体,所述聚合酶能够在没有模板链的情况下合成核酸分子,所述变体包含在选自由E457、T331、G332、G333、F334、R336、K338、H342、D343、V344、D345、F346、A397、D399、D434、V436、A446、L447、L448、G449、W450、G452、R454、Q455、F456、R458、R461、N474、E491、D501、Y502、I503、P505、R508、N509和A510组成的组的至少一个位置处的残基或功能等同的残基的至少一个突变,所示位置通过与SEQ ID NO:1比对确定。

2. 根据权利要求1所述的polX家族的DNA聚合酶的变体,所述变体能够合成DNA链和/或RNA链。

3. 根据权利要求1或2所述的polX家族的DNA聚合酶的变体,所述变体为Pol IV、Polμ或末端脱氧核糖核苷转移酶(TdT)的变体。

4. 根据前述权利要求之一所述的polX家族的DNA聚合酶的变体,所述变体与根据SEQ ID NO:1的序列具有至少60%的同一性,优选地与根据SEQ ID NO:1的序列具有至少70%、80%、85%、90%、95%、96%、97%、98%、99%的同一性。

5. 根据前述权利要求之一所述的polX家族的DNA聚合酶的变体,其中至少一个突变由一个或更多个氨基酸残基的取代、缺失或添加组成。

6. 根据前述权利要求之一所述的polX家族的DNA聚合酶的变体,所述变体包含在选自由T331、G332、G333、F334、R336、D343、L447、L448、G449、W450、G452、R454、Q455、E457、R461和R508组成的组的至少一个位置处的残基或功能等同的残基的至少一个突变,优选地在选自由R336、R454和E457组成的组的至少一个位置处的残基或功能等同的残基的至少一个突变,所示位置通过与SEQ ID NO:1比对确定。

7. 根据前述权利要求之一所述的polX家族的DNA聚合酶的变体,所述变体包含在选自由R336、R454和E457组成的组的至少两个位置处的残基的至少一个突变,优选地在所述三个位置R336、R454和E457处的残基的突变。

8. 根据前述权利要求之一所述的polX家族的DNA聚合酶的变体,所述变体具有在以下序列的至少一个半保守区域中的残基的至少一个突变

(i) X₁X₂GGFR₁R₂GKX₃X₄ (SEQ ID NO:4) ,

其中

X₁表示选自M、I、V、L的残基

X₂表示选自T、A、M、Q的残基

X₃表示选自M、K、E、Q、L、S、P、R、D的残基

X₄表示选自T、I、M、F、K、V、Y、E、Q、H、S、R、D的残基

(ii) X₁X₂LGX₃X₄GSR₁X₅X₆ER₂ (SEQ ID NO:5)

其中

X₁表示选自A、C、G、S的残基

X₂表示选自L、T、R的残基

X₃表示选自W、Y的残基

X₄表示选自T、S、I的残基

X₅表示选自Q、L、H、F、Y、N、E、D或Ø的残基

X₆表示选自F、Y的残基

(iii) LX₁YX₂X₃PX₄X₅RNA (SEQ ID NO:6)

X₁表示选自D、E、S、P、A、K的残基

X₂表示选自I、L、M、V、A、T的残基

X₃表示选自E、Q、P、Y、L、K、G、N的残基

X₄表示选自W、S、V、E、R、Q、T、C、K、H的残基

X₅表示选自E、Q、D、H、L的残基。

9. 根据权利要求8所述的polX家族的DNA聚合酶的变体，所述变体具有

-在序列SEQ ID NO:4的半保守区域的R₁、R₂和/或K中的至少一个位置处的残基的至少一个取代；和/或

-在序列SEQ ID NO:5的半保守区域的S、R1和/或E中的至少一个位置处的残基的至少一个取代；和/或

-在序列SEQ ID NO:6的半保守区域的位置X1处的残基的缺失和/或在位置R和/或N处的至少一个取代。

10. 根据前述权利要求之一所述的polX家族的DNA聚合酶的变体，所述变体包含在选自由R336、K338、H342、A397、S453、R454、E457、R461、N474、D501、Y502、I503、R508和N509组成的组的至少一个位置处的残基或功能等同的残基的取代，优选地在选自由R336、A397、R454、E457、R461、N474、D501、Y502和I503组成的组的至少一个位置处的残基或功能等同的残基的取代，更优选地在选自由R336、R454和E457组成的组的至少一个位置处的残基或功能等同的残基的取代，甚至更优选地在位置E457处的残基或功能等同的残基的至少一个取代，所示位置通过与SEQ ID NO:1比对确定。

11. 根据权利要求10所述的polX家族的DNA聚合酶的变体，其中在位置R336、K338、H342、A397、S453、R454、E457、R461、N474、D501、Y502、I503、R508和N509处的取代选自由R336K/H/N/G/D、K338A/C/G/S/T/N、H342A/C/G/S/T/N、A397R/H/K/D/E、S453A/C/G/S/T、R454F/Y/W/A、E457G/N/S/T、N474S/T/N/Q、D501A/G/X、Y502A/G/X、I503A/G/X、R508A/C/G/S/T、N509A/C/G/S/T组成的组。

12. 根据前述权利要求之一所述的polX家族的DNA聚合酶的变体，其中所述变体包含或具有表1中列出的取代、缺失、取代和/或缺失的组合，所示位置通过与SEQ ID NO:1比对确定。

13. 根据前述权利要求之一所述的polX家族的DNA聚合酶的变体，所述变体为序列SEQ ID NO:1的TdT的变体，且还包含位置C378至L406之间的残基或功能等同的残基被序列SEQ ID NO:2的聚合酶Polμ的残基H363至C390或功能等同的残基的取代。

14. 根据前述权利要求之一所述的polX家族的DNA聚合酶的变体，其中所述变体包含选自R336G-E457N、R336N-E457N、R336N-R454A-E457N、R336N-E457N、R336N-R454A-E457G、R336N-E457G、R336G-R454A-E457N、R336G-E457N的取代组合，所示位置通过与SEQ ID NO:1比对确定。

15. 一种核酸，所述核酸编码根据权利要求1至14中之一所述的polX家族的DNA聚合酶的变体。

16. 一种根据权利要求15所述的核酸的表达盒。

17. 一种载体，所述载体包含根据权利要求15所述的核酸或根据权利要求16所述的表

达盒。

18. 一种宿主细胞,所述宿主细胞包含根据权利要求15所述的核酸或根据权利要求16所述的表达盒或根据权利要求17所述的载体。

19. 根据权利要求15所述的核酸、根据权利要求16所述的表达盒、根据权利要求17所述的载体或根据权利要求18所述的细胞产生根据权利要求1-14中任一项所述的polX家族的DNA聚合酶的变体的用途。

20. 一种用于产生根据权利要求1-14中任一项所述的polX家族的DNA聚合酶的变体的方法,根据该方法,在允许编码所述变体的核酸表达的培养条件下培养根据权利要求18所述的宿主细胞,并任选地从培养基或从所述宿主细胞回收如此表达的所述变体。

21. 根据权利要求1-14中任一项所述的polX家族的DNA聚合酶的变体在没有模板链的情况下由3'-OH-修饰的核苷酸合成核酸分子的用途。

22. 根据权利要求21所述的用途,用于合成DNA链或RNA链。

23. 一种用于在没有模板链的情况下酶促合成核酸分子的方法,根据该方法,在根据权利要求1-14中任一项所述的polX家族的DNA聚合酶的变体的存在下,使引物链与至少一种核苷酸,优选地3'-OH-修饰的核苷酸接触。

24. 一种试剂盒,所述试剂盒用于在没有模板链的情况下酶促合成核酸分子,所述试剂盒包含至少一种根据权利要求1-14中任一项所述的polX家族的DNA聚合酶的变体、核苷酸,优选地3'-OH-修饰的核苷酸和任选地至少一种核苷酸引物。

polX家族的DNA聚合酶的变体

[0001] 引言

[0002] 本发明属于酶增强领域。本发明涉及polX家族的DNA聚合酶的增强型变体，编码所述变体的核酸，所述变体在宿主细胞中的产生，所述变体用于在没有模板链的情况下合成核酸分子的用途，和用于在没有模板链的情况下合成核酸分子的试剂盒。

[0003] 核酸片段的化学合成是在实验室中广泛使用的技术 (Adams 等人, 1983, J.Amer.Chem.Soc. 105:661; Froehler 等人, 1983, Tetrahedron Lett. 24:3171)。该技术能够快速产生包含所期望核苷酸序列的核酸分子。与在5'至3'方向上进行合成的酶不同，化学合成发生在3'至5'方向上。然而，化学合成具有某些缺陷。事实上，化学合成要求使用多于一种(multiple)溶剂和试剂。另外，仅能获得短核酸片段，然后必须对其进行组装以获得所期望的最终核酸链。

[0004] 已经开发了替代性解决方案，其使用能够从初始核酸片段(引物)且在不存在模板链的情况下进行核苷酸之间的偶联反应的酶。数种聚合酶似乎适合于此类合成方法。

[0005] 存在能够在有或没有模板链的情况下催化核酸链的合成的非常大量的DNA聚合酶。例如，polX家族的DNA聚合酶特别地以DNA修复机制或对DNA序列中出现的错误的校正机制参与广泛的生物学过程。这些酶能够将核苷酸引入已经历了在序列中的错误被鉴定之后的切除的核酸链。polX家族的DNA聚合酶包括DNA聚合酶β(Polβ)、λ(Polλ)、μ(Polμ)、来自酵母的IV(Pol IV)和末端脱氧核糖核苷转移酶(TdT)。特别是TdT被广泛用于酶促合成核酸分子的方法。

[0006] 然而，这些DNA聚合酶通常仅允许掺入天然核苷酸。在所有情况下，在非天然核苷酸特别是3'OH-修饰的核苷酸的存在下天然DNA聚合酶失去其催化活性，3'OH-修饰的核苷酸具有比天然核苷酸大的空间体积。

[0007] 然而，修饰的核苷酸的使用对于某些特定应用可能是有用的。因此，有必要开发能够催化通过掺入此类核苷酸来合成核酸链的酶。因此开发了DNA聚合酶的变体，目的是使用含有显著结构修饰的核苷酸。

[0008] 然而，目前可得的变体并不完全令人满意，特别是由于它们的低活性，仅与实验室规模的酶促合成相容。因此，如果可能的话，对于能够在工业规模上、在不存在模板链的情况下并使用修饰的核苷酸合成核酸的DNA聚合酶存在需求。

[0009] 发明概述

[0010] 本发明消除了某些技术障碍，这些技术障碍妨碍了DNA聚合酶在工业规模用于核酸的酶促合成。

[0011] 因此，本发明提出了能够在模板链的不存在下合成核酸并且可以使用修饰的核苷酸的polX家族的DNA聚合酶的变体。开发的变体具有修饰的核苷酸的掺入能力，远远优于它们所源自的天然DNA聚合酶的修饰的核苷酸的掺入能力。特别地，本发明涉及的DNA聚合酶的变体对于掺入具有糖修饰的核苷酸特别有效。事实上，本发明人开发了与其所源自的DNA聚合酶的催化口袋体积相比具有增加的催化口袋体积的变体，所述变体促进了比天然核苷酸体积大的修饰的核苷酸的掺入。更准确地说，本发明涉及的polX家族的DNA聚合酶的变体

包含氨基酸的至少一个突变,所述突变直接在酶的催化腔处干预或允许该腔的轮廓变形以便适应由于核苷酸中存在的修饰而变化的空间体积。例如,引入的突变使得扩大修饰的核苷酸的3'-OH末端所适合的酶的催化腔是可能的。可选地或另外地,所产生的突变使得可能使催化腔的体积膨大或增加,以增加3'-OH-修饰的核苷酸对催化口袋的接近和/或以赋予酶结构必要的柔性以允许其适应3'-OH-修饰的核苷酸的空间上显著的改变。由于此类突变,在聚合酶与待延长的核酸片段结合后,修饰的核苷酸进入催化口袋(其进入部位(access)被扩大)的核心并采取最佳空间构象,在核酸链的最后一个核苷酸的3'-OH末端与修饰的核苷酸的5'-三磷酸末端之间产生磷酸二酯键。

[0012] 因此,本发明涉及能够在没有模板链的情况下合成核酸分子的polX家族的DNA聚合酶的变体,或此类聚合酶的功能片段的变体,所述变体包含在选自由T331、G332、G333、F334、R336、K338、H342、D343、V344、D345、F346、A397、D399、D434、V436、A446、L447、L448、G449、W450、G452、R454、Q455、F456、E457、R458、R461、N474、E491、D501、Y502、I503、P505、R508、N509和A510组成的组的至少一个位置处的残基或功能等同的残基的至少一个突变,所示位置通过与SEQ ID N0:1比对确定。

[0013] 在一个具体实施方案中,变体能够合成DNA链或RNA链。

[0014] 本发明特别涉及polX家族的DNA聚合酶的变体,并且特别是酵母Pol IV、Polμ或野生型TdT的变体,并且所述变体包含所选择的突变。在一个具体实施方案中,根据本发明的变体是序列SEQ ID N0:1的TdT的变体或具有与SEQ ID N0:1的序列具有至少70%、80%、85%、90%、95%、96%、97%、98%、99%同一性的同源序列并且携带所选择的突变的变体。

[0015] 本发明还涉及编码根据本发明的polX家族的DNA聚合酶的变体的核酸、包含根据本发明的核酸的表达盒和包含根据本发明的核酸或表达盒的载体。编码本发明的变体的核酸可以是根据本发明的DNA聚合酶的成熟形式或前体形式的核酸。

