

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2024 年 7 月 4 日 (04.07.2024)



(10) 国际公布号
WO 2024/138762 A1

(51) 国际专利分类号:

C12N 9/12 (2006.01) *C12N 15/10* (2006.01)
C12N 15/52 (2006.01)

(21) 国际申请号: PCT/CN2022/144377

(22) 国际申请日: 2022 年 12 月 30 日 (30.12.2022)

(25) 申请语言: 中文

(26) 公布语言: 中文

(71) 申请人: 深圳华大智造科技股份有限公司
(**MGI TECH CO., LTD.**) [CN/CN]; 中国广东省深圳市盐田区北山工业区综合楼及 11 栋 2 楼, Guangdong 518083 (CN)。

(72) 发明人: 黄思谦(**HUANG, Siqian**); 中国广东省深圳市盐田区北山工业区综合楼及 11 栋 2 楼, Guangdong 518083 (CN)。刘月鹏(**LIU, Yuepeng**); 中国广东省深圳市盐田区北山工业区综合楼及 11 栋 2 楼, Guangdong 518083 (CN)。杨丽媛(**YANG, Liyuan**); 中国广东省深圳市盐田区北山工业区综合楼及 11 栋 2 楼, Guangdong 518083 (CN)。宋承威(**SONG, Chengwei**); 中国广东省深圳市盐田区北山工业区综合楼及 11 栋 2 楼, Guangdong 518083 (CN)。徐崇钧(**XU, Chongjun**); 中国广东省深圳市盐田区北山工业区综合楼及 11 栋 2 楼, Guangdong 518083 (CN)。刘芬(**LIU, Fen**); 中国广东省深圳市盐田区北山工业区综合楼及 11 栋 2 楼, Guangdong 518083 (CN)。

(74) 代理人: 北京法胜知识产权代理有限公司 (**FASHENG INTELLECTUAL PROPERTY COMPANY, LTD.**); 中国北京市海淀区北洼路 45 号 1 号楼 2 层 202 室, Beijing 100142 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU,

CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。

(84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

- 包括国际检索报告(条约第 21 条(3))。
- 包括说明书序列表部分(细则 5.2(a))。

(54) **Title:** RECOMBINANT PROTEIN AND USE THEREOF

(54) 发明名称: 重组蛋白及其应用

(57) **Abstract:** Provided are a recombinant protein having improved amplification performance, a nucleic acid encoding the recombinant protein, a vector comprising the nucleic acid, a kit comprising the protein, the nucleic acid or the vector, a method for preparing the recombinant protein, and a method for amplifying a target DNA by using the recombinant protein. As a chimeric family A DNA polymerase, the recombinant protein has improved processivity, the amplification performance of the recombinant protein is remarkably better than that of a wild-type family A DNA polymerase, and the thermal stability thereof is higher than that of a family B DNA polymerase.

(57) 摘要: 提供了具有提高的扩增性能的重组蛋白, 编码该重组蛋白的核酸, 包含该核酸的载体, # 包含上述蛋白、核酸或载体的试剂盒, 制备上述重组蛋白的方法, 以及利用该重组蛋白扩增目的 DNA 的方法。重组蛋白作为嵌合型 A 家族 DNA 聚合酶具有提高的持续合成能力, 其扩增性能显著优于野生型 A 家族 DNA 聚合酶, 而且热稳定性高于 B 家族 DNA 聚合酶。

WO 2024/138762 A1

重组蛋白及其应用

技术领域

本发明涉及生物技术领域，具体涉及重组蛋白，编码该重组蛋白的核酸，包含该核酸的载体，包含上述蛋白、核酸或载体的试剂盒，制备上述重组蛋白的方法，以及利用该重组蛋白扩增目的 DNA 的方法。

背景技术

DNB-SEQ 测序方法中通常使用 phi29 DNA 聚合酶进行滚环扩增（RCA）反应和多重置换扩增（MDA）反应。phi29 DNA 聚合酶是 B 家族聚合酶，尽管具有极高的链置换活性和可持续合成能力，但是 phi29 DNA 聚合酶的热稳定性较差以致容易失活，因此需要将其在严苛条件下低温保存且保质期较短。此外，phi29 DNA 聚合酶的非特异结合能力强且难以洗脱，导致了 phi29DNA 聚合酶在测序中产生较强荧光背景。

