(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 103025869 B
(45)授权公告日 2015.07.01

(21)申请号 201180038487.7
(22)申请日 2011.06.17
(30)优先权数据
61/356219 2010.06.18 US
(85)PCT国际申请进入国家阶段日
2013.01.23
(86)PCT国际申请的申请数据
PCT/EP2011/003002 2011.06.17
(87)PCT国际申请的公布数据
WO2011/157437 EN 2011.12.22
(73)专利权人 霍夫曼－拉罗奇有限公司
地址 瑞士巴塞尔
(72)发明人 F. 赖切特 K. 鲍尔
T. W. 迈尔斯
(74)专利代理机构 中国专利代理(香港)有限公司 72001
代理人 翟建伟 权陆军
(51)Int.Cl.
C12N 9/12(2006.01)

(56)对比文件
CN 101528919 A,2009.09.09, 全文.
CN 101528919 A,2009.09.09, 全文.

审查员 蔺娜

权利要求书1页 说明书37页
序列表61页 附图2页

(54)发明名称

具有增强的3'-错配辨别力的DNA聚合酶

(57)摘要

公开了相对于对应的未修饰的聚合酶具有增强的3′-错配辨识力的突变型DNA聚合酶。突变型聚合酶可用于多种公开的引物延伸方法中。还公开了相关的组合物,包括可用于,例如,突变型DNA聚合酶生产的重组核酸、载体以及宿主细胞。

```
                                    *
Z05     AWIEKTLEEGRKRGYVE TLF G RR R Y V PDLNARVK (SEQ ID NO:12)
Taq     AWIEKTLEEGRRRGYVE TLF G RR R Y V PDLEARVK (SEQ ID NO:13)
Tfi     AWIAKTLEEGRKKGYVE TLF G RR R Y V PDLNARVK (SEQ ID NO:14)
Tfl     AWIEGTLEEGRRRGYVE TLF G RR R Y V PDLNARVK (SEQ ID NO:15)
Sps17   AWIAKTLEEGRKKGYVE TLF G RR R Y V PDLNARVK (SEQ ID NO:16)
Tth     AWIEKTLEEGRKRGYVE TLF G RR R Y V PDLNARVK (SEQ ID NO:17)
Tca     AWIEKTLEEGRKRGYVE TLF G RR R Y V PDLNARVK (SEQ ID NO:18)
Tma     DYIQRVVSEAKEKGYVR TLF G RK R D I PQLMARDR (SEQ ID NO:19)
Tne     SYIQQVVAEAKEKGYVR TLF G RK R D I PQLMARDK (SEQ ID NO:20)
Taf     EYLKRMKDEARKKGYVT TLF G RR R Y I PQLRSKNG (SEQ ID NO:21)
Dra     RYINHTLDFGRTHGYVE TLY G RR R Y V PGLSSRNR (SEQ ID NO:23)
Bst     QYMDNIVQEAKQKGYVT TLL H RR R Y L PDITSRNF (SEQ ID NO:24)
Bca     RYMENIVQEAKQKGYVT TLL H RR R Y L PDITSRNF (SEQ ID NO:25)
        ----------GYVX1TLX2X3RX4X5X6X7P------- (SEQ ID NO:26)
```

CN 103025869 B

1. DNA 聚合酶，与对照 DNA 聚合酶相比，所述 DNA 聚合酶具有增强的 3'- 错配辨别力活性，其中所述 DNA 聚合酶由 SEQ ID NO:1 组成，除了对应于 SEQ ID NO:1 的 730 位的 DNA 聚合酶的氨基酸是 A、C、D、G 或 N 中的任一个和对应于 SEQ ID NO:1 的 580 位的 DNA 聚合酶的氨基酸是 G，所述对照 DNA 聚合酶由 SEQ ID NO:1 组成，除了对应于 SEQ ID NO:1 的 580 位的对照 DNA 聚合酶的氨基酸是 G。

2. 重组核酸，其编码权利要求 1 的 DNA 聚合酶。

3. 进行引物延伸的非诊断方法，包括

在适于引物延伸的条件下将权利要求 1 的 DNA 聚合酶与引物、多核苷酸模板和三磷酸核苷接触，从而产生延伸的引物。

4. 产生延伸的引物的试剂盒，包括：

至少一个包含权利要求 1 的 DNA 聚合酶的容器。

5. 权利要求 4 的试剂盒，进一步包括一个或更多个额外的容器，所述容器选自：

(a) 容器，其包含在引物延伸条件下与预定的多核苷酸模板可杂交的引物；

(b) 容器，其包含三磷酸核苷；以及

(c) 容器，其包含适合用于引物延伸的缓冲剂。

6. 反应混合物，包含权利要求 1 的 DNA 聚合酶，至少一种引物，多核苷酸模板和三磷酸核苷。

具有增强的3'-错配辨别力的DNA聚合酶

发明领域

[0001] 本发明提供了具有增强的3'-错配辨别力的DNA聚合酶及它们在各种应用，包括核酸多核苷酸延伸和扩增中的用途。

[0002] 发明背景

[0003] DNA聚合酶负责基因组的复制和维护，这是在世代间(from generation to generation)精确传递遗传信息的中心职责。DNA聚合酶在细胞中的功能是作为负责DNA合成的酶。它们在有金属活化剂如Mg^{2+}存在的情况下，以被复制的DNA模板或多核苷酸模板支配的顺序聚合脱氧核糖核苷三磷酸。在体内，DNA聚合酶参与一系列DNA合成过程，包括DNA复制、DNA修复、重组和基因扩增。在每一DNA合成过程期间，复制DNA模板一次或最多几次以产生相同的复制品。与此相反，在体外，DNA复制可以重复多次，例如在聚合酶链式反应期间(参见例如美国专利号4,683,202)。

[0004] 在聚合酶链式反应(PCR)的最初研究中，DNA聚合酶是在每一轮DNA复制的开始时添加的(参见上述美国专利号4,683,202)。后来，确认了从在高温下生长的细菌中可获得热稳定DNA聚合酶，这些酶只需要添加一次(参见美国专利号4,889,818和美国专利号4,965,188)。在PCR期间使用的高温下，这些酶没有不可逆地失活。因此，可以进行聚合酶链式反应的重复循环而不需要在每一合成加成过程开始时添加新鲜的酶。DNA聚合酶特别是热稳定聚合酶，是重组DNA研究和疾病的医疗诊断中大量技术的关键。特别是对于诊断应用，目标核酸序列可能仅仅是所考察(in question)DNA或RNA的一小部分，因此在不扩增的情况下可能难以检测到目标核酸序列的存在。

[0005] DNA聚合酶的总体折叠模式类似人的右手，包含手掌、手指和拇指三个独特的亚结构域。(参见beese等，*Science* 260:352-355, 1993);Patel等，*Biochemistry* 34:5351-5363, 1995)。尽管在大小和细胞功能不同的聚合酶之间手指和拇指亚结构域的结构变化很大，但是催化性的手掌亚结构域全部都是重叠的(superimposable)。例如，在化学催化作用期间与引入的dNTP相互作用并稳定过渡态的基序A在哺乳动物polα和原核生物的pol I家族DNA聚合酶之间是重叠的，平均偏差约为1Å(Wang等，*Cell* 89:1087-1099, 1997)。在结构上基序A以主要包含疏水性残基的反向平行β-折叠链起始，延续到α-螺旋。DNA聚合酶活性位点的主要氨基酸序列是特别保守的。在基序A的情况下，例如，序列DYSQIELR(SEQ ID NO:28)保留在从数百万年进化分离的生物体包括例如水生栖热菌(*Thermus aquaticus*)、沙眼衣原体(*Chlamydia trachomatis*)和大肠杆菌(*Escherichia coli*)的聚合酶中。

[0006] 除了高度保守以外，DNA聚合酶的活性位点也被证明是相对易变的，能够容纳某些氨基酸取代而不显著降低DNA聚合酶活性。(参见例如美国专利号6,602,695)。这种突变型DNA聚合酶可以在例如包括核酸合成反应的诊断和研究应用中提供各种选择优势。因此，本领域需要鉴别适于突变的氨基酸位置以产生改进的聚合酶活性。本发明，如同本文中所述，满足这些及其它需要。

[0007] 发明概述

[0008] 本文提供了相对于对应的未修饰的对照聚合酶具有增强的3'-错配辨别力的DNA聚合酶以及制备和使用这种DNA聚合酶的方法。在一些实施方案中，聚合酶是热稳定DNA聚合酶。在一些实施方案中，DNA聚合酶是热活性的DNA聚合酶。在一些实施方案中，DNA聚合酶源自栖热菌种(*Thermus* species)。在一些实施方案中，DNA聚合酶源自栖热袍菌种(*Thermotoga* species)。在一些实施方案中，对应于SEQ ID NO:1的730位的DNA聚合酶的氨基酸是除R以外的任何氨基酸，除了对应于SEQ ID NO:1的730位的对照DNA聚合酶的氨基酸是R，对照DNA聚合酶具有与所述DNA聚合酶相同的氨基酸序列。例如，在一些实施方案中，对应于SEQ ID NO:1的730位的位置处的氨基酸选自G, A, V, L, I, M, F, W, P, S, T, C, Y, N, Q, D, E, K，或H。在一些实施方案中，对应于SEQ ID NO:1的730位的位置处的氨基酸是A, C, D, G，或N。

[0009] 在一些实施方案中，对应于SEQ ID NO:1的730位的位置处的氨基酸是具有非极性，不带电荷的侧链的氨基酸(例如，G, A, L, M, W, P, F, C, V，或I)。在一些实施方案中，对应于SEQ ID NO:1的730位的位置处的氨基酸是具有极性，不带电荷的侧链的氨基酸(例如，N, Q, H, S, T，或Y)。在一些实施方案中，对应于SEQ ID NO:1的730位的位置处的氨基酸是具有极性，带负电荷的侧链的氨基酸(例如，D或E)。

[0010] 在一些实施方案中，具有增强的3'-错配辨别力的DNA聚合酶在聚合酶结构域中包含基序，所述基序包含

[0011] G-Y-V-X_1-T-L-X_2-X_3-R-X_4-X_5-X_6-X_7-P，其中：

[0012] X_1是E, R或T；

[0013] X_2是F, Y或L；

[0014] X_3是G或H；

[0015] X_4是R或K；

[0016] X_5是除R以外的任何氨基酸；

[0017] X_6是Y或D；和

[0018] X_7是V, I或L(SEQ ID NO:8)。

[0019] 在一些实施方案中，X_5选自G, A, L, M, W, P, S, T, F, Y, C, N, Q, D, E, K, V, I或H(SEQ ID NO:42)。

[0020] 在一些实施方案中，具有增强的3'-错配辨别力的DNA聚合酶在聚合酶结构域中包含基序，所述基序包含

[0021] G-Y-V-X_1-T-L-F-G-R-X_4-X_5-X_6-X_7-P，其中：

[0022] X_1是E或R；

[0023] X_4是R或K；

[0024] X_5是除R以外的任何氨基酸；

[0025] X_6是Y或D；和

[0026] X_7是V或I(SEQ ID NO:9)。

[0027] 在一些实施方案中，具有增强的3'-错配辨别力的DNA聚合酶在聚合酶结构域中包含基序，所述基序包含

[0028] G-Y-V-E-T-L-F-G-R-R-X_5-Y-V-P，其中：

[0029] X_5是除R以外的任何氨基酸(SEQ ID NO:10)。

[0030] 在一些实施方案中，$X_{5是}$具有非极性、不带电荷的侧链的氨基酸（例如，G，A，L，M，W，P，F，C，V，或 I），具有极性、不带电荷的侧链的氨基酸（例如，N，Q，H，S，T，或 Y）或具有极性、带负电荷的侧链的氨基酸（例如，D 或 E）。

[0031] 在一些实施方案中，X_5是 A，C，D，G，或 N（SEQ ID NO:11）。

[0032] 在一些实施方案中，对应于 SEQ ID NO:1 的 580 位的 DNA 聚合酶的氨基酸是除 D 或 E 以外的任何氨基酸。在一些实施方案中，对应于 SEQ ID NO:1 的 580 位的 DNA 聚合酶的氨基酸是除 D 以外的任何氨基酸。在一些实施方案中，对应于 SEQ ID NO:1 的 580 位的 DNA 聚合酶的氨基酸选自 L，G，T，Q，A，S，N，R 和 K。在一些实施方案中，对应于 SEQ ID NO:1 的 580 位的 DNA 聚合酶的氨基酸是 G。

[0033] 各种 DNA 聚合酶适于本发明的突变。尤其适合的是热稳定聚合酶，包括来自各种嗜热菌（thermophilic bacteria）的野生型或天然存在的热稳定聚合酶，以及通过氨基酸取代、插入或缺失或其它修饰而源自这种野生型或天然存在的酶的合成的热稳定聚合酶。示例性的聚合酶的未修饰形式包括例如 CS5（SEQ ID NO:29）或 CS6（SEQ ID NO:30）或 Z05 DNA 聚合酶（SEQ ID NO:1），或与其具有至少 80％、优选至少 90％或更优选至少 95％的序列同一性的功能性 DNA 聚合酶。其它未修饰的聚合酶包括例如来自于下列嗜热菌中任一种的 DNA 聚合酶（或与这种聚合酶具有至少 80％、优选至少 90％或更优选至少 95％的序列同一性的功能性 DNA 聚合酶）：海栖热袍菌（*Thermotoga maritima*）（SEQ ID NO:38）；水生栖热菌（*Thermus aquaticus*）（SEQ ID NO:2）；嗜热栖热菌（*Thermus thermophilus*）（SEQ ID NO:6）；黄栖热菌（*Thermus flavus*）（SEQ ID NO:4）；丝状栖热菌（*Thermus filiformis*）（SEQ ID NO:3）；栖热菌种 Sps17（*Thermus* sp. sps17）（SEQ ID NO:5）；栖热菌种 Z05（*Thermus* sp. Z05）（SEQ ID NO:1）；那不勒斯栖热袍菌（*Thermotoga neopolitana*）（SEQ ID NO:39）；非洲栖热腔菌（*Thermosipho africanus*）（SEQ ID NO:37）；*Thermus caldophilus*（SEQ ID NO:7），耐放射异常球菌（*Deinococcus radiodurans*）（SEQ ID NO:36），嗜热脂肪芽孢杆菌（*Bacillus stearothermophilus*）（SEQ ID NO:40）或热坚芽孢杆菌（*Bacillus caldotenax*）（SEQ ID NO:41）。合适的聚合酶还包括具有逆转录酶（RT）活性和 / 或能够掺入非常规核苷酸如核糖核苷酸或其它 2'- 修饰的核苷酸的那些。

[0034] 尽管具有有效的 3'- 错配辨别力活性的热稳定性 DNA 聚合酶尤其适合用于进行 PCR，具有有效的 3'- 错配辨别力活性的热活性的、但不是热稳定的 DNA 聚合酶也适于本发明的突变。

[0035] 在一些实施方案中，DNA 聚合酶是栖热菌属 DNA 聚合酶。例如，在一些实施方案中，DNA 聚合酶与选自下列的聚合酶具有至少 80％，优选至少 90％，更优选至少 95％的序列同一性：

[0036] (a) 栖热菌种 Z05DNA 聚合酶（Z05）（SEQ ID NO:1）；

[0037] (b) 水生栖热菌 DNA 聚合酶（Taq）（SEQ ID NO:2）；

[0038] (c) 丝状栖热菌 DNA 聚合酶（Tfi）（SEQ ID NO:3）；

[0039] (d) 黄栖热菌 DNA 聚合酶（Tfl）（SEQ ID NO:4）；

[0040] (e) 栖热菌种 Sps17DNA 聚合酶（Sps17）（SEQ ID NO:5）；

[0041] (f) 嗜热栖热菌 DNA 聚合酶（Tth）（SEQ ID NO:6）；和

[0042] (g) *Thermus caldophilus* DNA 聚合酶（Tca）（SEQ ID NO:7）。

[0043] 在一些实施方案中，DNA 聚合酶是栖热袍菌 DNA 聚合酶。例如，在一些实施方案中，DNA 聚合酶与选自下列的聚合酶具有至少 80％，优选至少 90％，更优选至少 95％的序列同一性：

[0044] (a) 海栖热袍菌 DNA 聚合酶 (Tma) (SEQ ID NO:38)；

[0045] (b) 那不勒斯栖热袍菌 DNA 聚合酶 (Tne) (SEQ ID NO:39)；

[0046] 在某些实施方案中，DNA 聚合酶与 SEQ ID NO:1 具有至少 80％，优选至少 90％，更优选至少 95％的序列同一性。在一些实施方案中，DNA 聚合酶是栖热菌种 Z05 DNA 聚合酶 (Z05)DNA 聚合酶(即 SEQ ID NO:1)，除了 730 位的氨基酸是除 R 以外的任何氨基酸。例如，在一些实施方案中，对应于 SEQ ID NO:1 的 730 位的位置处的氨基酸选自 G, A, V, L, I, M, F, W, P, S, T, C, Y, N, Q, D, E, K, 或 H。在一些实施方案中，DNA 聚合酶是 Z05 DNA 聚合酶，730 位的氨基酸是 A, C, D, G, 或 N。在一些实施方案中，DNA 聚合酶是进一步包含 580 位的取代的 Z05 DNA 聚合酶，580 位的氨基酸是除 D 或 E 以外的任何氨基酸。在一些实施方案中，DNA 聚合酶是 Z05 DNA 聚合酶，580 位的氨基酸是除 D 以外的任何氨基酸。在一些实施方案中，DNA 聚合酶是 Z05 DNA 聚合酶，580 位的氨基酸选自 L, G, T, Q, A, S, N, R 和 K。在一些实施方案中，DNA 聚合酶是 Z05 DNA 聚合酶，580 位的氨基酸是 G。

[0047] 突变的或改进的聚合酶可以包括其它非取代的修饰。一种这种修饰是热可逆的共价修饰，其使酶失活，但是在高温，如通常用于多核苷酸延伸的温度下孵育后其被逆转而激活酶。Birch 等的美国专利号 5,773,258 和 5,677,152 中描述了这种热可逆的修饰的示例性试剂。

[0048] 在一些实施方案中，通过在具有预定拷贝数的野生型 BRAF V600R 目标多核苷酸和预定拷贝数的突变型 BRAF V600 目标多核苷酸（数量上等于或少于野生型目标的拷贝数（例如，10,000 或更少拷贝））的一个或多个反应混合物中在正向引物（与突变型序列完美匹配并且与野生型序列具有一个单一的 3’A:C 错配）存在的情况下使用具有 SEQ ID NO:35(野生型 BRAF = SEQ ID NO:34)的核酸序列的突变型 BRAF V600R 目标多核苷酸而测定 3’- 错配活性。两个或更多的反应混合物可以具有滴定的拷贝数的突变型 BRAF V600R 目标多核苷酸（例如，在数个反应混合物中，1:5 滴定，1:10 滴定，例如，10,000 拷贝， 1000 拷贝， 100 拷贝， 10 拷贝， 1 拷贝， 0 拷贝）。如本文中所述，可以在预先选择的单位时间，将本发明的聚合酶的 3’- 错配辨别力与参照聚合酶（例如，天然存在的或未修饰的聚合酶）的 3’- 错配辨别力相比较。当与突变型等位基因完美匹配的引物接触时，具有增强的 3’- 错配辨别力的聚合酶将不能扩增野生型序列，或者，与天然存在的或未修饰的聚合酶相比，使用突变型等位基因特异性引物扩增野生型序列将需要更大的 PCR 循环数（即，表现出更高的 Cp 值）。

[0049] 在各种其它方面，本发明提供了编码本文所述的突变的或改进的 DNA 聚合酶的重组核酸，包含该重组核酸的载体，和 / 或用所述载体转化的宿主细胞。在某些实施方案中，所述载体是表达载体。包含这种表达载体的宿主细胞可用于本发明的生产突变的或改进的聚合酶的方法，所述方法通过在适于表达重组核酸的条件下培养宿主细胞。反应混合物和 / 或试剂盒中可以包含本发明的聚合酶。重组核酸，宿主细胞，载体，表达载体，反应混合物和试剂盒的实施方案如上面和本文中所述。

[0050] 在另一方面，提供了实施多核苷酸延伸的方法。该方法总体包括在适于引物延伸

的条件下将本文所述的具有增强的 3'- 错配辨别力的 DNA 聚合酶与引物、多核苷酸模板以及三磷酸核苷接触，从而产生延伸的引物。多核苷酸模板可以是例如 RNA 或 DNA 模板。三磷酸核苷可以包括非常规核苷酸如核糖核苷酸和 / 或标记的核苷酸。此外，引物和 / 或模板可以包括一个或更多个核苷酸类似物。在一些变型中，多核苷酸延伸方法是用于多核苷酸扩增的方法，包括在适于多核苷酸扩增的条件下将突变的或改进的 DNA 聚合酶与引物对、多核苷酸模板以及三磷酸核苷接触。多核苷酸延伸反应可以是，例如，PCR，等温延伸，或测序（例如，454 测序反应）。

[0051] 在一些实施方案中，引物延伸方法是用于实施聚合酶链式反应 (PCR) 的方法。

[0052] 本发明还提供了用于这种多核苷酸延伸方法中的试剂盒。通常，试剂盒包括至少一个提供本文所述的突变的或改进的 DNA 聚合酶的容器。在某些实施方案中，试剂盒进一步包括提供一种或更多种额外试剂的一个或更多个额外的容器。例如，在特定的变型中，一个或更多个额外的容器提供三磷酸核苷；适于多核苷酸延伸的缓冲剂；和 / 或在多核苷酸延伸条件下与预定的多核苷酸模板杂交的引物。

[0053] 进一步提供了包含本发明的聚合酶的反应混合物。反应混合物也可以包含模板核酸 (DNA 和 / 或 RNA)，一种或多种引物或探针多核苷酸，三磷酸核苷（包括，例如，脱氧核糖核苷酸，核糖核苷酸，标记的核苷酸，非常规的核苷酸），缓冲剂，盐，标记（例如，荧光团）。

[0054] 本文描述了本发明的进一步的实施方案。

[0055] 定义

[0056] 除非另外定义，本文中使用的所有技术和科学术语具有与本发明所属技术领域的普通技术人员通常理解的相同的意义。尽管基本上与本文所述的那些相似的任何方法和材料都可用于本发明的实践或试验，但是仅仅描述了示例性的方法和材料。对于本发明的目的，下列术语定义如下。

[0057] 术语“一 (a)，”“一 (an)，”和“该 (the)”包括复数对象，除非上下文清楚地指出别的含义。

[0058] “氨基酸”是指可以并入肽、多肽或蛋白的任何单体单位。如本文所用，术语“氨基酸”包括下列 20 种天然或遗传编码的 α- 氨基酸：丙氨酸 (Ala 或 A)、精氨酸 (Arg 或 R)、天冬酰胺 (Asn 或 N)、天冬氨酸 (Asp 或 D)、半胱氨酸 (Cys 或 C)、谷胺酰胺 (Gln 或 Q)、谷氨酸 (Glu 或 E)、甘氨酸 (Gly 或 G)、组氨酸 (His 或 H)、异亮氨酸 (Ile 或 I)、亮氨酸 (Leu 或 L)、赖氨酸 (Lys 或 K)、甲硫氨酸 (Met 或 M)、苯丙氨酸 (Phe 或 F)、脯氨酸 (Pro 或 P)、丝氨酸 (Ser 或 S)、苏氨酸 (Thr 或 T)、色氨酸 (Trp 或 W)、酪氨酸 (Tyr 或 Y) 和缬氨酸 (Val 或 V)。在其中“X”残基没有定义的情况下，这些应该定义为“任何氨基酸”。这 20 种天然氨基酸的结构见例如 Stryer 等，BioChemistry，5th ed.，Freeman and Company(2002)。额外的氨基酸如硒代半胱氨酸和吡咯赖氨酸也能够被遗传编码 (Stadtman (1996)“Selenocysteine，”Annu Rev Biochem.65:83-100 和 Ibba 等 .(2002)“Genetic code:introducing pyrrolysine，”Curr Biol.12(13):R464-R466)。术语“氨基酸”还包括非天然氨基酸、修饰的氨基酸（例如，具有修饰的侧链和 / 或骨架）和氨基酸类似物。参见，例如，Zhang 等 .(2004)“Selective incorporation of 5-hydroxytryptophan into proteins in mammalian cells，”Proc. Natl. Acad. Sci. U.S.A. 101(24):8882-8887，Anderson 等 .(2004)“An expanded genetic code with a functional quadruplet codon”Proc. Natl. Acad. Sci. U.S.A.

101(20):7566-7571, Ikeda等. (2003)“Synthesis of a novel histidine analogue and its efficient incorporation into a protein in vivo,”Protein Eng. Des. Sel. 16(9):699-706, Chin等. (2003)“An Expanded Eukaryotic Genetic Code,”Science 301(5635):964-967, James等. (2001)“Kinetic characterization of ribonuclease S mutants containing photoisomerizable phenylazophenylalanine residues,”Protein Eng. Des. Sel. 14(12):983-991, Kohrer等. (2001)“Import of amber and ochre suppressor tRNAs into mammalian cells:A general approach to site-specific insertion of amino acid analogues into proteins,”Proc. Natl. Acad. Sci. U.S.A. 98(25):14310-14315, Bacher 等 . (2001)“Selection and Characterization of Escherichia coli Variants Capable of Growth on an Otherwise Toxic Tryptophan Analogue,”J. Bacteriol. 183(18):5414-5425, Hamano-Takaku 等 . (2000)“A Mutant Escherichia coli Tyrosyl-tRNA Synthetase Utilizes the Unnatural Amino Acid Azatyrosine More Efficiently than Tyrosine,”J. Biol. Chem. 275(51):40324-40328,和Budisa等. (2001)“Proteins with {beta}-(thienopyrrolyl) alanines as alternative chromophores and pharmaceutically active amino acids,”Protein Sci. 10(7):1281-1292。

[0059] 为了进一步说明,氨基酸典型地为包括取代的或未取代的氨基、取代的或未取代的羧基和一种或多种侧链或基团,或这些基团的任一种的类似物的有机酸。示例性的侧链包括,例如巯基、硒基、磺酰基、烷基、芳基、酰基、酮基、叠氮基、羟基、肼、氰基、卤素、酰肼、链烯基、炔基、醚基、硼酸酯 (borate)、硼酸盐 (boronate)、二氧磷基、膦酰基、膦、杂环、烯酮、亚胺、醛、酯、硫代酸、羟胺或这些基团的任何组合。其它代表性的氨基酸包括但不限于,包含光敏交联剂的氨基酸、金属结合氨基酸、自旋标记的氨基酸、发荧光的氨基酸、包含金属的氨基酸、含新官能团的氨基酸、与其它分子共价或非共价相互作用的氨基酸、对光不稳 (photocaged) 和 / 或可光异构化的氨基酸、放射性氨基酸、包含生物素或生物素类似物的氨基酸、糖基化的氨基酸、其它碳水化合物修饰的氨基酸、包含聚乙二醇或聚醚的氨基酸、重原子取代的氨基酸、化学可裂的和 / 或光可裂的氨基酸、包含碳连接糖的氨基酸、氧化还原活性氨基酸、包含氨基硫代酸的氨基酸和包含一个或多个毒性部分的氨基酸。

[0060] 术语“适配子”是指识别并结合 DNA 聚合酶并有效抑制聚合酶活性的单链 DNA,如美国专利号 5,693,502 中所述。

[0061] 在本发明的 DNA 聚合酶上下文中的术语“突变体”是指相对于对应的天然存在的或未修饰的 DNA 聚合酶包含一个或多个氨基酸取代的多肽,典型为重组的。

[0062] 在突变型聚合酶上下文中的术语“未修饰的形式”在本文中是用来定义本发明的突变型 DNA 聚合酶的术语 :术语“未修饰的形式”是指具有除指定为表征突变型聚合酶的一个或多个氨基酸位置以外的突变型聚合酶的氨基酸序列的功能性 DNA 聚合酶。因此,在 (a) 其未修饰的形式和 (b) 一个或多个特定氨基酸取代的方面提到突变型 DNA 聚合酶是指除特定的氨基酸取代以外,突变型聚合酶在特定的基序中另外具有与未修饰的形式相同的氨基酸序列。“未修饰的聚合酶”(和因此还有具有增强的 3’- 错配辨别力的修饰的形式) 可以包含额外的突变以提供期望的功能性,例如双脱氧核糖核苷酸、核糖核苷酸、核糖核苷酸类似物、染料标记的核苷酸的改进的掺入,调节 5’- 核酸酶活性,调节 3’- 核酸酶 (或校正)

活性等。因此，在实施本文所述的本发明时，DNA 聚合酶的未修饰的形式是预先确定的。DNA 聚合酶的未修饰的形式可以是，例如野生型和 / 或天然存在的 DNA 聚合酶，或已经有意识地修饰过的 DNA 聚合酶。聚合酶的未修饰的形式优选热稳定 DNA 聚合酶，如来自各种嗜热菌的 DNA 聚合酶以及与野生型或天然存在的热稳定聚合酶具有基本上序列同一性的其功能性变体。这种变体可以包括，例如，嵌合 DNA 聚合酶，如美国专利号 6,228,628 和美国申请公开号 2004/0005599 中所述的嵌合 DNA 聚合酶。

[0063] 术语“热稳定聚合酶”是指对热稳定的酶，其是耐热的，当经受高温并持续影响双链核酸变性所必需的时间时，其保留足够的活性以影响随后的多核苷酸延伸反应，并且没有变得不可逆变性的（失活的）。核酸变性所需的加热条件是本领域公知的，在例如美国专利号 4,683,202、4,683,195 和 4,965,188 中举例说明。如本文所述，热稳定聚合酶适于在温度循环变化的反应如聚合酶链式反应（“PCR”）中使用。用于本文目的的不可逆变性指酶活性的永久且完全的损失。对于热稳定聚合酶，酶活性是指以适当的方式催化核苷酸的组合以形成与模板核酸链互补的多核苷酸延伸产物。来自嗜热菌的热稳定 DNA 聚合酶包括，例如来自海栖热袍菌，水生栖热菌，嗜热栖热菌，黄栖热菌，丝状栖热菌，栖热菌种 Sps17，栖热菌种 Z05，*Thermus caldophilus*，热坚芽孢杆菌，那不勒斯栖热袍菌和非洲栖热腔菌的 DNA 聚合酶。

[0064] 术语“热活性的”是指在通常用于 RT-PCR 和 / 或 PCR 反应中逆转录或退火 / 延伸步骤的温度（即，45-80℃）保持催化性能的酶。热稳定酶是当经受对于核酸变性所必需的高温时不会被不可逆地失活或变性的那些酶。热活性的酶可以是或可以不是热稳定的。热活性的 DNA 聚合酶可以是 DNA 或 RNA 依赖于嗜热种或依赖于嗜常温种，包括但不限于，大肠杆菌，莫洛尼小鼠白血病病毒和禽成髓细胞瘤病毒 *(Avian myoblastosis virus)*。

[0065] 如本文使用的，“嵌合”蛋白是指氨基酸序列代表来自至少两种不同蛋白的氨基酸序列的子序列的融合产物的蛋白。嵌合蛋白典型地不是通过氨基酸序列的直接操作产生，而是从编码该嵌合氨基酸序列的“嵌合”基因表达的。在某些实施方案中，例如，本发明的突变型 DNA 聚合酶的未修饰形式是由来源于栖热菌种 DNA 聚合酶的氨基末端（N- 末端）区域和来源于 Tma DNA 聚合酶的羧基末端（C- 末端）区域组成的嵌合蛋白。N- 末端区域是指从 N- 末端（氨基酸位置 1）延伸到内部氨基酸的区域。类似地，C- 末端区域是指从内部氨基酸延伸到 C- 末端的区域。