[0016] 本发明还涉及根据本发明的核酸、表达盒或载体转化或转染宿主细胞的用途。本发明还涉及包含编码根据本发明的polX家族的DNA聚合酶的核酸、表达盒或载体的宿主细胞。本发明涉及此类核酸、表达盒、载体或宿主细胞产生根据本发明的polX家族的DNA聚合酶的变体的用途。

[0017] 本发明还涉及用于产生根据本发明的polX家族的DNA聚合酶的变体的方法,所述方法包括用根据本发明的核酸、表达盒或载体转化或转染宿主细胞,在允许编码所述变体的核酸表达的培养条件下培养转化/转染的宿主细胞,以及任选地收集由宿主细胞产生的polX家族的DNA聚合酶的变体。

[0018] 宿主细胞可以是原核细胞或真核细胞。特别地,宿主细胞可以是微生物,优选地是细菌、酵母或真菌。在一个实施方案中,宿主细胞是细菌,优选地是大肠杆菌(E.coli)。在另一个实施方案中,宿主细胞是酵母,优选地是巴斯德毕赤酵母(P.pastoris)或乳酸克鲁维酵母(K.lactis)。在另一个实施方案中,宿主细胞是哺乳动物细胞,优选地是COS7或CHO细胞。

[0019] 本发明还涉及根据本发明的polX家族的DNA聚合酶的变体由3'-OH-修饰的核苷酸在没有模板链的情况下合成核酸分子的用途。当然,根据本发明的polX家族的DNA聚合酶的变体也可以在本发明的背景中从未修饰的核苷酸或修饰的和未修饰的核苷酸的混合物在没有模板链的情况下合成核酸分子。

[0020] 本发明还提出了一种用于在没有模板链的情况下酶促合成核酸分子的方法,根据该方法,在根据本发明的polX家族的DNA聚合酶的变体的存在下,使引物链与至少一种核苷酸,优选地3'-OH-修饰的核苷酸接触。该方法尤其可以使用纯化的变体、被转化以表达所述变体的宿主细胞的培养物(culture medium)和/或该宿主细胞的细胞提取物来实施。

[0021] 本发明还涉及一种用于在没有模板链的情况下酶促合成核酸分子的试剂盒,所述试剂盒包含至少一种根据本发明的polX家族的DNA聚合酶的变体、核苷酸,优选地3'-OH-修饰的核苷酸、和任选地至少一种引物链或核苷酸引物、和/或反应缓冲液。

[0022] 附图描述

[0023] 图1:根据本发明的实施方案的TdT变体的级分的SDS-PAGE凝胶(M:分子量标志;1:加载前的离心液(Centrifugate);2:加载后的离心液;3:加载后的洗涤缓冲液;4:洗脱级分3mL;5:洗脱级分30mL;6:洗脱峰组合(assembly);7:浓缩物);

[0024] 图2:通过Mutalin在线比对软件(<http://multalin.toulouse.inra.fr/multalin/multalin.html>)获得的对来自智人(Homo sapiens)的DNA聚合酶Pol μ (UniProtKB Q9NP87)、来自黑猩猩(Pan troglodytes)的Pol μ (UniProtKB H2QUI0),来自小鼠(Mus musculus)的Pol μ (UniProtKB Q924W4)、来自狗(Canis lupus familiaris)的TdT(UniProtKB F1P657)、来自小鼠的TdT(UniProtKB Q3UZ80)、来自原鸡(Gallus gallus)的TdT(UniProtKB P36195)和来自智人的TdT(UniProtKB P04053)的氨基酸序列的比对;

[0025] 图3:在先前5'-放射性标记的引物和用3'-0-氨基-2',3'-双脱氧腺苷-5'-三磷酸修饰的核苷酸(ONH2凝胶)或3'-biot-EDA-2',3'-双脱氧腺苷-5'-三磷酸修饰的核苷酸(Biot-EDA凝胶)的存在下,对序列SEQ ID NO:3的截短的野生型TdT与包含表1所示的不同取代的该截短的TdT的多种变体的活性在SDS-PAGE凝胶(No:无酶存在;wt:序列SEQ ID NO:3的截短的野生型TdT;DSi:表1中定义的变体i)上的比较;;

[0026] 图4:在先前5'-放射性标记的引物和用3'-0-氨基-2',3'-双脱氧腺苷-5'-三磷酸修饰的不同核苷酸的存在下,对根据本发明的变体DS124(参见表1)的活性在SDS-PAGE凝胶上的研究;

[0027] 图5:在先前5'-放射性标记的引物和用3'-0-氨基-2',3'-双脱氧腺苷-5'-三磷酸修饰的不同核苷酸的存在下,对变体DS22、DS24、DS124、DS125、DS126、DS127和DS128的活性在SDS-PAGE凝胶上的研究;

[0028] 图6:借助根据本发明的具有取代组合R336N-R454A-E457G的TdT的变体(DS125)在序列5'-AAAAAAAAGGGG-3'(SEQ ID NO:14)的引物之后合成序列5'-GTACGCTAGT-3'(SEQ ID NO:15)的DNA链。

[0029] 发明详述

[0030] 定义

[0031] 根据以下命名法,氨基酸在本文件中由单字母或三字母代码表示:A:Ala(丙氨酸);R:Arg(精氨酸);N:Asn(天冬酰胺);D:Asp(天冬氨酸);C:Cys(半胱氨酸);Q:Gln(谷氨酰胺);E:Glu(谷氨酸);G:Gly(甘氨酸);H:His(组氨酸);I:Ile(异亮氨酸);L:Leu(亮氨酸);K:Lys(赖氨酸);M:Met(甲硫氨酸);F:Phe(苯丙氨酸);P:Pro(脯氨酸);S:Ser(丝氨酸);T:Thr(苏氨酸);W:Trp(色氨酸);Y:Tyr(酪氨酸);V:Val(缬氨酸)。

[0032] 出于本发明的目的,两个核酸或氨基酸序列之间的“同一性百分比”意指在最佳比

对(best alignment)后获得的待比较的两个序列之间的相同核苷酸或氨基酸残基的百分比,该百分比是纯粹统计学的,且两个序列之间的差异随机地分布并分布在其整个长度上。如以下计算的,最佳比对(best alignment)或最优比对(optimal alignment)是,待比较的两个序列之间的同一性百分比最高的比对。传统地,两个核酸或氨基酸序列之间的序列比较通过在最优比对后比较这些序列来进行,所述比较通过区段或比较窗口进行以鉴定和比较具有序列相似性的局部区域。除了手动地,可以借助以下实现对于比较的最佳序列比对:借助Smith和Waterman(1981)(Ad.App.Math.2:482)的局部同源性算法、借助Needleman和Wunsch(1970)(J.Mol.Biol.48:443)的局部同源性算法、借助Pearson和Lipman(1988)(Proc.Natl.Acad.Sci.USA 85:2444)的相似性搜索方法、借助使用这些算法的计算机程序(Wisconsin遗传学软件包中的GAP、BESTFIT、FASTA和TFASTA,Genetics Computer Group,575 Science Dr.,Madison,WI)、借助Mutalin在线比对软件(<http://multalin.toulouse.inra.fr/multalin/multalin.html>;1988,Nucl.Acids Res.16(22),10881-10890)。通过由比较窗口比较这两个最佳比对序列来确定两个核酸或氨基酸序列之间的同一性百分比,在比较窗口中,对于这两个序列之间的最佳比对,待比较的核酸或氨基酸序列的区域可以包含相对于参考序列的添加或缺失。同一性百分比通过以下计算:确定两个序列之间核苷酸或氨基酸残基相同的相同位置的数目,将相同位置的这一数目除以比较窗口中的位置总数,并将获得的结果乘以100,获得这两个序列之间的同一性百分比。

[0033] 本发明涉及的变体根据它们在特定残基上的突变来描述,其位置通过与酶序列SEQ ID N0:1比对或参考酶序列SEQ ID N0:1来确定。在本发明的上下文中,还涉及在功能等同的残基上携带这些相同突变的任何变体。“功能等同的残基”意指在与SEQ ID N0:1序列同源的polX家族的DNA聚合酶的序列中的、并且具有相同的功能作用的残基。使用序列比对,例如使用Mutalin在线比对软件(<http://multalin.toulouse.inra.fr/multalin/multalin.html>;1988,Nucl.Acids Res.16(22),10881-10890)鉴定功能等同的残基。比对后,功能等同的残基处于所讨论的不同序列上的同源位置。可以在polX家族的任何DNA聚合酶与其天然变体,包括种间变体,之间进行序列比对和功能等同残基的鉴定。举例来说,人类TdT(UniProtKB P04053)的残基L40在功能上等同于鸡TdT(UniProtKB P36195)的残基M40和黑猩猩Polμ(UniProtKB H2QUI0)的残基V40,所述残基在序列比对后被考虑(图2)。

[0034] “功能片段”意指具有DNA聚合酶活性的polX家族的DNA聚合酶的片段。该片段可以包含polX家族的DNA聚合酶的100、200、300、310、320、330、340、350、360、370、380或更多个连续氨基酸。优选地,该片段包含polX家族的DNA聚合酶的组成所述酶的催化片段的380个连续氨基酸。

[0035] 术语“突变体”和“变体”可以互换使用,指来源于polX家族的DNA聚合酶,并且尤其是TdT,诸如根据序列SEQ ID N0:1的鼠(murine)TdT的多肽,或来源于此类DNA聚合酶的功能片段的多肽,并且包含在一个或更多个位置处的改变,即取代、插入和/或缺失,并具有DNA聚合酶活性。变体可以通过本领域熟知的各种技术获得。特别地,用于修饰编码野生型蛋白的DNA序列的示例性技术包括但不限于,定向诱变、随机诱变和合成寡核苷酸的构建。

[0036] 如本文关于氨基酸位置或残基使用的术语“修饰”或“突变”意指,相对于参考野生型蛋白的氨基酸修饰了所讨论位置处的氨基酸。此类修饰包括一个或更多个氨基酸的取代、缺失和/或插入,且特别是在一个或更多个位置,且特别是1、2、3、4、5或更多个位置处的

1至5、1至4、1至3、1至2个氨基酸的取代、缺失和/或插入。

[0037] 关于氨基酸位置或残基的术语“取代”意指，特定位置处的氨基酸已被野生型或亲本DNA聚合酶中的氨基酸以外的氨基酸代替。优选地，术语“取代”意指一个氨基酸残基被选自以下的另一个氨基酸残基代替：20个标准天然氨基酸残基、罕见的天然存在的氨基酸残基(例如，羟脯氨酸、羟赖氨酸、别羟基赖氨酸、6-N-甲基赖氨酸、N-乙基甘氨酸、N-甲基甘氨酸、N-乙基天冬酰胺、别-异亮氨酸、N-甲基异亮氨酸、N-甲基缬氨酸、焦谷氨酰胺、氨基丁酸、鸟氨酸)和通常经合成产生的罕见的非天然氨基酸残基(例如，正亮氨酸、正缬氨酸和环己基丙氨酸)。优选地，术语“取代”意指一个氨基酸残基被选自20个标准天然存在的氨基酸残基(G、P、A、V、L、I、M、C、F、Y、W、H、K、R、Q、N、E、D、S和T)的另一个氨基酸残基代替。取代可以是保守的或非保守的。保守取代在同一组氨基酸内进行，包括碱性氨基酸(精氨酸、赖氨酸和组氨酸)、酸性氨基酸(谷氨酸和天冬氨酸)、极性氨基酸(谷氨酰胺和天冬酰胺)、疏水性氨基酸(甲硫氨酸、亮氨酸、异亮氨酸和缬氨酸)、芳香族氨基酸(苯丙氨酸、色氨酸和酪氨酸)和小氨基酸(甘氨酸、丙氨酸、丝氨酸和苏氨酸)。在本文件中，以下术语用于表示取代：R454F表示在SEQ ID N0:1的位置454处的氨基酸残基(精氨酸，R)被苯丙氨酸(F)代替。N474S/T/N/Q意指，在位置474处的氨基酸残基(天冬酰胺，N)可以被丝氨酸(S)、苏氨酸(T)、天冬酰胺(N)或谷氨酰胺(Q)代替。“+”符号表示取代的组合。

[0038] 本发明涉及能够在没有模板链，且特别是DNA链或RNA链的情况下合成核酸分子的polX家族的DNA聚合酶(EC 2.7.7.7; Advances in Protein Chemistry, 第71卷, 401-440)的变体。polX家族的DNA聚合酶特别地包括DNA聚合酶Pol β (人类中为UniProt P06746; 小鼠中为Q8K409)、Pol σ 、Pol λ (人类中为UniProt Q9UGP5; 小鼠中为Q9QUG2和Q9QXE2)和Pol μ (人类中为UniProt Q9NP87; 小鼠中为Q9J1W4)、Pol4(多孢范氏酵母(Vanderwaltomya polyspora)中为UniProt A7TER5; 酵母酿酒酵母(Saccharomyces cerevisiae)中为P25615)、和末端脱氧核糖核苷转移酶或TdT(EC 2.7.7.31; 人类中为UniProt P04053; 小鼠中为P09838)。

[0039] 更准确地，本发明涉及能够在没有链模板链的情况下合成核酸分子的polX家族的DNA聚合酶的变体，或此类聚合酶的功能片段的变体，所述变体包含在选自由T331、G332、G333、F334、R336、K338、H342、D343、V344、D345、F346、A397、D399、D434、V436、A446、L447、L448、G449、W450、G452、R454、Q455、F456、E457、R458、R461、N474、E491、D501、Y502、I503、P505、R508、N509和A510组成的组的至少一个位置处的残基或功能等同的残基的至少一个突变，所示位置通过与序列SEQ ID N0:1比对或参考序列SEQ ID N0:1来确定。