Bst 聚合酶是 A 家族聚合酶，具有耐高温的特性，其最适反应温度在 60-65℃。因此无论在其制备还是应用过程中，A 家族聚合酶对温度具有更好的耐受性，且同样具有较高的链置换活性。在测序中，A 家族聚合酶相比于 phi29 聚合酶更容易被洗脱，并且在复杂体系表现更好。然而，相比于 B 家族的 DNA 聚合酶（如 phi29、Pfu、Kod 聚合酶），A 家族聚合酶的拇指结构域与 DNA 模板结合的区域较小，导致包括 Bst 聚合酶在内的 A 家族聚合酶的可持续扩增能力较弱。

因此，需要研究开发具有高持续合成能力的 A 家族聚合酶。

发明内容

本发明旨在至少在一定程度上解决相关技术中的技术问题之一。

为此，本发明的实施方案提供了具有在扩增反应中加强 DNA 聚合酶与目的 DNA 相互作用的氨基酸序列的重组蛋白，编码该重组蛋白的核酸，包含该核酸的载体，包含上述蛋白、核酸或载体的试剂盒，制备上述重组蛋白的方法，以及利用该重组蛋白扩增目的 DNA 的方法。本发明实施方案提供的重组蛋白为具有提高的持续合成能力以及加强的热稳定性、聚合活性和 DNA 双链亲和力的嵌合型 A 家族 DNA 聚合酶，可用于扩增目的 DNA。

第一方面，本发明的实施方案提供了一种重组蛋白，该重组蛋白的氨基酸序列是用在扩增反应中能够加强 DNA 聚合酶与目的 DNA 相互作用的氨基酸序列替换野生型的 A 家族 DNA 聚合酶中的同源序列获得的氨基酸序列。

在一些实施例中，A 家族 DNA 聚合酶选自大肠杆菌 DNA 聚合酶 I、T3 DNA 聚合酶、

T5 DNA 聚合酶、T7 DNA 聚合酶、Taq DNA 聚合酶、Bsu DNA 聚合酶和 Bst DNA 聚合酶组成的组。

在一些实施例中，A 家族 DNA 聚合酶选自具有 SEQ ID NO: 1 所示氨基酸序列的 Bst DNA 聚合酶、具有 SEQ ID NO: 9 所示氨基酸序列的 Taq DNA 聚合酶、具有 SEQ ID NO: 10 所示氨基酸序列的大肠杆菌 DNA 聚合酶和具有 SEQ ID NO: 11 所示氨基酸序列的 Bsu DNA 聚合酶中的至少一种。

在一些实施例中，A 家族 DNA 聚合酶中的同源序列来源于拇指结构域。

在一些实施例中，A 家族 DNA 聚合酶中的同源序列选自以下中的至少一种：

如 SEQ ID NO: 2 所示的 Bst DNA 聚合酶中的第 546 位至第 554 位氨基酸序列；

如 SEQ ID NO: 12 所示的 Taq DNA 聚合酶中的第 199 位至第 209 位氨基酸序列；

如 SEQ ID NO: 13 所示的大肠杆菌 DNA 聚合酶中的第 276 位至第 285 位氨基酸序列；

或

如 SEQ ID NO: 14 所示的大肠杆菌 Bsu 聚合酶中的第 249 位至第 257 位氨基酸序列。

在一些实施例中，A 家族 DNA 聚合酶中的同源序列被替换为包含 SEQ ID NO: 3 所示氨基酸序列的序列、基本上由 SEQ ID NO: 3 所示氨基酸序列组成的序列、或由 SEQ ID NO: 3 所示氨基酸序列组成的序列。

在一些实施例中，重组蛋白的氨基酸序列选自 SEQ ID NO: 4、SEQ ID NO: 15、SEQ ID NO: 16 和 SEQ ID NO: 17 中的至少一种。

在一些实施例中，重组蛋白还包含额外的保守突变、添加和缺失中的任一种。

在一些实施例中，重组蛋白作为嵌合型 DNA 聚合酶的持续合成能力高于野生型的所述 A 家族 DNA 聚合酶。

第二方面，本发明的实施方案提供了一种核酸，该核酸编码如第一方面任一实施方案所述的重组蛋白。

在一些实施例中，所述核酸分子的核苷酸序列选自 SEQ ID NO: 5、SEQ ID NO: 18、SEQ ID NO: 19 和 SEQ ID NO: 20 中的至少一种。