[0066] 在 DNA 聚合酶的上下文中，与另一条序列（例如区域、片段、核苷酸或氨基酸位置等）的“对应”是基于根据核苷酸或氨基酸位置数编号并且然后以最大化序列同一性百分比的方式比对序列的规则。因为不是给定的“对应区域”内的所有位置都需要是相同的，对应区域内的非匹配位置可以被认为是“对应位置”。因此，如本文使用的，特定 DNA 聚合酶的“对应于氨基酸位点 [X] 的氨基酸位置”是指，基于比对，在其它 DNA 聚合酶和结构同源物和家族中的等效位置。在本发明的一些实施方案中，氨基酸位置的“对应”是根据包含 SEQ ID NO:1, 2, 3, 4, 5, 6, 7, 36, 37, 38, 39, 40 或 41 的一个或多个基序的聚合酶区域来确定的。当聚合酶多肽序列不同于 SEQ ID NOS:1, 2, 3, 4, 5, 6, 7, 36, 37, 38, 39, 40 或 41（例如，通过氨基酸的变化或氨基酸的增加或缺失）时，可能与本文所讨论的改进的活性相关的特定突变将不是在与它在 SEQ ID NOS:1, 2, 3, 4, 5, 6, 7, 36, 37, 38, 39, 40 或 41 中相同的位置数。例如，在表 1 中说明了这一点。

[0067] 如本文使用的，“重组体”是指已经通过重组方法有意识地修饰过的氨基酸序列或核苷酸序列。本文中的术语“重组核酸”是指通常通过内切核酸酶的核酸操作，最初在体外形成的以自然中通常不存在的形式的核酸。因此，线性形式的分离的突变型 DNA 聚合酶核酸或通过连接通常不结合的 DNA 分子体外形成的表达载体都被认为是用于本发明的目的的重组体。可以理解的是，一旦重组核酸被制备并重新引入宿主细胞，它将非重组地复制，即，使用宿主细胞的体内细胞机制而不是体外操作；然而，这种核酸一旦重组产生，尽管随后非重组复制，仍然被认为是用于本发明的目的的重组体。“重组蛋白”是使用重组技术，即通过上述重组核酸的表达制备的蛋白。

[0068] 当被置于与另一种核酸序列的功能性关系时核酸是“可操作连接的”。例如，如果启动子或增强子影响编码序列的转录，那么它是与编码序列可操作连接的；或者如果核糖体结合位点被定位以促进翻译，那么它是与编码序列可操作连接的。

[0069] 术语“宿主细胞”是指单细胞的原核生物和真核生物有机体（例如细菌、酵母和放线菌）以及细胞培养物中生长的来自高等植物或者动物的单个细胞。

[0070] 术语“载体”指一条 DNA，典型是双链的，它可以是其中已经插入了一条外源 DNA 的。载体可以是例如质粒来源的。载体包含在宿主细胞中促进载体的自主复制的“复制子”多核苷酸序列。外源 DNA 定义为异源的 DNA，其是在宿主细胞中天然不存在的 DNA，例如，其复制载体分子，编码可选择或可筛选的标记，或者编码转基因。载体用于转运外源的或异源的 DNA 进入合适的宿主细胞。一旦在宿主细胞中，载体可以独立于宿主染色体 DNA 复制或与宿主染色体 DNA 同时复制，可以产生数个拷贝的载体及其插入的 DNA。另外，载体还可以包含容许插入的 DNA 转录为 mRNA 分子或者使插入的 DNA 复制为多拷贝的 RNA 的必要元件。一些表达载体还包括插入的 DNA 附近的序列元件，其增加表达的 mRNA 的半衰期，和 / 或允许所述 mRNA 翻译为蛋白分子。因此，可以迅速地合成插入的 DNA 编码的 mRNA 和多肽的许多分子。

[0071] 术语“核苷酸”，除了是指天然存在的核糖核苷酸或脱氧核糖核苷酸单体以外，本文中还应当理解为是指其相关的结构变体，包括衍生物和类似物，其对于使用该核苷酸的特定上下文（例如，与互补碱基杂交）在功能上是等效的，除非上下文明确另外指明。

[0072] 术语“核酸”或“多核苷酸”是指聚合物，其可以对应于核糖核酸 (RNA) 或脱氧核糖核酸 (DNA) 聚合物或其类似物。这包括核苷酸的聚合物如 RNA 和 DNA，以及其合成形式、修饰（例如化学或生物化学修饰的）形式，和混合聚合物（例如包括 RNA 和 DNA 亚单位）。示例性的修饰包括甲基化，用类似物取代一个或更多个天然存在的核苷酸，核苷酸间的修饰如不带电的键（例如膦酸甲酯、磷酸三酯、磷酸酰胺化物、氨基甲酸盐等），侧面部分（例如多肽），嵌入剂（例如吖啶、补骨脂素等），螯合剂，烷化剂，和修饰的键（例如 α 异头核酸等）。还包括模拟多核苷酸通过氢键及其他化学相互作用与指定序列结合的能力的合成分子。虽然合成形式的核酸可以包括其它键（例如，Nielsen 等 (Science 254:1497 - 1500, 1991) 描述的肽核酸），典型地核苷酸单体通过磷酸二酯键连接。核酸可以是或可以包括，例如，染色体或染色体区段、载体（例如表达载体）、表达盒、裸 DNA 或 RNA 聚合物、聚合酶链式反应 (PCR) 的产物、寡核苷酸、探针和引物。核酸可以是，例如，单链、双链、或三链的，并且不限于任何特定的长度。除非另外指出，除任何明确指出的序列以外，特定的核酸序列包含或编码互补序列。

[0073] 术语“寡核苷酸”是指包括至少 2 个核酸单体单位（例如核苷酸）的核酸。寡核苷酸典型地包括约 6- 约 175 个核酸单体单位，更典型地约 8- 约 100 个核酸单体单位，仍更典型地约 10- 约 50 个核酸单体单位（例如约 15，约 20，约 25，约 30，约 35，或更多的核酸单体单位）。寡核苷酸的确切大小取决于许多因素，包括该寡核苷酸的最终功能或用途。寡核苷酸任选地通过任何合适的方法制备，包括但不限于现有或天然序列的分离，DNA 复制或扩增，逆转录，适当序列的克隆和限制性消化，或通过例如，Narang 等的磷酸三酯法（*Meth. Enzymol.* 68:90-99, 1979）；Brown 等的磷酸二酯法（*Meth. Enzymol.* 68:109-151, 1979）；Beaucage 等的二乙基亚磷酰胺法（*Tetrahedron Lett.* 22:1859-1862, 1981）；Matteucci 等的三酯法（*J. Am. Chem. Soc.* 103:3185-3191, 1981）；自动合成法；或美国专利号 4,458,066 的固相支持法，或本领域技术人员已知的其它方法的方法直接化学合成。

[0074] 本文中使用的术语“引物”是指当置于起始多核苷酸延伸的条件下（例如，在包括在适当的缓冲液中存在所需的三磷酸核苷（在要拷贝的模板支配下）和聚合酶以及在合适的温度或温度的循环（例如在聚合酶链式反应中）的条件下）能够作为模板引导的核酸合成的起始点的多核苷酸。为了进一步说明，引物还可以用于多种其它寡核苷酸介导的合成过程，包括作为 RNA 从头合成和体外转录相关的过程（例如基于核酸序列的扩增 (NASBA)、转录介导的扩增 (TMA) 等）的引发剂。引物典型是单链寡核苷酸（例如寡脱氧核糖核苷酸）。引物的适当长度取决于该引物预期的用途，但是典型地为 6-40 个核苷酸，更典型地为 15-35 个核苷酸。短的引物分子通常需要更低的温度以与模板形成足够稳定的杂合复合物。引物不需要反映模板的精确序列，但是必须是足够互补的以便与模板杂交来发生引物延伸。在某些实施方案中，术语“引物对”表示一组引物，其包括与扩增的核酸序列的 5’ 末端的互补序列杂交的 5’ 有义引物（有时称为 ″ 正向 ″）以及与扩增的序列的 3’ 末端杂交的 3’ 反义引物（有时称为 ″ 反向 ″）（例如，如果目标序列作为 RNA 表达或者是 RNA）。如果需要，可以通过掺入用光谱学的、光化学的、生物化学的、免疫化学的或化学的方法可检测的标记来标记引物。例如，有用的标记包括 ^{32}P、荧光染料、电子致密试剂、酶（通常用于 ELISA 测定中）、生物素、或者存在抗血清或单克隆抗体的半抗原和蛋白。

[0075] 术语“5’- 核酸酶探针”是指包含至少一个发光标记部分并用于 5’- 核酸酶反应中以影响目标核酸检测的寡核苷酸。在一些实施方案中，例如，5’- 核酸酶探针仅包括单一的发光部分（例如，荧光染料等）。在某些实施方案中，5’- 核酸酶探针包括自我互补的区域，使得该探针能够在选择的条件下形成发夹结构。为了进一步说明，在一些实施方案中，5’- 核酸酶探针包括至少两个标记部分，在两个标记之一被裂解或者从寡核苷酸分离后发射增加强度的辐射。在某些实施方案中，用两种不同的荧光染料，例如，5’ 末端报告染料和 3’ 末端淬灭染料或部分标记 5’- 核酸酶探针。在一些实施方案中，在不同于或除了末端位置以外的一个或多个位置标记 5’- 核酸酶探针。当探针完整时，通常在两个荧光团之间发生能量转移，使得从报告染料的荧光发射至少部分淬灭。在聚合酶链式反应的延伸步骤期间，例如，通过，例如，Taq 聚合酶或具有这种活性的另一种聚合酶的 5’ 至 3’ 核酸酶活性裂解结合至模板核酸的 5’- 核酸酶探针，使得报告染料的荧光发射不再被淬灭。示例性的 5’- 核酸酶探针也描述在，例如，美国专利号 5,210,015；美国专利号 5,994,056；和美国专利号 6,171,785 中。在其他实施方案中，可以用两种或更多种不同的报告染料和 3’ 末端淬灭染料或部分标记 5’ 核酸酶探针。

[0076] 术语“FRET”或“荧光共振能量转移”或“Foerster 共振能量转移”是指至少两个发色团、供体发色团和受体发色团（被称为淬灭剂）之间的能量转移。当供体被具有合适波长的光辐射激发时，供体通常将能量转移至受体。受体通常以具有不同的波长的光辐射的形式重新发射转移的能量。当受体是“黑暗”淬灭剂时，它耗散除光以外的形式的转移的能量。特定的荧光团是否作为供体或受体起作用取决于 FRET 对其它成员的特性。常用的供体－受体对包括 FAM-TAMRA 对。常用的淬灭剂是 DABCYL 和 TAMRA。常用的黑暗淬灭剂包括 BlackHole Quenchers (BHQ), (Biosearch Technologies, Inc., Novato, Cal.), Iowa Black (Integrated DNA Tech., Inc., Coralville, Iowa) 和 BlackBerry Quencher 650 (BBQ-650) (Berry & Assoc., Dexter, Mich.)。

[0077] 当涉及核酸碱基、三磷酸核苷或核苷酸时，术语“常规的”或“天然的”是指描述的多核苷酸中天然存在的那些（即，对于 DNA，这些是 dATP、dGTP、dCTP 和 dTTP)。另外，在体外 DNA 合成反应如测序中，经常采用 dITP 和 7- 脱氮 -dGTP 来代替 dGTP，经常采用 7- 脱氮 -dATP 来代替 dATP。共同地，这些称为 dNTPs。

[0078] 当涉及核酸碱基、核苷或核苷酸时，术语“非常规的”或“修饰的”包括在特定的多核苷酸中天然存在的常规的碱基、核苷或核苷酸的修饰物、衍生物或类似物。与常规的 dNTPs 相比，某些非常规核苷酸在核糖的 2’ 位置修饰。因此，尽管对于 RNA，天然存在的核苷酸是核糖核苷酸（即 ATP、GTP、CTP、UTP，共称 rNTPs)，因为这些核苷酸具有在糖的 2’ 位的羟基，与之相比，dNTPs 中没有所述羟基，如本文使用的，作为 DNA 聚合酶的底物的核糖核苷酸可以是非常规的核苷酸。如本文使用的，非常规的核苷酸包括但不限于在核酸测序中用作终止剂的化合物。示例性的终止剂化合物包括但不限于具有 2’,3’ 双脱氧结构的称为双脱氧核苷三磷酸的那些化合物。双脱氧核苷三磷酸 ddATP、ddTTP、ddCTP 和 ddGTP 共同地称为 ddNTPs。终止剂化合物的其他例子包括核糖核苷酸的 2’-PO_4类似物（参见，例如美国申请公开号 2005/0037991 和 2005/0037398)。其它非常规的核苷酸包括硫代磷酸酯 dNTPs([[α]-S]dNTPs)、5’-[α]- 硼烷 (borano)-dNTPs、[α]- 膦酸甲酯 dNTPs 和核糖核苷三磷酸 (rNTPs)。可以用放射性同位素如 ^{32}P、^{33}P 或 ^{35}S；荧光标记；化学发光标记；生物发光标记；半抗原标记如生物素；或者酶标记如链霉亲和素或亲和素来标记非常规的碱基。荧光标记可以包括带负电荷的染料，如荧光素家族的染料，或者电荷中性的染料，如罗丹明家族的染料，或者带正电荷的染料，如花菁家族的染料。荧光素家族的染料包括，例如 FAM、HEX、TET、JOE、NAN 和 ZOE。罗丹明家族的染料包括 Texas Red、ROX、R110、R6G 和 TAMRA。Perkin-Elmer(Boston, MA)、Applied Biosystems(Foster City, CA) 或 Invitrogen/Molecular Probes(Eugene, OR) 销售各种染料或用 FAM、HEX、TET、JOE、NAN、ZOE、ROX、R110、R6G、Texas Red 和 TAMRA 标记的核苷酸。花菁家族的染料包括 Cy2、Cy3、Cy5 和 Cy7，由 GE Healthcare UK Limited(Amersham Place, Little Chalfont, Buckinghamshire, England) 销售。

[0079] 如本文使用的，通过在比较窗口比较两个最佳比对的序列来确定“序列同一性百分比”，其中对于两个序列的最佳比对，在比较窗口中的序列部分与参照序列（其不包括增加或缺失）相比可以包括增加或缺失（即缺口）。通过确定存在于两个序列中相同的核酸碱基或氨基酸残基的位置的数量以得到匹配位置的数量，用匹配位置的数量除以比较窗口中的位置总数并用 100 乘以该结果以得到序列同一性百分比来计算百分比。

[0080] 在两个或更多个核酸或多肽序列的上下文中的术语“相同的”或“同一性”是指相同的两个或更多个序列或子序列 (subsequences)。当如使用下列序列比较算法中的一种或通过手工比对并目视检查所测量而在比较窗口或者指定的区域比较并比对最大对应时，如果它们具有特定百分比的相同的核苷酸或氨基酸残基（例如，在特定区域，至少 20%，至少 25%，至少 30%，至少 35%，至少 40%，至少 45%，至少 50%，至少 55%，至少 60%，至少 65%，至少 70%，至少 75%，至少 80%，至少 85%，至少 90%或至少 95%同一性），序列是彼此“基本上相同的”。这些定义也是指测试序列的互补序列。任选地，同一性存在于至少约 50 个核苷酸长度的区域中，或更典型地在 100-500 个或 1000 个或更多核苷酸长度的区域中。

[0081] 在两个或更多多肽序列的上下文中，术语“相似性”或“百分比相似性”是指当如使用下列序列比较算法中的一种或通过手工比对并目视检查所测量而在比较窗口或者指定的区域比较并比对最大对应时，具有特定百分比相同的或通过保守氨基酸取代定义为相似的氨基酸残基（例如，在特定区域内 60% 相似性，任选地 65%、70%、75%、80%、85%、90% 或 95% 相似的）的两个或更多个序列或子序列。如果序列至少 20%，至少 25%，至少 30%，至少 35%，至少 40%，至少 45%，至少 50%，或者至少 55% 彼此相似，那么它们彼此是“基本上相似的”。任选地，这种相似性存在于至少约 50 个氨基酸长度的区域中，或更典型地在至少约 100-500 个或 1000 个或更多氨基酸长度的区域中。

[0082] 对于序列比较，通常以一个序列作为参照序列，用测试序列与之比较。当使用序列比较算法时，将试验和参照序列输入电脑，指定子序列坐标 (coordinate)，如果必要，指定序列算法程序参数。通常使用默认程序参数，或者可以指定另外的参数。然后序列比较算法基于程序参数计算测试序列相对于参照序列的百分比序列同一性或相似性。

[0083] 如本文使用的，“比较窗口”包括涉及选自 20-600，通常约 50- 约 200，更通常约 100- 约 150 的任一数量的连续位置的区段，其中在两个序列最佳比对后可以将序列与相同数量的连续位置的参照序列比较。用于比较的序列比对方法是本领域公知的。可以，例如，通过 Smith 和 Waterman 的局部同源性算法 (*Adv. Appl. Math.* 2:482, 1970)，通过 Needleman 和 Wunsch 的同源性比对算法 (*J. Mol. Biol.* 48:443, 1970)，通过 Pearson 和 Lipman 的搜索相似性法 (*Proc. Natl. Acad. Sci. USA* 85:2444, 1988)，通过这些算法的计算机化执行（例如，the Wisconsin Genetics Software Package 中的 GAP, BESTFIT, FASTA, 和 TFASTA, Genetics Computer Group, 575 Science Dr., Madison, Wis.），或者通过手工比对和目测检查（参见，例如，Ausubel 等, *Current Protocols in Molecular Biology* (1995 supplement)) 实施比较的最佳序列比对。

[0084] 适合用于确定百分比序列同一性和序列相似性的算法是 BLAST 和 BLAST 2.0 算法，其分别由 Altschul 等 (*Nuc. Acids Res.* 25:3389-402, 1977) 和 Altschul 等 (*J. Mol. Biol.* 215:403-10, 1990) 描述。用于执行 BLAST 分析的软件可通过国立生物技术信息中心 (http://www.ncbi.nlm.nih.gov/) 公开地获得。该算法涉及：首先通过在查询序列中鉴别长度 W 的短字段来鉴别高评分序列对 (HSPs)，其在与数据库序列中相同长度字段比对时匹配或满足一些阳性值得阈值评分 T。T 被认为是邻近字段评分阈值（上述 Altschul 等）。这些最初的邻近字段命中作为起始搜索的种子以发现包含它们的更长 HSPs。字段命中沿着各序列双向延伸，只要累计的比对评分增加。对于核苷酸序列，使用参数 M（一对匹配残

基的奖分；始终大于 0）和 N（错配残基的罚分；始终小于 0）来计算累计评分。对于氨基酸序列，使用评分矩阵来计算累计评分。字段命中在各方向的延伸停止于：累计比对评分从它最大获得值降低数量 X 时；由于一个或更多个负分残基比对的累积，累计分数达到或低于 0 时；或者到达任一序列的末端。BLAST 算法参数 W、T 和 X 决定比对的灵敏度和速度。BLASTN 程序（用于核苷酸序列）使用字长 (W)11，期望值 (E)10，M=5，N=-4 作为默认值，并比较两条链。对于氨基酸序列，BLASTP 程序使用字长 3，期望值 (E)10，和 BLOSUM62 评分矩阵（参见 Henikoff and Henikoff, ProC. Natl. Acad. Sci. USA 89:10915, 1989）比对 (B)50，期望值 (E)10，M=5，N=-4 作为默认值，并比较两条链。

[0085] BLAST 算法也执行两个序列之间相似性的统计分析（参见，例如，Karlin 和 Altschul, *ProC. Natl. Acad. Sci. USA* 90:5873-87, 1993）。BLAST 算法提供的一种相似性的量度是最小总和概率 (P((N))，其提供两个核苷酸或氨基酸序列之间匹配偶然发生的概率的指标。例如，如果在测试核酸与参照核酸的比较中最小总和概率低于约 0.2，典型地低于约 0.01，以及更典型地低于约 0.001，那么该核酸被认为与参照序列是相似的。

[0086] 术语“错配辨别力”是指当通过将一个或更多个核苷酸连接（例如共价地）于核酸而以模板依赖方式延伸核酸（例如，引物或其它寡核苷酸）时，生物催化剂（例如酶，如聚合酶、连接酶等）区分完全互补的序列与包含错配的序列的能力。术语“3’- 错配辨别力”是指当被延伸的核酸（例如，引物或其它寡核苷酸）与该核酸杂交的模板相比在核酸的 3’末端处有错配时，生物催化剂区分完全互补的序列与包含错配（接近互补）的序列的能力。在一些实施方案中，相对于完全互补的序列，被延伸的核酸在 3’末端包含错配。在一些实施方案中，相对于完全互补的序列，被延伸的核酸在倒数第二 (N-1) 3’位置和 / 在 N-2 位置包含错配。

[0087] 术语“Cp 值”或“交叉点”值是指允许定量投入的目标核酸的值。可以根据二阶导数最大法 (the second-derivative maximum method) 确定 Cp 值 (Van Luu-The, 等，“Improved real-time RT-PCR method for high-throughput measurements using second derivative calculation and double correction,” BioTechniques, Vol. 38, No. 2, February 2005, pp.287 - 293)。在二阶导数法中，Cp 对应二阶导数曲线的第一峰。此峰对应对数线性相 (log-linear phase) 的开始。二阶导数法计算实时荧光强度曲线的二阶导数值，仅获得一个值。初始 Cp 方法基于强度值的局部定义的可区分的近似值，例如，通过多项式函数。然后计算三阶导数。Cp 值是三阶导数的最小的根 (root)。可以使用拟合点法确定 Cp，其中通过在对数线性区域中平行线与阈值线的交叉确定 Cp(Van Luu-The, 等，BioTechniques, Vol. 38, No. 2, February 2005, pp.287 - 293)。任何本领域技术人员很容易进行这些计算。

[0088] 术语“PCR 效率”是指对于完美匹配的引物模板的循环至循环扩增效率的指标。使用公式：% PCR 效率 = $(10^{(-斜率)}-1)$ x 100 计算每种条件的 PCR 效率，其中通过 y- 轴上绘制的对数拷贝数和 x- 轴上绘制的 Cp 的线性回归而计算斜率。

[0089] 术语“多重”是指使用多于一组引物的扩增，或者在单一反应中多于一个多态性位点的扩增。

[0090] 附图简要说明

[0091] 图 1 显示来自各种细菌的示例性 DNA 聚合酶的聚合酶结构域的区域的氨基酸序

列比对：栖热菌种 Z05 (Z05) (SEQ ID NO:12)，水生栖热菌 (Taq) (SEQ ID NO:13)，丝状栖热菌 (Tfi) (SEQ ID NO:14)，黄栖热菌 (Tfl) (SEQ ID NO:15)，栖热菌种 Sps17 (Sps17) (SEQ ID NO:16)，嗜热栖热菌 (Tth) (SEQ ID NO:17)，*Thermus caldophilus* (Tca) (SEQ ID NO:18)，海栖热袍菌 (Tma) (SEQ ID NO:19)，那不勒斯栖热袍菌 (Tne) (SEQ ID NO:20)，非洲栖热腔菌 (Taf) (SEQ ID NO:21)，耐放射异常球菌 (Dra) (SEQ ID NO:23)，嗜热脂肪芽孢杆菌 (Bst) (SEQ ID NO:24) 和热坚芽孢杆菌 (Bca) (SEQ ID NO:25)。此外，显示的多肽区域包含氨基酸基序 G-Y-V-X_1-T-L-X_2-X_3-R-X_4-X_5-X_6-X_7-P (SEQ ID NO:26)，其可变位置在本文中进一步定义。对于每种聚合酶序列，这些基序用粗体突出显示。适于根据本发明的突变的氨基酸位置用星号 (*) 指出。

[0092] 图 2 提供了下列 DNA 聚合酶 I 间的序列同一性：栖热菌种 Z05 DNA 聚合酶 (Z05)；水生栖热菌 DNA 聚合酶 (Taq)；丝状栖热菌 DNA 聚合酶 (Tfi)；黄栖热菌 DNA 聚合酶 (Tfl)；栖热菌种 Sps17DNA 聚合酶 (Sps17)；嗜热栖热菌 DNA 聚合酶 (Tth)；*Thermus caldophilus* DNA 聚合酶 (Tca)；耐放射异常球菌 DNA 聚合酶 (Dra)；海栖热袍菌 DNA 聚合酶 (Tma)；那不勒斯栖热袍菌 DNA 聚合酶 (Tne)；非洲栖热腔菌 DNA 聚合酶 (Taf)；嗜热脂肪芽孢杆菌 DNA 聚合酶 (Bst)；和热坚芽孢杆菌 DNA 聚合酶 (Bca)。(A) 整个聚合酶 I 酶（对应于 Z05 的氨基酸 1-834）的序列同一性；和 (B) 对应 Z05 的氨基酸 420-834 的聚合酶亚结构域的序列同一性。

[0093] 发明详述

[0094] 本发明提供了改进的 DNA 聚合酶，其中聚合酶结构域中一个或多个氨基酸已经被鉴别为改进了一种或多种聚合酶活性或特征。相对于除了本文所述氨基酸差异外否则相同的聚合酶的未修饰形式，本发明的 DNA 聚合酶是具有增强的 3'- 错配辨别力活性的活性酶（即，本文所述的创造性的聚合酶延伸在引物的 3' 末端处或附近与模板错配的该引物的可能性更小）。DNA 聚合酶在涉及多核苷酸延伸或多核苷酸模板的扩增的多种应用（包括，例如在重组 DNA 研究和疾病的医疗诊断中的应用）中是有用的。

[0095] 本发明的聚合酶

[0096] 在一些实施方案中，本发明的 DNA 聚合酶的特征在于具有下列基序：

[0097] Gly-Tyr-Val-X_1-Thr-Leu-X_2-X_3-Arg-X_4-X_5-X_6-X_7-Pro（在本文中也用单字母代码称为 G-Y-V-X_1-T-L-X_2-X_3-R-X_4-X_5-X_6-X_7-P）；其中

[0098] X_1 是 Glu (E)，Arg (R) 或 Thr (T)；

[0099] X_2 是 Phe (F)，Tyr (Y) 或 Leu (L)；

[0100] X_3 是 Gly (G) 或 His (H)；

[0101] X_4 是 Arg (R) 或 Lys (K)；

[0102] X_5 是除 Arg (R) 以外的任何氨基酸；

[0103] X_6 是 Tyr (Y) 或 Asp (D)；和

[0104] X_7 是 Val (V)，Ile (I) 或 Leu (L) (SEQ ID NO:8)。

[0105] 在一些实施方案中，X_5 选自 G，A，L，M，W，P，S，T，F，Y，C，N，Q，D，E，K，V，I 或 H (SEQ ID NO:42)。

[0106] 在一些实施方案中，本发明的 DNA 聚合酶的特征在于具有下列基序（对应于栖热菌和栖热袍菌）：

[0107] Gly-Tyr-Val-X_1-Thr-Leu-Phe-Gly-Arg-X_4-X_5-X_6-X_7-Pro （在本文中也用单字母代码称为 G-Y-V-X_1-T-L-F-G-R-X_4-X_5-X_6-X_7-P）；其中

[0108] X_1是 Glu (E) 或 Arg (R)；

[0109] X_4是 Arg (R) 或 Lys (K)；

[0110] X_5是除 Arg (R) 以外的任何氨基酸；

[0111] X_6是 Tyr (Y) 或 Asp (D)；和

[0112] X_7是 Val (V) 或 Ile (I) (SEQ ID NO:9)。

[0113] 在一些实施方案中，本发明的 DNA 聚合酶的特征在于具有下列基序：

[0114] Gly-Tyr-Val-Glu-Thr-Leu-Phe-Gly-Arg-Arg-X_a-Tyr-Val-Pro （在本文中也用单字母代码称为 G-Y-V-E-T-L-F-G-R-R-X_a-Y-V-P）；其中

[0115] X_5是除 Arg (R) 以外的任何氨基酸 (SEQ ID NO:10)。

[0116] 在一些实施方案中，本发明的 DNA 聚合酶的特征在于具有下列基序：

[0117] Gly-Tyr-Val-Glu-Thr-Leu-Phe-Gly-Arg-Arg-X_5-Tyr-Val-Pro （在本文中也用单字母代码称为 G-Y-V-E-T-L-F-G-R-R-X_{a5}-Y-V-P）；其中

[0118] X_5是 Ala (A), Cys (C), Asp (D), Gly (G), 或 Asn (N) (SEQ ID NO:11)。

[0119] 这种基序存在于许多家族 A 型 DNA 依赖的 DNA 聚合酶，尤其是来自嗜热菌的热稳定 DNA 聚合酶的“手指”结构域中 (Li 等., *EMBO J.* 17:7514-7525, 1998)。例如，图 1 显示了包含对应于来自下列各种细菌的 DNA 聚合酶中上述基序的天然序列的氨基酸序列比对：大肠杆菌，热坚芽孢杆菌，嗜热脂肪芽孢杆菌，耐放射异常球菌，非洲栖热腔菌，海栖热袍菌，那不勒斯栖热袍菌，水生栖热菌，*Thermus caldophilus*，丝状栖热菌，黄栖热菌，栖热菌种 Sps17，栖热菌种 Z05 和嗜热栖热菌。如所示，SEQ ID NO:8 （除了其中 X_5是 R）的基序存在于这些聚合酶中间的每一个中，表明对于聚合酶的该区域的保守功能。图 2 提供了这些 DNA 聚合酶间的序列同一性。

[0120] 因此，在一些实施方案中，本发明提供了包含 SEQ ID NO:8,9,10,或 11（例如，其中X_5选自，基于共有序列中视情况，G, A, L, M, W, P, S, T, F, Y, C, N, Q, D, E, K, V, I 或 H）并具有本文所述的改进的活性和 / 或特征的聚合酶，其中所述 DNA 聚合酶或者是野生型或天然存在的 DNA 聚合酶，如，例如，来自上面列出的嗜热细菌的任一种的 DNA 聚合酶，或者与这种野生型或天然存在的 DNA 聚合酶基本上相同的 DNA 聚合酶。例如，在一些实施方案中，本发明的聚合酶包括 SEQ ID NO:8,9,10 或 11，与 SEQ ID NO:1, 2, 3, 4, 5, 6, 7, 36, 37, 38, 39, 40 或 41 至少 80％，85％，90％或 95％相同。在一个变型中，聚合酶的未修饰的形式来自于栖热菌种。在本发明的一些实施方案中，未修饰的聚合酶来自除栖热菌以外的嗜热菌种，例如栖热袍菌。许多热稳定 DNA 聚合酶的完整的核酸和氨基酸序列是可获得的。PCT 国际专利公开号 WO 92/06200 中已经公布了水生栖热菌 (Taq) (SEQ ID NO:2)、嗜热栖热菌 (Tth) (SEQ ID NO:6)、栖热菌种 Z05(SEQ ID NO:1)、栖热菌种 Sps17(SEQ ID NO:5)、海栖热袍菌 (Tma) (SEQ ID NO:38) 和非洲栖热腔菌 (Taf) (SEQ ID NO:37) 聚合酶中每一种的序列。Akhmetzjanov 和 Vakhitov(*Nucleic Acids Research* 20:5839, 1992) 已经公布了来自黄栖热菌的 DNA 聚合酶的序列 (SEQ ID NO:4)。来自 *Thermus caldophilus* 的热稳定 DNA 聚合酶的序列 (SEQ ID NO:7) 在 EMBL/GenBank 登录号 U62584 发现。来自丝状栖热菌的热稳定 DNA 聚合酶的序列可以使用例如美国专利号 4,889,818 提供的方法以及

表 1 中提供的序列信息从 ATCC 保藏号 42380 中重新获得。那不勒斯栖热袍菌 DNA 聚合酶的序列 (SEQ ID NO:39) 来自 GeneSeq 专利数据库登录号 R98144 和 PCT WO 97/09451。例如 Uemori 等 (*J BioChem(Tokyo)* 113(3):401-410, 1993 描述了来自热坚芽孢杆菌的热稳定 DNA 聚合酶的序列 (SEQ ID NO:41) ;也参见 Swiss-Prot 数据库登录号 Q04957 以及 GenBank 登录号 D12982 和 BAA02361)。例如,美国专利号 6,228,628; 6,346,379; 7,030,220; 6,881,559; 6,794,177; 6,468,775 ;和美国专利申请号 20040005599; 20020012970; 20060078928; 20040115639 中也描述了可以如本文所述修饰的 DNA 聚合酶的未修饰的形式的例子。序列表中还提供了代表性的全长聚合酶序列。