[0040] 在一个实施方案中，变体能够合成DNA链和/或RNA链。

[0041] 表述“包含至少一个突变(comprises at least one mutation)”或“包含至少一个突变(comprising at least one mutation)”意指，该变体相对于多肽序列SEQ ID N0:1具有所示的一个或更多个突变，但可以具有其他修饰，特别是取代、缺失或添加。

[0042] 通常，在以上位置处的一个或更多个残基的突变扩大了催化口袋(通过靶向例如位置W450、D434、D435、H342、D343、T331、R336、D399、R461、和/或R508)，增加了对催化口袋的接近(通过靶向例如位置R458、E455、R454、A397、K338、和/或N509)，和/或赋予酶结构更大的柔性，允许其接受具有较大的空间体积的修饰的核苷酸(通过靶向例如位置V436、F346、V344、F334、M330、L448、E491、E457和/或N474)。

[0043] 本发明涉及的变体可以是Pol IV、Pol μ 、Pol β 、Pol λ 或TdT的变体，优选地Pol IV、Pol μ 或TdT的变体。可选地，变体可以是组合了例如polX家族的至少两种DNA聚合酶的不同序列的部分的嵌合酶的变体。

[0044] 在一个具体实施方案中，变体与根据SEQ ID NO:1的序列具有至少60%的同一性，优选地与根据SEQ ID NO:1的序列具有至少70%、80%、85%、90%、95%、96%、97%、98%、99%和小于100%的同一性。

[0045] 根据本发明，突变可以由一个或更多个氨基酸残基的取代、缺失或添加组成。在缺失的情况下，使用标注X，其表示，编码所讨论残基的密码子被终止密码子代替；因此所有之后的氨基酸以及所讨论的残基被缺失。因此，突变D501X意指，酶终止于位置501处的天冬氨酸(D)之前的残基，即位置500处的亮氨酸(L)，其后的所有残基都已被缺失。另一方面，标注 \emptyset 表示所考虑的残基的简单点缺失。因此，突变D501 \emptyset 意指，位置501处的天冬氨酸(D)已被缺失。

[0046] 优选地，根据本发明的变体包含在选自由T331、G332、G333、F334、R336、D343、L447、L448、G449、W450、G452、R454、Q455、E457和R508组成的组的至少一个位置处的残基或功能等同的残基的至少一个突变，优选地在选自由R336、R454、E457组成的组的至少一个位置处的残基或功能等同的残基的至少一个突变，所示位置通过与SEQ ID NO:1比对确定。

[0047] 在一个具体实施方案中，所述变体包含在选自由R336、R454和E457组成的组的至少两个位置处的残基的至少一个突变，优选地在所述三个位置R336、R454和E457处的残基或功能等同的残基的突变，所示位置通过与SEQ ID NO:1比对确定。

[0048] 在一个具体实施方案中，变体还包含在序列X₁X₂GGFR₁R₂GKX₃X₄(SEQ ID NO:4)的至少半保守区域中的残基的至少一个突变，其中

[0049] X₁表示选自M、I、V、L的残基

[0050] X₂表示选自T、A、M、Q的残基

[0051] X₃表示选自M、K、E、Q、L、S、P、R、D的残基

[0052] X₄表示选自T、I、M、F、K、V、Y、E、Q、H、S、R、D的残基。

[0053] 优选地，所述变体具有在序列SEQ ID NO:4的半保守区域的R₁、R₂和/或K中的至少一个位置处的残基的至少一个取代。

[0054] 在另一个具体实施方案中，变体还包含在序列X₁X₂LGX₃X₄GSR₁X₅X₆ER₂(SEQ ID NO:5)的至少一个半保守区域中的残基的至少一个突变，其中

[0055] X₁表示选自A、C、G、S的残基

[0056] X₂表示选自L、T、R的残基

[0057] X₃表示选自W、Y的残基

[0058] X₄表示选自T、S、I的残基

[0059] X₅表示选自Q、L、H、F、Y、N、E、D或 \emptyset 的残基

[0060] X₆表示选自F、Y的残基。

[0061] 优选地，所述变体具有在序列SEQ ID NO:5的半保守区域的S、R₁和/或E中的至少一个位置处的残基的至少一个取代。

[0062] 在另一个具体实施方案中，变体还包含在序列LX₁YX₂X₃PX₄X₅RNA(SEQ ID NO:6)的至少一个半保守区域中的残基的至少一个突变，其中

[0063] X₁表示选自D、E、S、P、A、K的残基

[0064] X₂表示选自I、L、M、V、A、T的残基

[0065] X₃表示选自E、Q、P、Y、L、K、G、N的残基

[0066] X₄表示选自W、S、V、E、R、Q、T、C、K、H的残基

[0067] X₅表示选自E、Q、D、H、L的残基。

[0068] 优选地,所述变体具有在序列SEQ ID N0:6的半保守区域的位置X₁处的残基的至少一个残基缺失和/或在位置R和/或N处的至少一个取代。

[0069] 在一个具体实施方案中,变体包含在选自由R336、K338、H342、A397、S453、R454、E457、N474、D501、Y502、I503、R508和N509组成的组的至少一个位置处的残基或功能等同的残基的取代,优选地在选自由R336、A397、R454、E457、N474、D501、Y502和I503组成的组的至少一个位置处的残基或功能等同的残基的取代,更优选地在选自由R336、R454、E457组成的组的至少一个位置处的残基或功能等同的残基的至少一个取代,所示位置通过与SEQ ID N0:1比对确定。

[0070] 本发明优选地涉及polX家族的DNA聚合酶的变体,所述变体包含选自由R336K/H/G/N/D、K338A/C/G/S/T/N、H342A/C/G/S/T/N、A397R/H/K/D/E、S453A/C/G/S/T、R454F/Y/W/A、E457G/N/S/T、N474S/T/N/Q、D501A/G/X、Y502A/G/X、I503A/G/X、R508A/C/G/S/T、N509A/C/G/S/T组成的组的至少一个取代。在一个具体实施方案中,变体包含在选自由R336、R454、E457组成的组的至少两个位置处的残基或功能等同的残基的取代,优选地在所述三个位置处的残基或功能等同的残基的取代,所示位置通过与SEQ ID N0:1比对确定。特别地,取代选自由R336K/H/G/N/D、R454F/Y/W/A和E457N/D/G/S/T组成的组,优选地选自由R336N/G、R454A和E457G/N/S/T组成的组。

[0071] 在一个实施方案中,变体包含根据E457G/N/S/T的至少一个取代。

[0072] 有利地,变体包含选自以上提及的组的取代组合。组合可以由选自所述组的2、3、4、5、6、7、8、9、10或11个取代组成。

[0073] 本发明更具体地涉及能够在没有模板链诸如DNA链或RNA链的情况下合成核酸分子的polX家族的DNA聚合酶的变体,或涉及此类聚合酶的功能片段的变体,所述变体包含表1中描述的突变的至少一种组合,所示位置通过与SEQ ID N0:1比对确定。

[0074] 在一个实施方案中, polX家族的DNA聚合酶的变体包含以下中的取代组合:R336G-E457N;R336N-E457N;R336N-R454A-E457N;R336N-R454A-E457G;R336N-E457G;和R336G-R454A-E457N。

[0075] 表1:polX家族的DNA聚合酶的变体的示例性突变组合

[0076]

	突变的组合
DS1	R454F - E457N - A397D
DS2	R454F - E457N
DS3	R454Y - E457N - A397D
DS4	R454Y - E457N
DS5	R454W - E457N - A397D
DS6	R454W - E457N
DS7	R335A - E457N - A397D
DS8	R335A - E457N
DS9	R335G - E457N - A397D
DS10	R335G - E457N
DS11	R335N - E457N - A397D
DS12	R335N - E457N
DS13	R335D - E457N - A397D
DS14	R335D - E457N
DS15	R336K - E457N - A397D
DS16	R336K - E457N
DS17	R336H - E457N - A397D
DS18	R336H - E457N
DS21	R336G - E457N - A397D
DS22	R336G - E457N
DS23	R336N - E457N - A397D
DS24	R336N - E457N
DS25	R336D - E457N - A397D
DS26	R336D - E457N
DS27	R454A - E457N
DS28	R454A - E457A
DS29	R454A - E457G
DS30	R454A - E457D

[0077]

DS31	E457N
DS32	E457D
DS33	R454A - E457N - A397D
DS34	R454A - E457N - A397K
DS35	R454A - E457N - N474S
DS36	R454A - E457D - A397D
DS37	D501X
DS38	D501X - E457N
DS39	D501X - E457N - A397D
DS40	R454F - E457S - A397D
DS41	R454F - E457S
DS42	R454Y - E457S - A397D
DS43	R454Y - E457S
DS44	R454W - E457S - A397D
DS45	R454W - E457S
DS46	R335A - E457S - A397D
DS47	R335A - E457S
DS48	R335G - E457S - A397D
DS49	R335G - E457S
DS50	R335N - E457S - A397D
DS51	R335N - E457S
DS52	R335D - E457S - A397D
DS53	R335D - E457S
DS54	R336K - E457S - A397D
DS55	R336K - E457S
DS56	R336H - E457S - A397D
DS57	R336H - E457S
DS60	R336G - E457S - A397D
DS61	R336G - E457S
DS62	R336N - E457S - A397D
DS63	R336N - E457S
DS64	R336D - E457S - A397D

[0078]

DS65	R336D - E457S
DS66	R454A - E457S
DS70	E457S
DS72	R454A - E457S - A397D
DS73	R454A - E457S - A397K
DS74	R454A - E457S - N474S
DS75	D501X - E457S
DS76	D501X - E457S - A397D
DS77	R454F - E457T - A397D
DS78	R454F - E457T
DS79	R454Y - E457T - A397D
DS80	R454Y - E457T
DS81	R454W - E457T - A397D
DS82	R454W - E457T
DS83	R335A - E457T - A397D
DS84	R335A - E457T
DS85	R335G - E457T - A397D
DS86	R335G - E457T
DS87	R335N - E457T - A397D
DS88	R335N - E457T
DS89	R335D - E457T - A397D
DS90	R335D - E457T
DS91	R336K - E457T - A397D
DS92	R336K - E457T
DS93	R336H - E457T - A397D
DS94	R336H - E457T
DS97	R336G - E457T - A397D
DS98	R336G - E457T
DS99	R336N - E457T - A397D
DS100	R336N - E457T
DS101	R336D - E457T - A397D
DS102	R336D - E457T

[0079]

DS103	R454A - E457T
DS104	E457T
DS105	R454A - E457T - A397D
DS106	R454A - E457T - A397K
DS107	R454A - E457T - N474S
DS108	D501X - E457T
DS109	D501X - E457T - A397D
DS110	D502X
DS111	D502X - E457N
DS112	D502X - E457TN - A397D
DS113	D502X - E457S
DS114	D502X - E457S - A397D
DS115	D502X - E457T
DS116	D502X - E457T - A397D
DS117	D503X
DS118	D503X - E457N
DS119	D503X - E457TN - A397D
DS120	D503X - E457S
DS121	D503X - E457S - A397D
DS122	D503X - E457T
DS123	D503X - E457T - A397D
DS124	R336N – R454A – E457N
DS125	R336N – R454A – E457G
DS126	R336N – E457G
DS127	R336G – R454A – E457N

[0080] 在一个具体实施方案中,变体是polX家族的DNA聚合酶的嵌合构建体。“嵌合构建体”意指一种嵌合酶,其由代替所考虑的DNA聚合酶变体中的一个或更多个同源序列的属于polX家族的酶的一个或更多个明确的序列的连接,特别是融合或缀合组成。

[0081] 因此,本发明提出了序列SEQ ID NO:1的TdT的变体,除了包含在以上位置的一个和/或另一个处的一个或更多个点突变外,还包含位置C378至L406之间的残基或功能等同的残基被序列SEQ ID NO:2的聚合酶Polμ的残基H363至C390或功能等同的残基的取代。

[0082] 可选地或另外地,本发明涉及的变体可以具有在N末端部分处的一个或更多个连续氨基酸残基的缺失。这些缺失可以特别地针对参与与其他蛋白结合和/或参与细胞定位的一个或更多个酶结构域。例如,TdT多肽序列包含用于与其他蛋白诸如Ku70/80相互作用的N末端BRCT结构域和核定位结构域(NLS)。

[0083] 在本发明的一个具体实施方案中,变体是序列SEQ ID NO:1的TdT的变体,除了具有以上描述的一个或更多个突变外,还具有对应于野生型TdT的N-末端的残基1-129的缺失。

[0084] 在某些特定情况下,诱变策略可以通过已知信息指导,所述已知信息诸如天然变体的序列、与类似蛋白的序列比较、物理性质、三维结构的研究或涉及若干实体的计算机模拟的研究。

[0085] 本发明涉及编码根据本发明的polX家族的DNA聚合酶的变体的核酸,所述变体能够在没有模板链的情况下合成核酸分子。本发明还涉及用于根据本发明的核酸的表达盒。本发明还涉及包含根据本发明的核酸或表达盒的载体。载体可以选自质粒和病毒载体。