第三方面，本发明的实施方案提供了一种载体，该载体包含如第二方面任一实施方案所述的核酸。

第四方面，本发明的实施方案提供了一种试剂盒，包含如第一方面任一实施方案所述的重组蛋白、如第二方面任一实施方案所述的核酸、或如第三方面任一实施方案所述的载体。

第五方面，本发明的实施方案提供了一种制备第一方面任一实施方案所述的重组蛋白的方法，该方法包括：用在扩增反应中能够加强 DNA 聚合酶与目的 DNA 相互作用的氨基

酸序列替换野生型的 A 家族 DNA 聚合酶中的同源序列获得的氨基酸序列，以获得所述重组蛋白。

在一些实施例中，所述方法包括：对来自所述野生型 A 家族 DNA 聚合酶多个氨基酸序列进行序列比对，以确定嵌合序列；和用所述嵌合序列替换所述野生型 A 家族 DNA 聚合酶序列中的同源序列，以获得候选重组蛋白序列。

在一些实施例中，所述方法还包括：对所述候选重组蛋白序列进行计算机模拟分析，以确定重组蛋白序列，其中所示重组蛋白序列表现出减小的拇指结构域与手指结构域之间的空间构型。

在一些实施例中，对所述候选重组蛋白序列进行计算机模拟分析，以获得重组蛋白序列，包括：对所述候选重组蛋白序列进行结构预测，以获得重组蛋白的三级结构；对所述重组蛋白的三级结构进行模拟，以获得所述重组蛋白的分子对接结构；对所述分子对接结构进行基于配体的结构优化，以获得优化的重组蛋白结构；和对所述优化的重组蛋白结构进行动力学模拟，以确定重组蛋白序列。

在一些实施例中，所述方法还包括：表达并纯化所述重组蛋白序列。

在一些实施例中，A 家族 DNA 聚合酶选自大肠杆菌 DNA 聚合酶 I、T3 DNA 聚合酶、T5 DNA 聚合酶、T7 DNA 聚合酶、Taq DNA 聚合酶、Bsu DNA 聚合酶和 Bst DNA 聚合酶组成的组。

在一些实施例中，A 家族 DNA 聚合酶选自具有 SEQ ID NO: 1 所示氨基酸序列的 Bst DNA 聚合酶具有 SEQ ID NO: 9 所示的氨基酸序列的 Taq DNA 聚合酶、具有 SEQ ID NO: 10 所示氨基酸序列的大肠杆菌 DNA 聚合酶和具有 SEQ ID NO: 11 所示氨基酸序列的 Bsu DNA 聚合酶中的至少一种。

在一些实施例中，所述 A 家族 DNA 聚合酶中的同源序列来源于拇指结构域序列。

在一些实施例中，所述 A 家族 DNA 聚合酶中的同源序列选自以下中的至少一种：

如 SEQ ID NO: 2 所示的 Bst DNA 聚合酶中的第 546 位至第 554 位氨基酸序列；

如 SEQ ID NO: 12 所示的 Taq DNA 聚合酶中的第 199 位至第 209 位氨基酸序列；

如 SEQ ID NO: 13 所示的大肠杆菌 DNA 聚合酶中的第 276 位至第 285 位氨基酸序列；

或

如 SEQ ID NO: 14 所示的大肠杆菌 Bsu 聚合酶中的第 249 位至第 257 位氨基酸序列。

在一些实施例中，A 家族 DNA 聚合酶中的同源序列被替换为包含 SEQ ID NO: 3 所示氨基酸序列的序列、基本上由 SEQ ID NO: 3 所示氨基酸序列组成的序列、或由 SEQ ID NO: 3 所示氨基酸序列组成的序列。

在一些实施例中，重组蛋白的氨基酸序列选自 SEQ ID NO: 4、SEQ ID NO: 15、SEQ ID

NO: 16 和 SEQ ID NO: 17 中的至少一种。

第六方面，本发明的实施方案提供了一种扩增目的 DNA 的方法，使用第一方面任一实施方案所述的重组蛋白对目的 DNA 进行扩增。

在一些实施例中，所述扩增是等温扩增，所述等温扩增选自滚环扩增 RCA、多重置换扩增 MDA、重组酶聚合酶扩增反应 RPA、链置换扩增 SDA、环介导等温扩增 LAMP。

在一些实施例中，所述扩增用于构建 DNA 文库或测序。

本发明实施方案中的重组蛋白作为嵌合型 A 家族 DNA 聚合酶具有提高的持续合成能力，其扩增性能显著优于野生型 A 家族 DNA 聚合酶。此外，与 B 家族 DNA 聚合酶相比，本发明实施方案中的重组蛋白还具有以下优势：热稳定性高、方便洗脱且易于储存。