[0121] 在一些实施方案中,本发明的以及具有包含 SEQ ID NOS:8, 9, 10 或 11 的聚合酶结构域的聚合酶也包含核酸酶结构域 (例如,对应于 Z05 的 1-291 位)。

[0122] 在一些实施方案中,本发明的聚合酶是嵌合的聚合酶,即,包含来自两种或更多种酶的多肽区域。例如美国专利号 6,228,628 描述了这种嵌合 DNA 聚合酶的例子。特别合适的是嵌合 CS 家族 DNA 聚合酶,其包括 CS5(SEQ ID NO:29) 和 CS6(SEQ ID NO:30) 聚合酶及其与 SEQ ID NO:29 或 SEQ ID NO:30 具有基本上的序列同一性或相似性的变体 (典型地至少 80% 序列同一性,更典型地至少 90%,最典型地至少 95% 序列同一性),因此可以被修饰为包含 SEQ ID NO:8。CS5 和 CS6 DNA 聚合酶是源自栖热菌种 Z05 和海栖热袍菌 (*Tma*)DNA 聚合酶的嵌合酶。它们包括栖热菌酶的 N- 末端的 5'- 核酸酶结构域和 *Tma*酶的 C- 末端的 3'-5' 核酸外切酶以及聚合酶结构域。这些酶具有有效的逆转录酶活性,可以延伸包含核苷酸类似物的引物,并可以利用 α - 硫代磷酸酯 dNTPs、dUTP、dITP 以及荧光素和花菁染料家族标记的 dNTPs。CS5 和 CS6 聚合酶也是有效的 Mg^{2+}活化的 PCR 酶。例如,美国专利申请公开号 2004/0005599 中进一步描述了 CS5 和 CS6 嵌合聚合酶。

[0123] 在一些实施方案中,本发明的聚合酶包括 SEQ ID NO :8,9,10 或 11,并进一步包含与天然的聚合酶相比的一个或多个额外的氨基酸变化 (例如,通过氨基酸取代,增加或缺失)。在一些实施方案中,这种聚合酶保留了 SEQ ID NO :8 的氨基酸基序 (或 SEQ ID NO:9, 10 或 11 的基序),并进一步包含如下的 SEQ ID NO :27 的氨基酸基序 (对应于 Z05(SEQ ID NO :1) 的 D580X 突变) :

[0124] T-G-R-L-S-S-X_7-X_8-P-N-L-Q-N ;其中

[0125] X_7是 Ser (S) 或 Thr (T) ;和

[0126] X_8除 D 或 E 以外的任何氨基酸 (SEQ ID NO:27)。

[0127] 例如,美国专利公开号 2009/0148891 中更详细地讨论了由 SEQ ID NO :27 表征的突变。在一些实施方案中,这种功能性变体聚合酶通常与野生型或天然存在的聚合酶 (例如, SEQ ID NO:1, 2, 3, 4, 5, 6, 7, 39, 40, 41, 42, 43 或 44) 具有基本上的序列同一性或相似性,典型地至少 80%序列同一性,更典型地至少 90%,最典型地至少 95%序列同一性。

[0128] 在一些实施方案中,如 SEQ ID NO:8, 9, 10 或 11 中所述用氨基酸取代 X_5位的氨基酸,如 SEQ ID NO:27 中所述用氨基酸取代 X_8位的氨基酸。因此,在一些实施方案中, X_5位的氨基酸是除 Arg (R) 以外的任何氨基酸, X_8位的氨基酸是除 Asp (D) 或 Glu (E) 以外的任何氨基酸。在一些实施方案中,氨基酸取代包括在 SEQ ID NO:27 的 X_8位的亮氨酸 (L),甘氨酸 (G),苏氨酸 (T),谷氨酰胺 (Q),丙氨酸 (A),丝氨酸 (S),天冬酰胺 (N),精氨酸

(R) 和赖氨酸 (K)。在某些实施方案中，氨基酸取代独立地包括 X_5位的丙氨酸 (A)，半胱氨酸 (C)，天冬氨酸 (D)，甘氨酸 (G)，或天冬酰胺 (N)，以及 X_8位的甘氨酸 (G)。可以使用，例如，定点诱变的已知方法和本文中进一步描述的测定中多核苷酸延伸性能的确定方法或本领域技术人员已知的其它方法确定在一个或多个鉴别的位点的其它合适的氨基酸取代。

[0129] 因为DNA 聚合酶的精确长度变化，对应于 X_5 和 X_8的每一个的精确的氨基酸位置可以根据使用的特定聚合酶而变化。氨基酸和核酸序列比对程序是容易获得的（见，例如，上文提到的那些），考虑到本文鉴别的特定基序，所述程序用来帮助鉴别用于本发明一致的修饰的正确的氨基酸（和对应的密码子）。对于来自示例性的嗜热菌种的代表性的嵌合的热稳定 DNA 聚合酶和热稳定的 DNA 聚合酶，对应于 X_5 和 X_8中每一个的位置如表 1 中所示。

[0130] 表 1. 对应于示例性聚合酶中基序位置（例如，SEQ ID NOS:8, 9, 10 和 11 的）X_5 和 (SEQ ID NO:27 的)X_8的氨基酸位置。

[0131]

生物体或嵌合序列	氨基酸位置	
共有的 (SEQ ID NO:)	X_5	(SEQ ID NO:27) 的 X_8
嗜热栖热菌 (6)	730	580
T. caldophilus (7)	730	580
栖热菌种Z05 (1)	730	580
水生栖热菌 (2)	728	578
黄栖热菌 (4)	727	577
丝状栖热菌 (3)	726	576
栖热菌种Sps17 (5)	726	576
海栖热袍菌 (38)	791	640
那不勒斯栖热袍菌 (39)	791	640
非洲栖热腔菌 (37)	790	639
热坚芽孢杆菌 (41)	772	621
嗜热脂肪芽孢杆菌 (40)	771	620
CS5 (29)	791	640
CS6 (30)	791	640

[0132] 在一些实施方案中，本发明的 DNA 聚合酶源自栖热菌种 Z05 DNA 聚合酶 (SEQ ID NO:1) 或其变体（例如，携带 D580G 突变等）。如上面提到的，在栖热菌种 Z05 DNA 聚合酶中，X_5 位对应 730 位的精氨酸 (R)；X_8位对应 580 位的天冬氨酸 (D)。因此，在本发明的某些变型中，相对于栖热菌种 Z05 DNA 聚合酶，突变型聚合酶在 R730 和 D580 包含至少一个氨基酸取代。因此，典型地，730 位的氨基酸不是 R。在一些实施方案中，730 位的氨基酸选自 G, A, V, L, I, M, F, W, P, S, T, C, Y, N, Q, D, E, K, 或 H。在某些实施方案中，730 位的氨基酸残基是 A, C, D, G, 或 N。在某些实施方案中，D580 位的氨基酸残基可以选自亮氨酸 (L)，甘氨酸 (G)，苏氨酸 (T)，谷氨酰胺 (Q)，丙氨酸 (A)，丝氨酸 (S)，天冬酰胺 (N)，

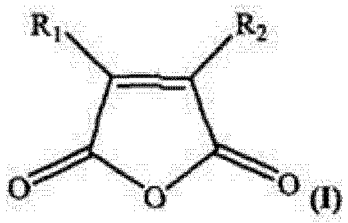
精氨酸 (R) 和赖氨酸 (K)。示例性的栖热菌种 Z05 DNA 聚合酶突变体包括包含氨基酸取代 R730A, R730C, R730D, R730G, 或 R730N 和 D580G 的那些。

[0133] 在一些实施方案中，本发明的DNA聚合酶包含对应于SEQ ID NO:1的730位的位置处的氨基酸，所述氨基酸在中性pH(例如，约pH 7.4)时，不具有极性、带正电荷的侧链(例如，R)。在一些实施方案中，本发明的DNA聚合酶源自栖热菌种，对应于SEQ ID NO:1的730位的位置处的氨基酸是在中性pH时，具有非极性、不带电荷的侧链的氨基酸(例如，G, A, L, M, W, P, F, C, V, 或I)，具有极性、不带电荷的侧链的氨基酸(例如，N, Q, H, S, T, 或Y)，具有极性、带负电荷的侧链的氨基酸(例如，D或E)。在一些实施方案中，对应于SEQ ID NO:1的730位的位置处的具有非极性、不带电荷的侧链的氨基酸是A, C, 或G。在一些实施方案中，对应于SEQ ID NO:1的730位的位置处的具有极性、不带电荷的侧链的氨基酸是N。在一些实施方案中，对应于SEQ ID NO:1的730位的位置处的具有极性、带负电荷的侧链的氨基酸是D。

[0134] 在一些实施方案中，本发明的DNA聚合酶也可以包括其它未取代的修饰。这种修饰可以包括，例如，本领域已知的共价修饰，以赋予包括引物延伸的应用中的额外的优势。例如，在某些实施方案中，突变型DNA聚合酶进一步包括热可逆的共价修饰。包括这种热可逆修饰的DNA聚合酶特别适合于热启动的应用，如，各种热启动PCR技术。例如，Birch等的美国专利号5,773,258描述了适于根据本发明的突变型DNA聚合酶应用的热可逆修饰物试剂。

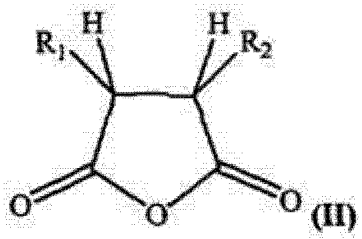
[0135] 例如，通过在碱性pH和低于约25℃的温度实施的热稳定性酶和具有下列式I或具有下述式II所述通式的二羧酸酸酐的混合物的反应而产生包含热可逆共价修饰的尤其合适的聚合酶：

[0136]



[0137] 其中R_1和R_2是氢或者可以连接有机基团；式II：

[0138]



[0139] 其中R_1和R_2是可以连接的有机基团，氢是顺式的，基本上如上面的Birch等中所述。

[0140] 可以通过突变编码对应的未修饰的聚合酶(例如，野生型聚合酶或者本发明的聚合酶来源自的对应变体)的DNA序列，如通过使用通常称为定点诱变的技术来构建本发明的DNA聚合酶。可以通过本领域普通技术人员公知的多种聚合酶链式反应(PCR)技术来

突变编码聚合酶的未修饰形式的核酸分子。(参见,例如,*PCR Strategies* (M. A. Innis, D. H. Gelfand 和 J. J. Sninsky eds., 1995, Academic Press, San Diego, CA) 的第 14 章 ; *PCR Protocols :A Guide to Methods and Applications* (M. A. Innis, D. H. Gelfand, J. J. Sninsky 和 T. J. White eds., Academic Press, NY, 1990)。

[0141] 作为非限制性例子,来自 Clontech 的转化子定点诱变试剂盒 (Transformer Site-Directed Mutagenesis kit) 使用的双引物系统可以用来将定点突变引入编码聚合酶的未修饰的形式的多核苷酸。该系统中的目标质粒变性之后,两条引物同时退火到质粒 ; 这些引物之一包含期望的定点突变,另一条包含在该质粒中引起限制性位点消除的另一个位点的突变。然后进行第二条链合成,紧密连接这两个突变,然后将获得的质粒转化入大肠杆菌的 mutS 菌株。从转化的细菌中分离质粒 DNA,用相关的限制酶限制性酶切 (从而线性化未突变的质粒),然后重新转化进入大肠杆菌。该系统允许直接在表达质粒中产生突变,而不必亚克隆或者产生单链的噬菌粒 (phagemids)。两个突变的紧密连接和随后的未突变质粒的线性化导致高突变效率并允许最少的筛选。在初始限制性位点引物的合成之后,对每个突变位点,该方法只需要使用一种新的引物类型。不是分别准备各个位点突变体,可以合成一组“设计的简并”寡核苷酸引物,用于同时在给定的位点引入所有期望的突变。可以通过对质粒 DNA 的诱变的区域测序来筛选转化体以鉴别并分类突变型克隆。然后可以限制性酶切每条突变型 DNA 并通过电泳分析,如,在突变检测增强凝胶 (Mallinckrodt Baker, Inc., Phillipsburg, NJ) 上进行以验证在该序列中没有发生其它的改变 (通过与未诱变处理的对照进行条带迁移比较)。或者可以对整个 DNA 区域测序以验证在目标区域之外没有额外的突变事件发生。

[0142] 在 pET(或者其它) 过表达载体中验证的突变型双链体可以用于转化大肠杆菌,如大肠杆菌菌株 BL21(DE3)pLysS,用于突变型蛋白的高水平生产,并用标准方案纯化。FAB-MS 作图的方法,例如,可用于快速检查突变体表达的保真度。该技术提供遍及整个蛋白的片段的测序并提供序列分配 (sequence assignment) 中必要的置信度。在这种类型的作图实验中,用蛋白酶消化蛋白 (选择将取决于待修饰的特定区域,因为该区段是主要感兴趣的而剩余的图谱应该与未诱变处理蛋白的图谱一致)。通过,例如微内径 HPLC(反相或离子交换,同样取决于待修饰的特定区域) 分级分离裂解片段的组以提供在各部分的数种肽,通过标准方法如 FAB-MS 测定肽的分子量。然后将各片段测定的质量与从预测序列的消化期望的肽的分子量比较,并快速确定序列的正确性。因为进行蛋白修饰的这种诱变方法是定向的,如果 MS 数据与预测一致,改变的肽的测序不是必要的。如果需要验证变化的残基,可以采用 CAD- 串联 MS/MS 对所讨论的混合物的肽测序,或者可以纯化目标肽用于根据修饰的位置的消减性 Edman 降解或羧肽酶 Y 消化。

[0143] 可以以多种方式产生具有多于一个氨基酸取代的突变型 DNA 聚合酶。在氨基酸紧密在一起地位于多肽链中的情况下,可以使用编码所有期望的氨基酸取代的一条寡核苷酸来同时突变它们。然而如果氨基酸彼此隔开一些距离定位 (例如隔开多于 10 个氨基酸),则产生编码所有期望的变化的单一寡核苷酸更困难。作为替代,可以用两种可选方法中的一种。在第一种方法中,对每个待取代的氨基酸产生单独的寡核苷酸。然后将寡核苷酸同时与单链模板 DNA 退火,而从模板合成的 DNA 第二条链将编码所有期望的氨基酸取代。一种可选方法涉及两个或更多轮的诱变以产生期望的突变体。第一轮如同单突变体描述的那

样：编码未修饰的聚合酶的 DNA 作为模板，将编码第一个期望的氨基酸取代的寡核苷酸与该模板退火，然后产生异源双链 DNA 分子。第二轮的诱变利用第一轮的诱变产生的突变 DNA 作为模板。因此，该模板已经包含一个或更多个突变。然后将编码额外的期望的氨基酸取代的寡核苷酸与该模板退火，产生的 DNA 链现在编码来自第一轮和第二轮的诱变两者的突变。可以使用这种产生的 DNA 作为第三轮的诱变中的模板，等等。或者，可以利用 Seyfang 和 Jin 的多位点诱变法（*Anal. BioChem.* 324:285-291. 2004）。

[0144] 因此，也提供编码本发明的 DNA 聚合酶中任一种（例如，包括 SEQ ID NOS:8, 9, 10 或 11 中任一种的聚合酶）的重组核酸。使用本发明的编码本发明的 DNA 聚合酶的核酸，可以产生多种载体。在本发明的实施中可以使用包含源自与宿主细胞相容的物种的复制子和控制序列的任何载体。通常，表达载体包括与编码突变型 DNA 聚合酶的核酸可操纵连接的转录和翻译调控核酸区域。术语“控制序列”是指在特定的宿主生物体中为可操纵连接的编码序列的表达所必需的 DNA 序列。适用于原核生物的控制序列，例如，包括启动子，任选地操纵子序列，和核糖体结合位点。此外，载体可以包含正性反向调控元件 (Positive Retroregulatory Element,PRE) 以增加转录的 mRNA 的半衰期（参见 Gelfand 等的美国专利号 4,666,848）。转录和翻译调控核酸区域通常适合于用来表达聚合酶的宿主细胞。用于各种宿主细胞的许多类型的适当的表达载体和合适的调控序列是本领域已知的。通常，转录和翻译调控序列可以包括，例如，启动子序列、核糖体结合位点、转录起始和终止序列、翻译起始和终止序列和增强子或激活子序列。在典型的实施方案中，调控序列包括启动子和转录起始和终止序列。载体还典型地包括包含数个用于外源 DNA 插入的限制性位点的多接头区域。在某些实施方案中，使用“融合标志”以帮助纯化和随后标签 / 标志序列的去除（如果需要的话），例如“组氨酸标签”（“His-Tag”）。然而，当从嗜常温宿主（例如大肠杆菌）中纯化热活性和 / 或热稳定蛋白（其中可以使用“热步骤”(heat-step)）时，这些通常是不必要的。使用标准重组 DNA 操作制备包含编码复制序列，调控序列，表型选择基因，和感兴趣的突变型聚合酶的 DNA 的合适的载体的构建。以特定的顺序将分离的质粒、病毒载体和 DNA 片段切割、裁剪并连接到一起以产生期望的载体，这是本领域公知的（参见，例如，Sambrook 等, *Molecular Cloning:A Laboratory Manual*(Cold Spring Harbor Laboratory Press, New York, NY, 2nd ed. 1989)）。

[0145] 在某些实施方案中，表达载体包含选择性标记基因以允许选择转化的宿主细胞。选择基因是本领域公知的，随使用的宿主细胞而变化。合适的选择基因可以包括，例如编码氨苄青霉素和 / 或四环素抗性的基因，其使以这些载体转化的细胞在这些抗生素存在的情况下生长。

[0146] 在本发明的一个方面，将编码本发明的 DNA 聚合酶的核酸以单独或与载体组合地引入细胞。此处“引入”或语法上的等价物表示核酸以一种适于随后的核酸的整合、扩增和 / 或表达的方式进入细胞。引入的方法主要由目标细胞类型支配。示例性方法包括 $CaPO_4$ 沉淀，脂质体融合，LIPOFECTIN®，电穿孔，病毒感染等。

[0147] 在一些实施方案中，使用原核生物作为用于本发明初始的克隆步骤的宿主细胞。它们对于大量 DNA 的迅速生产，对于用于定点诱变的单链 DNA 模板的生产，对于同时筛选许多突变体以及对于产生的突变体的 DNA 测序特别有用。合适的原核宿主细胞包括大肠杆菌 K12 菌株 94(ATCC No. 31,446)、大肠杆菌菌株 W3110(ATCC No. 27,325)、大肠杆菌

K12 菌株 DG116(ATCC No. 53,606)、大肠杆菌 X1776(ATCC No. 31,537) 和大肠杆菌 B ;然而大肠杆菌的许多其它菌株如 HB101、JM101、NM522、NM538、NM539,以及许多其它原核生物的种和属包括杆菌如枯草芽胞杆菌 (*Bacillus subtilis*)、其它肠杆菌科如鼠伤寒沙门氏杆菌 (*Salmonella typhimurium*) 或粘质沙雷氏菌 (*Serratia marcesans*) 和各种假单胞菌属 (*Pseudomonas* species) 的菌种全部都可以用作宿主。典型地使用如上文的 Sambrook 等在 1.82 部分描述的氯化钙方法转化原核宿主细胞或其它具有刚性细胞壁的宿主细胞。或者,可以使用电穿孔用于这些细胞的转化。例如,Dower, 在 *Genetic Engineering, Principles and Methods* 12:275-296(Plenum Publishing Corp., 1990) ;Hanahan 等, *Meth. Enzymol.*, 204:63, 1991 中描述了原核生物转化技术。典型用于大肠杆菌转化的质粒包括 pBR322、pUCI8、pUCI9、pUCI18、pUC119 和 Bluescript M13,它们所有都描述于上文的 Sambrook 等的 1.12-1.20 部分。然而,许多其它合适的载体也是可用的。

[0148] 在一些实施方案中,通过在诱导或引起突变型 DNA 聚合酶表达的适当条件下培养用包含编码 DNA 聚合酶的核酸的表达载体转化的宿主细胞而产生本发明的 DNA 聚合酶。在适于蛋白表达的条件下培养转化的宿主细胞的方法是本领域公知的 (参见,例如,上文的 Sambrook 等)。用于从包含 λ pL 启动子的质粒载体产生聚合酶的合适的宿主细胞包括大肠杆菌菌株 DG116(ATCC No. 53606)(参见美国专利号 5,079,352 和 Lawyer, F.C. 等, *PCR Methods and Applications* 2:275-87, 1993)。表达之后,可以收获并分离聚合酶。例如,上文的 Lawyer 等中描述了用于纯化热稳定 DNA 聚合酶的方法。

[0149] 一旦纯化,可以测定 DNA 聚合酶的 3’ 错配辨别力。例如,在一些实施方案中,通过比较与引物完美匹配的目标序列的扩增和在引物的 3’ 末端具有单个碱基错配的目标的扩增而测定 3’ 错配辨别力。可以,例如,通过使用 TaqManCps 而评价聚合酶在两个目标序列之间区分的能力。任选地,可以在作为对照的第二个光学通道中同时实施并检测每个孔中第二个目标基因的扩增。“ΔCp 值”是指与错配的模板相关的 Cp 减去匹配的目标的 Cp 之间的值的差异 (参见,例如,实施例)。在一些实施方案中,与在 SEQ ID NO:8 的 X_5 位具有天然的氨基酸 (例如, R) 的另外相同的对照聚合酶相比,本发明的改进的聚合酶具有至少 1,2,3,4,5 或更大的 ΔCp 值。在一些实施方案中,以实施例中所述的准确的材料和条件进行这种测定。

[0150] 本发明的方法

[0151] 本发明的改进的 DNA 聚合酶可以用于其中需要或期望这种酶活性的任何目的。改进的 DNA 聚合酶可以是如本文所述的热活性或热稳定的 DNA 聚合酶。因此,在本发明的一个方面,提供了使用本发明的聚合酶的多核苷酸延伸方法,包括 PCR。在一些实施方案中,本发明提供了热活性的 DNA 聚合酶,当不需要扩增模板核酸时,例如,当需要立即检测目标核酸的存在时,所述热活性的 DNA 聚合酶可用于延伸 RNA 或 DNA 模板。在一些实施方案中,本发明提供了热稳定的 DNA 聚合酶,当需要延伸和 / 或扩增目标核酸时,所述热稳定的 DNA 聚合酶是有用的。适合用于多核苷酸延伸的条件是本领域公知的 (参见,例如,上文的 Sambrook 等。也参见 Ausubel 等, *Short Protocols in Molecular Biology*(4th ed., John Wiley & Sons 1999))。通常,使引物与目标核酸退火,即杂交以形成引物 - 模板复合物。在合适的环境中将引物 - 模板复合物与突变型 DNA 聚合酶和三磷酸核苷接触,以允许一个或更多个核苷酸增加至引物的 3’- 末端,从而产生与目标核酸互补的延伸的引物。引

物可以包括，例如，一个或更多个核苷酸类似物。此外，三磷酸核苷可以是常规核苷酸、非常规核苷酸（例如核糖核苷酸或标记的核苷酸）或其混合物。在一些变型中，多核苷酸延伸反应包括目标核酸的扩增。适于使用 DNA 聚合酶和引物对进行核酸扩增的条件也是本领域已知的（例如，PCR 扩增方法）。（参见，例如，上文的 Sambrook 等；上文的 Ausubel 等；PCR *Applications: Protocols for Functional Genomics*(Innis 等 eds., Academic Press 1999)。

[0152] 在一些实施方案中，提供增强的 3’ 错配辨别力的本发明的聚合酶的使用允许，例如，稀有等位基因的检测。例如，特定聚合酶的 3’ 错配辨别力的保真度设定了它的灵敏度（在更大数量的不同但相关的非目标序列存在的情况下准确地检测少量目标序列的能力）。因此，增强的 3’- 错配辨别力导致稀有等位基因的检测的更大灵敏度。例如，当筛选活检样品或其它样品的稀有的遗传变化，例如，在大量正常细胞中的携带癌等位基因的细胞时，稀有等位基因的检测是有用的。

[0153] 在一些实施方案中，改进的聚合酶用于等位基因特异性 PCR 或单核苷酸多态性 (SNP) 检测的背景中的多核苷酸延伸。在 Chen 等，"Single nucleotide polymorphism genotyping:biochemistry, protocol, cost and throughput" Pharmacogenomics J. 3(2):77-96 (2003)；Kwok 等，"Detection of single nucleotide polymorphisms" Curr.Issues Mol. Biol. 5(2):43-60 (April 2003)；Shi, "Technologies for individual genotyping:detection of genetic polymorphisms in drug targets and disease genes" Am. J. Pharmacogenomics 2(3):197-205 (2002)；和 Kwok, "Methods for genotyping single nucleotide polymorphisms" Annu. Rev. Genomics Hum. Genet. 2:235-58 (2001) 中描述了示例性的 SNP 检测方法。Marnellos, "High-throughput SNP analysis for genetic association studies" Curr. Opin. Drug Discov. Devel. 6(3):317-21 (May 2003) 中描述了用于高通量 SNP 检测的示例性技术。通常的 SNP 检测方法包括但不限于，TaqMan 测定，分子信标测定，核酸阵列，等位基因特异性引物延伸，等位基因特异性 PCR，阵列的引物延伸，同质引物延伸测定 (homogeneous primer extension assays)，具有质谱法检测的引物延伸 (primer extension with detection by mass spectrometry)，焦磷酸测序，在基因阵列上分选的多重引物延伸，具有滚环扩增的连接，同质连接 (homogeneous ligation)，OLA（美国专利号 4,988,167），在基因阵列上分选的多重连接反应，限制性片段长度多态性，单碱基延伸标签测定和侵入测定 (Invader assay)。这种方法也可以组合使用检测机理如，发光或化学发光检测，荧光检测，时间分辨的荧光检测，荧光共振能量转移，荧光偏振，质谱法和电检测。

[0154] 也可以使用多重反应（其允许在单个反应中检测多个不同的等位基因）完成多个不同的等位基因的检测。在多重反应中，使用两个或更多的等位基因特异性引物来延伸和扩增 SNP 或多核苷酸多态性或等位基因。美国专利公开号 2006/0172324 中描述了单核苷酸多态性和多核苷酸多态性的多重检测的示例性方法。

[0155] 使用本文所述的改进的聚合酶检测延伸产物或扩增产物的其它方法包括使用荧光双链核苷酸结合染料或荧光双链核苷酸插入染料。荧光双链 DNA 结合染料的例子包括 SYBR-green (Molecular Probes)。荧光双链插入染料的例子包括溴化乙锭。可以连同熔解曲线分析使用双链 DNA 结合染料以测定引物延伸产物和 / 或扩增产物。可以在实时 PCR

仪器，如具有机载软件（SDS 2.1）的ABI 5700/7000(96孔格式）或ABI 7900(384孔格式）仪器上进行熔解曲线分析。或者，可以进行熔解曲线分析作为终点分析。美国专利公开号2006/0172324中描述了熔点分析的示例性方法。

[0156] 在其它实施方案中，本发明的聚合酶用于在DNA测序、DNA标记或引物延伸产物标记的背景中的引物延伸。例如，对于能够掺入非常规的链终止核苷酸的改进的聚合酶，本发明改进了用Sanger双脱氧核苷酸法(Sanger 等, *ProC. Natl. Acad. Sci. USA* 74:5463, 1977)的DNA测序。基础的Sanger等方法的发展提供了新的载体(Yanisch-Perron 等, *Gene* 33:103-119, 1985)和碱基类似物(Mills 等, *ProC. Natl. Acad. Sci. USA* 76:2232-2235, 1979；和Barr 等, *Biotechniques* 4:428-432, 1986)。通常，DNA测序需要在链终止碱基类似物存在的情况下的模板依赖的引物延伸，产生随后按大小分离的部分片段的分布。基础的双脱氧测序操作涉及（i）将任选地标记的寡核苷酸引物与模板退火；(ii)在4个分开的反应中用DNA聚合酶延伸引物，每个反应包含未标记的dNTPs和限制数量的一种任选地被标记的链终止剂如ddNTP的混合物；以及（iii）在高分辨率的变性聚丙烯酰胺/尿素凝胶上分离4组反应产物。根据使用的标记，可以通过放射自显影或通过荧光检测在凝胶中检测反应产物，可以检查图像以推断核苷酸序列。这些方法利用DNA聚合酶如大肠杆菌Pol I的Klenow片段或修饰的T7 DNA聚合酶。

[0157] 热稳定聚合酶（如Taq DNA聚合酶）的可用性，导致了用热稳定DNA聚合酶的测序的改进的方法（参见Innis等, *ProC. Natl. Acad. Sci. USA* 85:9436, 1988)及其被称为“循环测序”的修改形式(Murray, *Nuc Acids Res.* 17:8889, 1989)。因此，可以连同这种方法使用本发明的热稳定聚合酶。作为基础的双脱氧法测序的替代，循环测序是在链终止剂存在的情况下与模板序列互补的目标序列的线性的、不对称的扩增。单个循环产生所有可能长度的延伸产物的家族。从DNA模板的延伸反应产物的变性之后，在终止剂如ddNTPs存在的情况下发生多循环的引物退火和引物延伸。循环测序需要比常规的链终止测序更少的模板DNA。热稳定DNA聚合酶在循环测序中具有若干优势；它们耐受引物与核酸目标的特异性杂交所需的严谨的退火温度以及耐受多循环的在每个循环中发生的高温变性，例如90-95℃。为此，在公司如Perkin-Elmer(Norwalk, CT)和Applied Biosystems(Foster City, CA)商业化的Taq循环测序试剂盒中已经包括AMPLITAQ® DNA聚合酶及其衍生物和后代产品，例如AmpliTaq CS DNA聚合酶和AmpliTaq FS DNA聚合酶。

[0158] 改进的聚合酶用于454测序(Roche) (Margulies, M等.2005, Nature, 437, 376-380)中。454测序涉及两个步骤。在第一个步骤中，将DNA剪切为约300-800碱基对的片段，使片段末端钝化。然后将寡核苷酸适配子连接至片段的末端。适配子作为片段的扩增和测序的引物。可以将片段连接至DNA捕获珠粒，例如，使用，例如，包含5’-生物素标记的适配子B的链霉亲和素包被的珠粒。在油-水乳液的液滴内PCR扩增连接至珠粒的片段。结果是在每个珠粒上多个拷贝的克隆扩增的DNA片段。在第二个步骤中，在孔（皮升大小）中捕获珠粒。平行地对每个DNA片段进行焦磷酸测序。一个或多个核苷酸的增加产生光信号，通过测序仪中的CCD相机记录该光信号。信号强度与掺入的核苷酸的数量成比例。

[0159] 焦磷酸测序利用在核苷酸增加后释放的焦磷酸(PPi)。在腺苷5’磷酸硫酸(adenosine 5’ phosphosulfate)存在的情况下，ATP硫酸化酶将PPi转化为ATP。荧光素

酶使用 ATP 将荧光素转化为氧化荧光素，该反应产生光，检测并分析所述光。

[0160] 链终止测序法的变型包括染料 - 引物测序和染料 - 终止剂测序。在染料 - 引物测序中，ddNTP 终止剂是未标记的，利用标记的引物来检测延伸产物 (Smith 等, *Nature* 32:674-679, 1986)。在染料 - 终止剂 DNA 测序中，使用 DNA 聚合酶将 dNTPs 和荧光标记的 ddNTPs 掺入 DNA 引物的末端上 (Lee 等, *Nuc. Acids. Res.* 20:2471, 1992)。该方法提供了不必合成染料标记的引物的优势。而且，染料 - 终止剂反应更方便，因为可以在相同管中进行所有 4 个反应。