[0086] 编码DNA聚合酶的变体的核酸可以是DNA (cDNA或gDNA)、RNA、两者的混合物。它可以是单链或双链形式或两者的混合物。它可以包括修饰的核苷酸,包括例如修饰的键、修饰的嘌呤或嘧啶碱基、或修饰的糖。它可以通过技术人员已知的所有方法制备,包括化学合成、重组、诱变等。

[0087] 表达盒包含表达能够在没有模板链的情况下合成核酸分子的polX家族的DNA聚合酶的根据本发明的变体所需的所有元件,特别是在宿主细胞中转录和翻译必需的元件。宿主细胞可以是原核细胞或真核细胞。特别地,表达盒包含启动子和终止子,任选地扩增子 (amplifier)。启动子可以是原核的或真核的。优选的原核启动子的实例是:LacI、LacZ、pLacT、ptac、pARA、pBAD、T3或T7噬菌体RNA聚合酶启动子、多角体蛋白启动子、λ噬菌体启动子PR或PL。优选的真核启动子的实例是:CMV早期启动子、HSV胸苷激酶启动子、SV40早期或晚期启动子、小鼠金属硫蛋白-L启动子、和某些逆转录病毒的LTR区。通常,为了选择合适的启动子,技术人员可以有利地参考Sambrook等人(1989)的工作抑或Fuller等人(1996; Immunology in Current Protocols in Molecular Biology)描述的技术。

[0088] 本发明涉及携带编码能够在没有模板链的情况下合成核酸分子的polX家族的DNA聚合酶的根据本发明的变体的核酸或表达盒的载体。载体优选地是表达载体,即,它包含在宿主细胞中表达变体必需的元件。宿主细胞可以是原核生物,例如大肠杆菌,或真核生物。真核生物可以是低等真核生物,诸如酵母(例如,巴斯德毕赤酵母或乳酸克鲁维酵母)或真菌(例如属曲霉属(Aspergillus))或高等真核生物诸如昆虫细胞(例如Sf9或Sf21)、哺乳动物细胞或植物细胞。细胞可以是哺乳动物细胞,例如COS(绿猴细胞系)(例如,COS 1(ATCC CRL-1650)、COS 7(ATCC CRL-1651)、CHO (US 4,889,803; US 5,047,335)、CHO-K1 (ATCC CCL-61))、小鼠细胞和人类细胞。在一个具体实施方案中,细胞是非人类的和非胚胎的。载体可以是质粒、噬菌体、噬菌粒、粘粒、病毒、YAC、BAC、土壤杆菌属(Agrobacterium) pTi质粒等。载体可以优选地包含选自以下的一个或更多个元件:复制起点、多克隆位点和可选择的基因。在优选的实施方案中,载体是质粒。原核载体的一些非详尽实例如下:pQE70、pQE60、pQE-9(Qiagen)、pbs、pD10、phagescript、psiX174、pbluescript SK、pbsks、pNH8A、pNH16A、pNH18A、pNH46A(Stratagene);ptrc99a、pKK223-3、pKK233-3、pDR540、pBR322、和pRIT5(Pharmacia)、pET(Novagen)。真核载体的一些非详尽实例如下:pWLNE0、pSV2CAT、pPICZ、pcDNA3.1 (+) Hyg(Invitrogen)、pOG44、pXT1、pSG(Stratagene);pSVK3、pBPV、pCI-neo(Stratagene)、pMSG、pSVL(Pharmacia);和pQE-30(QLAexpress)。病毒载体可包括但不限于,腺病毒、AAV、HSV、慢病毒等。优选地,表达载体是质粒或病毒载体。

[0089] 编码根据本发明的变体的序列可以任选地包含信号肽。如果不是如此，则可以任选地将甲硫氨酸添加到N-末端。在另一个可选方案中，可以引入异源信号肽。该异源信号肽可以来源于原核生物诸如大肠杆菌或来自真核生物，特别是哺乳动物、昆虫或酵母细胞。

[0090] 本发明涉及根据本发明的多核苷酸、表达盒或载体用于转化或转染细胞的用途。本发明涉及包含编码能够在没有模板链的情况下合成核酸分子的polX家族的DNA聚合酶的变体的核酸、表达盒或载体的宿主细胞，及其用于产生根据本发明的能够在没有重组模板链的情况下合成核酸分子的polX家族的DNA聚合酶的变体的用途。术语“宿主细胞”包括由所述细胞的培养或生长产生的子细胞。在一个具体实施方案中，细胞是非人类的和非胚胎的。本发明还涉及一种用于产生根据本发明的能够在没有重组模板链的情况下合成核酸分子的polX家族的DNA聚合酶的变体的方法，所述方法包括用根据本发明的多核苷酸、表达盒或载体转化或转染细胞；培养转染/转化的细胞；并且收集由细胞产生的能够在没有模板链的情况下合成核酸分子的polX家族的DNA聚合酶的变体。在一个可选的实施方案中，用于产生根据本发明的能够在没有重组模板链的情况下合成核酸分子的polX家族的DNA聚合酶的变体的方法，包括提供包含根据本发明的多核苷酸、表达盒或载体的细胞；培养转染/转化的细胞；并且收集由细胞产生的能够在没有模板链的情况下合成核酸分子的polX家族的DNA聚合酶的变体。特别地，可以用编码变体的核酸瞬时或稳定地转化/转染细胞。该核酸可以作为附加体被包含在细胞中或被包含在染色体中。用于产生重组蛋白的方法是本领域技术人员熟知的。例如，可以提及US 5,004,689、EP 446582、Wang等人(Sci.Sin.B 24:1076-1084, 1994)和Nature 295, page 503)中描述的具体方法用于在大肠杆菌中产生，和提及JAMES等人.(Protein Science (1996), 5:331-340)中描述的具体方法用于在哺乳动物细胞中产生。

[0091] 根据本发明的DNA聚合酶的变体对于在没有模板链的情况下合成核酸特别感兴趣。更具体地，根据本发明的变体具有增加的催化口袋，特别适于使用具有比天然核苷酸更大体积的修饰的核苷酸的核酸合成。根据本发明的变体可以特别地使得可能将修饰的核苷酸，诸如申请W02016/034807中描述的那些掺入到核酸链中。

[0092] 与野生型DNA聚合酶的掺入的动力学相比，具有以上描述的特定突变或突变组合的根据本发明的DNA聚合酶的变体，特别是TdT的变体的掺入动力学大大增强。这些变体可以有利地用于高效酶促DNA合成方法的背景下。

[0093] 因此，本发明还涉及根据本发明的polX家族的DNA聚合酶的变体用于由3'-OH-修饰的核苷酸，特别地申请W02016034807中描述的那些在没有模板链的情况下合成核酸分子的用途。

[0094] 本发明还涉及一种用于在没有模板链的情况下酶促合成核酸分子的方法，根据该方法，在根据本发明的polX家族的DNA聚合酶的变体的存在下，使引物链与至少一种核苷酸，优选地3'-OH-修饰的核苷酸接触。

[0095] 有利地，根据本发明的变体可以用于实施申请W02015/159023中描述的合成方法。

[0096] 本发明还涉及一种用于在没有模板链的情况下酶促合成核酸分子的试剂盒，所述试剂盒包含根据本发明的polX家族的DNA聚合酶的至少一种变体、核苷酸优选地3'-OH-修饰的核苷酸、和任选地至少一种核苷酸引物。

[0097] 本说明书中引用的所有参考文献通过引用并入本说明书中。通过阅读以下实施

例,本发明的其他特征和优点将变得更加容易明显,以下实施例当然是说明性的而非限制性的。

实施例

[0098] 实施例1-根据本发明的polX家族的DNA聚合酶的变体的产生、生产和纯化

[0099] 生产菌株的产生

[0100] 截短的小鼠TdT基因由质粒pET28b生成,质粒pET28b的构建描述于[Boulé等人,1998,Mol.Biotechnol.,10,199-208]。使用常规PCR扩增和分子生物学技术,使用以下引物扩增对应的序列SEQ ID NO:3(对应于前120个氨基酸被截去的SEQ ID NO:1):

[0101] ♦T7-pro:TAATACGACTCACTATAGGG (SEQ ID NO:7)

[0102] ♦T7-ter:GCTAGTTATTGCTCAGCGG (SEQ ID NO:8)。

[0103] 将其克隆到质粒pET32中,得到载体pET32-SEQ ID NO:3。

[0104] 先将质粒pET32-SEQ ID NO:3测序,然后转化到商业化大肠杆菌菌株BL21 (DE3) (Novagen) 中。将能够在卡那霉素/氯霉素平板上生长的菌落分离并标记为Ec-SEQ ID NO:3。

[0105] 变体的产生

[0106] 载体pET32-SEQ ID NO:3用作起始载体。由Agilent的在线工具(<http://www.genomics.agilent.com/primerDesignProgram.jsp>)产生包含点突变(或在某些情况下是更多个点突变,如果它们足够接近的话)的引物。

[0107] 使用QuickChange II试剂盒(Agilent)产生具有所期望突变的变体的质粒。严格遵循制造商提供的诱变方案以获得质粒pET32-DS*i*(*i*是表1中所示的所讨论变体的编号)。在该程序结束时,先将质粒pET32-DSx测序,然后转化到商业化大肠杆菌菌株BL21 (DE3) (Novagen) 中。将能够在卡那霉素/氯霉素平板上生长的菌落分离并标记为Ec-DSx。

[0108] 生产

[0109] 将Ec-SEQ ID NO:3和Ec-DSx细胞在含有50mL加入了适量的卡那霉素和氯霉素的LB培养基的250mL Erlenmeyer烧瓶中预培养。将培养物在37℃振荡孵育过夜。然后使用预培养物接种含有2L补充有适量卡那霉素和氯霉素的LB培养基的5L Erlenmeyer烧瓶中。初始光密度(OD)为0.01。在37℃伴随振荡地孵育培养物。定期测量OD直至其达到0.6与0.9之间的值。一旦达到该值,向培养基中加入1mL 1M异丙基β-D-1-硫代吡喃半乳糖苷。将培养物再次在37℃孵育直至第二天。然后通过不超过5 000rpm的离心收集细胞。在用裂解缓冲液(20mM Tris-HCl, pH 8.3, 0.5M NaCl)洗涤期间,将获得的多个沉淀物合并为单个沉淀物。将细胞沉淀物冷冻在-20℃。它可以照此储存几个月。

[0110] 提取

[0111] 将在先前步骤中冷冻的细胞沉淀物在25℃至37℃的水浴中解冻。完全解冻后,将其重悬于约100mL裂解缓冲液中。要特别注意重悬,其必须得到非常均一的溶液,且特别是完全不存在聚集物。如此重悬后,使用French压力机在14 000psi的压力下裂解细胞。将收集的裂解物以高速10000g离心1小时至1小时30分钟。将离心液过滤通过0.2μM过滤器,并收集在足够体积的管中。

[0112] 纯化

[0113] 在亲和柱上纯化TdT。将5mL His-Trap粗提柱(GE Life Sciences)与蠕动泵(MINIPULS® Evolution, Gilson)一起使用。首先, 使用2至3个CV(柱体积)的裂解缓冲液平衡柱。然后将来自先前步骤的离心液以约0.5mL/min与5mL/min之间的速率加载到柱上。在加载了全部离心液后, 用3CV裂解缓冲液和3CV洗涤缓冲液(20mM Tris-HCl, pH 8.3, 0.5M NaCl, 60mM咪唑)洗涤柱。在该步骤结束时, 将洗脱缓冲液(20mM Tris-HCl, pH 8.3, 0.5M NaCl, 1M咪唑)以约0.5ml/min至1ml/min注入到柱中, 总体积为3CV。在整个洗脱阶段, 以1mL级分收集柱输出物。通过SDS-PAGE分析这些级分以确定哪些级分含有洗脱峰。在确定后, 将这些级分合并为单个级分并用透析缓冲液(20mM Tris-HCl, pH 6.8, 200mM NaCl, 50mM MgOAc, 100mM [NH₄]₂SO₄)透析。然后将TdT浓缩(Amicon Ultra-30Centrifugal Filters, Merk Millipore)至终浓度为5mg/mL至15mg/mL。将浓缩的TdT在加入50%甘油后于-20℃冷冻以长期储存。在整个纯化阶段, 收集多种样品的等分试样(约5μL)用于SDS-PAGE凝胶分析, 其结果显示在图1中。

[0114] 实施例2-可能用于创建根据本发明的变体的Pol X家族的多种聚合酶之间的序列比对

[0115] 使用Mutalin在线比对软件(<http://multalin.toulouse.inra.fr/multalin/multalin.html>, accessed 04/04/2016)比对Pol X家族的多种DNA聚合酶。

[0116] 表2:比对的序列

[0117]

标识符	DNA聚合酶	物种	长度
Q9NP87	Polμ (SEQ ID NO:2)	智人	494
H2QUI0	Polμ (SEQ ID NO:9)	黑猩猩	494
Q924W4	Polμ (SEQ ID NO:10)	小鼠	496
F1P657	TdT (SEQ ID NO:11)	狗	509
Q3UZ80	TdT (SEQ ID NO:1)	小鼠	510
P36195	TdT (SEQ ID NO:12)	原鸡	506
P04053	TdT (SEQ ID NO:13)	智人	509

[0118] 所得的比对示于图2中。

[0119] 实施例3-在非天然底物的存在下对变体活性的研究

[0120] 通过以下测试确定根据本发明的不同变体的活性。将结果与通过每种变体所源自的天然酶获得的那些结果比较。

[0121] 活性测试

[0122] 表3:反应混合物

[0123]