附图说明

为了更清楚地说明本公开实施例的技术方案，下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本公开的一些实施例，对于本领域普通技术人员来讲，还可以根据这些附图获得其他的附图。

图 1 为 Bst DNA 聚合酶大片段的结构示意图。

图 2 为野生型 Bst DNA 聚合酶大片段与根据本发明实施例的嵌合型 DNA 聚合酶的序列比对示意图。

图 3 为根据本发明实施例的嵌合型 DNA 聚合酶 Taq-HS-1(a)、Ecoli-HS-1(b)、Bsu-HS-1(c) 和 Bst-HS-1(d) 与 DNA 分子对接后的结构示意图。

图 4 为根据本发明实施例的嵌合型 DNA 聚合酶与 DNA 分子复合物的动力学模拟后的结构示意图。

具体实施方式

下面详细描述本发明的实施例，所述实施例的示例在附图中示出。下面通过参考附图描述的实施例是示例性的，旨在用于解释本发明，而不能理解为对本发明的限制。

A 家族 DNA 聚合酶包括诸如大肠杆菌聚合酶 I、Bst DNA 聚合酶、T7 DNA 聚合酶和 Taq DNA 聚合酶，后者在 PCR 中已被广泛应用。来自嗜热脂肪地芽孢杆菌 (*Geobacillus stearothermophilus*) 的 Bst DNA 聚合酶是一种多功能酶，以耐高温、强大的链置换活性和可实现环介导的等温扩增为特点，然而其可持续扩增能力仍有待改进。Bst DNA 聚合酶属于 A 家族 DNA 聚合酶，具有典型的右手结构，分为拇指结构域、手掌结构域、手指结构域和 5'-3' 外切酶结构域，其中在聚合反应中目的 DNA 和部分新合成的 DNA 夹在拇指结构域和手指结构域之间。在生物技术应用中，Bst DNA 聚合酶的 5'-3' 外切酶结构域通常会被敲除，只保

留包含拇指结构域、手掌结构域、手指结构域的Bst聚合酶大片段（Bst-LF），其结构如图1所示。

目前应用于测序的DNA聚合酶主要是以phi29聚合酶为主的B家族链置换酶。phi29聚合酶是来源于枯草芽孢杆菌（*Bacillus subtilis*）的噬菌体phi29中的聚合酶，具有较高的连续合成能力和链置换活性，常应用于基于RCA的DNB制备及基于MDA的单细胞全基因组扩增。phi29聚合酶是一种中温恒温酶，最适反应温度在30℃左右，但在65℃孵育10分钟即可完全失活。phi29聚合酶具有强链置换能力及高持续合成能力主要是因为它的TPR2（terminal protein region 2）结构域。TPR2结构域在链合成过程中可以将模板的双链解链，只允许模板链进入并会把模板一直夹在Exo结构域和TPR2结构域之间，使酶和模板不能分离，从而使phi29聚合酶具有>70000nt的持续扩增能力。但是，phi29聚合酶存在保存条件严苛、保质期较短、测序时荧光背景较大等缺陷。

相比于B家族的DNA聚合酶（如phi29、Pfu、Kod聚合酶），A家族聚合酶的拇指结构域与DNA模板结合的区域较小，且与新生DNA链的相互作用也较小，导致包括Bst聚合酶在内的A家族聚合酶的可持续扩增能力较弱。此外，由于5'-3'外切结构域具有DNA结合作用，能阻止聚合酶与模板的彻底解离，在敲除5'-3'外切结构域的情况下，A家族聚合酶大片段的持续合成能力会下降，通常不能很好的合成长度大于10 kb的DNA链。

基于以上问题，发明人以Bst聚合酶、Taq DNA聚合酶、大肠杆菌DNA聚合酶和Bsu DNA聚合酶为例对A家族聚合酶的结构进行了探索，以期提供一种对DNA双链具有高亲和力且持续合成能力高的聚合酶变体。根据本申请实施方案提供的重组蛋白，即嵌合型A家族聚合酶，其在结构方面实现了使得拇指结构域向手指结构域方向延伸的改造，缩小了拇指结构域与手指结构域之间的空间，使模板链和部分新合成的DNA链保持在拇指结构域和手指结构域之间，从而起到类似于phi29聚合酶中TPR2结构域的效果，加强了拇指结构域与DNA双链的相互作用并且形成环绕DNA模板链的环形滑动夹，提高了A家族DNA聚合酶的持续合成能力。