[0161] 使用 Applied Biosystems(Foster City, CA) 生产的自动测序仪（美国专利号 5,171,534) 可以使染料 - 引物和染料 - 终止剂法自动化。当使用该仪器时，在变性聚丙烯酰胺凝胶或该仪器中安装的毛细管上分级分离完成的测序反应混合物。当荧光产物根据大小电泳分离通过凝胶时，在仪器底部的激光检测所述荧光产物。

[0162] 通常使用两种类型的荧光染料来标记用于染料 - 终止剂测序的终止剂 - 带负电荷的和两性离子的荧光染料。带负电荷的荧光染料包括荧光素和 BODIPY 家族的那些。国际专利公布 WO 97/00967 描述了 BODIPY 染料 (4,4- 二氟 -4- 硼 -3a,4a- 二氮杂 -s-indacene)。两性离子荧光染料包括罗丹明家族的那些。商业上可获得的循环测序试剂盒使用罗丹明衍生物标记的终止剂。然而，罗丹明标记的终止剂是相当昂贵的并且产物在上样至凝胶之前必须与未掺入的染料 -ddNTPs 分离，因为它们与测序产物共同迁移。罗丹明染料家族终止剂似乎使富含 GC 的区域中的发夹结构稳定，这引起产物反常迁移。这可以涉及使用 dITP，其使二级结构松散并且还影响终止剂掺入的效率。

[0163] 相反，荧光素标记的终止剂除去了在凝胶上样前的分离步骤，因为它们具有更大的净负电荷并且比测序产物迁移得更快。此外，荧光素标记的测序产物比罗丹明标记的测序产物具有更好的电泳迁移。尽管野生型 Taq DNA 聚合酶不能有效地掺入荧光素家族染料标记的终止剂，现在可以通过利用如美国专利申请公开号 2002/0142333 描述的修饰的酶有效地完成这一点。因此，可以在本发明的背景中使用如 US2002/0142333 描述的修饰以产生具有改进的引物延伸速率的荧光素家族染料掺入性热稳定聚合酶。例如，在某些实施方案中，本发明的未修饰的 DNA 聚合酶是如 US2002/0142333 所述的修饰的热稳定聚合酶，其具有 SEQ ID NO:8 中所述的基序（或 SEQ ID NO:9, 10 或 11 的基序）以及任选的 SEQ ID NO:27 的基序。

[0164] 其中可以使用本发明的突变型 DNA 聚合酶的其它示例性的核酸测序形式包括涉及包括核糖核苷酸的 2'-PO_4类似物的终止剂化合物的那些形式（参见，例如，美国申请公开号 2005/0037991 和 2005/0037398，和美国专利申请号 12/174,488)。

[0165] 试剂盒

[0166] 在本发明的另一个方面，提供了用于本文所述的引物延伸方法的试剂盒。在一些实施方案中，为了使用的便利将试剂盒区室化，该试剂盒包含至少一个提供本发明的 DNA 聚合酶的容器，所述 DNA 聚合酶具有本发明的增强的 3' 错配辨别力。还可以包括一个或更多个提供额外的试剂的额外的容器。这种额外的容器可以包括本领域技术人员所知的用于根据上述方法的引物延伸操作中的任何试剂或其它元件，包括用于，例如，核酸扩增操作 (例如 PCR、RT-PCR)，DNA 测序操作或 DNA 标记操作中的试剂。例如，在某些实施方案中，试剂盒进一步包括容器，所述容器提供在引物延伸条件下可与预定的多核苷酸模板杂交的 5'

有义引物，或包括 5’ 有义引物和对应的 3’ 反义引物的引物对。在一些实施方案中，试剂盒包括一个或更多个容器，所述容器包含与单核苷酸多态性或多核苷酸多态性完全互补的一种或更多种引物，其中所述引物可用于如上所述的多重反应。在其它非互相排斥的变型中，试剂盒包括一个或更多个提供三磷酸核苷（常规的和 / 或非常规的）的容器。在特定实施方案中，试剂盒包括 α - 硫代磷酸酯（α -phophorothioate)dNTPs、dUTP、dITP，和 / 或标记的 dNTPs，如荧光素 - 或花菁 - 染料家族 dNTPs。在其它非互相排斥的实施方案中，试剂盒包括一个或更多个容器，所述容器提供适于引物延伸反应的缓冲剂。在一些实施方案中，试剂盒包括一个或更多个标记的或未标记的探针。探针的例子包括双重标记的 FRET（荧光共振能量转移）探针和分子信标探针。在另一个实施方案中，试剂盒中包含适配子，例如，用于热启动 PCR 测定的适配子。

[0167] 反应混合物

[0168] 在本发明的另一个方面，提供了包含本文所述的具有增强的 3’- 错配辨别力活性的聚合酶的反应混合物。反应混合物可以进一步包括用于，例如，核酸扩增操作（例如，PCR，RT-PCR），DNA 测序操作，或 DNA 标记操作中的试剂。例如，在某些实施方案中，反应混合物包括适合用于引物延伸反应的缓冲剂。反应混合物也可以包含模板核酸 (DNA 和 / 或 RNA)，一种或多种引物或探针多核苷酸，三磷酸核苷（包括，例如，脱氧核糖核苷酸，核糖核苷酸，标记的核苷酸，非常规的核苷酸），盐（例如，Mn^{2+}, Mg^{2+}）和标记（例如，荧光团）。在一些实施方案中，反应混合物进一步包括双链 DNA 结合染料，如 SYBR 绿，或双链 DNA 嵌入性染料，如溴化乙锭。在一些实施方案中，反应混合物包含在引物延伸条件下可与预定的多核苷酸模板杂交的 5’- 有义引物，或包含所述 5’- 有义引物和对应的 3’- 反义引物的引物对。在某些实施方案中，反应混合物进一步包括用于检测扩增的模板核酸的发荧光的 FRET 水解探针，例如，Taqman® 探针。在一些实施方案中，反应混合物包含两种或更多种与单核苷酸多态性或多核苷酸多态性完全互补的引物。在一些实施方案中，反应混合物包含 α - 硫代磷酸酯（α -phophorothioate)dNTPs、dUTP、dITP，和 / 或标记的 dNTPs，如荧光素 - 或花菁 - 染料家族 dNTPs。

实施例

[0169] 下列实施例用于说明，但不是限制要求保护的本发明。

[0170] 实施例 1 ：具有增强的 3’- 错配辨别力的突变型 DNA 聚合酶的鉴别

[0171] 本实施例显示，在 Z05 DNA 聚合酶的 R730 位的各种氨基酸取代导致了增强的 3’- 错配辨别力。

[0172] 本实施例的对照 DNA 聚合酶是除了 580 位的氨基酸是甘氨酸（例如，D580G 取代）的 SEQ ID NO:1 的栖热菌种 Z05 DNA 聚合酶（以下被称为 Z05 D580G 聚合酶）。

[0173] 鉴别了在 Z05 D580G 聚合酶中的突变，该酶提供了延伸具有与模板 3’- 错配的寡核苷酸引物的降低的能力。简而言之，该筛选过程中的步骤包括突变型酶的文库生成、表达和部分纯化，筛选酶的期望的性质，DNA 测序，克隆纯化，和进一步表征选择的候选突变体。下面进一步描述这些步骤中的每一个。

[0174] 克隆文库生成 ：编码 Z05 D580G DNA 聚合酶的聚合酶结构域的核酸经受在包括该核酸序列的质粒的 Blp I 和 Bgl II 限制性位点之间的易错（诱变）PCR。扩增的序列如 SEQ

ID NO:33 所提供。下面给出了用于这个的引物：

[0175] 正向引物:5'- CTACCTCCTGGACCCCTCCAA-3' (SEQ ID NO:31); 和,

[0176] 反向引物:5'- ATAACCAACTGGTAGTGGCGTGTAA-3' (SEQ ID NO:32)。

[0177] 使用 1.8-3.6mM 的 Mg^{2+}浓度范围进行 PCR,以便产生具有突变率范围的文库。缓冲剂条件为 50 mM Bicine pH 8.2, 115 mM KOAc, 8% w/v 甘油和 0.2 mM 每种 dNTPs。在 0.15 U/ μL 使用 GeneAmp® AccuRT 热启动 PCR 酶。以每 50 μL 反应体积 $5x10^5$拷贝的线性化 Z05 D580G 质粒 DNA 起始,用 60 秒 94℃温度使反应变性,然后使用 15 秒 94℃的变性温度,15 秒 60℃的退火温度,120 秒 72℃的延伸温度以及随后 5 分钟在 72℃温度的最终延伸进行 30 循环的扩增。

[0178] 用QIAquick PCR纯化试剂盒(Qiagen, Inc., Valencia, CA, USA)纯化产生的扩增子,用 Blp I 和 Bgl II 切割,然后用 QIAquick PCR 纯化试剂盒重新纯化。通过用相同的两种限制性酶切割,用重组的碱性磷酸酶 (RAS,目录# 03359123001) 处理,并用 QIAquick PCR纯化试剂盒纯化而制备 Z05 D580G 载体质粒。将切割的载体和突变的插入片段以 1:3 的比例混合并用 T4 DNA 连接酶在室温处理 5 分钟 (NEB Quick Ligation Kit)。用 QIAquick PCR 纯化试剂盒纯化连接产物并通过电穿孔法将其转化入大肠杆菌宿主菌株。

[0179] 将表达的培养物的等分样品涂覆在氨苄青霉素选择性培养基上以测定在每个转化中独特转化体的数量。在作为冷冻保护剂的甘油存在的情况下,将转化物存储于 -70℃到 -80℃。

[0180] 然后将每个文库涂布在大号氨苄青霉素选择性琼脂平板上。用自动菌落挑选机(QPix2, Genetix Ltd) 将单菌落转移至包含具有氨苄青霉素和 10% w/v 甘油的 2X Luria 液体培养基的 384 孔板。在 30℃将这些平板孵育过夜以允许培养物生长,然后存储于 -70℃到 -80℃。加入 2X Luria 液体培养基的甘油量足够低以允许培养物生长,然而足够高以提供冷冻保护。以这种方式制备了若干诱变 (Mg^{2+}) 水平的数千个菌落用于随后使用。

[0181] 提取物文库制备部分 1—发酵:从上述克隆文库制备适于筛选目的的部分纯化提取物的对应文库。该过程的第一步是制备每个克隆的小规模表达培养物。这些培养物以 96 孔形式生长;因此对于每块 384 孔文库平板有 4 块表达培养平板。从克隆文库平板的每个孔将 0.5 μL 转移至包含 150 μL 培养基 A 的 96 孔种子平板的一个孔(参见下面的表 3)。这种种子平板在 iEMS 平板孵育器 / 振荡器 (ThermoElectron) 中于 30℃下在 1150 rpm 振荡过夜。然后使用这些种子培养物接种相同的培养基,这次将 20 μL 接种进大号 96 孔板 (Nunc # 267334) 中的 250 μL 培养基 A 中。这些平板于 37℃振荡孵育过夜。表达质粒包含转录控制元件,其允许在 37℃表达而不能在 30℃表达。过夜孵育后,培养物典型地以总细胞蛋白的 1-10% 表达克隆蛋白。通过离心从这些培养物中收获细胞。这些细胞或冻存 (-20℃) 或如下所述立即处理。

[0182] 表 2. 培养基 A(使用前过滤灭菌)

组分	浓度
$MgSO_4 \cdot 7H_2O$	0.2 g/L
柠檬酸 $\cdot H_2O$	2 g/L
K_2HPO_4	10 g/L
$NaNH_4PO_4 \cdot 4H_2O$	3.5 g/L
$MgSO_4$	2 mM
酪蛋白氨基酸	2.5 g/L

葡萄糖	2 g/L
硫胺素·HCl	10 mg/L
氨苄青霉素	100 mg/L

[0183] 提取物文库制备部分2—提取：将来自发酵步骤的细胞团重悬浮在25 μL裂解缓冲液（下表3）中，转移至384孔热循环仪平板并密封。注意该缓冲液包含溶菌酶以帮助细胞裂解，并包含DNase以便从提取物中除去DNA。为了裂解细胞，将板在37℃孵育15分钟，在-20℃冷冻过夜并再次在37℃下孵育15分钟。添加硫酸铵（1.5 μL的2M溶液）并在75℃孵育平板15分钟以沉淀并失活污染蛋白，包括外源添加的核酸酶。将平板在4℃ 3000 x g离心15分钟，然后将上清液转移至新的384孔热循环仪平板。在-20℃冻存这些提取物平板用于随后的筛选中。每孔包含约0.5-3 μM的突变型文库聚合酶。

[0184] 表3. 裂解缓冲液

组分	浓度或百分比
Tris pH 7.5	50 mM
EDTA	1 mM
$MgCl_2$	6 mM
Tween 20	0.5% v/v
溶菌酶（来自粉末）	1 mg/mL
DNase I	0.05单位/μL

[0185] 筛选提取物文库的降低的3’引物错配延伸率：通过比较与寡核苷酸模板完美匹配的引物的延伸率比具有3’G:T错配的引物的延伸率而筛选提取物文库。

[0186] 通过在384孔热循环仪平板中将2.5μl提取物与22.5μl包含20 mM Tris-HCl, pH 8, 100 mM KCl, 0.1 mM EDTA和0.2% Tween-20的缓冲液混合，覆盖并在90℃加热10分钟而将上述酶提取物稀释10倍用于引物延伸反应中。具有完美匹配引物的对照反应在384孔PCR平板中将0.5 μl稀释的提取物与15μl母液混合物混合。在改进的动态热循环仪中用CCD照相机每10秒监测已经引发的模板的延伸（参见上面的Watson）。母液混合物包含50 nM已经引发的引物模板，25 mM Tricine, pH 8.3, 100 mM KOAc, 0.6X SYBR Green I, 200 μM每种dNTP, 100 nM适配子和2.5 mM醋酸镁。为了从背景荧光中区分延伸来源的荧光，实验中包括平行孔，其中通过从反应母液混合物中省略(leaving out)核苷酸而阻止引物链延伸。如上所述进行3’-错配引物的反应，除了将1.5μl稀释的提取物加入每个反应以及1.5mM醋酸锰取代醋酸镁。将提取物的量增加至3倍并使用锰作为金属活化剂既使得错配延伸的可能性更大，而且因此改进了筛选具有针对3’-错配延伸辨别的最大能力的那些酶的选择性。

[0187] 使用上述方案筛选约5000突变体提取物。基于在假定截止值以上的完美匹配的引物延伸值和低错配与完美匹配延伸比值而选择初始合并物的约7%用于再次筛选。将对应最佳提取物的培养孔取样至新鲜生长培养基中并再次生长以产生包含最佳突变体的新的培养板以及用于比较的许多亲本培养物。然后使用这些培养板制备新鲜提取物，再次筛选该新鲜提取物以验证初始筛选表型。将具有完美的3’-匹配的引物和3’-错配的引物的反应的引物延伸率计算为荧光随时间的上升的曲线的线性部分的斜率。使用错配延伸斜率除以完美匹配的延伸斜率的比值来排序并选择最佳候选。下表中包括了来自再次筛选选择的克隆以及用于比较的亲本克隆Z05 D580G的各自的基因型和表型。

[0188] 表4

酶	完美匹配斜率	错配斜率	MM 斜率 / PM 斜率
Z05 D580G	8.29	8.04	0.97
Z05 D580G R730C	14.50	1.18	0.08

[0189] *实施例 2：* 在野生型 BRAF 人基因组模板的背景中扩增突变型 BRAF 质粒模板

[0190] 检查了在 Z05 DNA 聚合酶的 R730 位的各种取代对等位基因特异性 PCR 中的错配辨别力的影响。通过将合成的基因片段克隆进质粒载体中而在 Z05 D580G DNA 聚合酶中都产生这些取代，纯化并定量表达的突变型酶。在等位基因特异性 PCR 测定中将 Z05 D580G 突变体 A（丙氨酸），D （天冬氨酸），G（甘氨酸），和 N （天冬酰胺）与它们亲本酶进行比较。

[0191] 本实施例的对照 DNA 聚合酶是除了 580 位的氨基酸是甘氨酸（例如，D580G 取代）的 SEQ ID NO:1 的栖热菌种 Z05 DNA 聚合酶（以下被称为 Z05 D580G 聚合酶）。

[0192] 使用的引物扩增人 BRAF 基因的区域，当所述目标携带 BRAF 的密码子 600 的突变 V600K 时，所述引物与目标完美匹配。针对人基因组 DNA 中存在的野生型 BRAF 目标，等位基因选择性引物导致在 3’ 末端的单一的 A:C 错配。通用引物与 BRAF 基因完美匹配，这是允许扩增的实时 TaqMan 检测的探针序列。每个反应在 16μl 的最终体积中具有 10,000 拷贝 (33 ng) 的野生型人基因组细胞系 DNA，或 10,000 或 100 拷贝的包含 BRAF V600R 突变序列的线性化质粒。为了允许酶的不同盐条件最佳化，使用 25-130 mM 的 KCl 浓度范围进行扩增。缓冲液条件为 50 mM Tris-HCl pH 8.0, 2.5 mM $MgCl_2$, 0.2 mM 每种 dNTP, 0.02 U/μl UNG 和 200 nM 适配子。正向引物和反向引物为 100 nM，探针为 25 nM。在 20 nM 测定所有 DNA 聚合酶，将 2 %(v/v) 酶储存缓冲液 (50% v/v 甘油 , 100 mM KCl, 20 mM Tris pH 8.0, 0.1 mM EDTA, 1 mM DTT, 0.5% Tween 20) 加入反应。在 Roche LightCycler 480 热循环仪中进行反应并使用 95℃温度 60 秒进行变性，然后使用 92℃的变性温度 10 秒钟和 62℃的退火温度 30 秒进行 99 个循环的扩增。

[0193] 反应一式两份进行，通过 Abs Quant/ 二阶导数最大法 (2^{nd} derivative Max method) 计算交叉点（“Cps”）并将 Cps 取平均值。从 100 和 10,000 拷贝完美匹配质粒反应的斜率计算 PCR 效率，所述反应的 KCl 浓度导致最早的 10,000 拷贝完美匹配质粒 Cp。高拷贝 ΔCp 等于 10,000 拷贝的 3’- 错配的野生型基因组目标的反应的 Cp 值与 10,000 拷贝的完美匹配的质粒目标的反应的 Cp 值之间的差异。

[0194] 下表包含在对于每种酶导致最早的高拷贝质粒 Cp 的 KCl 浓度时的平均 Cp 值和计算的 PCR 效率以及高拷贝 ΔCp。包括如上所述在初始突变体筛选中鉴别的 Z05 D580G R730C DNA 聚合酶用于参照。这些数据表明，Z05 DNA 聚合酶的 R730 位的若干氨基酸取代导致在等位基因选择性 PCR 中的引物错配的改进的辨别力。

[0195] 表 5. BRAF V600K 突变型质粒比人基因组 DNA 的扩增的 Cps

[0196]

酶	100 拷贝 突变型 质粒	10,000 拷贝 突变型 质粒	10,000 拷贝人 gDNA	最佳 KCl	PCR 效率	高拷贝 Δ Cp
Z05 D580G*	31.2	24.3	26.3	115	96.3	2.0
Z05 D580G R730A	35.2	27.8	36.1	70	85.6	8.3
Z05 D580G R730C	35.9	28.2	37.2	70	81.4	9.0
Z05 D580G R730D	43.7	35.0	43.6	40	69.7	8.6
Z05 D580G R730G	38.1	30.1	39.9	55	78.2	9.8
Z05 D580G R730N	35.6	28.1	37.7	70	84.7	9.6

[0197] * 4 次实验的平均值。

[0198] 本实施例表明，相对于对照亲本聚合酶 Z05 D580G，R730A, R730C, R730D, R730G, 和 R730N 突变型酶具有改进的稀有等位基因检测。

[0199] 可以理解的是，此处所述的实施例和实施方案仅仅是用于说明性目的，基于所述实施例和实施方案的各种修改或改变将提示于本领域技术人员。

[0200] 序列表信息

[0201] SEQ ID NO:1 栖热菌种 Z05 DNA 聚合酶 (Z05)

[0202]

MKAMLPLFEPKGRVLLVDGHHLAYRTFFALKGLTTSRGEPVQAVYGFAKSLLKALKEDGYKAVF
VVFDAKAPSFRHEAYEAYKAGRAPTPEDFPRQLALIKELVDLLGFTRLEVPGFEADDVLATLAK
KAEREGYEVRILTADRDLYQLVSDRVAVLHPEGHLITPEWLWEKYGLKPEQWVDFRALVGDPSD
NLPGVKGIGEKTALKLLKEWGSLENILKNLDRVKPESVRERIKAHLEDLKLSLELSRVRSDLPL
EVDFARRREPDREGLRAFLERLEFGSLLHEFGLLEAPAPLEEAPWPPPEGAFVGFVLSRPEPMW
AELKALAACKEGRVHRAKDPLAGLKDLKEVRGLLAKDLAVLALREGLDLAPSDDPMLLAYLLDP
SNTTPEGVARRYGGEWTEDAAHRALLAERLQQNLLERLKGEEKLLWLYQEVEKPLSRVLAHMEA
TGVRLDVAYLKALSLELAEEIRRLEEEVFRLAGHPFNLNSRDQLERVLFDELRLPALGKTQKTG
KRSTSAAVLEALREAHPIVEKILQHRELTKLKNTYVDPLPGLVHPRTGRLHTRFNQTATATGRL
SSSDPNLQNIPIRTPLGQRIRRAFVAEAGWALVALDYSQIELRVLAHLSGDENLIRVFQEGKDI
HTQTASWMFGVSPEAVDPLMRRAAKTVNFGVLYGMSAHRLSQELAIPYEEAVAFIERYFQSFPK
VRAWIEKTLEEGRKRGYVETLFGRRRYVPDLNARVKSVREAAERMAFNMPVQGTAADLMKLAMV
KLFPHLREMGARMLLQVHDELLLEAPQARAEEVAALAKEAMEKAYPLAVPLEVEVGIGEDWLSA
KG

[0203] SEQ ID NO:2 水生栖热菌 DNA 聚合酶 (Taq)

[0204]

MRGMLPLFEPKGRVLLVDGHHLAYRTFHALKGLTTSRGEPVQAVYGFAKSLLKALKEDGDAVIV
VFDAKAPSFRHEAYGGYKAGRAPTPEDFPRQLALIKELVDLLGLARLEVPGYEADDVLASLAKK
AEKEGYEVRILTADKDLYQLLSDRIHVLHPEGYLITPAWLWEKYGLRPDQWADYRALTGDESDN
LPGVKGIGEKTARKLLEEWGSLEALLKNLDRLKPAIREKILAHMDDLKLSWDLAKVRTDLPLEV
DFAKRREPDRERLRAFLERLEFGSLLHEFGLLESPKALEEAPWPPPEGAFVGFVLSRKEPMWAD
LLALAAARGGRVHRAPEPYKALRDLKEARGLLAKDLSVLALREGLGLPPGDDPMLLAYLLDPSN
TTPEGVARRYGGEWTEEAGERAALSERLFANLWGRLEGEERLLWLYREVERPLSAVLAHMEATG
VRLDVAYLRALSLEVAEEIARLEAEVFRLAGHPFNLNSRDQLERVLFDELGLPAIGKTEKTGKR
STSAAVLEALREAHPIVEKILQYRELTKLKSTYIDPLPDLIHPRTGRLHTRFNQTATATGRLSS
SDPNLQNIPVRTPLGQRIRRAFIAEEGWLLVALDYSQIELRVLAHLSGDENLIRVFQEGRDIHT
ETASWMFGVPREAVDPLMRRAAKTINFGVLYGMSAHRLSQELAIPYEEAQAFIERYFQSFPKVR
AWIEKTLEEGRRRGYVETLFGRRRYVPDLEARVKSVREAAERMAFNMPVQGTAADLMKLAMVKL
FPRLEEMGARMLLQVHDELVLEAPKERAEAVARLAKEVMEGVYPLAVPLEVEVGIGEDWLSAKE

[0205] SEQ ID NO:3 丝状栖热菌 DNA 聚合酶 (Tfi)

[0206]

MLPLLEPKGRVLLVDGHHLAYRTFFALKGLTTSRGEPVQAVYGFAKSLLKALKEDGEVAIVVFD
AKAPSFRHEAYEAYKAGRAPTPEDFPRQLALIKELVDLLGLVRLEVPGFEADDVLATLARKAER
EGYEVRILSADRDLYQLLSDRIHLLHPEGEVLTPGWLQERYGLSPERWVEYRALVGDPSDNLPG
VPGIGEKTALKLLKEWGSLEAILKNLDQVKPERVREAIRNNLDKLQMSLELSRLRTDLPLEVDF
AKRREPTGKGLKAFLERLEFGSLLHEFGLLEAPKEAEEAPWPPPGGAFLGFLLSRPEPMWAELL
ALAGAKEGRVHRAEDPVGALKDLKEIRGLLAKDLSVLALREGREIPPGDDPMLLAYLLDPGNTN
PEGVARRYGGEWKEDAAARALLSERLWQALYPRVAEEERLLWLYREVERPLAQVLAHMEATGVR
LDVPYLEALSQEVAFELERLEAEVHRLAGHPFNLNSRDQLERVLFDELGLPPIGKTEKTGKRST
SAAVLELLREAHPIVGRILEYRELMKLKSTYIDPLPRLVHPKTGRLHTRFNQTATATGRLSSSD
PNLQNIPVRTPLGQRIRKAFIAEEGHLLVALDYSQIELRVLAHLSGDENLIRVFREGKDIHTET
AAWMFGVPPEGVDGAMRRAAKTVNFGVLYGMSAHRLSQELSIPYEEAAAFIERYFQSFPKVRAW
IAKTLEEGRKKGYVETLFGRRRYVPDLNARVKSVREAAERMAFNMPVQGTAADLMKLAMVKLFP
RLRPLGVRILLQVHDELVLEAPKARAEEAAQLAKETMEGVYPLSVPLEVEVGMGEDWLSAKE

[0207] SEQ ID NO:4 黄栖热菌 DNA 聚合酶 (Tfl)

[0208]

MAMLPLFEPKGRVLLVDGHHLAYRTFFALKGLTTSRGEPVQAVYGFAKSLLKALKEDGDVVVVV
FDAKAPSFRHEAYEAYKAGRAPTPEDFPRQLALIKELVDLLGLVRLEVPGFEADDVLATLAKRA
EKEGYEVRILTADRDLYQLLSERIAILHPEGYLITPAWLYEKYGLRPEQWVDYRALAGDPSDNI
PGVKGIGEKTAQRLIREWGSLENLFQHLDQVKPSLREKLQAGMEALALSRKLSQVHTDLPLEVD
FGRRRTPNLEGLRAFLERLEFGSLLHEFGLLEGPKAAEEAPWPPPEGAFLGFSFSRPEPMWAEL
LALAGAWEGRLHRAQDPLRGLRDLKGVRGILAKDLAVLALREGLDLFPEDDPMLLAYLLDPSNT
TPEGVARRYGGEWTEDAGERALLAERLFQTLKERLKGEERLLWLYEEVEKPLSRVLARMEATGV
RLDVAYLQALSLEVEAEVRQLEEEVFRLAGHPFNLNSRDQLERVLFDELGLPAIGKTEKTGKRS
TSAAVLEALREAHPIVDRILQYRELTKLKNTYIDPLPALVHPKTGRLHTRFNQTATATGRLSSS
DPNLQNIPVRTPLGQRIRRAFVAEEGWVLVVLDYSQIELRVLAHLSGDENLIRVFQEGRDIHTQ
TASWMFGVSPEGVDPLMRRAAKTINFGVLYGMSAHRLSGELSIPYEEAVAFIERYFQSYPKVRA
WIEGTLEEGRRRGYVETLFGRRRYVPDLNARVKSVREAAERMAFNMPVQGTAADLMKLAMVRLF
PRLQELGARMLLQVHDELVLEAPKDRAERVAALAKEVMEGVWPLQVPLEVEVGLGEDWLSAKE

MLPLFEPKGRVLLVDGHHLAYRTFFALKGLTTSRGEPVQAVYGFAKSLLKALKEDGEVAIVVFD
AKAPSFRHEAYEAYKAGRAPTPEDFPRQLALIKELVDLLGLVRLEVPGFEADDVLATLAKKAER
EGYEVRILSADRDLYQLLSDRIHLLHPEGEVLTPGWLQERYGLSPERWVEYRALVGDPSDNLPG
VPGIGEKTALKLLKEWGSLEAILKNLDQVKPERVREAIRNNLDKLQMSLELSRLRTDLPLEVDF
AKRREPDWEGLKAFLERLEFGSLLHEFGLLEAPKEAEEAPWPPPGGAFLGFLLSRPEPMWAELL
ALAGAKEGRVHRAEDPVGALKDLKEIRGLLAKDLSVLALREGREIPPGDDPMLLAYLLDPGNTN
PEGVARRYGGEWKEDAAARALLSERLWQALYPRVAEEERLLWLYREVERPLAQVLAHMEATGVR
LDVPYLEALSQEVAFELERLEAEVHRLAGHPFNLNSRDQLERVLFDELGLPPIGKTEKTGKRST
SAAVLELLREAHPIVGRILEYRELMKLKSTYIDPLPRLVHPKTGRLHTRFNQTATATGRLSSSD
PNLQNIPVRTPLGQRIRKAFIAEEGHLLVALDYSQIELRVLAHLSGDENLIRVFREGKDIHTET
AAWMFGVPPEGVDGAMRRAAKTVNFGVLYGMSAHRLSQELSIPYEEAAAFIERYFQSFPKVRAW
IAKTLEEGRKKGYVETLFGRRRYVPDLNARVKSVREAAERMAFNMPVQGTAADLMKLAMVKLFP
RLRPLGVRILLQVHDELVLEAPKARAEEAAQLAKETMEGVYPLSVPLEVEVGMGEDWLSAKA

[0211] SEQ ID NO:6 嗜热栖热菌 DNA 聚合酶 (Tth)

[0212]

MEAMLPLFEPKGRVLLVDGHHLAYRTFFALKGLTTSRGEPVQAVYGFAKSLLKALKEDGYKAVF
VVFDAKAPSFRHEAYEAYKAGRAPTPEDFPRQLALIKELVDLLGFTRLEVPGYEADDVLATLAK
KAEKEGYEVRILTADRDLYQLVSDRVAVLHPEGHLITPEWLWEKYGLRPEQWVDFRALVGDPSD
NLPGVKGIGEKTALKLLKEWGSLENLLKNLDRVKPENVREKIKAHLEDLRLSLELSRVRTDLPL
EVDLAQGREPDREGLRAFLERLEFGSLLHEFGLLEAPAPLEEAPWPPPEGAFVGFVLSRPEPMW
AELKALAACRDGRVHRAADPLAGLKDLKEVRGLLAKDLAVLASREGLDLVPGDDPMLLAYLLDP
SNTTPEGVARRYGGEWTEDAAHRALLSERLHRNLLKRLEGEEKLLWLYHEVEKPLSRVLAHMEA
TGVRRDVAYLQALSLELAEEIRRLEEEVFRLAGHPFNLNSRDQLERVLFDELRLPALGKTQKTG
KRSTSAAVLEALREAHPIVEKILQHRELTKLKNTYVDPLPSLVHPRTGRLHTRFNQTATATGRL
SSSDPNLQNIPVRTPLGQRIRRAFVAEAGWALVALDYSQIELRVLAHLSGDENLIRVFQEGKDI
HTQTASWMFGVPPEAVDPLMRRAAKTVNFGVLYGMSAHRLSQELAIPYEEAVAFIERYFQSFPK
VRAWIEKTLEEGRKRGYVETLFGRRRYVPDLNARVKSVREAAERMAFNMPVQGTAADLMKLAMV
KLFPRLREMGARMLLQVHDELLLEAPQARAEEVAALAKEAMEKAYPLAVPLEVEVGMGEDWLSA
KG