试剂	浓度	体积
H ₂ O	-	15μL
引物	500nM	2.5μL
缓冲液	10x	2.5μL
修饰的核苷酸	250μM	2.5μL
酶	20μM	2.5μL

[0124] 预先将所用的序列为5'-AAAAAAAAGGGG-3' (SEQ ID NO:14) 的引物使用酶PNK (NEB) 和放射性ATP (PerkinElmer) 通过标准标记方案进行5'-放射性标记。

[0125] 使用由250mM Tris-HCl pH7.2、80mM MgCl₂、3.3mM ZnSO₄组成的10×缓冲液。

[0126] 所用的修饰核苷酸是3'-0-氨基-2', 3'-双脱氧核苷酸-5'-三磷酸 (ONH₂, Firebird Biosciences) 或3'-biot-EDA-2', 3'-双脱氧核苷酸-5'-三磷酸 (Biot-EDA, Jena Biosciences), 诸如3'-0-氨基-2', 3'-双脱氧腺苷-5'-三磷酸或例如3'-biot-EDA-2', 3'-双脱氧腺苷-5'-三磷酸。3'-0-氨基基团是与3'-OH末端附接的体积更大的基团。3'-biot-EDA基团是与3'-OH末端附接的极其庞大且无柔性 (inflexible) 的基团。

[0127] 表1中列出的变体提供的相对于天然TdT (SEQ ID NO:3) 掺入修饰的核苷酸的性能通过同时进行的其中仅酶不同的活性测试来评价。

[0128] 按上表3中提供的顺序加入试剂, 然后在37°C 孵育90min。然后通过加入甲酰胺蓝 (100% 甲酰胺, 1mg至5mg溴酚蓝; Sigma) 终止反应。

[0129] 凝胶和射线照相术

[0130] 使用16% 变性聚丙烯酰胺凝胶 (Biorad) 分析先前的活性测试。预先倾倒凝胶并静置至聚合。然后将其放置在填充有TBE缓冲液 (Sigma) 的适当大小的电泳槽上。将无需预处理的多种样品直接上样到凝胶上。

[0131] 然后使凝胶经历500V至2 000V的电位差持续3至6小时。在迁移令人满意后, 释放凝胶, 然后转移到孵育盒中。使用磷光屏 (phosphor screen) (Amersham) 10min至60min以便可视化, 可视化使用预先配置有适当检测模式的Typhoon仪器 (GE Life Sciences) 进行。

[0132] 结果

[0133] 所使用的两种酶的比较结果显示在图3中。

[0134] 更准确地说, 在第一个凝胶 (ONH₂掺入) 上, 天然TdT (wt列) 不能掺入3'-0-氨基-2', 3'-双脱氧腺苷-5'-三磷酸修饰的核苷酸, 如通过与阴性对照 (No列) 比较所示。

[0135] 在不同的变体之间, 可以观察到3个不同的组:

[0136] 第一组变体 (DS7至DS34列) 能够掺入约50%。

[0137] 第二组变体 (DS46至DS73列) 能够掺入大于95%, 有时大于98%。

[0138] 第三组变体 (DS83至DS106列) 能够掺入60%至80%。

[0139] 在第二个凝胶 (Biot-EDA掺入) 上, 天然TdT (wt列) 也不能掺入3'-biot-EDA-2', 3'-双脱氧腺苷-5'-三磷酸修饰的核苷酸, 如通过与阴性对照 (No列) 比较所示。

[0140] 在不同的变体之间, 可以观察到3个不同的组:

[0141] 第一组变体 (DS7至DS34列) 能够掺入约5%至10%。

[0142] 第二组变体 (DS46至DS73列) 能够掺入大于30%, 有时大于40%。

[0143] 第三组变体 (DS83至DS106列) 能够掺入10%至25%。

[0144] 这些结果确认, 与野生型酶不同, 根据本发明的TdT的变体都能够使用修饰的核苷酸, 特别是3'-OH-修饰的核苷酸作为底物。特别有利的是, 某些变体具有非常高的掺入率并且即使在携带倾向于大大增加所述核苷酸的空间体积的修饰的核苷酸的存在下也是如此。

[0145] 实施例4-根据本发明的变体的动力学的研究

[0146] 按照前述实施例1产生并生产具有取代组合R336N-R454A-E457N的突变体 (DS124)。

[0147] 活性测试

[0148] 在活性测试中,将酶置于ONH2修饰的核苷酸的存在下,并在37°C孵育不同的时间长度。终止反应以观察DS124的掺入动力学并将其与天然WT酶(SEQ ID NO:3)的动力学比较。

[0149] 表4:反应混合物

[0150]

试剂	浓度	体积
H ₂ O	-	15μL
缓冲液	10x	2.5μL
核苷酸	2.5μM	2.5μL
酶	80μM	2.5μL
引物	1μM	2.5μL

[0151] 使用的引物和缓冲液与实施例3相同。

[0152] 使用的修饰的核苷酸是3'-0-氨基-2',3'-双脱氧核苷酸-5'-三磷酸(ONH₂, Firebird Biosciences):3'-0-氨基-2',3'-双脱氧鸟苷-5'-三磷酸、3'-0-氨基-2',3'-双脱氧胞苷-5'-三磷酸和3'-0-氨基-2',3'-双脱氧胸苷-5'-三磷酸。3'-0-氨基基团是与3'-OH末端附接的体积更大的基团。

[0153] 通过进行活性测试评价酶DS124对核苷酸混合物的掺入性能,制备含有除了引物外的所有试剂(按表4所示顺序加入)的预混合物用于这些活性测试。将它们分配在不同的反应孔中。在初始时间t=0时,将引物同时加入到所有孔中。在不同时间t=2min、t=5min、t=10min、t=15min、t=30min和t=90min时,通过加入甲酰胺蓝(100%甲酰胺、1mg至5mg溴酚蓝;Sigma)终止反应。

[0154] 凝胶和射线照相术

[0155] 根据实施例3中描述的方案,通过多种样品在聚丙烯酰胺凝胶中的迁移来进行活性测试的分析。

[0156] 结果

[0157] 两种酶(DS124和WT)的比较结果显示在图4中。

[0158] 更确切地说,在该凝胶上,阴性对照(No列)提供了所使用的引物尚未被延伸时即当核苷酸尚未被掺入时该引物的预期大小。天然TdT(WT列)不能掺入修饰的核苷酸(此处为ONH₂-dGTP):在与列No相同的高度(level)上观察到条带。

[0159] 对于所有测试的核苷酸以及对于范围从90min(此处用作阳性对照)到2min的所有时间,即孵育时间减少45倍,变体DS124能够以100%的表观效率掺入修饰的核苷酸。

[0160] 这些结果确认,在掺入效率和掺入速度方面,根据本发明的TdT的变体能够具有比天然TdT高得多的掺入性能。通过本发明描述的特定突变或突变的组合极大地增强了根据本发明的TdT的变体的动力学。

[0161] 实施例5-根据本发明的变体的特异性的研究

[0162] 根据实施例1产生并生产具有根据下表5的取代组合的突变体。

[0163] 表5:使用的酶变体的列表

[0164]

#	突变的组合
DS124	R336N-R454A-E457N
DS24	R336N-E457N
DS125	R336N-R454A-E457G
DS126	R336N-E457G
DS127	R336G-R454A-E457N
DS22	R336G-E457N
DS128	R336A-R454A-E457G
WT	SEQ ID NO:3

[0165] 活性测试

[0166] 在该活性测试中,将不同的变体与天然核苷酸和修饰的核苷酸的高度浓缩的混合物汇集到一起。还增加酶浓度以缩短孵育时间并实现定量添加(参见实施例4)。

[0167] 通过以下测试确定产生的不同变体的活性:

[0168] 在两种条件下测试每种变体:(1)在无核苷酸存在下(用H₂O代替)或(2)在核苷酸混合物存在下。将不同变体的结果彼此比较。加入对照样品,其既不含核苷酸也不含酶(被H₂O代替)。

[0169] 表6:反应混合物

[0170]

试剂	浓度	体积
H ₂ O	-	15μL
引物	1μM	2.5μL
缓冲液	10x	2.5μL
核苷酸混合物(10:90)	2.5μM	2.5μL
酶	80μM	2.5μL

[0171] 使用的引物和缓冲液与实施例3相同。

[0172] 核苷酸混合物当存在时由以下组成:天然2'-脱氧核苷酸5'-三磷酸核苷酸(Nuc, Sigma-Aldrich)诸如2'-脱氧鸟苷5'-三磷酸(dGTP)和3'-0-氨基-2',3'-双脱氧核苷酸-5'-三磷酸修饰的核苷酸(ONH2,Firebird Biosciences)诸如例如3'-0-氨基-2',3'-双脱氧鸟苷-5'-三磷酸。3'-0-氨基基团是与3'-OH末端交接的体积更大的基团。混合物由90%修饰的核苷酸ONH2-dGTP和10%天然dGTP核苷酸组成。

[0173] 通过同时进行的活性测试(对于这些活性测试,仅酶不同)来评价表5中列出的变体相对于彼此对核苷酸混合物的掺入性能。

[0174] 按以上表6中提供的顺序加入试剂,然后在37℃孵育15min。然后通过加入甲酰胺蓝(100%甲酰胺,1mg至5mg溴酚蓝;Sigma)终止反应。

[0175] 凝胶和射线照相术

[0176] 根据实施例3中描述的方案,通过多种样品在聚丙烯酰胺凝胶中的迁移来进行活性测试的分析。

[0177] 结果

[0178] 所使用的酶的比较结果显示在图5中。

[0179] 更确切地说，在该凝胶上，阴性对照(No列)提供了所使用的引物尚未被延伸时即当核苷酸尚未被掺入时该引物的预期大小。以下样品成对出现，每对对应于在两种条件下测试的同一酶变体：在核苷酸不存在和存在下(当其存在时呈混合物的形式)。

[0180] 在测试的不同的变体之间，可以观察到3个不同的组：

[0181] 第一组是变体DS128，构成阴性对照。该变体具有极低的核苷酸掺入率：当存在核苷酸混合物时观察到5%与10%之间的掺入，该掺入率对应混合物中存在的天然核苷酸的比例。

[0182] 第二组由变体DS127和DS22组成。这些变体具有高核苷酸掺入率：当存在核苷酸混合物时，观察到50%与60%之间的掺入。仍然在这种情况下，对于这两种变体，观察到对应于连续掺入两个核苷酸的过度添加条带(over-addition band)。该条带的强度对应核苷酸混合物中存在的天然核苷酸的比例。

[0183] 最后一组由变体DS124、DS24、DS125和DS126组成。这些变体具有极高的核苷酸掺入率：当存在核苷酸混合物时，对于DS124和DS125，观察到80%与100%之间的掺入。在这种情况下，不存在过度添加条带。在变体DS24和DS126的情况下，无掺入的比例类似于混合物中存在的天然核苷酸的比例。

[0184] 这些结果确认，根据本发明的TdT的变体能够优先使用修饰的核苷酸和天然核苷酸的混合物中的修饰的核苷酸。特别有利的是，这些变体具有极高的修饰的核苷酸的掺入率，并且能够区分天然核苷酸以不掺入它们，并且从而通过避免过度添加而大大增强DNA合成的质量。

[0185] 实施例6-在没有模板链的情况下合成DNA链的实例

[0186] 根据实施例1产生并生产具有取代组合R336N-R454A-E457G的TdT的变体(DS125)。

[0187] 使用变体DS125在序列为5'-AAAAAAAAGGGG-3' (SEQ ID NO:14) 的引物之后合成序列：5'-GTACGCTAGT-3' (SEQ ID NO:15)。将引物预先使用酶PNK (NEB) 和放射性ATP (PerkinElmer) 通过标准标记方案进行5'-放射性标记。

[0188] 通过与序列5'-CCTTTTTTTT-3' (SEQ ID NO:16) 的互补捕获片段相互作用将引物附接至固体支持物。捕获片段在其3'末端上具有使其能够与附接于表面的反应基团共价反应的基团。例如，该基团可以是NH₂，反应基团可以是N-羟基琥珀酰亚胺，且表面可以是磁珠(Dynabeads, Thermofisher)。引物与捕获片段的相互作用在标准DNA片段杂交条件下进行。

[0189] 所用的修饰核苷酸是3'-0-氨基-2', 3'-双脱氧核苷酸-5'-三磷酸(ONH₂, Firebird Biosciences)，诸如3'-0-氨基-2', 3'-双脱氧鸟苷-5'-三磷酸、3'-0-氨基-2', 3'-双脱氧胞苷-5'-三磷酸、3'-0-氨基-2', 3'-双脱氧胸苷-5'-三磷酸或3'-0-氨基-2', 3'-双脱氧腺苷-5'-三磷酸。3'-0-氨基基团是与3'-OH末端附接的体积更大的基团。

[0190] 合成

[0191] 表7:反应混合物

[0192]

试剂	浓度	体积
H ₂ O	-	210μL

缓冲液	10x	70μL
核苷酸	2.5μM	35μL
酶	80μM	35μL
固体支持物上的引物	1μM	-

[0193] 使用由250mM Tris-HCl pH 7.2、80mM MgCl₂、3.3mM ZnSO₄组成的10×缓冲液。

[0194] 使用的洗涤缓冲液W由pH 7.2的25mM Tris-HCl组成。

[0195] 使用的脱保护缓冲液D由10mM MgCl₂的存在下的pH 5.5的50mM乙酸钠组成。

[0196] 在合成开始之前,将构成固体支持物的对于35pmol的等同引物总量的珠用缓冲液W洗涤几次,引物杂交于所述珠上。在这些洗涤后,将珠保持在磁体上并将上清液全部去除。

[0197] 制备由以表7中所示顺序加入的不同试剂组成的几种预混合物。这些预混合物中的每一种含有不同的核苷酸,如下表8所示。

[0198] 表8:预混合物组成

[0199]

预混合物编号	预混合物核苷酸
1	G
2	T
3	A
4	C
5	G
6	C
7	T
8	A
9	G
10	T

[0200] 当将预混合物1加入到预先被洗涤并清除上清液的珠时,合成开始。以下表9中所示的合成步骤以所示顺序进行以产生新序列5'-GTACGCTAGT-3'。

[0201] 表9:在没有模板链的情况下合成DNA链的方法的步骤

[0202]

步骤	行动	体积	持续时间
延伸1	加入预混合物1	350 μL	15min
取样1	取样	5 μL	<1min
第1次洗涤1	加入缓冲液W	350 μL	5min
第1次脱保护1	加入缓冲液D	350 μL	15min
第2次脱保护1	加入缓冲液D	350 μL	15min
第2次洗涤1	加入缓冲液W	350 μL	5min
延伸2	加入预混合物2	350 μL	15min
取样2	取样	5 μL	<1min
第1次洗涤2	加入缓冲液W	350 μL	5min
第1次脱保护2	加入缓冲液D	350 μL	15min
第2次脱保护2	加入缓冲液D	350 μL	15min
第2次洗涤2	加入缓冲液W	350 μL	5min
延伸3	加入预混合物3	350 μL	15min
取样3	取样	5 μL	<1min
第1次洗涤3	加入缓冲液W	350 μL	5min
第1次脱保护3	加入缓冲液D	350 μL	15min
第2次脱保护3	加入缓冲液D	350 μL	15min
第2次洗涤3	加入缓冲液W	350 μL	5min
延伸4	加入预混合物4	350 μL	15min
取样4	取样	5 μL	<1min
第1次洗涤4	加入缓冲液W	350 μL	5min
第1次脱保护4	加入缓冲液D	350 μL	15min
第2次脱保护4	加入缓冲液D	350 μL	15min
第2次洗涤4	加入缓冲液W	350 μL	5min
延伸5	加入预混合物5	350 μL	15min
取样5	取样	5 μL	<1min
第1次洗涤5	加入缓冲液W	350 μL	5min
第1次脱保护5	加入缓冲液D	350 μL	15min
第2次脱保护5	加入缓冲液D	350 μL	15min
第2次洗涤5	加入缓冲液W	350 μL	5min
延伸6	加入预混合物6	350 μL	15min
取样6	取样	5 μL	<1min
第1次洗涤6	加入缓冲液W	350 μL	5min
第1次脱保护6	加入缓冲液D	350 μL	15min
第2次脱保护6	加入缓冲液D	350 μL	15min
第2次洗涤6	加入缓冲液W	350 μL	5min
延伸7	加入预混合物7	350 μL	15min
取样7	取样	5 μL	<1min
第1次洗涤7	加入缓冲液W	350 μL	5min
第1次脱保护7	加入缓冲液D	350 μL	15min
第2次脱保护7	加入缓冲液D	350 μL	15min
第2次洗涤7	加入缓冲液W	350 μL	5min
延伸8	加入预混合物8	350 μL	15min
取样8	取样	5 μL	<1min

[0203]

第1次洗涤8	加入缓冲液W	350 μL	5min
第1次脱保护8	加入缓冲液D	350 μL	15min
第2次脱保护8	加入缓冲液D	350 μL	15min
第2次洗涤8	加入缓冲液W	350 μL	5min
延伸9	加入预混合物9	350 μL	15min
取样9	取样	5 μL	<1min
第1次洗涤9	加入缓冲液W	350 μL	5min
第1次脱保护9	加入缓冲液D	350 μL	15min
第2次脱保护9	加入缓冲液D	350 μL	15min
第2次洗涤9	加入缓冲液W	350 μL	5min
延伸10	加入预混合物10	350 μL	15min
取样10	取样	5 μL	<1min

[0204] 除了取样步骤外,在每个步骤之间,通过磁体收集珠并且并将上清液全部去除。

[0205] 将每种样品加入到15μL甲酰胺蓝(100%甲酰胺、1mg至5mg溴酚蓝;Sigma)溶液中以终止反应并准备分析。

[0206] 凝胶和射线照相术

[0207] 根据实施例3中描述的方案,通过多种样品在聚丙烯酰胺凝胶中的迁移来进行活性测试的分析。

[0208] 结果

[0209] 该合成的结果显示在图6中。

[0210] 第0列(No,无核苷酸)提供了所使用的引物尚未被延伸时即当核苷酸尚未被掺入时该引物的预期大小。

[0211] 第1至10列对应于合成期间的样品1至10。核苷酸的每次掺入通过具有最大性能的酶进行。不进行另外的纯化步骤。

[0212] 用天然TdT进行类似的合成实验。因为后者不能掺入修饰的核苷酸,因此它不可能合成所期望的序列。

序列表

<110> DNA斯克瑞普特公司

巴斯德研究所

<120> polX家族的DNA聚合酶的变体

<130> PR1898

<160> 16

<170> PatentIn version 3.5

<210> 1

<211> 510

<212> PRT

<213> 人工序列 (artificial sequence)

<220>

<223> 小鼠TdT

<400> 1

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

Arg Ala Ser Ser Val Leu Lys Ser Leu Pro Phe Pro Ile Thr Ser Met
 195 200 205
 Lys Asp Thr Glu Gly Ile Pro Cys Leu Gly Asp Lys Val Lys Ser Ile
 210 215 220
 Ile Glu Gly Ile Ile Glu Asp Gly Glu Ser Ser Glu Ala Lys Ala Val
 225 230 235 240
 Leu Asn Asp Glu Arg Tyr Lys Ser Phe Lys Leu Phe Thr Ser Val Phe
 245 250 255
 Gly Val Gly Leu Lys Thr Ala Glu Lys Trp Phe Arg Met Gly Phe Arg
 260 265 270
 Thr Leu Ser Lys Ile Gln Ser Asp Lys Ser Leu Arg Phe Thr Gln Met
 275 280 285
 Gln Lys Ala Gly Phe Leu Tyr Tyr Glu Asp Leu Val Ser Cys Val Asn
 290 295 300
 Arg Pro Glu Ala Glu Ala Val Ser Met Leu Val Lys Glu Ala Val Val
 305 310 315 320
 Thr Phe Leu Pro Asp Ala Leu Val Thr Met Thr Gly Gly Phe Arg Arg
 325 330 335
 Gly Lys Met Thr Gly His Asp Val Asp Phe Leu Ile Thr Ser Pro Glu
 340 345 350
 Ala Thr Glu Asp Glu Glu Gln Gln Leu Leu His Lys Val Thr Asp Phe
 355 360 365
 Trp Lys Gln Gln Gly Leu Leu Tyr Cys Asp Ile Leu Glu Ser Thr
 370 375 380
 Phe Glu Lys Phe Lys Gln Pro Ser Arg Lys Val Asp Ala Leu Asp His
 385 390 395 400
 Phe Gln Lys Cys Phe Leu Ile Leu Lys Leu Asp His Gly Arg Val His
 405 410 415
 Ser Glu Lys Ser Gly Gln Gln Glu Gly Lys Trp Lys Ala Ile Arg
 420 425 430
 Val Asp Leu Val Met Cys Pro Tyr Asp 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 Arg
 465 470 475 480
 Thr Lys Arg Val Phe Leu Glu 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> 2		
<211> 494		
<212> PRT		
<213> 人工序列 (artificial sequence)		
<220>		
<223> 人类Polμ		
<400> 2		
Met Leu Pro Lys Arg Arg Arg Ala Arg Val Gly Ser Pro Ser Gly Asp		
1	5	10
Ala Ala Ser Ser Thr Pro Pro Ser Thr Arg Phe Pro Gly Val Ala Ile		
	20	25
		30
Tyr Leu Val Glu Pro Arg Met Gly Arg Ser Arg Arg Ala Phe Leu Thr		
	35	40
		45
Gly Leu Ala Arg Ser Lys Gly Phe Arg Val Leu Asp Ala Cys Ser Ser		
	50	55
		60
Glu Ala Thr His Val Val Met Glu Glu Thr Ser Ala Glu Glu Ala Val		
	65	70
		75
		80
Ser Trp Gln Glu Arg Arg Met Ala Ala Ala Pro Pro Gly Cys Thr Pro		
	85	90
		95
Pro Ala Leu Leu Asp Ile Ser Trp Leu Thr Glu Ser Leu Gly Ala Gly		
	100	105
		110
Gln Pro Val Pro Val Glu Cys Arg His Arg Leu Glu Val Ala Gly Pro		
	115	120
		125
Arg Lys Gly Pro Leu Ser Pro Ala Trp Met Pro Ala Tyr Ala Cys Gln		
	130	135
		140
Arg Pro Thr Pro Leu Thr His His Asn Thr Gly Leu Ser Glu Ala Leu		
	145	150
		155
		160
Glu Ile Leu Ala Glu Ala Ala Gly Phe Glu Gly Ser Glu Gly Arg Leu		
	165	170
		175
Leu Thr Phe Cys Arg Ala Ala Ser Val Leu Lys Ala Leu Pro Ser Pro		
	180	185
		190
Val Thr Thr Leu Ser Gln Leu Gln Gly Leu Pro His Phe Gly Glu His		
	195	200
		205
Ser Ser Arg Val Val Gln Glu Leu Leu Glu His Gly Val Cys Glu Glu		
	210	215
		220
Val Glu Arg Val Arg Arg Ser Glu Arg Tyr Gln Thr Met Lys Leu Phe		
	225	230
		235
		240
Thr Gln Ile Phe Gly Val Gly Val Lys Thr Ala Asp Arg Trp Tyr Arg		

245	250	255
Glu Gly Leu Arg Thr Leu Asp Asp Leu Arg Glu Gln Pro Gln Lys Leu		
260	265	270
Thr Gln Gln Gln Lys Ala Gly Leu Gln His His Gln Asp Leu Ser Thr		
275	280	285
Pro Val Leu Arg Ser Asp Val Asp Ala Leu Gln Gln Val Val Glu Glu		
290	295	300
Ala Val Gly Gln Ala Leu Pro Gly Ala Thr Val Thr Leu Thr Gly Gly		
305	310	315
Phe Arg Arg Gly Lys Leu Gln Gly His Asp Val Asp Phe Leu Ile Thr		
325	330	335
His Pro Lys Glu Gly Gln Glu Ala Gly Leu Leu Pro Arg Val Met Cys		
340	345	350
Arg Leu Gln Asp Gln Gly Leu Ile Leu Tyr His Gln His Gln His Ser		
355	360	365
Cys Cys Glu Ser Pro Thr Arg Leu Ala Gln Gln Ser His Met Asp Ala		
370	375	380
Phe Glu Arg Ser Phe Cys Ile Phe Arg Leu Pro Gln Pro Pro Gly Ala		
385	390	395
Ala Val Gly Gly Ser Thr Arg Pro Cys Pro Ser Trp Lys Ala Val Arg		
405	410	415
Val Asp Leu Val Val Ala Pro Val Ser Gln Phe Pro Phe Ala Leu Leu		
420	425	430
Gly Trp Thr Gly Ser Lys Leu Phe Gln Arg Glu Leu Arg Arg Phe Ser		
435	440	445
Arg Lys Glu Lys Gly Leu Trp Leu Asn Ser His Gly Leu Phe Asp Pro		
450	455	460
Glu Gln Lys Thr Phe Phe Gln Ala Ala Ser Glu Glu Asp Ile Phe Arg		
465	470	475
His Leu Gly Leu Glu Tyr Leu Pro Pro Glu Gln Arg Asn Ala		
485	490	
<210> 3		
<211> 401		
<212> PRT		
<213> 人工序列 (artificial sequence)		
<220>		
<223> 截短的小鼠TdT		
<400> 3		
Thr Met Gly Ser Ser His His His His His Ser Ser Gly Leu Val		

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

Ala Ile Arg Val Asp Leu Val Met Cys Pro Tyr Asp Arg Arg Ala Phe
 325 330 335
 Ala Leu Leu Gly Trp Thr Gly Ser Arg Gln Phe Glu Arg Asp Leu Arg
 340 345 350
 Arg Tyr Ala Thr His Glu Arg Lys Met Met Leu Asp Asn His Ala Leu
 355 360 365
 Tyr Asp Arg Thr Lys Arg Val Phe Leu Glu Ala Glu Ser Glu Glu Glu
 370 375 380
 Ile Phe Ala His Leu Gly Leu Asp Tyr Ile Glu Pro Trp Glu Arg Asn
 385 390 395 400
 Ala
 <210> 4
 <211> 11
 <212> PRT
 <213> 人工序列 (artificial sequence)
 <220>
 <223> 半保守序列
 <220>
 <221> MISC_FEATURE
 <222> (1) .. (1)
 <223> X=M, I, V, L
 <220>
 <221> MISC_FEATURE
 <222> (2) .. (2)
 <223> X=T, A, M, Q
 <220>
 <221> MISC_FEATURE
 <222> (10) .. (10)
 <223> X=M, K, E, Q, L, S, P, R, D
 <220>
 <221> MISC_FEATURE
 <222> (11) .. (11)
 <223> X=T, I, M, F, K, V, Y, E, Q, H, S, R, D
 <400> 4
 Xaa Xaa Gly Gly Phe Arg Arg Gly Lys Xaa Xaa
 1 5 10
 <210> 5
 <211> 13
 <212> PRT