[0213] SEQ ID NO:7 *Thermus caldophilus* DNA 聚合酶 (Tca)

[0214]

MEAMLPLFEPKGRVLLVDGHHLAYRTFFALKGLTTSRGEPVQAVYGFAKSLLKALKEDGYKAVF
VVFDAKAPSFRHEAYEAYKAGRAPTPEDFPRQLALIKELVDLLGFTRLEVPGYEADDVLATLAK
NPEKEGYEVRILTADRDLDQLVSDRVAVLHPEGHLITPEWLWQKYGLKPEQWVDFRALVGDPSD
NLPGVKGIGEKTALKLLKEWGSLENLLKNLDRVKPENVREKIKAHLEDLRLSLELSRVRTDLPL
EVDLAQGREPDREGLRAFLERLEFGSLLHEFGLLEAPAPLEEAPWPPPEGAFVGFVLSRPEPMW
AELKALAACRDGRVHRAADPLAGLKDLKEVRGLLAKDLAVLASREGLDLVPGDDPMLLAYLLDP
SNTTPEGVARRYGGEWTEDAAHRALLSERLHRNLLKRLQGEEKLLWLYHEVEKPLSRVLAHMEA
TGVRLDVAYLQALSLELAEEIRRLEEEVFRLAGHPFNLNSRDQLERVLFDELRLPALGKTQKTG
KRSTSAAVLEALREAHPIVEKILQHRELTKLKNTYVDPLPSLVHPNTGRLHTRFNQTATATGRL
SSSDPNLQNIPVRTPLGQRIRRAFVAEAGWALVALDYSQIELRVLAHLSGDENLIRVFQEGKDI
HTQTASWMFGVPPEAVDPLMRRAAKTVNFGVLYGMSAHRLSQELAIPYEEAVAFIERYFQSFPK
VRAWIEKTLEEGRKRGYVETLFGRRRYVPDLNARVKSVREAAERMAFNMPVQGTAADLMKLAMV
KLFPRLREMGARMLLQVHDELLLEAPQAGAEEVAALAKEAMEKAYPLAVPLEVEVGMGEDWLSA
KG

[0215] SEQ ID NO:8

[0216] Gly Tyr Val X4 Thr Leu X7 X8 Arg X10 X11 X12 X13 Pro, 其中 X4 是 Glu, Arg 或 Thr ;X7 是 Phe, Tyr 或 Leu ;X8 是 Gly 或 His ;X10 是 Arg 或 Lys ;X11 是除 Arg 以外

的任何氨基酸；X12 是 Tyr 或 Asp；以及 X13 是 Val, Ile 或 Leu。

[0217] SEQ ID NO:9

[0218] Gly Tyr Val X4 Thr Leu Phe Gly Arg X10 X11 X12 X13 Pro, 其中 X4 是 Glu 或 Arg；X10 是 Arg 或 Lys；X11 是除 Arg 以外的任何氨基酸；X12 是 Tyr 或 Asp；以及 X13 是 Val 或 Ile。

[0219] SEQ ID NO:10

[0220] Gly Tyr Val Glu Thr Leu Phe Gly Arg Arg X11 Tyr Val Pro, 其中 X11 是除 Arg 以外的任何氨基酸。

[0221] SEQ ID NO:11

[0222] Gly Tyr Val Glu Thr Leu Phe Gly Arg Arg X11 Tyr Val Pro, 其中 X11 是 Ala, Cys, Asp, Gly 或 Asn。

[0223] SEQ ID NO:12 Z05

[0224]

AWIEKTLEEGRKRGYVETLFGRRRYVPDLNARVK

[0225] SEQ ID NO:13 Taq

[0226]

AWIEKTLEEGRRRGYVETLFGRRRYVPDLEARVK

[0227] SEQ ID NO:14 Tfi

[0228]

AWIAKTLEEGRKKGYVETLFGRRRYVPDLNARVK

[0229] SEQ ID NO:15 Tfl

[0230]

AWIEGTLEEGRRRGYVETLFGRRRYVPDLNARVK

[0231] SEQ ID NO:16 Sps17

[0232]

AWIAKTLEEGRKKGYVETLFGRRRYVPDLNARVK

[0233] SEQ ID NO:17 Tth

[0234]

AWIEKTLEEGRKRGYVETLFGRRRYVPDLNARVK

[0235] SEQ ID NO:18 Tca

[0236]

AWIEKTLEEGRKRGYVETLFGRRRYVPDLNARVK

[0237] SEQ ID NO:19 Tma

[0238]

DYIQRVVSEAKEKGYVRTLFGRKRDIPQLMARDR

[0239] SEQ ID NO:20 Tne

[0240]

SYIQQVVAEAKEKGYVRTLFGRKRDIPQLMARDK

[0241] SEQ ID NO:21 Taf

[0242]

EYLKRMKDEARKKGYVTTLFGRRRYIPQLRSKNG

[0243] SEQ ID NO:23 Dra

[0244]

RYINHTLDFGRTHGYVETLYGRRRYVPGLSSRNR

[0245] SEQ ID NO:24 Bst

[0246]

QYMDNIVQEAKQKGYVTTLLHRRRYLPDITSRNF

[0247] SEQ ID NO:25 Bca

[0248]

RYMENIVQEAKQKGYVTTLLHRRRYLPDITSRNF

[0249] SEQ ID NO:26 天然的共有基序

[0250] Gly Tyr Val X4 Thr Leu X7 X8 Arg X10 Arg X12 X13 Pro, 其中 X4 是 Glu, Arg 或 Thr ;X7 是 Phe, Tyr 或 Leu ;X8 是 Gly 或 His ;X10 是 Arg 或 Lys ;X12 是 Tyr 或 Asp ; 以及 X13 是 Val, Ile 或 Leu。

[0251] SEQ ID NO:27 修饰的 Z05 D580 基序

[0252] T-G-R-L-S-S-X7-X8-P-N-L-Q-N

[0253] 其中 X7 是 Ser (S) 或 Thr (T) ;以及 X8 是除 Asp (D), 或 Glu (E) 以外的任何氨基酸。

[0254] SEQ ID NO:28 保守的 DNA 聚合酶活性位点

[0255]

DYSQIELR

[0256] SEQ ID NO:29 CS5 DNA 聚合酶

[0257]

MKAMLPLFEPKGRVLLVDGHHLAYRTFFALKGLTTSRGEPVQAVYGFAKSLLKALKEDGYKAVF
VVFDAKAPSFRHEAYEAYKAGRAPTPEDFPRQLALIKELVDLLGFTRLEVPGFEADDVLATLAK
KAEREGYEVRILTADRDLYQLVSDRVAVLHPEGHLITPEWLWEKYGLKPEQWVDFRALVGDPSD
NLPGVKGIGEKTALKLLKEWGSLENILKNLDRVKPESVRERIKAHLEDLKLSLELSRVRSDLPL
EVDFARRREPDREGLRAFLERLEFGSLLHEFGLLEESEPVGYRIVKDLVEFEKLIEKLRESPSF
AIDLETSSLDPFDCDIVGISVSFKPKEAYYIPLHHRNAQNLDEKEVLKKLKEILEDPGAKIVGQ
NLKFDYKVLMVKGVEPVPPYFDTMIAAYLLEPNEKKFNLDDLALKFLGYKMTSYQELMSFSFPL
FGFSFADVPVEKAANYSCEDADITYRLYKTLSLKLHEADLENVFYKIEMPLVNVLARMELNGVY
VDTEFLKKLSEEYGKKLEELAEEIYRIAGEPFNINSPKQVSRILFEKLGIKPRGKTTKTGDYST
RIEVLEELAGEHEIIPLILEYRKIQKLKSTYIDALPKMVNPKTGRIHASFNQTGTATGRLSSSD
PNLQNLPTKSEEGKEIRKAIVPQDPNWWIVSADYSQIELRILAHLSGDENLLRAFEEGIDVHTL
TASRIFNVKPEEVTEEMRRAGKMVNFSIIYGVTPYGLSVRLGVPVKEAEKMIVNYFVLYPKVRD
YIQRVVSEAKEKGYVRTLFGRKRDIPQLMARDRNTQAEGERIAINTPIQGTAADIIKLAMIEID
RELKERKMRSKMIIQVHDELVFEVPNEEKDALVELVKDRMTNVVKLSVPLEVDVTIGKTWS

[0258] SEQ ID NO:30 CS6 DNA 聚合酶

[0259]

MKAMLPLFEPKGRVLLVDGHHLAYRTFFALKGLTTSRGEPVQAVYGFAKSLLKALKEDGYKAVF
VVFDAKAPSFRHEAYEAYKAGRAPTPEDFPRQLALIKELVDLLGFTRLEVPGFEADDVLATLAK
KAEREGYEVRILTADRDLYQLVSDRVAVLHPEGHLITPEWLWEKYGLKPEQWVDFRALVGDPSD
NLPGVKGIGEKTALKLLKEWGSLENILKNLDRVKPESVRERIKAHLEDLKLSLELSRVRSDLPL
EVDFARRREPDREGLRAFLERLEFGSLLHEFGLLEESEPVGYRIVKDLVEFEKLIEKLRESPSF
AIALATSSLDPFDCDIVGISVSFKPKEAYYIPLHHRNAQNLDEKEVLKKLKEILEDPGAKIVGQ
NLKFDYKVLMVKGVEPVPPYFDTMIAAYLLEPNEKKFNLDDLALKFLGYKMTSYQELMSFSFPL
FGFSFADVPVEKAANYSCEDADITYRLYKTLSLKLHEADLENVFYKIEMPLVNVLARMELNGVY
VDTEFLKKLSEEYGKKLEELAEEIYRIAGEPFNINSPKQVSRILFEKLGIKPRGKTTKTGDYST
RIEVLEELAGEHEIIPLILEYRKIQKLKSTYIDALPKMVNPKTGRIHASFNQTGTATGRLSSSD
PNLQNLPTKSEEGKEIRKAIVPQDPNWWIVSADYSQIELRILAHLSGDENLLRAFEEGIDVHTL
TASRIFNVKPEEVTEEMRRAGKMVNFSIIYGVTPYGLSVRLGVPVKEAEKMIVNYFVLYPKVRD
YIQRVVSEAKEKGYVRTLFGRKRDIPQLMARDRNTQAEGERIAINTPIQGTAADIIKLAMIEID
RELKERKMRSKMIIQVHDELVFEVPNEEKDALVELVKDRMTNVVKLSVPLEVDVTIGKTWS

[0260] SEQ ID NO:31 正向引物

[0261]

5'- CTACCTCCTGGACCCCTCCAA-3'

[0262] SEQ ID NO:32 反向引物

[0263]

5'- ATAACCAACTGGTAGTGGCGTGTAA-3'

[0264] SEQ ID NO:33 Z05 D580G DNA 聚合酶的聚合酶结构域

[0265]

CTACCTCCTGGACCCCTCCAACACCACCCCCGAGGGGGTGGCCCGGCGCTACGGGGGGGAGTGG
ACGGAGGACGCCGCCCACCGGGCCCTCCTCGCTGAGCGGCTCCAGCAAAACCTCTTGGAACGCC
TCAAGGGAGAGGAAAAGCTCCTTTGGCTCTACCAAGAGGTGGAAAAGCCCCTCTCCCGGGTCCT
GGCCCACATGGAGGCCACCGGGGTAAGGCTGGACGTGGCCTATCTAAAGGCCCTTTCCCTGGAG
CTTGCGGAGGAGATTCGCCGCCTCGAGGAGGAGGTCTTCCGCCTGGCGGGCCACCCCTTCAACC
TGAACTCCCGTGACCAGCTAGAGCGGGTGCTCTTTGACGAGCTTAGGCTTCCCGCCCTGGGCAA
GACGCAAAAGACGGGGAAGCGCTCCACCAGCGCCGCGGTGCTGGAGGCCCTCAGGGAGGCCCAC
CCCATCGTGGAGAAGATCCTCCAGCACCGGGAGCTCACCAAGCTCAAGAACACCTACGTAGACC
CCCTCCCGGGCCTCGTCCACCCGAGGACGGGCCGCCTCCACACCCGCTTCAACCAGACAGCCAC
GGCCACGGGAAGGCTCTCTAGCTCCGGGCCCAACCTGCAGAACATCCCCATCCGCACCCCCTTG
GGCCAGAGGATCCGCCGGGCCTTCGTGGCCGAGGCGGGATGGGCGTTGGTGGCCCTGGACTATA
GCCAGATAGAGCTCCGGGTCCTCGCCCACCTCTCCGGGGACGAGAACCTGATCAGGGTCTTCCA
GGAGGGGAAGGACATCCACACCCAGACCGCAAGCTGGATGTTCGGCGTCTCCCCGGAGGCCGTG
GACCCCCTGATGCGCCGGGCGGCCAAGACGGTGAACTTCGGCGTCCTCTACGGCATGTCCGCCC
ATAGGCTCTCCCAGGAGCTTGCCATCCCCTACGAGGAGGCGGTGGCCTTTATAGAGCGCTACTT
CCAAAGCTTCCCCAAGGTGCGGGCCTGGATAGAAAAGACCCTGGAGGAGGGGAGGAAGCGGGGC
TACGTGGAAACCCTCTTCGGAAGAAGGCGCTACGTGCCCGACCTCAACGCCCGGGTGAAGAGCG
TCAGGGAGGCCGCGGAGCGCATGGCCTTCAACATGCCCGTCCAGGGCACCGCCGCCGACCTCAT
GAAGCTCGCCATGGTGAAGCTCTTCCCCCACCTCCGGGAGATGGGGGCCCGCATGCTCCTCCAG
GTCCACGACGAGCTCCTCCTGGAGGCCCCCCAAGCGCGGGCCGAGGAGGTGGCGGCTTTGGCCA
AGGAGGCCATGGAGAAGGCCTATCCCCTCGCCGTGCCCCTGGAGGTGGAGGTGGGGATCGGGGA
GGACTGGCTTTCCGCCAAGGGCTGATATCAGATCTCCCTGATTATGCGTCAGTCTATGAAGAAA
AATCGTATACAGATGGACGAAGAGAGAATCCTTGTGAATTTAACAGAGGGTATAGGGATTACAC
GCCACTACCAGTTGGTTAT

[0266] SEQ ID NO:34 - BRAF 野生型序列

[0267]

AGTAAAAATAGGTGATTTTGGTCTAGCTACAGTGAAATCTCGATGGAGTGGGTCCCATCAGTTT
GAACAGTTGTCTGGATCCATTTTGTGGATGGTAAGAATTGAGGCTA

[0268] SEQ ID NO:35 - BRAF V600R 突变型序列

[0269]

AGTAAAAATAGGTGATTTTGGTCTAGCTACAAGGAAATCTCGATGGAGTGGGTCCCATCAGTTT
GAACAGTTGTCTGGATCCATTTTGTGGATGGTAAGAATTGAGGCTA

[0270] SEQ ID NO:36 耐放射异常球菌 DNA 聚合酶 (Dra)

[0271]

MADASPDPSKPDALVLIDGHALAFRSYFALPPLNNSKGEMTDAIVGFMKLLLRLARQKSNQVIV
VFDPPVKTLRHEQYEGYKSGRAQTPEDLRGQINRIRALVDALGFPRLEEPGYEADDVIASLTRM
AEGKGYEVRIVTSDRDAYQLLDEHVKVIANDFSLIGPAQVEEKYGVTVRQWVDYRALTGDASDN
IPGAKGIGPKTAAKLLQEYGTLEKVYEAAHAGTLKPDGTRKKLLDSEENVKFSHDLSCMVTDLP
LDIEFGVRRLPDNPLVTEDLLTELELHSLRPMILGLNGPEQDGHAPDDLLEREHAQTPEEDEAA
ALPAFSAPELAEWQTPAEGAVWGYVLSREDDLTAALLAAATFEDGVARPARVSEPDEWAQAEAP
ENLFGELLPSDKPLTKKEQKALEKAQKDAEKARAKLREQFPATVDEAEFVGQRTVTAAAAKALA
AHLSVRGTVVEPGDDPLLYAYLLDPANTNMPVVAKRYLDREWPADAPTRAAITGHLVRELPPLL
DDARRKMYDEMEKPLSGVLGRMEVRGVQVDSDFLQTLSIQAGVRLADLESQIHEYAGEEFHIRS
PKQLETVLYDKLELASSKKTKLTGQRSTAVSALEPLRDAHPIIPLVLEFRELDKLRGTYLDPIP
NLVNPHTGRLHTTFAQTAVATGRLSSLNPNLQNIPIRSELGREIRKGFIAEDGFTLIAADYSQI
ELRLLAHIADDPLMQQAFVEGADIHRRTAAQVLGLDEATVDANQRRAAKTVNFGVLYGMSAHRL
SNDLGIPYAEAATFIEIYFATYPGIRRYINHTLDFGRTHGYVETLYGRRRYVPGLSSRNRVQRE
AEERLAYNMPIQGTAADIMKLAMVQLDPQLDAIGARMLLQVHDELLIEAPLDKAEQVAALTKKV
MENVVQLKVPLAVEVGTGPNWFDTK

[0272] SEQ ID NO:37 非洲栖热腔菌 DNA 聚合酶 (Taf)

[0273]

MGKMFLFDGTGLVYRAFYAIDQSLQTSSGLHTNAVYGLTKMLIKFLKEHISIGKDACVFVLDSKG
GSKKRKDILETYKANRPSTPDLLLEQIPYVEELVDALGIKVLKIEGFEADDIIATLSKKFESDFE
KVNIITGDKDLLQLVSDKVFVWRVERGITDLVLYDRNKVIEKYGIYPEQFKDYLSLVGDQIDNIP
GVKGIGKKTAVSLLKKYNSLENVLKNINLLTEKLRRLLEDSKEDLQKSIELVELIYDVPMDVEKD
EIIYRGYNPDKLLKVLKKYEFSSIIKELNLQEKLEKEYILVDNEDKLKKLAEEIEKYKTFSIDTE
TTSLDPFEAKLVGISISTMEGKAYYIPVSHFGAKNISKSLIDKFLKQILQEKDYNIVGQNLKFDY
EIFKSMGFSPNVPHFDTMIAAYLLNPDEKRFNLEELSLKYLGYKMISFDELVNENVPLFGNDFSY
VPLERAVEYSCEDADVTYRIFRKLGRKIYENEMEKLFYEIEMPLIDVLSEMELNGVYFDEEYLKE
LSKKYQEKMDGIKEKVFEIAGETFNLNSSTQVAYILFEKLNIAPYKKTATGKFSTNAEVLEELSK
EHEIAKLLLEYRKYQKLKSTYIDSIPLSINRKTNRVHTTFHQTGTSTGRLSSSNPNLQNLPTRSE
EGKEIRKAVRPQRQDWWILGADYSQIELRVLAHVSKDENLLKAFKEDLDIHTITAAKIFGVSEMF
VSEQMRRVGKMVNFAIIYGVSPYGLSKRIGLSVSETKKIIDNYFRYYKGVFEYLKRMKDEARKKG
YVTTLFGRRRYIPQLRSKNGNRVQEGERIAVNTPIQGTAADIIKIAMINIHNRLKKENLRSKMIL
QVHDELVFEVPDNELEIVKDLVRDEMENAVKLDVPLKVDVYYGKEWE

[0274] SEQ ID NO:38 海栖热袍菌 DNA 聚合酶 (Tma)

[0275]

MARLFLFDGTALAYRAYYALDRSLSTSTGIPTNATYGVARMLVRFIKDHIIVGKDYVAVAFDKKA
ATFRHKLLETYKAQRPKTPDLLIQQLPYIKKLVEALGMKVLEVEGYEADDIIATLAVKGLPLFDE
IFIVTGDKDMLQLVNEKIKVWRIVKGISDLELYDAQKVKEKYGVEPQQIPDLLALTGDEIDNIPG
VTGIGEKTAVQLLEKYKDLEDILNHVRELPQKVRKALLRDRENAILSKKLAILETNVPIEINWEE
LRYQGYDREKLLPLLKELEFASIMKELQLYEESEPVGYRIVKDLVEFEKLIEKLRESPSFAIDLE
TSSLDPFDCDIVGISVSFKPKEAYYIPLHHRNAQNLDEKEVLKKLKEILEDPGAKIVGQNLKFDY
KVLMVKGVEPVPPYFDTMIAAYLLEPNEKKFNLDDLALKFLGYKMTSYQELMSFSFPLFGFSFAD
VPVEKAANYSCEDADITYRLYKTLSLKLHEADLENVFYKIEMPLVNVLARMELNGVYVDTEFLKK
LSEEYGKKLEELAEEIYRIAGEPFNINSPKQVSRILFEKLGIKPRGKTTKTGDYSTRIEVLEELA
GEHEIIPLILEYRKIQKLKSTYIDALPKMVNPKTGRIHASFNQTGTATGRLSSSDPNLQNLPTKS
EEGKEIRKAIVPQDPNWWIVSADYSQIELRILAHLSGDENLLRAFEEGIDVHTLTASRIFNVKPE
EVTEEMRRAGKMVNFSIIYGVTPYGLSVRLGVPVKEAEKMIVNYFVLYPKVRDYIQRVVSEAKEK
GYVRTLFGRKRDIPQLMARDRNTQAEGERIAINTPIQGTAADIIKLAMIEIDRELKERKMRSKMI
IQVHDELVFEVPNEEKDALVELVKDRMTNVVKLSVPLEVDVTIGKTWS

MARLFLFDGTALAYRAYYALDRSLSTSTGIPTNAVYGVARMLVKFIKEHIIPEKDYAAVAFDKKA
ATFRHKLLVSDKAQRPKTPALLVQQLPYIKRLIEALGFKVLELEGYEADDIIATLAVRAARFLMR
FSLITGDKDMLQLVNEKIKVWRIVKGISDLELYDSKKVKERYGVEPHQIPDLLALTGDDIDNIPG
VTGIGEKTAVQLLGKYRNLEYILEHARELPQRVRKALLRDREVAILSKKLATLVTNAPVEVDWEE
MKYRGYDKRKLLPILKELEFASIMKELQLYEEAEPTGYEIVKDHKTFEDLIEKLKEVPSFALDLE
TSSLDPFNCEIVGISVSFKPKTAYYIPLHHRNAHNLDETLVLSKLKEILEDPSSKIVGQNLKYDY
KVLMVKGISPVYPHFDTMIAAYLLEPNEKKFNLEDLSLKFLGYKMTSYQELMSFSSPLFGFSFAD
VPVDKAAEYSCEDADITYRLYKILSMKLHEAELENVFYRIEMPLVNVLARMEFNWVYVDTEFLKK
LSEEYGKKLEELAEKIYQIAGEPFNINSPKQVSNILFEKLGIKPRGKTTKTGDYSTRIEVLEEIA
NEHEIVPLILEFRKILKLKSTYIDTLPKLVNPKTGRFHASFHQTGTATGRLSSSDPNLQNLPTKS
EEGKEIRKAIVPQDPDWWIVSADYSQIELRILAHLSGDENLVKAFEEGIDVHTLTASRIYNVKPE
EVNEEMRRVGKMVNFSIIYGVTPYGLSVRLGIPVKEAEKMIISYFTLYPKVRSYIQQVVAEAKEK
GYVRTLFGRKRDIPQLMARDKNTQSEGERIAINTPIQGTAADIIKLAMIDIDEELRKRNMKSRMI
IQVHDELVFEVPDEEKEELVDLVKNKMTNVVKLSVPLEVDISIGKSWS

[0278] SEQ ID NO:40 嗜热脂肪芽孢杆菌 DNA 聚合酶 (Bst)

[0279]

MKNKLVLIDGNSVAYRAFFALPLLHNDKGIHTNAVYGFTMMLNKILAEEQPTHILVAFDAGKTTF
RHETFQDYKGGRQQTPPELSEQFPLLRELLKAYRIPAYELDHYEADDIIGTMAARAEREGFAVKV
ISGDRDLTQLASPQVTVEITKKGITDIESYTPETVVEKYGLTPEQIVDLKGLMGDKSDNIPGVPG
IGEKTAVKLLKQFGTVENVLASIDEIKGEKLKENLRQYRDLALLSKQLAAICRDAPVELTLDDIV
YKGEDREKVVALFQELGFQSFLDKMAVQTDEGEKPLAGMDFAIADSVTDEMLADKAALVVEVVGD
NYHHAPIVGIALANERGRFFLRPETALADPKFLAWLGDETKKKTMFDSKRAAVALKWKGIELRGV
VFDLLLAAYLLDPAQAAGDVAAVAKMHQYEAVRSDEAVYGKGAKRTVPDEPTLAEHLARKAAAIW
ALEEPLMDELRRNEQDRLLTELEQPLAGILANMEFTGVKVDTKRLEQMGAELTEQLQAVERRIYE
LAGQEFNINSPKQLGTVLFDKLQLPVLKKTKTGYSTSADVLEKLAPHHEIVEHILHYRQLGKLQS
TYIEGLLKVVHPVTGKVHTMFNQALTQTGRLSSVEPNLQNIPIRLEEGRKIRQAFVPSEPDWLIF
AADYSQIELRVLAHIAEDDNLIEAFRRGLDIHTKTAMDIFHVSEEDVTANMRRQAKAVNFGIVYG
ISDYGLAQNLNITRKEAAEFIERYFASFPGVKQYMDNIVQEAKQKGYVTTLLHRRRYLPDITSRN
FNVRSFAERTAMNTPIQGSAADIIKKAMIDLSVRLREERLQARLLLQVHDELILEAPKEEIERLC
RLVPEVMEQAVALRVPLKVDYHYGPTWYDAK

[0280] SEQ ID NO:41 热坚芽孢杆菌 DNA 聚合酶 (Bca)

[0281]

MKKKLVLIDGSSVAYRAFFALPLLHNDKGIHTNAVYGFTMMLNKILAEEEPTHMLVAFDAGKTTF
RHEAFQEYKGGRQQTPPELSEQFPLLRELLRAYRIPAYELENYEADDIIGTLAARAEQEGFEVKV
ISGDRDLTQLASPHVTVDITKKGITDIEPYTPEAVREKYGLTPEQIVDLKGLMGDKSDNIPGVPG
IGEKTAVKLLRQFGTVENVLASIDEIKGEKLKETLRQHREMALLSKKLAAIRRDAPVELSLDDIA
YQGEDREKVVALFKELGFQSFLEKMESPSSEEEKPLAKMAFTLADRVTEEMLADKAALVVEVVEE
NYHDAPIVGIAVVNEHGRFFLRPETALADPQFVAWLGDETKKKSMFDSKRAAVALKWKGIELCGV
SFDLLLAAYLLDPAQGVDDVAAAAKMKQYEAVRPDEAVYGKGAKRAVPDEPVLAEHLVRKAAAIW
ALERPFLDELRRNEQDRLLVELEQPLSSILAEMEFAGVKVDTKRLEQMGEELAEQLRTVEQRIYE
LAGQEFNINSPKQLGVILFEKLQLPVLKKSKTGYSTSADVLEKLAPYHEIVENILQHYRQLGKLQ
STYIEGLLKVVRPDTKKVHTIFNQALTQTGRLSSTEPNLQNIPIRLEEGRKIRQAFVPSESDWLI
FAADYSQIELRVLAHIAEDDNLMEAFRRDLDIHTKTAMDIFQVSEDEVTPNMRRQAKAVNFGIVY
GISDYGLAQNLNISRKEAAEFIERYFESFPGVKRYMENIVQEAKQKGYVTTLLHRRRYLPDITSR
NFNVRSFAERMAMNTPIQGSAADIIKKAMIDLNARLKEERLQARLLLQVHDELILEAPKEEMERL
CRLVPEVMEQAVTLRVPLKVDYHYGSTWYDAK

[0282] SEQ ID NO:42

[0283] Gly Tyr Val X4 Thr Leu X7 X8 Arg X10 X11 X12 X13 Pro, 其中 X4 是 Glu, Arg 或 Thr ;X7 是 Phe, Tyr 或 Leu ;X8 是 Gly 或 His ;X10 是 Arg 或 Lys ;X11 是 Gly, Ala, Leu, Met, Trp, Pro, Ser, Thr, Phe, Tyr, Cys, Asn, Gln, Asp, Glu, Lys, Val, Ile

或 His ;X12 是 Tyr 或 Asp ;以及 X13 是 Val, Ile 或 Leu。

序列表

<110> F. Hoffmann-La Roche AG
Roche Diagnostics GmbH

<120> 具有增强的 3'- 错配辨别力的 DNA 聚合酶

<130> 26748 WO-HS

<140> 还没有指定
<141> 还没有指定

<150> US 61/356,219
<151> 2010-06-18

<160> 42

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 834
<212> PRT
<213> 栖热菌种

<220>
<223> 栖热菌种 Z05 DNA 聚合酶 (Z05)

<400> 1

```
Met Lys Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu
 1               5                  10                  15
Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu Lys Gly
            20                  25                  30
Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala
        35                  40                  45
Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala Val Phe
    50                  55                  60
Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Glu
65                  70                  75                  80
```

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

385 390 395 400
Thr Glu Asp Ala Ala His Arg Ala Leu Leu Ala Glu Arg Leu Gln Gln
405 410 415
Asn Leu Leu Glu Arg Leu Lys Gly Glu Glu Lys Leu Leu Trp Leu Tyr
420 425 430
Gln Glu Val Glu Lys Pro Leu Ser Arg Val Leu Ala His Met Glu Ala
435 440 445
Thr Gly Val Arg Leu Asp Val Ala Tyr Leu Lys Ala Leu Ser Leu Glu
450 455 460
Leu Ala Glu Glu Ile Arg Arg Leu Glu Glu Glu Val Phe Arg Leu Ala
465 470 475 480
Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu
485 490 495
Phe Asp Glu Leu Arg Leu Pro Ala Leu Gly Lys Thr Gln Lys Thr Gly
500 505 510
Lys Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His
515 520 525
Pro Ile Val Glu Lys Ile Leu Gln His Arg Glu Leu Thr Lys Leu Lys
530 535 540
Asn Thr Tyr Val Asp Pro Leu Pro Gly Leu Val His Pro Arg Thr Gly
545 550 555 560
Arg Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu
565 570 575
Ser Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Ile Arg Thr Pro Leu
580 585 590
Gly Gln Arg Ile Arg Arg Ala Phe Val Ala Glu Ala Gly Trp Ala Leu
595 600 605
Val Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu
610 615 620
Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Lys Asp Ile
625 630 635 640
His Thr Gln Thr Ala Ser Trp Met Phe Gly Val Ser Pro Glu Ala Val
645 650 655
Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly Val Leu
660 665 670
Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr
675 680 685
Glu Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys
690 695 700

Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys Arg Gly
705 710 715 720
Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Asn
725 730 735
Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn
740 745 750
Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val
755 760 765
Lys Leu Phe Pro His Leu Arg Glu Met Gly Ala Arg Met Leu Leu Gln
770 775 780
Val His Asp Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala Glu Glu
785 790 795 800
Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro Leu Ala
805 810 815
Val Pro Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala
820 825 830
Lys Gly

<210> 2
<211> 832
<212> PRT
<213> 水生栖热菌

<220>
<223> 水生栖热菌 DNA 聚合酶 (Taq)

<400> 2
Met Arg Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu
1 5 10 15
Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys Gly
20 25 30
Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala
35 40 45
Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile Val
50 55 60
Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly Gly
65 70 75 80
Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln Leu

85 90 95

Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu Glu
100 105 110

Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys Lys
115 120 125

Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys Asp
130 135 140

Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu Gly
145 150 155 160

Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg Pro
165 170 175

Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp Asn
180 185 190

Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu Leu
195 200 205

Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg Leu
210 215 220

Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu Lys
225 230 235 240

Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu Val
245 250 255

Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala Phe
260 265 270

Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu Leu
275 280 285

Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu Gly
290 295 300

Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala Asp
305 310 315 320

Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala Pro
325 330 335

Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu Leu
340 345 350

Ala Lys Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Leu Gly Leu Pro
355 360 365

Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser Asn
370 375 380

Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr Glu
385 390 395 400

Glu Ala Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu Phe Ala Asn Leu
405 410 415
Trp Gly Arg Leu Glu Gly Glu Glu Arg Leu Leu Trp Leu Tyr Arg Glu
420 425 430
Val Glu Arg Pro Leu Ser Ala Val Leu Ala His Met Glu Ala Thr Gly
435 440 445
Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val Ala
450 455 460
Glu Glu Ile Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly His
465 470 475 480
Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe Asp
485 490 495
Glu Leu Gly Leu Pro Ala Ile Gly Lys Thr Glu Lys Thr Gly Lys Arg
500 505 510
Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro Ile
515 520 525
Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser Thr
530 535 540
Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg Leu
545 550 555 560
His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser Ser
565 570 575
Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly Gln
580 585 590
Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val Ala
595 600 605
Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser Gly
610 615 620
Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His Thr
625 630 635 640
Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp Pro
645 650 655
Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr Gly
660 665 670
Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu Glu
675 680 685
Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val Arg
690 695 700
Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr Val

705 710 715 720
Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala Arg
725 730 735
Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met Pro
740 745 750
Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys Leu
755 760 765
Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val His
770 775 780
Asp Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val Ala
785 790 795 800
Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val Pro
805 810 815
Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys Glu
820 825 830

<210> 3
<211> 830
<212> PRT
<213> 丝状栖热菌

<220>
<223> 丝状栖热菌 DNA 聚合酶 (Tfi)

<400> 3
Met Leu Pro Leu Leu Glu Pro Lys Gly Arg Val Leu Leu Val Asp Gly
1 5 10 15
His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu Lys Gly Leu Thr Thr
20 25 30
Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala Lys Ser Leu
35 40 45
Leu Lys Ala Leu Lys Glu Asp Gly Glu Val Ala Ile Val Val Phe Asp
50 55 60
Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Glu Ala Tyr Lys Ala
65 70 75 80
Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln Leu Ala Leu Ile
85 90 95
Lys Glu Leu Val Asp Leu Leu Gly Leu Val Arg Leu Glu Val Pro Gly

100 105 110

Phe Glu Ala Asp Asp Val Leu Ala Thr Leu Ala Arg Lys Ala Glu Arg

115 120 125

Glu Gly Tyr Glu Val Arg Ile Leu Ser Ala Asp Arg Asp Leu Tyr Gln

130 135 140

Leu Leu Ser Asp Arg Ile His Leu Leu His Pro Glu Gly Glu Val Leu

145 150 155 160

Thr Pro Gly Trp Leu Gln Glu Arg Tyr Gly Leu Ser Pro Glu Arg Trp

165 170 175

Val Glu Tyr Arg Ala Leu Val Gly Asp Pro Ser Asp Asn Leu Pro Gly

180 185 190

Val Pro Gly Ile Gly Glu Lys Thr Ala Leu Lys Leu Leu Lys Glu Trp

195 200 205

Gly Ser Leu Glu Ala Ile Leu Lys Asn Leu Asp Gln Val Lys Pro Glu

210 215 220

Arg Val Trp Glu Ala Ile Arg Asn Asn Leu Asp Lys Leu Gln Met Ser

225 230 235 240

Leu Glu Leu Ser Arg Leu Arg Thr Asp Leu Pro Leu Glu Val Asp Phe

245 250 255

Ala Lys Arg Arg Glu Pro Thr Gly Lys Gly Leu Lys Ala Phe Leu Glu

260 265 270

Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu Leu Glu Ala

275 280 285

Pro Lys Glu Ala Glu Glu Ala Pro Trp Pro Pro Pro Gly Gly Ala Phe

290 295 300

Leu Gly Phe Leu Leu Ser Arg Pro Glu Pro Met Trp Ala Glu Leu Leu

305 310 315 320

Ala Leu Ala Gly Ala Lys Glu Gly Arg Val His Arg Ala Glu Asp Pro

325 330 335

Val Gly Ala Leu Lys Asp Leu Lys Glu Ile Arg Gly Leu Leu Ala Lys

340 345 350

Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Arg Glu Ile Pro Pro Gly

355 360 365

Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Gly Asn Thr Asn

370 375 380

Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Lys Glu Asp Ala

385 390 395 400

Ala Ala Arg Ala Leu Leu Ser Glu Arg Leu Trp Gln Ala Leu Tyr Pro

405 410 415

Arg Val Ala Glu Glu Glu Arg Leu Leu Trp Leu Tyr Arg Glu Val Glu
420 425 430
Arg Pro Leu Ala Gln Val Leu Ala His Met Glu Ala Thr Gly Val Arg
435 440 445
Leu Asp Val Pro Tyr Leu Glu Ala Leu Ser Gln Glu Val Ala Phe Glu
450 455 460
Leu Glu Arg Leu Glu Ala Glu Val His Arg Leu Ala Gly His Pro Phe
465 470 475 480
Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe Asp Glu Leu
485 490 495
Gly Leu Pro Pro Ile Gly Lys Thr Glu Lys Thr Gly Lys Arg Ser Thr
500 505 510
Ser Ala Ala Val Leu Glu Leu Leu Arg Glu Ala His Pro Ile Val Gly
515 520 525
Arg Ile Leu Glu Tyr Arg Glu Leu Met Lys Leu Lys Ser Thr Tyr Ile
530 535 540
Asp Pro Leu Pro Arg Leu Val His Pro Lys Thr Gly Arg Leu His Thr
545 550 555 560
Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser Ser Ser Asp
565 570 575
Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly Gln Arg Ile
580 585 590
Arg Lys Ala Phe Ile Ala Glu Glu Gly His Leu Leu Val Ala Leu Asp
595 600 605
Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser Gly Asp Glu
610 615 620
Asn Leu Ile Arg Val Phe Arg Glu Gly Lys Asp Ile His Thr Glu Thr
625 630 635 640
Ala Ala Trp Met Phe Gly Val Pro Pro Glu Gly Val Asp Gly Ala Met
645 650 655
Arg Arg Ala Ala Lys Thr Val Asn Phe Gly Val Leu Tyr Gly Met Ser
660 665 670
Ala His Arg Leu Ser Gln Glu Leu Ser Ile Pro Tyr Glu Glu Ala Ala
675 680 685
Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val Arg Ala Trp
690 695 700
Ile Ala Lys Thr Leu Glu Glu Gly Arg Lys Lys Gly Tyr Val Glu Thr
705 710 715 720
Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Asn Ala Arg Val Lys

725 730 735
Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met Pro Val Gln
740 745 750
Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys Leu Phe Pro
755 760 765
Arg Leu Arg Pro Leu Gly Val Arg Ile Leu Leu Gln Val His Asp Glu
770 775 780
Leu Val Leu Glu Ala Pro Lys Ala Arg Ala Glu Glu Ala Ala Gln Leu
785 790 795 800
Ala Lys Glu Thr Met Glu Gly Val Tyr Pro Leu Ser Val Pro Leu Glu
805 810 815
Val Glu Val Gly Met Gly Glu Asp Trp Leu Ser Ala Lys Glu
820 825 830

<210> 4
<211> 831
<212> PRT
<213> 黄栖热菌

<220>
<223> 黄栖热菌 DNA 聚合酶 (Tfl)

<400> 4
Met Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu Val
1 5 10 15
Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu Lys Gly Leu
20 25 30
Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala Lys
35 40 45
Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Val Val Val Val Val
50 55 60
Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Glu Ala Tyr
65 70 75 80
Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln Leu Ala
85 90 95
Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Val Arg Leu Glu Val
100 105 110
Pro Gly Phe Glu Ala Asp Asp Val Leu Ala Thr Leu Ala Lys Arg Ala

115 120 125

Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Arg Asp Leu

130 135 140

Tyr Gln Leu Leu Ser Glu Arg Ile Ala Ile Leu His Pro Glu Gly Tyr

145 150 155 160

Leu Ile Thr Pro Ala Trp Leu Tyr Glu Lys Tyr Gly Leu Arg Pro Glu

165 170 175

Gln Trp Val Asp Tyr Arg Ala Leu Ala Gly Asp Pro Ser Asp Asn Ile

180 185 190

Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Gln Arg Leu Ile Arg

195 200 205

Glu Trp Gly Ser Leu Glu Asn Leu Phe Gln His Leu Asp Gln Val Lys

210 215 220

Pro Ser Leu Arg Glu Lys Leu Gln Ala Gly Met Glu Ala Leu Ala Leu

225 230 235 240

Ser Arg Lys Leu Ser Gln Val His Thr Asp Leu Pro Leu Glu Val Asp

245 250 255

Phe Gly Arg Arg Arg Thr Pro Asn Leu Glu Gly Leu Arg Ala Phe Leu

260 265 270

Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu Leu Glu

275 280 285

Gly Pro Lys Ala Ala Glu Glu Ala Pro Trp Pro Pro Pro Glu Gly Ala

290 295 300

Phe Leu Gly Phe Ser Phe Ser Arg Pro Glu Pro Met Trp Ala Glu Leu

305 310 315 320

Leu Ala Leu Ala Gly Ala Trp Glu Gly Arg Leu His Arg Ala Gln Asp

325 330 335

Pro Leu Arg Gly Leu Arg Asp Leu Lys Gly Val Arg Gly Ile Leu Ala

340 345 350

Lys Asp Leu Ala Val Leu Ala Leu Arg Glu Gly Leu Asp Leu Phe Pro

355 360 365

Glu Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser Asn Thr

370 375 380

Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr Glu Asp

385 390 395 400

Ala Gly Glu Arg Ala Leu Leu Ala Glu Arg Leu Phe Gln Thr Leu Lys

405 410 415

Glu Arg Leu Lys Gly Glu Glu Arg Leu Leu Trp Leu Tyr Glu Glu Val

420 425 430

Glu Lys Pro Leu Ser Arg Val Leu Ala Arg Met Glu Ala Thr Gly Val
435 440 445
Arg Leu Asp Val Ala Tyr Leu Gln Ala Leu Ser Leu Glu Val Glu Ala
450 455 460
Glu Val Arg Gln Leu Glu Glu Glu Val Phe Arg Leu Ala Gly His Pro
465 470 475 480
Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe Asp Glu
485 490 495
Leu Gly Leu Pro Ala Ile Gly Lys Thr Glu Lys Thr Gly Lys Arg Ser
500 505 510
Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro Ile Val
515 520 525
Asp Arg Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Asn Thr Tyr
530 535 540
Ile Asp Pro Leu Pro Ala Leu Val His Pro Lys Thr Gly Arg Leu His
545 550 555 560
Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser Ser Ser
565 570 575
Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly Gln Arg
580 585 590
Ile Arg Arg Ala Phe Val Ala Glu Glu Gly Trp Val Leu Val Val Leu
595 600 605
Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser Gly Asp
610 615 620
Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His Thr Gln
625 630 635 640
Thr Ala Ser Trp Met Phe Gly Val Ser Pro Glu Gly Val Asp Pro Leu
645 650 655
Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr Gly Met
660 665 670
Ser Ala His Arg Leu Ser Gly Glu Leu Ser Ile Pro Tyr Glu Glu Ala
675 680 685
Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Tyr Pro Lys Val Arg Ala
690 695 700
Trp Ile Glu Gly Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr Val Glu
705 710 715 720
Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Asn Ala Arg Val
725 730 735
Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met Pro Val

740 745 750

Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Arg Leu Phe
755 760 765

Pro Arg Leu Gln Glu Leu Gly Ala Arg Met Leu Leu Gln Val His Asp
770 775 780

Glu Leu Val Leu Glu Ala Pro Lys Asp Arg Ala Glu Arg Val Ala Ala
785 790 795 800

Leu Ala Lys Glu Val Met Glu Gly Val Trp Pro Leu Gln Val Pro Leu
805 810 815

Glu Val Glu Val Gly Leu Gly Glu Asp Trp Leu Ser Ala Lys Glu
820 825 830

<210> 5
<211> 830
<212> PRT
<213> 栖热菌种

<220>
<223> 栖热菌种 Sps17 DNA 聚合酶 (Sps17)

<400> 5
Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu Val Asp Gly
1 5 10 15

His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu Lys Gly Leu Thr Thr
20 25 30

Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala Lys Ser Leu
35 40 45

Leu Lys Ala Leu Lys Glu Asp Gly Glu Val Ala Ile Val Val Phe Asp
50 55 60

Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Glu Ala Tyr Lys Ala
65 70 75 80

Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln Leu Ala Leu Ile
85 90 95

Lys Glu Leu Val Asp Leu Leu Gly Leu Val Arg Leu Glu Val Pro Gly
100 105 110

Phe Glu Ala Asp Asp Val Leu Ala Thr Leu Ala Lys Lys Ala Glu Arg
115 120 125

Glu Gly Tyr Glu Val Arg Ile Leu Ser Ala Asp Arg Asp Leu Tyr Gln

130 135 140

Leu Leu Ser Asp Arg Ile His Leu Leu His Pro Glu Gly Glu Val Leu
145 150 155 160

Thr Pro Gly Trp Leu Gln Glu Arg Tyr Gly Leu Ser Pro Glu Arg Trp
165 170 175

Val Glu Tyr Arg Ala Leu Val Gly Asp Pro Ser Asp Asn Leu Pro Gly
180 185 190

Val Pro Gly Ile Gly Glu Lys Thr Ala Leu Lys Leu Leu Lys Glu Trp
195 200 205

Gly Ser Leu Glu Ala Ile Leu Lys Asn Leu Asp Gln Val Lys Pro Glu
210 215 220

Arg Val Arg Glu Ala Ile Arg Asn Asn Leu Asp Lys Leu Gln Met Ser
225 230 235 240

Leu Glu Leu Ser Arg Leu Arg Thr Asp Leu Pro Leu Glu Val Asp Phe
245 250 255

Ala Lys Arg Arg Glu Pro Asp Trp Glu Gly Leu Lys Ala Phe Leu Glu
260 265 270

Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu Leu Glu Ala
275 280 285

Pro Lys Glu Ala Glu Glu Ala Pro Trp Pro Pro Pro Gly Gly Ala Phe
290 295 300

Leu Gly Phe Leu Leu Ser Arg Pro Glu Pro Met Trp Ala Glu Leu Leu
305 310 315 320

Ala Leu Ala Gly Ala Lys Glu Gly Arg Val His Arg Ala Glu Asp Pro
325 330 335

Val Gly Ala Leu Lys Asp Leu Lys Glu Ile Arg Gly Leu Leu Ala Lys
340 345 350

Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Arg Glu Ile Pro Pro Gly
355 360 365

Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Gly Asn Thr Asn
370 375 380

Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Lys Glu Asp Ala
385 390 395 400

Ala Ala Arg Ala Leu Leu Ser Glu Arg Leu Trp Gln Ala Leu Tyr Pro
405 410 415

Arg Val Ala Glu Glu Glu Arg Leu Leu Trp Leu Tyr Arg Glu Val Glu
420 425 430

Arg Pro Leu Ala Gln Val Leu Ala His Met Glu Ala Thr Gly Val Arg
435 440 445

Leu Asp Val Pro Tyr Leu Glu Ala Leu Ser Gln Glu Val Ala Phe Glu
450 455 460
Leu Glu Arg Leu Glu Ala Glu Val His Arg Leu Ala Gly His Pro Phe
465 470 475 480
Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe Asp Glu Leu
485 490 495
Gly Leu Pro Pro Ile Gly Lys Thr Glu Lys Thr Gly Lys Arg Ser Thr
500 505 510
Ser Ala Ala Val Leu Glu Leu Leu Arg Glu Ala His Pro Ile Val Gly
515 520 525
Arg Ile Leu Glu Tyr Arg Glu Leu Met Lys Leu Lys Ser Thr Tyr Ile
530 535 540
Asp Pro Leu Pro Arg Leu Val His Pro Lys Thr Gly Arg Leu His Thr
545 550 555 560
Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser Ser Ser Asp
565 570 575
Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly Gln Arg Ile
580 585 590
Arg Lys Ala Phe Ile Ala Glu Glu Gly His Leu Leu Val Ala Leu Asp
595 600 605
Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser Gly Asp Glu
610 615 620
Asn Leu Ile Arg Val Phe Arg Glu Gly Lys Asp Ile His Thr Glu Thr
625 630 635 640
Ala Ala Trp Met Phe Gly Val Pro Pro Glu Gly Val Asp Gly Ala Met
645 650 655
Arg Arg Ala Ala Lys Thr Val Asn Phe Gly Val Leu Tyr Gly Met Ser
660 665 670
Ala His Arg Leu Ser Gln Glu Leu Ser Ile Pro Tyr Glu Glu Ala Ala
675 680 685
Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val Arg Ala Trp
690 695 700
Ile Ala Lys Thr Leu Glu Glu Gly Arg Lys Lys Gly Tyr Val Glu Thr
705 710 715 720
Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Asn Ala Arg Val Lys
725 730 735
Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met Pro Val Gln
740 745 750
Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys Leu Phe Pro

755 760 765

Arg Leu Arg Pro Leu Gly Val Arg Ile Leu Leu Gln Val His Asp Glu

770 775 780

Leu Val Leu Glu Ala Pro Lys Ala Arg Ala Glu Glu Ala Ala Gln Leu

785 790 795 800

Ala Lys Glu Thr Met Glu Gly Val Tyr Pro Leu Ser Val Pro Leu Glu

805 810 815

Val Glu Val Gly Met Gly Glu Asp Trp Leu Ser Ala Lys Ala

820 825 830

<210> 6

<211> 834

<212> PRT

<213> 嗜热栖热菌

<220>

<223> 嗜热栖热菌 DNA 聚合酶 (Tth)

<400> 6

Met Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu

1 5 10 15

Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu Lys Gly

20 25 30

Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala

35 40 45

Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala Val Phe

50 55 60

Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Glu

65 70 75 80

Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln

85 90 95

Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr Arg Leu

100 105 110

Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu Ala Lys

115 120 125

Lys Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Arg

130 135 140

Asp Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His Pro Glu

145 150 155 160
Gly His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly Leu Arg
165 170 175
Pro Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro Ser Asp
180 185 190
Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Leu Lys Leu
195 200 205
Leu Lys Glu Trp Gly Ser Leu Glu Asn Leu Leu Lys Asn Leu Asp Arg
210 215 220
Val Lys Pro Glu Asn Val Arg Glu Lys Ile Lys Ala His Leu Glu Asp
225 230 235 240
Leu Arg Leu Ser Leu Glu Leu Ser Arg Val Arg Thr Asp Leu Pro Leu
245 250 255
Glu Val Asp Leu Ala Gln Gly Arg Glu Pro Asp Arg Glu Gly Leu Arg
260 265 270
Ala Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly
275 280 285
Leu Leu Glu Ala Pro Ala Pro Leu Glu Glu Ala Pro Trp Pro Pro Pro
290 295 300
Glu Gly Ala Phe Val Gly Phe Val Leu Ser Arg Pro Glu Pro Met Trp
305 310 315 320
Ala Glu Leu Lys Ala Leu Ala Ala Cys Arg Asp Gly Arg Val His Arg
325 330 335
Ala Ala Asp Pro Leu Ala Gly Leu Lys Asp Leu Lys Glu Val Arg Gly
340 345 350
Leu Leu Ala Lys Asp Leu Ala Val Leu Ala Ser Arg Glu Gly Leu Asp
355 360 365
Leu Val Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro
370 375 380
Ser Asn Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp
385 390 395 400
Thr Glu Asp Ala Ala His Arg Ala Leu Leu Ser Glu Arg Leu His Arg
405 410 415
Asn Leu Leu Lys Arg Leu Glu Gly Glu Glu Lys Leu Leu Trp Leu Tyr
420 425 430
His Glu Val Glu Lys Pro Leu Ser Arg Val Leu Ala His Met Glu Ala
435 440 445
Thr Gly Val Arg Arg Asp Val Ala Tyr Leu Gln Ala Leu Ser Leu Glu
450 455 460

Leu Ala Glu Glu Ile Arg Arg Leu Glu Glu Glu Val Phe Arg Leu Ala
465 470 475 480
Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu
485 490 495
Phe Asp Glu Leu Arg Leu Pro Ala Leu Gly Lys Thr Gln Lys Thr Gly
500 505 510
Lys Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His
515 520 525
Pro Ile Val Glu Lys Ile Leu Gln His Arg Glu Leu Thr Lys Leu Lys
530 535 540
Asn Thr Tyr Val Asp Pro Leu Pro Ser Leu Val His Pro Arg Thr Gly
545 550 555 560
Arg Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu
565 570 575
Ser Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu
580 585 590
Gly Gln Arg Ile Arg Arg Ala Phe Val Ala Glu Ala Gly Trp Ala Leu
595 600 605
Val Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu
610 615 620
Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Lys Asp Ile
625 630 635 640
His Thr Gln Thr Ala Ser Trp Met Phe Gly Val Pro Pro Glu Ala Val
645 650 655
Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly Val Leu
660 665 670
Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr
675 680 685
Glu Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys
690 695 700
Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys Arg Gly
705 710 715 720
Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Asn
725 730 735
Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn
740 745 750
Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val
755 760 765
Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu Leu Gln

770 775 780
Val His Asp Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala Glu Glu
785 790 795 800
Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro Leu Ala
805 810 815
Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu Ser Ala
820 825 830
Lys Gly

<210> 7
<211> 834
<212> PRT
<213> Thermus caldophilus

<220>
<223> Thermus caldophilus DNA 聚合酶 (Tca)

<400> 7
Met Glu Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu
1 5 10 15
Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu Lys Gly
20 25 30
Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala
35 40 45
Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala Val Phe
50 55 60
Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Glu
65 70 75 80
Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
85 90 95
Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr Arg Leu
100 105 110
Glu Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu Ala Lys
115 120 125
Asn Pro Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Arg
130 135 140
Asp Leu Asp Gln Leu Val Ser Asp Arg Val Ala Val Leu His Pro Glu
145 150 155 160

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

465 470 475 480
Gly His Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu
485 490 495
Phe Asp Glu Leu Arg Leu Pro Ala Leu Gly Lys Thr Gln Lys Thr Gly
500 505 510
Lys Arg Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His
515 520 525
Pro Ile Val Glu Lys Ile Leu Gln His Arg Glu Leu Thr Lys Leu Lys
530 535 540
Asn Thr Tyr Val Asp Pro Leu Pro Ser Leu Val His Pro Asn Thr Gly
545 550 555 560
Arg Leu His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu
565 570 575
Ser Ser Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu
580 585 590
Gly Gln Arg Ile Arg Arg Ala Phe Val Ala Glu Ala Gly Trp Ala Leu
595 600 605
Val Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu
610 615 620
Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Lys Asp Ile
625 630 635 640
His Thr Gln Thr Ala Ser Trp Met Phe Gly Val Pro Pro Glu Ala Val
645 650 655
Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly Val Leu
660 665 670
Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr
675 680 685
Glu Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys
690 695 700
Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys Arg Gly
705 710 715 720
Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Asn
725 730 735
Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn
740 745 750
Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val
755 760 765
Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu Leu Gln
770 775 780

Val His Asp Glu Leu Leu Leu Glu Ala Pro Gln Ala Gly Ala Glu Glu
785 790 795 800
Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro Leu Ala
805 810 815
Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu Ser Ala
820 825 830
Lys Gly

<210> 8
<211> 14
<212> PRT
<213> 人工序列

<220>
<223> 合成的 DNA 聚合酶结构域基序

<220>
<221> 变体
<222> (4)...(4)
<223> Xaa = Glu, Arg 或 Thr

<220>
<221> 变体
<222> (7)...(7)
<223> Xaa = Phe, Tyr 或 Leu

<220>
<221> 变体
<222> (8)...(8)
<223> Xaa = Gly 或 His

<220>
<221> 变体
<222> (10)...(10)
<223> Xaa = Arg 或 Lys

<220>
<221> 变体

<222> (11)...(11)
<223> Xaa = 除 Arg 以外的任何氨基酸

<220>
<221> 变体
<222> (12)...(12)
<223> Xaa = Tyr 或 Asp

<220>
<221> 变体
<222> (13)...(13)
<223> Xaa = Val, Ile 或 Leu

<400> 8
Gly Tyr Val Xaa Thr Leu Xaa Xaa Arg Xaa Xaa Xaa Xaa Pro
1 5 10

<210> 9
<211> 14
<212> PRT
<213> 人工序列

<220>
<223> 合成的 DNA 聚合酶结构域基序

<220>
<221> 变体
<222> (4)...(4)
<223> Xaa = Glu 或 Arg

<220>
<221> 变体
<222> (10)...(10)
<223> Xaa = Arg 或 Lys

<220>
<221> 变体
<222> (11)...(11)

<223> Xaa = 除 Arg 以外的任何氨基酸

<220>
<221> 变体
<222> (12)...(12)
<223> Xaa = Tyr 或 Asp

<220>
<221> 变体
<222> (13)...(13)
<223> Xaa = Val 或 Ile

<400> 9
Gly Tyr Val Xaa Thr Leu Phe Gly Arg Xaa Xaa Xaa Xaa Pro
1 5 10

<210> 10
<211> 14
<212> PRT
<213> 人工序列

<220>
<223> 合成的 DNA 聚合酶结构域基序

<220>
<221> 变体
<222> (11)...(11)
<223> Xaa = 除 Arg 以外的任何氨基酸

<400> 10
Gly Tyr Val Glu Thr Leu Phe Gly Arg Arg Xaa Tyr Val Pro
1 5 10

<210> 11
<211> 14
<212> PRT
<213> 人工序列

<220>
<223> 合成的 DNA 聚合酶结构域基序

<220>
<221> 变体
<222> (11)...(11)
<223> Xaa = Ala, Cys, Asp, Gly 或 Asn

<400> 11
Gly Tyr Val Glu Thr Leu Phe Gly Arg Arg Xaa Tyr Val Pro
1 5 10

<210> 12
<211> 34
<212> PRT
<213> 人工序列

<220>
<223> 来自栖热菌种 Z05 DNA 聚合酶 (Z05) 的聚合酶结构域的合成的区域

<400> 12
Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys Arg Gly Tyr Val
1 5 10 15
Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Asn Ala Arg
20 25 30
Val Lys

<210> 13
<211> 34
<212> PRT
<213> 人工序列

<220>
<223> 来自水生栖热菌 DNA 聚合酶 (Taq) 的聚合酶结构域的合成的区域

<400> 13

Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr Val
1 5 10 15
Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala Arg
20 25 30
Val Lys

<210> 14
<211> 34
<212> PRT
<213> 人工序列

<220>
<223> 来自丝状栖热菌 DNA 聚合酶 (Tfi) 的聚合酶结构域的合成的区域

<400> 14
Ala Trp Ile Ala Lys Thr Leu Glu Glu Gly Arg Lys Lys Gly Tyr Val
1 5 10 15
Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Asn Ala Arg
20 25 30
Val Lys

<210> 15
<211> 34
<212> PRT
<213> 人工序列

<220>
<223> 来自黄栖热菌 DNA 聚合酶 (Tfl) 的聚合酶结构域的合成的区域

<400> 15
Ala Trp Ile Glu Gly Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr Val
1 5 10 15
Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Asn Ala Arg
20 25 30
Val Lys

<210> 16
<211> 34
<212> PRT
<213> 人工序列

<220>
<223> 来自栖热菌种 Sps17 DNA 聚合酶 (Sps17) 的聚合酶结构域的合成的区域

<400> 16
Ala Trp Ile Ala Lys Thr Leu Glu Glu Gly Arg Lys Lys Gly Tyr Val
1 5 10 15
Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Asn Ala Arg
20 25 30
Val Lys

<210> 17
<211> 34
<212> PRT
<213> 人工序列

<220>
<223> 来自嗜热栖热菌 DNA 聚合酶 (Tth) 的聚合酶结构域的合成的区域

<400> 17
Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys Arg Gly Tyr Val
1 5 10 15
Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Asn Ala Arg
20 25 30
Val Lys

<210> 18
<211> 34
<212> PRT
<213> 人工序列

<220>
<223> 来自 Thermus caldophilus DNA 聚合酶 (Tca) 的聚合酶结构域的合成的区域

<400> 18
Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys Arg Gly Tyr Val
1 5 10 15
Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Asn Ala Arg
20 25 30
Val Lys

<210> 19
<211> 34
<212> PRT
<213> 人工序列

<220>
<223> 来自海栖热袍菌 DNA 聚合酶 (Tma) 的聚合酶结构域的合成的区域

<400> 19
Asp Tyr Ile Gln Arg Val Val Ser Glu Ala Lys Glu Lys Gly Tyr Val
1 5 10 15
Arg Thr Leu Phe Gly Arg Lys Arg Asp Ile Pro Gln Leu Met Ala Arg
20 25 30
Asp Arg

<210> 20
<211> 34
<212> PRT
<213> 人工序列

<220>
<223> 来自那不勒斯栖热袍菌 DNA 聚合酶 (Tne) 的聚合酶结构域的合成的区域

<400> 20
Ser Tyr Ile Gln Gln Val Val Ala Glu Ala Lys Glu Lys Gly Tyr Val
1 5 10 15
Arg Thr Leu Phe Gly Arg Lys Arg Asp Ile Pro Gln Leu Met Ala Arg
20 25 30
Asp Lys

<210> 21
<211> 34
<212> PRT
<213> 人工序列

<220>
<223> 来自非洲栖热腔菌 DNA 聚合酶 (Taf) 的聚合酶结构域的合成的区域

<400> 21
Glu Tyr Leu Lys Arg Met Lys Asp Glu Ala Arg Lys Lys Gly Tyr Val
1 5 10 15
Thr Thr Leu Phe Gly Arg Arg Arg Tyr Ile Pro Gln Leu Arg Ser Lys
20 25 30
Asn Gly

<210> 22

<400> 22
000

<210> 23
<211> 34
<212> PRT
<213> 人工序列

<220>
<223> 来自耐放射异常球菌 DNA 聚合酶 (Dra) 的聚合酶结构域的合成的区域

<400> 23
Arg Tyr Ile Asn His Thr Leu Asp Phe Gly Arg Thr His Gly Tyr Val
1 5 10 15
Glu Thr Leu Tyr Gly Arg Arg Arg Tyr Val Pro Gly Leu Ser Ser Arg
20 25 30
Asn Arg

<210> 24
<211> 34
<212> PRT
<213> 人工序列

<220>
<223> 来自嗜热脂肪芽孢杆菌 DNA 聚合酶 (Bst) 的聚合酶结构域的合成的区域

<400> 24
Gln Tyr Met Asp Asn Ile Val Gln Glu Ala Lys Gln Lys Gly Tyr Val
1 5 10 15
Thr Thr Leu Leu His Arg Arg Arg Tyr Leu Pro Asp Ile Thr Ser Arg
20 25 30
Asn Phe

<210> 25
<211> 34
<212> PRT
<213> 人工序列

<220>
<223> 来自热坚芽孢杆菌 DNA 聚合酶 (Bca) 的聚合酶结构域的合成的区域

<400> 25
Arg Tyr Met Glu Asn Ile Val Gln Glu Ala Lys Gln Lys Gly Tyr Val
1 5 10 15
Thr Thr Leu Leu His Arg Arg Arg Tyr Leu Pro Asp Ile Thr Ser Arg
20 25 30
Asn Phe