<213> 人工序列 (artificial sequence)
<220>
<223> 半保守区域
<220>
<221> MISC_FEATURE
<222> (1) .. (1)
<223> X=A, C, G, S
<220>
<221> MISC_FEATURE
<222> (2) .. (2)
<223> X=L, T, R
<220>
<221> MISC_FEATURE
<222> (5) .. (5)
<223> X=W, Y
<220>
<221> MISC_FEATURE
<222> (6) .. (6)
<223> X=T, S, I
<220>
<221> MISC_FEATURE
<222> (10) .. (10)
<223> X=Q, L, H, F, Y, N, E, D,
<220>
<221> MISC_FEATURE
<222> (11) .. (11)
<223> X=F, Y
<400> 5
Xaa Xaa Leu Gly Xaa Xaa Gly Ser Arg Xaa Xaa Glu Arg
1 5 10
<210> 6
<211> 11
<212> PRT
<213> 人工序列 (artificial sequence)
<220>
<223> 半保守区域
<220>
<221> MISC_FEATURE
<222> (2) .. (2)

<223> X=D, E, S, P, A, K
<220>
<221> MISC_FEATURE
<222> (4) .. (4)
<223> X=I, L, M, V, A, T
<220>
<221> MISC_FEATURE
<222> (5) .. (5)
<223> X=E, Q, P, Y, L, K, G, N
<220>
<221> MISC_FEATURE
<222> (7) .. (7)
<223> X=W, S, V, E, R, Q, T, C, K, H
<220>
<221> MISC_FEATURE
<222> (8) .. (8)
<223> X=E, Q, D, H, L
<400> 6
Leu Xaa Tyr Xaa Xaa Pro Xaa Xaa Arg Asn Ala
1 5 10
<210> 7
<211> 20
<212> DNA
<213> 人工序列 (artificial sequence)
<220>
<223> 引物 (amorce)
<400> 7
taatacgact cactatagg 20
<210> 8
<211> 19
<212> DNA
<213> 人工序列 (artificial sequence)
<220>
<223> 引物 (amorce)
<400> 8
gctagttatt gtcagcgg 19
<210> 9
<211> 494
<212> PRT

<213> 黑猩猩(Pan troglodytes)

<400> 9

Met	Leu	Pro	Lys	Arg	Arg	Arg	Ala	Arg	Val	Gly	Ser	Pro	Ser	Gly	Asp
1				5					10					15	
Ala	Ala	Ser	Ser	Thr	Pro	Pro	Ser	Thr	Arg	Phe	Pro	Gly	Val	Ala	Ile
				20				25					30		
Tyr	Leu	Val	Glu	Pro	Arg	Met	Gly	Arg	Ser	Arg	Arg	Ala	Phe	Leu	Thr
					35			40			45				
Arg	Leu	Thr	Arg	Ser	Lys	Gly	Phe	Arg	Val	Leu	Asp	Ala	Cys	Ser	Ser
					50			55			60				
Glu	Ala	Thr	His	Val	Val	Met	Glu	Glu	Thr	Ser	Ala	Glu	Glu	Ala	Val
					65			70		75			80		
Ser	Trp	Gln	Glu	Arg	Arg	Met	Ala	Ala	Ala	Pro	Pro	Gly	Cys	Thr	Pro
					85				90				95		
Pro	Ala	Leu	Leu	Asp	Ile	Ser	Trp	Leu	Thr	Glu	Ser	Leu	Gly	Ala	Gly
					100			105			110				
Gln	Pro	Val	Pro	Val	Glu	Cys	Arg	His	Arg	Leu	Glu	Val	Ala	Gly	Pro
					115			120			125				
Arg	Lys	Gly	Pro	Leu	Ser	Pro	Ala	Trp	Met	Pro	Ala	Tyr	Val	Cys	Gln
					130			135			140				
Arg	Pro	Thr	Pro	Leu	Thr	His	His	Asn	Thr	Gly	Leu	Ser	Glu	Ala	Leu
					145			150		155			160		
Glu	Thr	Leu	Ala	Glu	Ala	Ala	Gly	Phe	Glu	Gly	Ser	Glu	Gly	Arg	Leu
					165				170			175			
Leu	Thr	Phe	Cys	Arg	Ala	Ala	Ser	Val	Leu	Lys	Ala	Leu	Pro	Ser	Pro
					180				185			190			
Val	Thr	Thr	Leu	Ser	Gln	Leu	Gln	Gly	Leu	Pro	His	Phe	Gly	Glu	His
					195			200			205				
Ser	Ser	Arg	Val	Val	Gln	Glu	Leu	Leu	Glu	His	Gly	Val	Cys	Glu	Glu
					210			215			220				
Val	Glu	Arg	Val	Gln	Arg	Ser	Glu	Arg	Tyr	Gln	Thr	Met	Lys	Leu	Phe
					225			230		235			240		
Thr	Gln	Ile	Phe	Gly	Val	Gly	Val	Lys	Thr	Ala	Asp	Arg	Trp	Tyr	Arg
					245				250			255			
Glu	Gly	Leu	Arg	Thr	Leu	Asp	Asp	Leu	Arg	Glu	Gln	Pro	Gln	Lys	Leu
					260			265			270				
Thr	Gln	Gln	Gln	Lys	Ala	Gly	Leu	Gln	His	His	Gln	Asp	Leu	Ser	Thr
					275			280			285				
Pro	Val	Leu	Arg	Ser	Asp	Val	Asp	Ala	Leu	Gln	Gln	Val	Val	Glu	Glu

290	295	300
Ala Val Gly Gln Ala Leu Pro Gly Ala Thr Val Thr Leu Thr Gly Gly		
305	310	315
Phe Arg Arg Gly Lys Leu Gln Gly His Asp Val Asp Phe Leu Ile Thr		
325	330	335
His Pro Lys Glu Gly Gln Glu Ala Gly Leu Leu Pro Arg Val Met Cys		
340	345	350
Arg Leu Gln Asp Gln Gly Leu Ile Leu Tyr His Gln His Ser		
355	360	365
Cys Trp Glu Ser Pro Thr Arg Leu Ala Gln Gln Ser His Met Asp Ala		
370	375	380
Phe Glu Arg Ser Phe Cys Ile Phe Arg Leu Pro Gln Pro Pro Gly Ala		
385	390	395
Ala Val Gly Gly Ser Thr Arg Pro Cys Pro Ser Trp Lys Ala Val Arg		
405	410	415
Val Asp Leu Val Val Ala Pro Val Ser Gln Phe Pro Phe Ala Leu Leu		
420	425	430
Gly Trp Thr Gly Ser Lys Leu Phe Gln Arg Glu Leu Arg Arg Phe Ser		
435	440	445
Arg Lys Glu Lys Gly Leu Trp Leu Asn Ser His Gly Leu Phe Asp Pro		
450	455	460
Glu Gln Lys Thr Phe Phe Gln Ala Ala Ser Glu Glu Asp Ile Phe Arg		
465	470	475
His Leu Gly Leu Glu Tyr Leu Pro Pro Glu Gln Arg Asn Ala		
485	490	
<210> 10		
<211> 496		
<212> PRT		
<213> 小鼠 (Mus musculus)		
<400> 10		
Met Leu Pro Lys Arg Arg Val Arg Ala Gly Ser Pro His Ser Ala		
1	5	10
Val Ala Ser Ser Thr Pro Pro Ser Val Val Arg Phe Pro Asp Val Ala		
20	25	30
Ile Tyr Leu Ala Glu Pro Arg Met Gly Arg Ser Arg Arg Ala Phe Leu		
35	40	45
Thr Arg Leu Ala Arg Ser Lys Gly Phe Arg Val Leu Asp Ala Tyr Ser		
50	55	60
Ser Lys Val Thr His Val Val Met Glu Gly Thr Ser Ala Lys Glu Ala		

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

Asp Ala Phe Glu Arg Ser Phe Cys Ile Leu Gly Leu Pro Gln Pro Gln
 385 390 395 400
 Gln Ala Ala Leu Ala Gly Ala Leu Pro Pro Cys Pro Thr Trp Lys Ala
 405 410 415
 Val Arg Val Asp Leu Val Val Thr Pro Ser Ser Gln Phe Pro Phe Ala
 420 425 430
 Leu Leu Gly Trp Thr Gly Ser Gln Phe Phe Glu Arg Glu Leu Arg Arg
 435 440 445
 Phe Ser Arg Gln Glu Lys Gly Leu Trp Leu Asn Ser His Gly Leu Phe
 450 455 460
 Asp Pro Glu Gln Lys Arg Val Phe His Ala Thr Ser Glu Glu Asp Val
 465 470 475 480
 Phe Arg Leu Leu Gly Leu Lys Tyr Leu Pro Pro Glu Gln Arg Asn Ala
 485 490 495
 <210> 11
 <211> 509
 <212> PRT
 <213> 狗 (Canis lupus)
 <400> 11
 Met Asp Pro Leu Gln Met Ala His Ser Gly Pro Arg Lys Lys Arg Pro
 1 5 10 15
 Arg Gln Met Gly Ala Pro Met Val Ser Pro Pro His Asn Ile Lys Phe
 20 25 30
 Gln Asp Leu Val Leu Tyr 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 Asp
 50 55 60
 Asn Glu Phe Ser Asp Ser Ile Thr His Ile Val Ala Glu Asn Asn Ser
 65 70 75 80
 Gly Ser Asp Val Leu Glu Trp Leu Gln Val Gln Asn Ile Lys Ala Ser
 85 90 95
 Ser Gln Leu Glu Leu Leu Asp Ile Ser Trp Leu Ile Glu Ser Met Gly
 100 105 110
 Ala Gly Lys Pro Val Glu Met Thr Gly Lys His Gln Leu Met Arg Arg
 115 120 125
 Asp Tyr Thr Ala Ser Pro Asn Pro Glu Leu Gln Lys Thr Leu Pro Val
 130 135 140
 Ala Val Lys Lys Ile Ser Gln Tyr Ala Cys Gln Arg Arg Thr Thr Leu
 145 150 155 160

Asn Asn Tyr Asn Asn Val Phe Thr Asp Ala Phe Glu Val Leu Ala Glu
 165 170 175
 Asn Tyr Glu Phe Arg Glu Asn Glu Val Phe Ser Leu Thr Phe Met Arg
 180 185 190
 Ala Ala Ser Val Leu Lys Ser Leu Pro Phe Thr Ile Ile Ser Met Lys
 195 200 205
 Asp Thr Glu Gly Ile Pro Cys Leu Gly Asp Gln Val Lys Cys Ile Ile
 210 215 220
 Glu Glu Ile Ile Glu Asp Gly Glu Ser Ser Glu Val Lys Ala Val Leu
 225 230 235 240
 Asn Asp Glu Arg Tyr Gln Ser Phe Lys Leu Phe Thr Ser Val Phe Gly
 245 250 255
 Val Gly Leu Lys Thr Ser Glu Lys Trp Phe Arg Met Gly Phe Arg Thr
 260 265 270
 Leu Ser Lys Ile Lys Ser Asp Lys Ser Leu Lys Phe Thr Pro Met Gln
 275 280 285
 Lys Ala Gly Phe Leu Tyr Tyr Glu Asp Leu Val Ser Cys Val Thr Arg
 290 295 300
 Ala Glu Ala Glu Ala Val Gly Val Leu Val Lys Glu Ala Val Gly Ala
 305 310 315 320
 Phe Leu Pro Asp Ala Phe Val Thr Met Thr Gly Gly Phe Arg Arg Gly
 325 330 335
 Lys Lys Met Gly His Asp Val Asp Phe Leu Ile Thr Ser Pro Gly Ser
 340 345 350
 Thr Asp Glu Asp Glu Glu Gln Leu Leu Pro Lys Val Ile Asn Leu Trp
 355 360 365
 Glu Arg Lys Gly Leu Leu Leu Tyr Cys Asp Leu Val Glu Ser Thr Phe
 370 375 380
 Glu Lys Leu Lys Leu Pro Ser Arg Lys Val Asp Ala Leu Asp His Phe
 385 390 395 400
 Gln Lys Cys Phe Leu Ile Leu Lys Leu His His Gln Arg Val Asp Gly
 405 410 415
 Gly Lys Cys Ser Gln Gln Glu Gly Lys Thr Trp Lys Ala Ile Arg Val
 420 425 430
 Asp Leu Val Met Cys Pro Tyr Glu Arg Arg Ala Phe Ala Leu Leu Gly
 435 440 445
 Trp Thr Gly Ser Arg Gln Phe Glu Arg Asp Leu Arg Arg Tyr Ala Ser
 450 455 460
 His Glu Arg Lys Met Ile Leu Asp Asn His Ala Leu Tyr Asp Lys Thr