<210> 26
<211> 14
<212> PRT
<213> 人工序列

<220>

<223> 来自细菌 DNA 聚合酶的聚合酶结构域的区域的合成的天然的共有基序

<220>
<221> 变体
<222> (4)...(4)
<223> Xaa = Glu, Arg 或 Thr

<220>
<221> 变体
<222> (7)...(7)
<223> Xaa = Phe, Tyr 或 Leu

<220>
<221> 变体
<222> (8)...(8)
<223> Xaa = Gly 或 His

<220>
<221> 变体
<222> (10)...(10)
<223> Xaa = Arg 或 Lys

<220>
<221> 变体
<222> (12)...(12)
<223> Xaa = Tyr 或 Asp

<220>
<221> 变体
<222> (13)...(13)
<223> Xaa = Val, Ile 或 Leu

<400> 26
Gly Tyr Val Xaa Thr Leu Xaa Xaa Arg Xaa Arg Xaa Xaa Pro
1 5 10

<210> 27
<211> 13

<212> PRT
<213> 人工序列

<220>
<223> 对应于 Z05 的 D580X 突变的合成的聚合酶基序,修饰的 Z05 D580 基序

<220>
<221> 变体
<222> (7)...(7)
<223> Xaa = Ser 或 Thr

<220>
<221> 变体
<222> (8)...(8)
<223> Xaa = 除 Asp 或 Glu 以外的任何氨基酸

<400> 27
Thr Gly Arg Leu Ser Ser Xaa Xaa Pro Asn Leu Gln Asn
1 5 10

<210> 28
<211> 8
<212> PRT
<213> 人工序列

<220>
<223> 合成的保守的 DNA 聚合酶活性位点基序 A

<400> 28
Asp Tyr Ser Gln Ile Glu Leu Arg
1 5

<210> 29
<211> 893
<212> PRT
<213> 人工序列

<220>

<223> 源自栖热菌种 Z05 的 N- 末端的 5'- 核酸酶结构域和海栖热袍菌 DNA 聚合酶的 C- 末端的 3'-5' 核酸外切酶和聚合酶结构域的合成的嵌合的 CS5 DNA 聚合酶

<400> 29

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

Ala Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly
275 280 285
Leu Leu Glu Glu Ser Glu Pro Val Gly Tyr Arg Ile Val Lys Asp Leu
290 295 300
Val Glu Phe Glu Lys Leu Ile Glu Lys Leu Arg Glu Ser Pro Ser Phe
305 310 315 320
Ala Ile Asp Leu Glu Thr Ser Ser Leu Asp Pro Phe Asp Cys Asp Ile
325 330 335
Val Gly Ile Ser Val Ser Phe Lys Pro Lys Glu Ala Tyr Tyr Ile Pro
340 345 350
Leu His His Arg Asn Ala Gln Asn Leu Asp Glu Lys Glu Val Leu Lys
355 360 365
Lys Leu Lys Glu Ile Leu Glu Asp Pro Gly Ala Lys Ile Val Gly Gln
370 375 380
Asn Leu Lys Phe Asp Tyr Lys Val Leu Met Val Lys Gly Val Glu Pro
385 390 395 400
Val Pro Pro Tyr Phe Asp Thr Met Ile Ala Ala Tyr Leu Leu Glu Pro
405 410 415
Asn Glu Lys Lys Phe Asn Leu Asp Asp Leu Ala Leu Lys Phe Leu Gly
420 425 430
Tyr Lys Met Thr Ser Tyr Gln Glu Leu Met Ser Phe Ser Phe Pro Leu
435 440 445
Phe Gly Phe Ser Phe Ala Asp Val Pro Val Glu Lys Ala Ala Asn Tyr
450 455 460
Ser Cys Glu Asp Ala Asp Ile Thr Tyr Arg Leu Tyr Lys Thr Leu Ser
465 470 475 480
Leu Lys Leu His Glu Ala Asp Leu Glu Asn Val Phe Tyr Lys Ile Glu
485 490 495
Met Pro Leu Val Asn Val Leu Ala Arg Met Glu Leu Asn Gly Val Tyr
500 505 510
Val Asp Thr Glu Phe Leu Lys Lys Leu Ser Glu Glu Tyr Gly Lys Lys
515 520 525
Leu Glu Glu Leu Ala Glu Glu Ile Tyr Arg Ile Ala Gly Glu Pro Phe
530 535 540
Asn Ile Asn Ser Pro Lys Gln Val Ser Arg Ile Leu Phe Glu Lys Leu
545 550 555 560
Gly Ile Lys Pro Arg Gly Lys Thr Thr Lys Thr Gly Asp Tyr Ser Thr
565 570 575
Arg Ile Glu Val Leu Glu Glu Leu Ala Gly Glu His Glu Ile Ile Pro

580 585 590

Leu Ile Leu Glu Tyr Arg Lys Ile Gln Lys Leu Lys Ser Thr Tyr Ile

595 600 605

Asp Ala Leu Pro Lys Met Val Asn Pro Lys Thr Gly Arg Ile His Ala

610 615 620

Ser Phe Asn Gln Thr Gly Thr Ala Thr Gly Arg Leu Ser Ser Ser Asp

625 630 635 640

Pro Asn Leu Gln Asn Leu Pro Thr Lys Ser Glu Glu Gly Lys Glu Ile

645 650 655

Arg Lys Ala Ile Val Pro Gln Asp Pro Asn Trp Trp Ile Val Ser Ala

660 665 670

Asp Tyr Ser Gln Ile Glu Leu Arg Ile Leu Ala His Leu Ser Gly Asp

675 680 685

Glu Asn Leu Leu Arg Ala Phe Glu Glu Gly Ile Asp Val His Thr Leu

690 695 700

Thr Ala Ser Arg Ile Phe Asn Val Lys Pro Glu Glu Val Thr Glu Glu

705 710 715 720

Met Arg Arg Ala Gly Lys Met Val Asn Phe Ser Ile Ile Tyr Gly Val

725 730 735

Thr Pro Tyr Gly Leu Ser Val Arg Leu Gly Val Pro Val Lys Glu Ala

740 745 750

Glu Lys Met Ile Val Asn Tyr Phe Val Leu Tyr Pro Lys Val Arg Asp

755 760 765

Tyr Ile Gln Arg Val Val Ser Glu Ala Lys Glu Lys Gly Tyr Val Arg

770 775 780

Thr Leu Phe Gly Arg Lys Arg Asp Ile Pro Gln Leu Met Ala Arg Asp

785 790 795 800

Arg Asn Thr Gln Ala Glu Gly Glu Arg Ile Ala Ile Asn Thr Pro Ile

805 810 815

Gln Gly Thr Ala Ala Asp Ile Ile Lys Leu Ala Met Ile Glu Ile Asp

820 825 830

Arg Glu Leu Lys Glu Arg Lys Met Arg Ser Lys Met Ile Ile Gln Val

835 840 845

His Asp Glu Leu Val Phe Glu Val Pro Asn Glu Glu Lys Asp Ala Leu

850 855 860

Val Glu Leu Val Lys Asp Arg Met Thr Asn Val Val Lys Leu Ser Val

865 870 875 880

Pro Leu Glu Val Asp Val Thr Ile Gly Lys Thr Trp Ser

885 890

<210> 30
<211> 893
<212> PRT
<213> 人工序列

<220>
<223> 源自栖热菌种 Z05 的 N- 末端的 5'- 核酸酶结构域和海栖热袍菌 DNA 聚合酶的 C- 末端的 3'-5' 核酸外切酶和聚合酶结构域的合成的嵌合的 CS6 DNA 聚合酶

<400> 30

```
Met Lys Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu
 1               5                  10                  15
Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu Lys Gly
            20                  25                  30
Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala
        35                  40                  45
Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Tyr Lys Ala Val Phe
    50                  55                  60
Val Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Glu
65                  70                  75                  80
Ala Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln
                85                  90                  95
Leu Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Phe Thr Arg Leu
            100                 105                 110
Glu Val Pro Gly Phe Glu Ala Asp Asp Val Leu Ala Thr Leu Ala Lys
        115                 120                 125
Lys Ala Glu Arg Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Arg
    130                 135                 140
Asp Leu Tyr Gln Leu Val Ser Asp Arg Val Ala Val Leu His Pro Glu
145                 150                 155                 160
Gly His Leu Ile Thr Pro Glu Trp Leu Trp Glu Lys Tyr Gly Leu Lys
                165                 170                 175
Pro Glu Gln Trp Val Asp Phe Arg Ala Leu Val Gly Asp Pro Ser Asp
            180                 185                 190
Asn Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Leu Lys Leu
        195                 200                 205
Leu Lys Glu Trp Gly Ser Leu Glu Asn Ile Leu Lys Asn Leu Asp Arg
```

210 215 220

Val Lys Pro Glu Ser Val Arg Glu Arg Ile Lys Ala His Leu Glu Asp
225 230 235 240

Leu Lys Leu Ser Leu Glu Leu Ser Arg Val Arg Ser Asp Leu Pro Leu
245 250 255

Glu Val Asp Phe Ala Arg Arg Arg Glu Pro Asp Arg Glu Gly Leu Arg
260 265 270

Ala Phe Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly
275 280 285

Leu Leu Glu Glu Ser Glu Pro Val Gly Tyr Arg Ile Val Lys Asp Leu
290 295 300

Val Glu Phe Glu Lys Leu Ile Glu Lys Leu Arg Glu Ser Pro Ser Phe
305 310 315 320

Ala Ile Ala Leu Ala Thr Ser Ser Leu Asp Pro Phe Asp Cys Asp Ile
325 330 335

Val Gly Ile Ser Val Ser Phe Lys Pro Lys Glu Ala Tyr Tyr Ile Pro
340 345 350

Leu His His Arg Asn Ala Gln Asn Leu Asp Glu Lys Glu Val Leu Lys
355 360 365

Lys Leu Lys Glu Ile Leu Glu Asp Pro Gly Ala Lys Ile Val Gly Gln
370 375 380

Asn Leu Lys Phe Asp Tyr Lys Val Leu Met Val Lys Gly Val Glu Pro
385 390 395 400

Val Pro Pro Tyr Phe Asp Thr Met Ile Ala Ala Tyr Leu Leu Glu Pro
405 410 415

Asn Glu Lys Lys Phe Asn Leu Asp Asp Leu Ala Leu Lys Phe Leu Gly
420 425 430

Tyr Lys Met Thr Ser Tyr Gln Glu Leu Met Ser Phe Ser Phe Pro Leu
435 440 445

Phe Gly Phe Ser Phe Ala Asp Val Pro Val Glu Lys Ala Ala Asn Tyr
450 455 460

Ser Cys Glu Asp Ala Asp Ile Thr Tyr Arg Leu Tyr Lys Thr Leu Ser
465 470 475 480

Leu Lys Leu His Glu Ala Asp Leu Glu Asn Val Phe Tyr Lys Ile Glu
485 490 495

Met Pro Leu Val Asn Val Leu Ala Arg Met Glu Leu Asn Gly Val Tyr
500 505 510

Val Asp Thr Glu Phe Leu Lys Lys Leu Ser Glu Glu Tyr Gly Lys Lys
515 520 525

Leu Glu Glu Leu Ala Glu Glu Ile Tyr Arg Ile Ala Gly Glu Pro Phe
530 535 540
Asn Ile Asn Ser Pro Lys Gln Val Ser Arg Ile Leu Phe Glu Lys Leu
545 550 555 560
Gly Ile Lys Pro Arg Gly Lys Thr Thr Lys Thr Gly Asp Tyr Ser Thr
565 570 575
Arg Ile Glu Val Leu Glu Glu Leu Ala Gly Glu His Glu Ile Ile Pro
580 585 590
Leu Ile Leu Glu Tyr Arg Lys Ile Gln Lys Leu Lys Ser Thr Tyr Ile
595 600 605
Asp Ala Leu Pro Lys Met Val Asn Pro Lys Thr Gly Arg Ile His Ala
610 615 620
Ser Phe Asn Gln Thr Gly Thr Ala Thr Gly Arg Leu Ser Ser Ser Asp
625 630 635 640
Pro Asn Leu Gln Asn Leu Pro Thr Lys Ser Glu Glu Gly Lys Glu Ile
645 650 655
Arg Lys Ala Ile Val Pro Gln Asp Pro Asn Trp Trp Ile Val Ser Ala
660 665 670
Asp Tyr Ser Gln Ile Glu Leu Arg Ile Leu Ala His Leu Ser Gly Asp
675 680 685
Glu Asn Leu Leu Arg Ala Phe Glu Glu Gly Ile Asp Val His Thr Leu
690 695 700
Thr Ala Ser Arg Ile Phe Asn Val Lys Pro Glu Glu Val Thr Glu Glu
705 710 715 720
Met Arg Arg Ala Gly Lys Met Val Asn Phe Ser Ile Ile Tyr Gly Val
725 730 735
Thr Pro Tyr Gly Leu Ser Val Arg Leu Gly Val Pro Val Lys Glu Ala
740 745 750
Glu Lys Met Ile Val Asn Tyr Phe Val Leu Tyr Pro Lys Val Arg Asp
755 760 765
Tyr Ile Gln Arg Val Val Ser Glu Ala Lys Glu Lys Gly Tyr Val Arg
770 775 780
Thr Leu Phe Gly Arg Lys Arg Asp Ile Pro Gln Leu Met Ala Arg Asp
785 790 795 800
Arg Asn Thr Gln Ala Glu Gly Glu Arg Ile Ala Ile Asn Thr Pro Ile
805 810 815
Gln Gly Thr Ala Ala Asp Ile Ile Lys Leu Ala Met Ile Glu Ile Asp
820 825 830
Arg Glu Leu Lys Glu Arg Lys Met Arg Ser Lys Met Ile Ile Gln Val

835 840 845

His Asp Glu Leu Val Phe Glu Val Pro Asn Glu Glu Lys Asp Ala Leu

850 855 860

Val Glu Leu Val Lys Asp Arg Met Thr Asn Val Val Lys Leu Ser Val

865 870 875 880

Pro Leu Glu Val Asp Val Thr Ile Gly Lys Thr Trp Ser

885 890

<210> 31

<211> 21

<212> DNA

<213> 人工序列

<220>

<223> 合成的易错（诱变）PCR 扩增正向引物

<400> 31

ctacctcctg gacccctcca a 21

<210> 32

<211> 25

<212> DNA

<213> 人工序列

<220>

<223> 合成的易错（诱变）PCR 扩增反向引物

<400> 32

ataaccaact ggtagtggcg tgtaa 25

<210> 33

<211> 1491

<212> DNA

<213> 人工序列

<220>

<223> 通过易错（诱变）PCR 在 BlpI 和 BglII 限制性位点之间扩增的编码 Z05 D580G DNA 聚合酶的聚合酶结构域的合成的扩增子

<400> 33
ctacctcctg gacccctcca acaccacccc cgaggggggtg gcccggcgct acgggggggga 60
gtggacggag gacgccgccc accgggccct cctcgctgag cggctccagc aaaacctctt 120
ggaacgcctc aagggagagg aaaagctcct ttggctctac caagaggtgg aaaagcccct 180
ctcccgggtc ctggcccaca tggaggccac cggggtaagg ctggacgtgg cctatctaaa 240
ggcccctttcc ctggagcttg cggaggagat tcgccgcctc gaggaggagg tcttccgcct 300
ggcgggccac cccttcaacc tgaactcccg tgaccagcta gagcgggtgc tctttgacga 360
gcttaggctt cccgccctgg gcaagacgca aaagacgggg aagcgctcca ccagcgccgc 420
ggtgctggag gccctcaggg aggcccaccc catcgtggag aagatcctcc agcaccggga 480
gctcaccaag ctcaagaaca cctacgtaga ccccctcccg ggcctcgtcc acccgaggac 540
gggccgcctc cacacccgct tcaaccagac agccacggcc acgggaaggc tctctagctc 600
cgggcccaac ctgcagaaca tccccatccg cacccccttg ggccagagga tccgccgggc 660
cttcgtggcc gaggcgggat gggcgttggt ggccctggac tatagccaga tagagctccg 720
ggtcctcgcc cacctctccg gggacgagaa cctgatcagg gtcttccagg aggggaagga 780
catccacacc cagaccgcaa gctggatgtt cggcgtctcc ccggaggccg tggacccccct 840
gatgcgccgg gcggccaaga cggtgaactt cggcgtcctc tacggcatgt ccgcccatag 900
gctctcccag gagcttgcca tcccctacga ggaggcggtg gcctttatag agcgctactt 960
ccaaagcttc cccaaggtgc gggcctggat agaaaagacc ctggaggagg ggaggaagcg 1020
gggctacgtg gaaaccctct tcggaagaag gcgctacgtg cccgacctca acgcccgggt 1080
gaagagcgtc agggaggccg cggagcgcat ggccttcaac atgcccgtcc agggcaccgc 1140
cgccgacctc atgaagctcg ccatggtgaa gctcttcccc cacctccggg agatgggggc 1200
ccgcatgctc ctccaggtcc acgacgagct cctcctggag gcccccccaag cgcgggccga 1260
ggaggtggcg gctttggcca aggaggccat ggagaaggcc tatcccctcg ccgtgccccct 1320
ggaggtggag gtggggatcg gggaggactg gctttccgcc aagggctgat atcagatctc 1380
cctgattatg cgtcagtcta tgaagaaaaa tcgtatacag atggacgaag agagaatcct 1440
tgtgaattta acagagggta tagggattac acgccactac cagttggtta t 1491

<210> 34
<211> 110
<212> DNA
<213> 人工序列

<220>
<223> 合成的野生型 BRAF V600K 目标多核苷酸

<400> 34
agtaaaaata ggtgattttg gtctagctac agtgaaatct cgatggagtg ggtcccatca 60
gtttgaacag ttgtctggat ccattttgtg gatggtaaga attgaggcta 110

<210> 35
<211> 110
<212> DNA
<213> 人工序列

<220>
<223> 合成的突变型 BRAF V600K 目标多核苷酸

<400> 35
agtaaaaata ggtgattttg gtctagctac aaggaaatct cgatggagtg ggtcccatca 60
gtttgaacag ttgtctggat ccattttgtg gatggtaaga attgaggcta 110

<210> 36
<211> 921
<212> PRT
<213> 耐放射异常球菌

<220>
<223> 耐放射异常球菌 DNA 聚合酶 (Dra)

<400> 36
Met Ala Asp Ala Ser Pro Asp Pro Ser Lys Pro Asp Ala Leu Val Leu
1 5 10 15
Ile Asp Gly His Ala Leu Ala Phe Arg Ser Tyr Phe Ala Leu Pro Pro
20 25 30
Leu Asn Asn Ser Lys Gly Glu Met Thr Asp Ala Ile Val Gly Phe Met
35 40 45
Lys Leu Leu Leu Arg Leu Ala Arg Gln Lys Ser Asn Gln Val Ile Val
50 55 60
Val Phe Asp Pro Pro Val Lys Thr Leu Arg His Glu Gln Tyr Glu Gly
65 70 75 80
Tyr Lys Ser Gly Arg Ala Gln Thr Pro Glu Asp Leu Arg Gly Gln Ile
85 90 95
Asn Arg Ile Arg Ala Leu Val Asp Ala Leu Gly Phe Pro Arg Leu Glu
100 105 110
Glu Pro Gly Tyr Glu Ala Asp Asp Val Ile Ala Ser Leu Thr Arg Met
115 120 125
Ala Glu Gly Lys Gly Tyr Glu Val Arg Ile Val Thr Ser Asp Arg Asp

130 135 140

Ala Tyr Gln Leu Leu Asp Glu His Val Lys Val Ile Ala Asn Asp Phe
145 150 155 160

Ser Leu Ile Gly Pro Ala Gln Val Glu Glu Lys Tyr Gly Val Thr Val
165 170 175

Arg Gln Trp Val Asp Tyr Arg Ala Leu Thr Gly Asp Ala Ser Asp Asn
180 185 190

Ile Pro Gly Ala Lys Gly Ile Gly Pro Lys Thr Ala Ala Lys Leu Leu
195 200 205

Gln Glu Tyr Gly Thr Leu Glu Lys Val Tyr Glu Ala Ala His Ala Gly
210 215 220

Thr Leu Lys Pro Asp Gly Thr Arg Lys Lys Leu Leu Asp Ser Glu Glu
225 230 235 240

Asn Val Lys Phe Ser His Asp Leu Ser Cys Met Val Thr Asp Leu Pro
245 250 255

Leu Asp Ile Glu Phe Gly Val Arg Arg Leu Pro Asp Asn Pro Leu Val
260 265 270

Thr Glu Asp Leu Leu Thr Glu Leu Glu Leu His Ser Leu Arg Pro Met
275 280 285

Ile Leu Gly Leu Asn Gly Pro Glu Gln Asp Gly His Ala Pro Asp Asp
290 295 300

Leu Leu Glu Arg Glu His Ala Gln Thr Pro Glu Glu Asp Glu Ala Ala
305 310 315 320

Ala Leu Pro Ala Phe Ser Ala Pro Glu Leu Ala Glu Trp Gln Thr Pro
325 330 335

Ala Glu Gly Ala Val Trp Gly Tyr Val Leu Ser Arg Glu Asp Asp Leu
340 345 350

Thr Ala Ala Leu Leu Ala Ala Ala Thr Phe Glu Asp Gly Val Ala Arg
355 360 365

Pro Ala Arg Val Ser Glu Pro Asp Glu Trp Ala Gln Ala Glu Ala Pro
370 375 380

Glu Asn Leu Phe Gly Glu Leu Leu Pro Ser Asp Lys Pro Leu Thr Lys
385 390 395 400

Lys Glu Gln Lys Ala Leu Glu Lys Ala Gln Lys Asp Ala Glu Lys Ala
405 410 415

Arg Ala Lys Leu Arg Glu Gln Phe Pro Ala Thr Val Asp Glu Ala Glu
420 425 430

Phe Val Gly Gln Arg Thr Val Thr Ala Ala Ala Ala Ala Lys Ala Leu Ala
435 440 445

Ala His Leu Ser Val Arg Gly Thr Val Val Glu Pro Gly Asp Asp Pro
450 455 460
Leu Leu Tyr Ala Tyr Leu Leu Asp Pro Ala Asn Thr Asn Met Pro Val
465 470 475 480
Val Ala Lys Arg Tyr Leu Asp Arg Glu Trp Pro Ala Asp Ala Pro Thr
485 490 495
Arg Ala Ala Ile Thr Gly His Leu Val Arg Glu Leu Pro Pro Leu Leu
500 505 510
Asp Asp Ala Arg Arg Lys Met Tyr Asp Glu Met Glu Lys Pro Leu Ser
515 520 525
Gly Val Leu Gly Arg Met Glu Val Arg Gly Val Gln Val Asp Ser Asp
530 535 540
Phe Leu Gln Thr Leu Ser Ile Gln Ala Gly Val Arg Leu Ala Asp Leu
545 550 555 560
Glu Ser Gln Ile His Glu Tyr Ala Gly Glu Glu Phe His Ile Arg Ser
565 570 575
Pro Lys Gln Leu Glu Thr Val Leu Tyr Asp Lys Leu Glu Leu Ala Ser
580 585 590
Ser Lys Lys Thr Lys Leu Thr Gly Gln Arg Ser Thr Ala Val Ser Ala
595 600 605
Leu Glu Pro Leu Arg Asp Ala His Pro Ile Ile Pro Leu Val Leu Glu
610 615 620
Phe Arg Glu Leu Asp Lys Leu Arg Gly Thr Tyr Leu Asp Pro Ile Pro
625 630 635 640
Asn Leu Val Asn Pro His Thr Gly Arg Leu His Thr Thr Phe Ala Gln
645 650 655
Thr Ala Val Ala Thr Gly Arg Leu Ser Ser Leu Asn Pro Asn Leu Gln
660 665 670
Asn Ile Pro Ile Arg Ser Glu Leu Gly Arg Glu Ile Arg Lys Gly Phe
675 680 685
Ile Ala Glu Asp Gly Phe Thr Leu Ile Ala Ala Asp Tyr Ser Gln Ile
690 695 700
Glu Leu Arg Leu Leu Ala His Ile Ala Asp Asp Pro Leu Met Gln Gln
705 710 715 720
Ala Phe Val Glu Gly Ala Asp Ile His Arg Arg Thr Ala Ala Gln Val
725 730 735
Leu Gly Leu Asp Glu Ala Thr Val Asp Ala Asn Gln Arg Arg Ala Ala
740 745 750
Lys Thr Val Asn Phe Gly Val Leu Tyr Gly Met Ser Ala His Arg Leu

755 760 765

Ser Asn Asp Leu Gly Ile Pro Tyr Ala Glu Ala Ala Thr Phe Ile Glu
770 775 780

Ile Tyr Phe Ala Thr Tyr Pro Gly Ile Arg Arg Tyr Ile Asn His Thr
785 790 795 800

Leu Asp Phe Gly Arg Thr His Gly Tyr Val Glu Thr Leu Tyr Gly Arg
805 810 815

Arg Arg Tyr Val Pro Gly Leu Ser Ser Arg Asn Arg Val Gln Arg Glu
820 825 830

Ala Glu Glu Arg Leu Ala Tyr Asn Met Pro Ile Gln Gly Thr Ala Ala
835 840 845

Asp Ile Met Lys Leu Ala Met Val Gln Leu Asp Pro Gln Leu Asp Ala
850 855 860

Ile Gly Ala Arg Met Leu Leu Gln Val His Asp Glu Leu Leu Ile Glu
865 870 875 880

Ala Pro Leu Asp Lys Ala Glu Gln Val Ala Ala Leu Thr Lys Lys Val
885 890 895

Met Glu Asn Val Val Gln Leu Lys Val Pro Leu Ala Val Glu Val Gly
900 905 910

Thr Gly Pro Asn Trp Phe Asp Thr Lys
915 920

<210> 37
<211> 892
<212> PRT
<213> 非洲栖热腔菌

<220>
<223> 非洲栖热腔菌 DNA 聚合酶 (Taf)

<400> 37

Met Gly Lys Met Phe Leu Phe Asp Gly Thr Gly Leu Val Tyr Arg Ala
1 5 10 15

Phe Tyr Ala Ile Asp Gln Ser Leu Gln Thr Ser Ser Gly Leu His Thr
20 25 30

Asn Ala Val Tyr Gly Leu Thr Lys Met Leu Ile Lys Phe Leu Lys Glu
35 40 45

His Ile Ser Ile Gly Lys Asp Ala Cys Val Phe Val Leu Asp Ser Lys

50 55 60

Gly Gly Ser Lys Lys Arg Lys Asp Ile Leu Glu Thr Tyr Lys Ala Asn
65 70 75 80

Arg Pro Ser Thr Pro Asp Leu Leu Leu Glu Gln Ile Pro Tyr Val Glu
85 90 95

Glu Leu Val Asp Ala Leu Gly Ile Lys Val Leu Lys Ile Glu Gly Phe
100 105 110

Glu Ala Asp Asp Ile Ile Ala Thr Leu Ser Lys Lys Phe Glu Ser Asp
115 120 125

Phe Glu Lys Val Asn Ile Ile Thr Gly Asp Lys Asp Leu Leu Gln Leu
130 135 140

Val Ser Asp Lys Val Phe Val Trp Arg Val Glu Arg Gly Ile Thr Asp
145 150 155 160

Leu Val Leu Tyr Asp Arg Asn Lys Val Ile Glu Lys Tyr Gly Ile Tyr
165 170 175

Pro Glu Gln Phe Lys Asp Tyr Leu Ser Leu Val Gly Asp Gln Ile Asp
180 185 190

Asn Ile Pro Gly Val Lys Gly Ile Gly Lys Lys Thr Ala Val Ser Leu
195 200 205

Leu Lys Lys Tyr Asn Ser Leu Glu Asn Val Leu Lys Asn Ile Asn Leu
210 215 220

Leu Thr Glu Lys Leu Arg Arg Leu Leu Glu Asp Ser Lys Glu Asp Leu
225 230 235 240

Gln Lys Ser Ile Glu Leu Val Glu Leu Ile Tyr Asp Val Pro Met Asp
245 250 255

Val Glu Lys Asp Glu Ile Ile Tyr Arg Gly Tyr Asn Pro Asp Lys Leu
260 265 270

Leu Lys Val Leu Lys Lys Tyr Glu Phe Ser Ser Ile Ile Lys Glu Leu
275 280 285

Asn Leu Gln Glu Lys Leu Glu Lys Glu Tyr Ile Leu Val Asp Asn Glu
290 295 300

Asp Lys Leu Lys Lys Leu Ala Glu Glu Ile Glu Lys Tyr Lys Thr Phe
305 310 315 320

Ser Ile Asp Thr Glu Thr Thr Ser Leu Asp Pro Phe Glu Ala Lys Leu
325 330 335

Val Gly Ile Ser Ile Ser Thr Met Glu Gly Lys Ala Tyr Tyr Ile Pro
340 345 350

Val Ser His Phe Gly Ala Lys Asn Ile Ser Lys Ser Leu Ile Asp Lys
355 360 365

Phe Leu Lys Gln Ile Leu Gln Glu Lys Asp Tyr Asn Ile Val Gly Gln
370 375 380
Asn Leu Lys Phe Asp Tyr Glu Ile Phe Lys Ser Met Gly Phe Ser Pro
385 390 395 400
Asn Val Pro His Phe Asp Thr Met Ile Ala Ala Tyr Leu Leu Asn Pro
405 410 415
Asp Glu Lys Arg Phe Asn Leu Glu Glu Leu Ser Leu Lys Tyr Leu Gly
420 425 430
Tyr Lys Met Ile Ser Phe Asp Glu Leu Val Asn Glu Asn Val Pro Leu
435 440 445
Phe Gly Asn Asp Phe Ser Tyr Val Pro Leu Glu Arg Ala Val Glu Tyr
450 455 460
Ser Cys Glu Asp Ala Asp Val Thr Tyr Arg Ile Phe Arg Lys Leu Gly
465 470 475 480
Arg Lys Ile Tyr Glu Asn Glu Met Glu Lys Leu Phe Tyr Glu Ile Glu
485 490 495
Met Pro Leu Ile Asp Val Leu Ser Glu Met Glu Leu Asn Gly Val Tyr
500 505 510
Phe Asp Glu Glu Tyr Leu Lys Glu Leu Ser Lys Lys Tyr Gln Glu Lys
515 520 525
Met Asp Gly Ile Lys Glu Lys Val Phe Glu Ile Ala Gly Glu Thr Phe
530 535 540
Asn Leu Asn Ser Ser Thr Gln Val Ala Tyr Ile Leu Phe Glu Lys Leu
545 550 555 560
Asn Ile Ala Pro Tyr Lys Lys Thr Ala Thr Gly Lys Phe Ser Thr Asn
565 570 575
Ala Glu Val Leu Glu Glu Leu Ser Lys Glu His Glu Ile Ala Lys Leu
580 585 590
Leu Leu Glu Tyr Arg Lys Tyr Gln Lys Leu Lys Ser Thr Tyr Ile Asp
595 600 605
Ser Ile Pro Leu Ser Ile Asn Arg Lys Thr Asn Arg Val His Thr Thr
610 615 620
Phe His Gln Thr Gly Thr Ser Thr Gly Arg Leu Ser Ser Ser Asn Pro
625 630 635 640
Asn Leu Gln Asn Leu Pro Thr Arg Ser Glu Glu Gly Lys Glu Ile Arg
645 650 655
Lys Ala Val Arg Pro Gln Arg Gln Asp Trp Trp Ile Leu Gly Ala Asp
660 665 670
Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Val Ser Lys Asp Glu

675 680 685

Asn Leu Leu Lys Ala Phe Lys Glu Asp Leu Asp Ile His Thr Ile Thr

690 695 700

Ala Ala Lys Ile Phe Gly Val Ser Glu Met Phe Val Ser Glu Gln Met

705 710 715 720

Arg Arg Val Gly Lys Met Val Asn Phe Ala Ile Ile Tyr Gly Val Ser

725 730 735

Pro Tyr Gly Leu Ser Lys Arg Ile Gly Leu Ser Val Ser Glu Thr Lys

740 745 750

Lys Ile Ile Asp Asn Tyr Phe Arg Tyr Tyr Lys Gly Val Phe Glu Tyr

755 760 765

Leu Lys Arg Met Lys Asp Glu Ala Arg Lys Lys Gly Tyr Val Thr Thr

770 775 780

Leu Phe Gly Arg Arg Arg Tyr Ile Pro Gln Leu Arg Ser Lys Asn Gly

785 790 795 800

Asn Arg Val Gln Glu Gly Glu Arg Ile Ala Val Asn Thr Pro Ile Gln

805 810 815

Gly Thr Ala Ala Asp Ile Ile Lys Ile Ala Met Ile Asn Ile His Asn

820 825 830

Arg Leu Lys Lys Glu Asn Leu Arg Ser Lys Met Ile Leu Gln Val His

835 840 845

Asp Glu Leu Val Phe Glu Val Pro Asp Asn Glu Leu Glu Ile Val Lys

850 855 860

Asp Leu Val Arg Asp Glu Met Glu Asn Ala Val Lys Leu Asp Val Pro

865 870 875 880

Leu Lys Val Asp Val Tyr Tyr Gly Lys Glu Trp Glu

885 890

<210> 38

<211> 893

<212> PRT

<213> 海栖热袍菌

<220>

<223> 海栖热袍菌 DNA 聚合酶 (Tma)

<400> 38

Met Ala Arg Leu Phe Leu Phe Asp Gly Thr Ala Leu Ala Tyr Arg Ala

1 5 10 15

Tyr Tyr Ala Leu Asp Arg Ser Leu Ser Thr Ser Thr Gly Ile Pro Thr
20 25 30

Asn Ala Thr Tyr Gly Val Ala Arg Met Leu Val Arg Phe Ile Lys Asp
35 40 45

His Ile Ile Val Gly Lys Asp Tyr Val Ala Val Ala Phe Asp Lys Lys
50 55 60

Ala Ala Thr Phe Arg His Lys Leu Leu Glu Thr Tyr Lys Ala Gln Arg
65 70 75 80

Pro Lys Thr Pro Asp Leu Leu Ile Gln Gln Leu Pro Tyr Ile Lys Lys
85 90 95

Leu Val Glu Ala Leu Gly Met Lys Val Leu Glu Val Glu Gly Tyr Glu
100 105 110

Ala Asp Asp Ile Ile Ala Thr Leu Ala Val Lys Gly Leu Pro Leu Phe
115 120 125

Asp Glu Ile Phe Ile Val Thr Gly Asp Lys Asp Met Leu Gln Leu Val
130 135 140

Asn Glu Lys Ile Lys Val Trp Arg Ile Val Lys Gly Ile Ser Asp Leu
145 150 155 160

Glu Leu Tyr Asp Ala Gln Lys Val Lys Glu Lys Tyr Gly Val Glu Pro
165 170 175

Gln Gln Ile Pro Asp Leu Leu Ala Leu Thr Gly Asp Glu Ile Asp Asn
180 185 190

Ile Pro Gly Val Thr Gly Ile Gly Glu Lys Thr Ala Val Gln Leu Leu
195 200 205

Glu Lys Tyr Lys Asp Leu Glu Asp Ile Leu Asn His Val Arg Glu Leu
210 215 220

Pro Gln Lys Val Arg Lys Ala Leu Leu Arg Asp Arg Glu Asn Ala Ile
225 230 235 240

Leu Ser Lys Lys Leu Ala Ile Leu Glu Thr Asn Val Pro Ile Glu Ile
245 250 255

Asn Trp Glu Glu Leu Arg Tyr Gln Gly Tyr Asp Arg Glu Lys Leu Leu
260 265 270

Pro Leu Leu Lys Glu Leu Glu Phe Ala Ser Ile Met Lys Glu Leu Gln
275 280 285

Leu Tyr Glu Glu Ser Glu Pro Val Gly Tyr Arg Ile Val Lys Asp Leu
290 295 300

Val Glu Phe Glu Lys Leu Ile Glu Lys Leu Arg Glu Ser Pro Ser Phe
305 310 315 320

Ala Ile Asp Leu Glu Thr Ser Ser Leu Asp Pro Phe Asp Cys Asp Ile
325 330 335
Val Gly Ile Ser Val Ser Phe Lys Pro Lys Glu Ala Tyr Tyr Ile Pro
340 345 350
Leu His His Arg Asn Ala Gln Asn Leu Asp Glu Lys Glu Val Leu Lys
355 360 365
Lys Leu Lys Glu Ile Leu Glu Asp Pro Gly Ala Lys Ile Val Gly Gln
370 375 380
Asn Leu Lys Phe Asp Tyr Lys Val Leu Met Val Lys Gly Val Glu Pro
385 390 395 400
Val Pro Pro Tyr Phe Asp Thr Met Ile Ala Ala Tyr Leu Leu Glu Pro
405 410 415
Asn Glu Lys Lys Phe Asn Leu Asp Asp Leu Ala Leu Lys Phe Leu Gly
420 425 430
Tyr Lys Met Thr Ser Tyr Gln Glu Leu Met Ser Phe Ser Phe Pro Leu
435 440 445
Phe Gly Phe Ser Phe Ala Asp Val Pro Val Glu Lys Ala Ala Asn Tyr
450 455 460
Ser Cys Glu Asp Ala Asp Ile Thr Tyr Arg Leu Tyr Lys Thr Leu Ser
465 470 475 480
Leu Lys Leu His Glu Ala Asp Leu Glu Asn Val Phe Tyr Lys Ile Glu
485 490 495
Met Pro Leu Val Asn Val Leu Ala Arg Met Glu Leu Asn Gly Val Tyr
500 505 510
Val Asp Thr Glu Phe Leu Lys Lys Leu Ser Glu Glu Tyr Gly Lys Lys
515 520 525
Leu Glu Glu Leu Ala Glu Glu Ile Tyr Arg Ile Ala Gly Glu Pro Phe
530 535 540
Asn Ile Asn Ser Pro Lys Gln Val Ser Arg Ile Leu Phe Glu Lys Leu
545 550 555 560
Gly Ile Lys Pro Arg Gly Lys Thr Thr Lys Thr Gly Asp Tyr Ser Thr
565 570 575
Arg Ile Glu Val Leu Glu Glu Leu Ala Gly Glu His Glu Ile Ile Pro
580 585 590
Leu Ile Leu Glu Tyr Arg Lys Ile Gln Lys Leu Lys Ser Thr Tyr Ile
595 600 605
Asp Ala Leu Pro Lys Met Val Asn Pro Lys Thr Gly Arg Ile His Ala
610 615 620
Ser Phe Asn Gln Thr Gly Thr Ala Thr Gly Arg Leu Ser Ser Ser Asp

625 630 635 640
Pro Asn Leu Gln Asn Leu Pro Thr Lys Ser Glu Glu Gly Lys Glu Ile
645 650 655
Arg Lys Ala Ile Val Pro Gln Asp Pro Asn Trp Trp Ile Val Ser Ala
660 665 670
Asp Tyr Ser Gln Ile Glu Leu Arg Ile Leu Ala His Leu Ser Gly Asp
675 680 685
Glu Asn Leu Leu Arg Ala Phe Glu Glu Gly Ile Asp Val His Thr Leu
690 695 700
Thr Ala Ser Arg Ile Phe Asn Val Lys Pro Glu Glu Val Thr Glu Glu
705 710 715 720
Met Arg Arg Ala Gly Lys Met Val Asn Phe Ser Ile Ile Tyr Gly Val
725 730 735
Thr Pro Tyr Gly Leu Ser Val Arg Leu Gly Val Pro Val Lys Glu Ala
740 745 750
Glu Lys Met Ile Val Asn Tyr Phe Val Leu Tyr Pro Lys Val Arg Asp
755 760 765
Tyr Ile Gln Arg Val Val Ser Glu Ala Lys Glu Lys Gly Tyr Val Arg
770 775 780
Thr Leu Phe Gly Arg Lys Arg Asp Ile Pro Gln Leu Met Ala Arg Asp
785 790 795 800
Arg Asn Thr Gln Ala Glu Gly Glu Arg Ile Ala Ile Asn Thr Pro Ile
805 810 815
Gln Gly Thr Ala Ala Asp Ile Ile Lys Leu Ala Met Ile Glu Ile Asp
820 825 830
Arg Glu Leu Lys Glu Arg Lys Met Arg Ser Lys Met Ile Ile Gln Val
835 840 845
His Asp Glu Leu Val Phe Glu Val Pro Asn Glu Glu Lys Asp Ala Leu
850 855 860
Val Glu Leu Val Lys Asp Arg Met Thr Asn Val Val Lys Leu Ser Val
865 870 875 880
Pro Leu Glu Val Asp Val Thr Ile Gly Lys Thr Trp Ser
885 890

<210> 39
<211> 893
<212> PRT
<213> 那不勒斯栖热袍菌

<220>

<223> 那不勒斯栖热袍菌 DNA 聚合酶 (Tne)

<400> 39

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

Pro Ile Leu Lys Glu Leu Glu Phe Ala Ser Ile Met Lys Glu Leu Gln
275 280 285
Leu Tyr Glu Glu Ala Glu Pro Thr Gly Tyr Glu Ile Val Lys Asp His
290 295 300
Lys Thr Phe Glu Asp Leu Ile Glu Lys Leu Lys Glu Val Pro Ser Phe
305 310 315 320
Ala Leu Asp Leu Glu Thr Ser Ser Leu Asp Pro Phe Asn Cys Glu Ile
325 330 335
Val Gly Ile Ser Val Ser Phe Lys Pro Lys Thr Ala Tyr Tyr Ile Pro
340 345 350
Leu His His Arg Asn Ala His Asn Leu Asp Glu Thr Leu Val Leu Ser
355 360 365
Lys Leu Lys Glu Ile Leu Glu Asp Pro Ser Ser Lys Ile Val Gly Gln
370 375 380
Asn Leu Lys Tyr Asp Tyr Lys Val Leu Met Val Lys Gly Ile Ser Pro
385 390 395 400
Val Tyr Pro His Phe Asp Thr Met Ile Ala Ala Tyr Leu Leu Glu Pro
405 410 415
Asn Glu Lys Lys Phe Asn Leu Glu Asp Leu Ser Leu Lys Phe Leu Gly
420 425 430
Tyr Lys Met Thr Ser Tyr Gln Glu Leu Met Ser Phe Ser Ser Pro Leu
435 440 445
Phe Gly Phe Ser Phe Ala Asp Val Pro Val Asp Lys Ala Ala Glu Tyr
450 455 460
Ser Cys Glu Asp Ala Asp Ile Thr Tyr Arg Leu Tyr Lys Ile Leu Ser
465 470 475 480
Met Lys Leu His Glu Ala Glu Leu Glu Asn Val Phe Tyr Arg Ile Glu
485 490 495
Met Pro Leu Val Asn Val Leu Ala Arg Met Glu Phe Asn Trp Val Tyr
500 505 510
Val Asp Thr Glu Phe Leu Lys Lys Leu Ser Glu Glu Tyr Gly Lys Lys
515 520 525
Leu Glu Glu Leu Ala Glu Lys Ile Tyr Gln Ile Ala Gly Glu Pro Phe
530 535 540
Asn Ile Asn Ser Pro Lys Gln Val Ser Asn Ile Leu Phe Glu Lys Leu
545 550 555 560
Gly Ile Lys Pro Arg Gly Lys Thr Thr Lys Thr Gly Asp Tyr Ser Thr
565 570 575
Arg Ile Glu Val Leu Glu Glu Ile Ala Asn Glu His Glu Ile Val Pro

580 585 590

Leu Ile Leu Glu Phe Arg Lys Ile Leu Lys Leu Lys Ser Thr Tyr Ile

595 600 605

Asp Thr Leu Pro Lys Leu Val Asn Pro Lys Thr Gly Arg Phe His Ala

610 615 620

Ser Phe His Gln Thr Gly Thr Ala Thr Gly Arg Leu Ser Ser Ser Asp

625 630 635 640

Pro Asn Leu Gln Asn Leu Pro Thr Lys Ser Glu Glu Gly Lys Glu Ile

645 650 655

Arg Lys Ala Ile Val Pro Gln Asp Pro Asp Trp Trp Ile Val Ser Ala

660 665 670

Asp Tyr Ser Gln Ile Glu Leu Arg Ile Leu Ala His Leu Ser Gly Asp

675 680 685

Glu Asn Leu Val Lys Ala Phe Glu Glu Gly Ile Asp Val His Thr Leu

690 695 700

Thr Ala Ser Arg Ile Tyr Asn Val Lys Pro Glu Glu Val Asn Glu Glu

705 710 715 720

Met Arg Arg Val Gly Lys Met Val Asn Phe Ser Ile Ile Tyr Gly Val

725 730 735

Thr Pro Tyr Gly Leu Ser Val Arg Leu Gly Ile Pro Val Lys Glu Ala

740 745 750

Glu Lys Met Ile Ile Ser Tyr Phe Thr Leu Tyr Pro Lys Val Arg Ser

755 760 765

Tyr Ile Gln Gln Val Val Ala Glu Ala Lys Glu Lys Gly Tyr Val Arg

770 775 780

Thr Leu Phe Gly Arg Lys Arg Asp Ile Pro Gln Leu Met Ala Arg Asp

785 790 795 800

Lys Asn Thr Gln Ser Glu Gly Glu Arg Ile Ala Ile Asn Thr Pro Ile

805 810 815

Gln Gly Thr Ala Ala Asp Ile Ile Lys Leu Ala Met Ile Asp Ile Asp

820 825 830

Glu Glu Leu Arg Lys Arg Asn Met Lys Ser Arg Met Ile Ile Gln Val

835 840 845

His Asp Glu Leu Val Phe Glu Val Pro Asp Glu Glu Lys Glu Glu Leu

850 855 860

Val Asp Leu Val Lys Asn Lys Met Thr Asn Val Val Lys Leu Ser Val

865 870 875 880

Pro Leu Glu Val Asp Ile Ser Ile Gly Lys Ser Trp Ser

885 890

<210> 40
<211> 876
<212> PRT
<213> 嗜热脂肪芽孢杆菌

<220>
<223> 嗜热脂肪芽孢杆菌 DNA 聚合酶 (Bst)

<400> 40
Met Lys Asn Lys Leu Val Leu Ile Asp Gly Asn Ser Val Ala Tyr Arg
1 5 10 15
Ala Phe Phe Ala Leu Pro Leu Leu His Asn Asp Lys Gly Ile His Thr
20 25 30
Asn Ala Val Tyr Gly Phe Thr Met Met Leu Asn Lys Ile Leu Ala Glu
35 40 45
Glu Gln Pro Thr His Ile Leu Val Ala Phe Asp Ala Gly Lys Thr Thr
50 55 60
Phe Arg His Glu Thr Phe Gln Asp Tyr Lys Gly Gly Arg Gln Gln Thr
65 70 75 80
Pro Pro Glu Leu Ser Glu Gln Phe Pro Leu Leu Arg Glu Leu Leu Lys
85 90 95
Ala Tyr Arg Ile Pro Ala Tyr Glu Leu Asp His Tyr Glu Ala Asp Asp
100 105 110
Ile Ile Gly Thr Met Ala Ala Arg Ala Glu Arg Glu Gly Phe Ala Val
115 120 125
Lys Val Ile Ser Gly Asp Arg Asp Leu Thr Gln Leu Ala Ser Pro Gln
130 135 140
Val Thr Val Glu Ile Thr Lys Lys Gly Ile Thr Asp Ile Glu Ser Tyr
145 150 155 160
Thr Pro Glu Thr Val Val Glu Lys Tyr Gly Leu Thr Pro Glu Gln Ile
165 170 175
Val Asp Leu Lys Gly Leu Met Gly Asp Lys Ser Asp Asn Ile Pro Gly
180 185 190
Val Pro Gly Ile Gly Glu Lys Thr Ala Val Lys Leu Leu Lys Gln Phe
195 200 205
Gly Thr Val Glu Asn Val Leu Ala Ser Ile Asp Glu Ile Lys Gly Glu
210 215 220

Lys Leu Lys Glu Asn Leu Arg Gln Tyr Arg Asp Leu Ala Leu Leu Ser
225 230 235 240
Lys Gln Leu Ala Ala Ile Cys Arg Asp Ala Pro Val Glu Leu Thr Leu
245 250 255
Asp Asp Ile Val Tyr Lys Gly Glu Asp Arg Glu Lys Val Val Ala Leu
260 265 270
Phe Gln Glu Leu Gly Phe Gln Ser Phe Leu Asp Lys Met Ala Val Gln
275 280 285
Thr Asp Glu Gly Glu Lys Pro Leu Ala Gly Met Asp Phe Ala Ile Ala
290 295 300
Asp Ser Val Thr Asp Glu Met Leu Ala Asp Lys Ala Ala Leu Val Val
305 310 315 320
Glu Val Val Gly Asp Asn Tyr His His Ala Pro Ile Val Gly Ile Ala
325 330 335
Leu Ala Asn Glu Arg Gly Arg Phe Phe Leu Arg Pro Glu Thr Ala Leu
340 345 350
Ala Asp Pro Lys Phe Leu Ala Trp Leu Gly Asp Glu Thr Lys Lys Lys
355 360 365
Thr Met Phe Asp Ser Lys Arg Ala Ala Val Ala Leu Lys Trp Lys Gly
370 375 380
Ile Glu Leu Arg Gly Val Val Phe Asp Leu Leu Leu Ala Ala Tyr Leu
385 390 395 400
Leu Asp Pro Ala Gln Ala Ala Gly Asp Val Ala Ala Val Ala Lys Met
405 410 415
His Gln Tyr Glu Ala Val Arg Ser Asp Glu Ala Val Tyr Gly Lys Gly
420 425 430
Ala Lys Arg Thr Val Pro Asp Glu Pro Thr Leu Ala Glu His Leu Ala
435 440 445
Arg Lys Ala Ala Ala Ile Trp Ala Leu Glu Glu Pro Leu Met Asp Glu
450 455 460
Leu Arg Arg Asn Glu Gln Asp Arg Leu Leu Thr Glu Leu Glu Gln Pro
465 470 475 480
Leu Ala Gly Ile Leu Ala Asn Met Glu Phe Thr Gly Val Lys Val Asp
485 490 495
Thr Lys Arg Leu Glu Gln Met Gly Ala Glu Leu Thr Glu Gln Leu Gln
500 505 510
Ala Val Glu Arg Arg Ile Tyr Glu Leu Ala Gly Gln Glu Phe Asn Ile
515 520 525
Asn Ser Pro Lys Gln Leu Gly Thr Val Leu Phe Asp Lys Leu Gln Leu

530 535 540

Pro Val Leu Lys Lys Thr Lys Thr Gly Tyr Ser Thr Ser Ala Asp Val
545 550 555 560

Leu Glu Lys Leu Ala Pro His His Glu Ile Val Glu His Ile Leu His
565 570 575

Tyr Arg Gln Leu Gly Lys Leu Gln Ser Thr Tyr Ile Glu Gly Leu Leu
580 585 590

Lys Val Val His Pro Val Thr Gly Lys Val His Thr Met Phe Asn Gln
595 600 605

Ala Leu Thr Gln Thr Gly Arg Leu Ser Ser Val Glu Pro Asn Leu Gln
610 615 620

Asn Ile Pro Ile Arg Leu Glu Glu Gly Arg Lys Ile Arg Gln Ala Phe
625 630 635 640

Val Pro Ser Glu Pro Asp Trp Leu Ile Phe Ala Ala Asp Tyr Ser Gln
645 650 655

Ile Glu Leu Arg Val Leu Ala His Ile Ala Glu Asp Asp Asn Leu Ile
660 665 670

Glu Ala Phe Arg Arg Gly Leu Asp Ile His Thr Lys Thr Ala Met Asp
675 680 685

Ile Phe His Val Ser Glu Glu Asp Val Thr Ala Asn Met Arg Arg Gln
690 695 700

Ala Lys Ala Val Asn Phe Gly Ile Val Tyr Gly Ile Ser Asp Tyr Gly
705 710 715 720

Leu Ala Gln Asn Leu Asn Ile Thr Arg Lys Glu Ala Ala Glu Phe Ile
725 730 735

Glu Arg Tyr Phe Ala Ser Phe Pro Gly Val Lys Gln Tyr Met Asp Asn
740 745 750

Ile Val Gln Glu Ala Lys Gln Lys Gly Tyr Val Thr Thr Leu Leu His
755 760 765

Arg Arg Arg Tyr Leu Pro Asp Ile Thr Ser Arg Asn Phe Asn Val Arg
770 775 780

Ser Phe Ala Glu Arg Thr Ala Met Asn Thr Pro Ile Gln Gly Ser Ala
785 790 795 800

Ala Asp Ile Ile Lys Lys Ala Met Ile Asp Leu Ser Val Arg Leu Arg
805 810 815

Glu Glu Arg Leu Gln Ala Arg Leu Leu Leu Gln Val His Asp Glu Leu
820 825 830

Ile Leu Glu Ala Pro Lys Glu Glu Ile Glu Arg Leu Cys Arg Leu Val
835 840 845

Pro Glu Val Met Glu Gln Ala Val Ala Leu Arg Val Pro Leu Lys Val
850 855 860
Asp Tyr His Tyr Gly Pro Thr Trp Tyr Asp Ala Lys
865 870 875

<210> 41
<211> 877
<212> PRT
<213> 热坚芽孢杆菌

<220>
<223> 热坚芽孢杆菌 DNA 聚合酶 (Bca)

<400> 41
Met Lys Lys Lys Leu Val Leu Ile Asp Gly Ser Ser Val Ala Tyr Arg
1 5 10 15
Ala Phe Phe Ala Leu Pro Leu Leu His Asn Asp Lys Gly Ile His Thr
20 25 30
Asn Ala Val Tyr Gly Phe Thr Met Met Leu Asn Lys Ile Leu Ala Glu
35 40 45
Glu Glu Pro Thr His Met Leu Val Ala Phe Asp Ala Gly Lys Thr Thr
50 55 60
Phe Arg His Glu Ala Phe Gln Glu Tyr Lys Gly Gly Arg Gln Gln Thr
65 70 75 80
Pro Pro Glu Leu Ser Glu Gln Phe Pro Leu Leu Arg Glu Leu Leu Arg
85 90 95
Ala Tyr Arg Ile Pro Ala Tyr Glu Leu Glu Asn Tyr Glu Ala Asp Asp
100 105 110
Ile Ile Gly Thr Leu Ala Ala Arg Ala Glu Gln Glu Gly Phe Glu Val
115 120 125
Lys Val Ile Ser Gly Asp Arg Asp Leu Thr Gln Leu Ala Ser Pro His
130 135 140
Val Thr Val Asp Ile Thr Lys Lys Gly Ile Thr Asp Ile Glu Pro Tyr
145 150 155 160
Thr Pro Glu Ala Val Arg Glu Lys Tyr Gly Leu Thr Pro Glu Gln Ile
165 170 175
Val Asp Leu Lys Gly Leu Met Gly Asp Lys Ser Asp Asn Ile Pro Gly
180 185 190

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

500 505 510

Thr Val Glu Gln Arg Ile Tyr Glu Leu Ala Gly Gln Glu Phe Asn Ile

515 520 525

Asn Ser Pro Lys Gln Leu Gly Val Ile Leu Phe Glu Lys Leu Gln Leu

530 535 540

Pro Val Leu Lys Lys Ser Lys Thr Gly Tyr Ser Thr Ser Ala Asp Val

545 550 555 560

Leu Glu Lys Leu Ala Pro Tyr His Glu Ile Val Glu Asn Ile Leu Gln

565 570 575

His Tyr Arg Gln Leu Gly Lys Leu Gln Ser Thr Tyr Ile Glu Gly Leu

580 585 590

Leu Lys Val Val Arg Pro Asp Thr Lys Lys Val His Thr Ile Phe Asn

595 600 605

Gln Ala Leu Thr Gln Thr Gly Arg Leu Ser Ser Thr Glu Pro Asn Leu

610 615 620

Gln Asn Ile Pro Ile Arg Leu Glu Glu Gly Arg Lys Ile Arg Gln Ala

625 630 635 640

Phe Val Pro Ser Glu Ser Asp Trp Leu Ile Phe Ala Ala Asp Tyr Ser

645 650 655

Gln Ile Glu Leu Arg Val Leu Ala His Ile Ala Glu Asp Asp Asn Leu

660 665 670

Met Glu Ala Phe Arg Arg Asp Leu Asp Ile His Thr Lys Thr Ala Met

675 680 685

Asp Ile Phe Gln Val Ser Glu Asp Glu Val Thr Pro Asn Met Arg Arg

690 695 700

Gln Ala Lys Ala Val Asn Phe Gly Ile Val Tyr Gly Ile Ser Asp Tyr

705 710 715 720

Gly Leu Ala Gln Asn Leu Asn Ile Ser Arg Lys Glu Ala Ala Glu Phe

725 730 735

Ile Glu Arg Tyr Phe Glu Ser Phe Pro Gly Val Lys Arg Tyr Met Glu

740 745 750

Asn Ile Val Gln Glu Ala Lys Gln Lys Gly Tyr Val Thr Thr Leu Leu

755 760 765

His Arg Arg Arg Tyr Leu Pro Asp Ile Thr Ser Arg Asn Phe Asn Val

770 775 780

Arg Ser Phe Ala Glu Arg Met Ala Met Asn Thr Pro Ile Gln Gly Ser

785 790 795 800

Ala Ala Asp Ile Ile Lys Lys Ala Met Ile Asp Leu Asn Ala Arg Leu

805 810 815

Lys Glu Glu Arg Leu Gln Ala Arg Leu Leu Leu Gln Val His Asp Glu
820 825 830
Leu Ile Leu Glu Ala Pro Lys Glu Glu Met Glu Arg Leu Cys Arg Leu
835 840 845
Val Pro Glu Val Met Glu Gln Ala Val Thr Leu Arg Val Pro Leu Lys
850 855 860
Val Asp Tyr His Tyr Gly Ser Thr Trp Tyr Asp Ala Lys
865 870 875

<210> 42
<211> 14
<212> PRT
<213> 人工序列

<220>
<223> 合成的 DNA 聚合酶结构域基序

<220>
<221> 变体
<222> (4)...(4)
<223> Xaa = Glu, Arg 或 Thr

<220>
<221> 变体
<222> (7)...(7)
<223> Xaa = Phe, Tyr 或 Leu

<220>
<221> 变体
<222> (8)...(8)
<223> Xaa = Gly 或 His

<220>
<221> 变体
<222> (10)...(10)
<223> Xaa = Arg 或 Lys

<220>

<221> 变体
<222> (11)...(11)
<223> Xaa = Gly, Ala, Leu, Met, Trp, Pro, Ser, Thr, Phe,
Tyr, Cys, Asn, Gln, Asp, Glu, Lys, Val, Ile 或 His

<220>
<221> 变体
<222> (12)...(12)
<223> Xaa = Tyr 或 Asp

<220>
<221> 变体
<222> (13)...(13)
<223> Xaa = Val, Ile 或 Leu

<400> 42
Gly Tyr Val Xaa Thr Leu Xaa Xaa Arg Xaa Xaa Xaa Xaa Pro
1 5 10

```
                                  *
Z05   AWIEKTLEEGRKRGYVETLFGRRRYVPDLNARVK (SEQ ID NO:12)
Taq   AWIEKTLEEGRRRGYVETLFGRRRYVPDLEARVK (SEQ ID NO:13)
Tfi   AWIAKTLEEGRKKGYVETLFGRRRYVPDLNARVK (SEQ ID NO:14)
Tfl   AWIEGTLEEGRRRGYVETLFGRRRYVPDLNARVK (SEQ ID NO:15)
Sps17 AWIAKTLEEGRKKGYVETLFGRRRYVPDLNARVK (SEQ ID NO:16)
Tth   AWIEKTLEEGRKRGYVETLFGRRRYVPDLNARVK (SEQ ID NO:17)
Tca   AWIEKTLEEGRKRGYVETLFGRRRYVPDLNARVK (SEQ ID NO:18)
Tma   DYIQRVVSEAKEKGYVRTLFGRKRDIPQLMARDR (SEQ ID NO:19)
Tne   SYIQQVVAEAKEKGYVRTLFGRKRDIPQLMARDK (SEQ ID NO:20)
Taf   EYLKRMKDEARKKGYVTTLFGRRRYIPQLRSKNG (SEQ ID NO:21)
Dra   RYINHTLDFGRTHGYVETLYGRRRYVPGLSSRNR (SEQ ID NO:23)
Bst   QYMDNIVQEAKQKGYVTTLLHRRRYLPDITSRNF (SEQ ID NO:24)
Bca   RYMENIVQEAKQKGYVTTLLHRRRYLPDITSRNF (SEQ ID NO:25)
      ----------GYVX1TLX2X3RX4X5X6X7P------ (SEQ ID NO:26)
```

图 1

A. 与整个聚合酶I酶(对应于Z05的氨基酸1-834)的序列同一性

名称	Z05	Taq	Tfi	Tfl	Sps17	Tth	Tca	Dra	Tma	Tne	Taf	Bst	Bca
Z05		0.864	0.833	0.859	0.839	0.962	0.958	0.459	0.374	0.368	0.359	0.407	0.408
Taq	0.864		0.831	0.854	0.836	0.872	0.864	0.468	0.382	0.368	0.351	0.397	0.397
Tfi	0.833	0.831		0.82	0.991	0.829	0.824	0.45	0.371	0.375	0.353	0.405	0.397
Tfl	0.859	0.854	0.82		0.824	0.853	0.848	0.462	0.381	0.374	0.356	0.397	0.398
Sps17	0.839	0.836	0.991	0.824		0.835	0.83	0.452	0.375	0.377	0.355	0.407	0.399
Tth	0.962	0.872	0.829	0.853	0.835		0.989	0.463	0.373	0.367	0.358	0.406	0.406
Tca	0.958	0.864	0.824	0.848	0.83	0.989		0.46	0.371	0.365	0.356	0.404	0.404
Dra	0.459	0.468	0.45	0.462	0.452	0.463	0.46		0.334	0.325	0.314	0.338	0.339
Tma	0.374	0.382	0.371	0.381	0.375	0.373	0.371	0.334		0.854	0.567	0.37	0.377
Tne	0.368	0.368	0.375	0.374	0.377	0.367	0.365	0.325	0.854		0.558	0.377	0.376
Taf	0.359	0.351	0.353	0.356	0.355	0.358	0.356	0.314	0.567	0.558		0.356	0.364
Bst	*0.407*	*0.397*	*0.405*	*0.397*	*0.407*	*0.406*	*0.404*	*0.338*	*0.37*	*0.377*	*0.356*		*0.881*
Bca	0.408	0.397	0.397	0.398	0.399	0.406	0.404	0.339	0.377	0.376	0.364	0.881	

B. 与仅仅聚合酶亚结构域(对应于Z05的氨基酸420-834)的序列同一性

名称	Z05	Taq	Tfi	Tfl	Sps17	Tth	Tca	Dra	Tma	Tne	Taf	Bst	Bca
Z05		0.901	0.845	0.891	0.845	0.975	0.973	0.563	0.483	0.478	0.44	0.498	0.49
Taq	0.901		0.879	0.901	0.877	0.906	0.901	0.561	0.488	0.473	0.44	0.503	0.495
Tfi	0.845	0.879		0.857	0.997	0.853	0.853	0.566	0.495	0.49	0.449	0.512	0.49
Tfl	0.891	0.901	0.857		0.855	0.889	0.889	0.571	0.492	0.48	0.444	0.494	0.485
Sps17	0.845	0.877	0.997	0.855		0.853	0.853	0.566	0.495	0.49	0.449	0.512	0.49
Tth	0.975	0.906	0.853	0.889	0.853		0.99	0.563	0.478	0.473	0.437	0.496	0.488
Tca	0.973	0.901	0.853	0.889	0.853	0.99		0.563	0.478	0.473	0.437	0.496	0.488
Dra	0.563	0.561	0.566	0.571	0.566	0.563	0.563		0.45	0.448	0.426	0.474	0.454
Tma	0.483	0.488	0.495	0.492	0.495	0.478	0.478	0.45		0.883	0.622	0.474	0.475
Tne	0.478	0.473	0.49	0.48	0.49	0.473	0.473	0.448	0.883		0.615	0.476	0.473
Taf	0.44	0.44	0.449	0.444	0.449	0.437	0.437	0.426	0.622	0.615		0.46	0.473
Bst	0.498	0.503	0.512	0.494	0.512	0.496	0.496	0.474	0.474	0.476	0.46		0.898
Bca	0.49	0.495	0.49	0.485	0.49	0.488	0.488	0.454	0.475	0.473	0.473	0.898	

图 2