465	470	475	480
Lys Lys Ile Phe Leu Lys Ala Glu Ser Glu Glu Glu Ile Phe Ala His			
485	490	495	
Leu Gly Leu Asp Tyr Ile Glu Pro Trp Glu Arg Asn Ala			
500	505		
<210> 12			
<211> 506			
<212> PRT			
<213> 原鸡 (Gallus gallus)			
<400> 12			
Met Glu Arg Ile Arg Pro Pro Thr Val Val Ser Gln Arg Lys Arg Gln			
1	5	10	15
Lys Gly Met Tyr Ser Pro Lys Leu Ser Cys Gly Tyr Glu Ile Lys Phe			
20	25	30	
Asn Lys Leu Val Ile Phe Ile Met Gln Arg Lys Met Gly Met Thr Arg			
35	40	45	
Arg Thr Phe Leu Met Glu Leu Ala Arg Ser Lys Gly Phe Arg Val Glu			
50	55	60	
Ser Glu Leu Ser Asp Ser Val Thr His Ile Val Ala Glu Asn Asn Ser			
65	70	75	80
Tyr Pro Glu Val Leu Asp Trp Leu Lys Gly Gln Ala Val Gly Asp Ser			
85	90	95	
Ser Arg Phe Glu Ile Leu Asp Ile Ser Trp Leu Thr Ala Cys Met Glu			
100	105	110	
Met Gly Arg Pro Val Asp Leu Glu Lys Lys Tyr His Leu Val Glu Gln			
115	120	125	
Ala Gly Gln Tyr Pro Thr Leu Lys Thr Pro Glu Ser Glu Val Ser Ser			
130	135	140	
Phe Thr Ala Ser Lys Val Ser Gln Tyr Ser Cys Gln Arg Lys Thr Thr			
145	150	155	160
Leu Asn Asn Cys Asn Lys Lys Phe Thr Asp Ala Phe Glu Ile Met Ala			
165	170	175	
Glu Asn Tyr Glu Phe Lys Glu Asn Glu Ile Phe Cys Leu Glu Phe Leu			
180	185	190	
Arg Ala Ala Ser Val Leu Lys Ser Leu Pro Phe Pro Val Thr Arg Met			
195	200	205	
Lys Asp Ile Gln Gly Leu Pro Cys Met Gly Asp Arg Val Arg Asp Val			
210	215	220	
Ile Glu Glu Ile Ile Glu Glu Gly Glu Ser Ser Arg Ala Lys Asp Val			

225	230	235	240
Leu Asn Asp Glu Arg Tyr Lys Ser Phe Lys Glu Phe Thr Ser Val Phe			
245	250	255	
Gly Val Gly Val Lys Thr Ser Glu Lys Trp Phe Arg Met Gly Leu Arg			
260	265	270	
Thr Val Glu Val Lys Ala Asp Lys Thr Leu Lys Leu Ser Lys Met			
275	280	285	
Gln Arg Ala Gly Phe Leu Tyr Tyr Glu Asp Leu Val Ser Cys Val Ser			
290	295	300	
Lys Ala Glu Ala Asp Ala Val Ser Ser Ile Val Lys Asn Thr Val Cys			
305	310	315	320
Thr Phe Leu Pro Asp Ala Leu Val Thr Ile Thr Gly Gly Phe Arg Arg			
325	330	335	
Gly Lys Lys Ile Gly His Asp Ile Asp Phe Leu Ile Thr Ser Pro Gly			
340	345	350	
Gln Arg Glu Asp Asp Glu Leu Leu His Lys Gly Leu Leu Tyr Cys			
355	360	365	
Asp Ile Ile Glu Ser Thr Phe Val Lys Glu Gln Ile Pro Ser Arg His			
370	375	380	
Val Asp Ala Met Asp His Phe Gln Lys Cys Phe Ala Ile Leu Lys Leu			
385	390	395	400
Tyr Gln Pro Arg Val Asp Asn Ser Ser Tyr Asn Met Ser Lys Lys Cys			
405	410	415	
Asp Met Ala Glu Val Lys Asp Trp Lys Ala Ile Arg Val Asp Leu Val			
420	425	430	
Ile Thr Pro Phe Glu Gln Tyr Ala Tyr Ala Leu Leu Gly Trp Thr Gly			
435	440	445	
Ser Arg Gln Phe Gly Arg Asp Leu Arg Arg Tyr Ala Thr His Glu Arg			
450	455	460	
Lys Met Met Leu Asp Asn His Ala Leu Tyr Asp Lys Arg Lys Arg Val			
465	470	475	480
Phe Leu Lys Ala Gly Ser Glu Glu Glu Ile Phe Ala His Leu Gly Leu			
485	490	495	
Asp Tyr Val Glu Pro Trp Glu Arg Asn Ala			
500	505		
<210> 13			
<211> 509			
<212> PRT			
<213> 智人 (Homo sapiens)			

<400> 13

Met	Asp	Pro	Pro	Arg	Ala	Ser	His	Leu	Ser	Pro	Arg	Lys	Lys	Arg	Pro
1					5				10						15
Arg	Gln	Thr	Gly	Ala	Leu	Met	Ala	Ser	Ser	Pro	Gln	Asp	Ile	Lys	Phe
		20						25						30	
Gln	Asp	Leu	Val	Val	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	Ser	Asp	Val	Leu	Glu	Trp	Leu	Gln	Ala	Gln	Lys	Val	Gln	Val	Ser
		85					90				95				
Ser	Gln	Pro	Glu	Leu	Leu	Asp	Val	Ser	Trp	Leu	Ile	Glu	Cys	Ile	Arg
		100				105					110				
Ala	Gly	Lys	Pro	Val	Glu	Met	Thr	Gly	Lys	His	Gln	Leu	Val	Val	Arg
		115				120				125					
Arg	Asp	Tyr	Ser	Asp	Ser	Thr	Asn	Pro	Gly	Pro	Pro	Lys	Thr	Pro	Pro
		130				135				140					
Ile	Ala	Val	Gln	Lys	Ile	Ser	Gln	Tyr	Ala	Cys	Gln	Arg	Arg	Thr	Thr
		145			150				155				160		
Leu	Asn	Asn	Cys	Asn	Gln	Ile	Phe	Thr	Asp	Ala	Phe	Asp	Ile	Leu	Ala
		165				170				175					
Glu	Asn	Cys	Glu	Phe	Arg	Glu	Asn	Glu	Asp	Ser	Cys	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	Ser	Lys	Val	Lys	Gly	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						
Thr	Leu	Ser	Lys	Val	Arg	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 Val Leu Val Lys Glu Ala Val Trp
 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 Lys Met Gly His Asp Val Asp Phe Leu Ile Thr Ser Pro Gly
 340 345 350
 Ser Thr Glu Asp Glu Glu Gln Leu Leu Gln Lys Val Met Asn Leu Trp
 355 360 365
 Glu Lys Lys Gly Leu Leu Tyr Tyr Asp Leu Val Glu Ser Thr Phe
 370 375 380
 Glu Lys Leu Arg Leu Pro Ser Arg Lys Val Asp Ala Leu Asp His Phe
 385 390 395 400
 Gln Lys Cys Phe Leu Ile Phe Lys Leu Pro Arg Gln Arg Val Asp Ser
 405 410 415
 Asp Gln Ser Ser Trp Gln Glu Gly Lys Thr Trp Lys Ala Ile Arg Val
 420 425 430
 Asp Leu Val Leu Cys Pro Tyr Glu Arg Arg Ala Phe Ala Leu Leu Gly
 435 440 445
 Trp Thr Gly Ser Arg Gln Phe Glu Arg Asp Leu Arg Arg Tyr Ala Thr
 450 455 460
 His Glu Arg Lys Met Ile Leu Asp Asn His Ala Leu Tyr Asp Lys Thr
 465 470 475 480
 Lys Arg Ile Phe Leu Lys Ala Glu Ser Glu Glu Glu Ile Phe Ala His
 485 490 495
 Leu Gly Leu Asp Tyr Ile Glu Pro Trp Glu Arg Asn Ala
 500 505

<210> 14

<211> 14

<212> DNA

<213> 人工序列 (artificial sequence)

<220>

<223> 引物 (amorce)

<400> 14

aaaaaaaaaaa gggg 14

<210> 15

<211> 10

<212> DNA

<213> 人工序列 (artificial sequence)

<220>

<223> 合成的序列
<400> 15
gtacgctagt 10
<210> 16
<211> 12
<212> DNA
<213> 人工序列 (artificial sequence)
<220>
<223> 捕获片段
<400> 16
cctttttttt tt 12

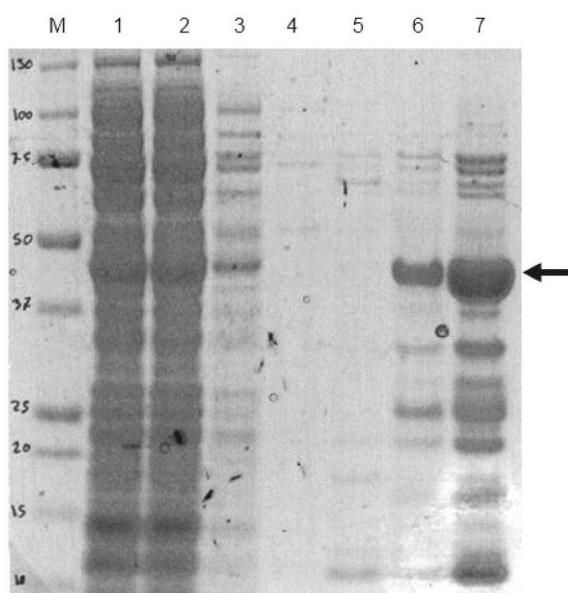
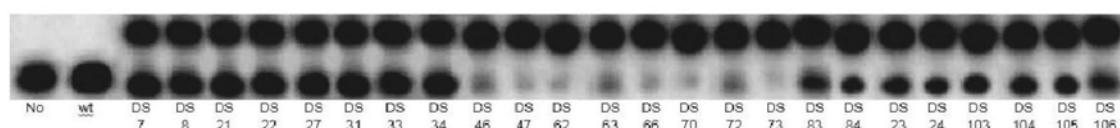


图1

ONH2 掺入



Biot-EDA 掺入

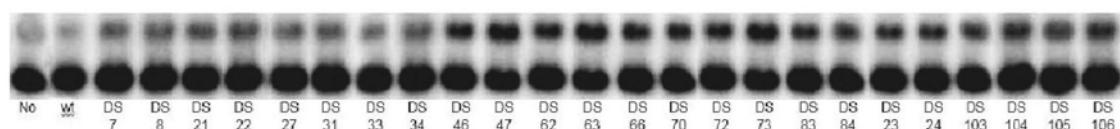
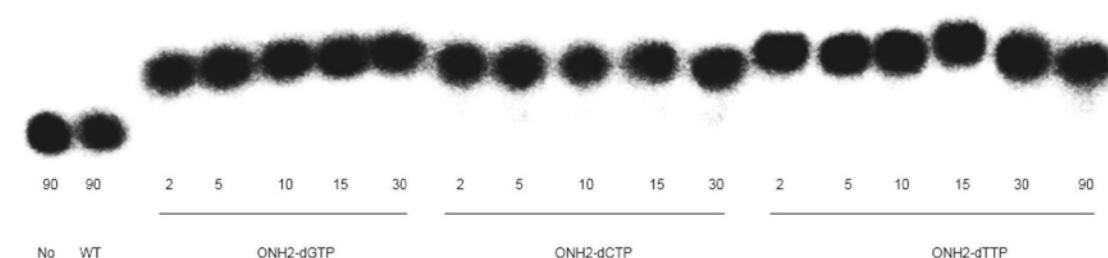


图3

图2



冬 4

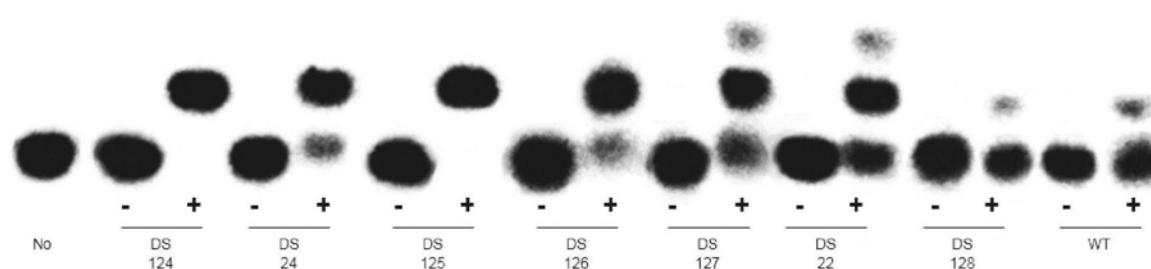


图5

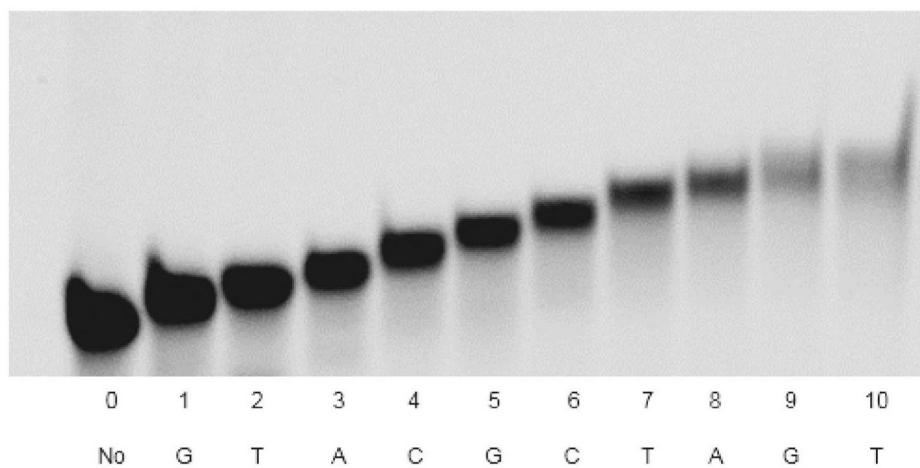


图6