根据本发明的一个方面，本发明提供了一种重组蛋白，该重组蛋白的氨基酸序列是用在扩增反应中能够加强DNA聚合酶与目的DNA相互作用的氨基酸序列替换野生型的A家族DNA聚合酶中的同源序列获得的氨基酸序列。

需要说明的是，本发明的实施例提供的重组蛋白的拇指结构域中替换为一段嵌合蛋白序列。该序列加强了拇指结构域与DNA双链的互作，并且形成环绕DNA模板链的环形滑动夹，使聚合酶对DNA双链的亲和力提高，从而提高持续合成能力的效果。该重组蛋白也由此具有提高的持续合成能力，加强的热稳定性、聚合活性、DNA双链亲和力，可用于建库和测序中的多种DNA等温扩增反应。

需要说明的是,聚合酶的持续合成能力可定义为在一次结合中处理的核苷酸数量。DNA 聚合酶的合成能力通常反应了合成率和合成速度,以及酶与底物的亲和力。合成能力高的 DNA 聚合酶适用于扩增长模板、具有二级结构和富含 GC 的序列,以及存在肝素、木聚糖和腐殖酸等 PCR 抑制剂的血液和植物组织样品。

在一些实施例中,A 家族 DNA 聚合酶选自大肠杆菌 DNA 聚合酶 I、T3 DNA 聚合酶、T5 DNA 聚合酶、T7 DNA 聚合酶、Taq DNA 聚合酶、Bsu DNA 聚合酶和 Bst DNA 聚合酶组成的组。

在一些实施例中,A 家族 DNA 聚合酶选自具有 SEQ ID NO: 1 所示氨基酸序列的 Bst DNA 聚合酶、具有 SEQ ID NO: 9 所示氨基酸序列的 Taq DNA 聚合酶、具有 SEQ ID NO: 10 所示氨基酸序列的大肠杆菌 DNA 聚合酶和具有 SEQ ID NO: 11 所示氨基酸序列的 Bsu DNA 聚合酶中的至少一种。

SEQ ID NO: 1 的序列为:

AEGEKPLEEMEF AIVDVITEEMLADKAALVVEVMEENYHDAPIVGIALVNEHGRFF
MRPETALADSQFLAWLADETKKKSMFDAKRAVVALKWKGIELRGVAFDLLLLAAYLLN
PAQDAGDIAAVAKMKQYEAVRSDEAVYGKGVKRSPLDEQTLAEHLVRKAAAIWALEQ
PFMDDLRNNEQDQLLTKLEQPLAAILAEMEFTGVNVDTKRLEQMGSELAEQLRAIEQRI
YELAGQEFNINSPKQLGVILFEKLQLP**VLKKTKTKGY**STSADVLEKLAPHHEIVENILHYR
QLGKLQSTYIEGLLKVVRPDTGKVHTMFNQALTQTGRLSSAEPNLQNIPIRLEEGRKIRQ
AFVPSEPDWLIFAADYSQIELRVLAHIADDDNLIEAFQRDLDIHTKTAMDIFHVSEEEVTA
NMRRQAKAVNFGIVYGISDYGLAQNLNITRKEAAEFIERYFASFPGVKQYMENIVQEAK
QKGYVTLLHRRRYLPDITSRNFNVRSAERTAMNTPIQGSAAIIKKAMIDLAARLKEE
QLQARLLLQVHDELILEAPKEEIERLCELVPEVMEQAVTLRVPLKVDYHYGPTWYDAK。

(加粗下划线为同源序列 SEQ ID NO: 2)

SEQ ID NO: 9 的序列为:

GAFVGFVLSRKEPMWADLLALAAARGGRVHRAPEPYKALRDLKEARGLLAKDLS
VLALREGLGLPPGDDPMLLAYLLDPSNTTPEGVARRYGGEWTEEAGERAAALSERLFANL
WGRLEGEERLLWLYREVERPLSAVLAHVEATGVRLDVAYLRALSLEVAEEIARLEAEVF
RLAGHPFNLNSRDQLERVLFDELGLP**AIGKTEKTGKR**STSAAVLEALREAHAIVEKILQ
YRELTKLKSTYIDPLPELIHPRTGRLHTRFNQTATATGRLSSSDPNLQNIPIVRTPLGQRIRR
AFIAEEGWLLVALDYSQIELRVLAHLSGDENLIRVFQEGRDIHTETASWMFGVPREAVDP
LMRRAAKTINFGVLYGMSAHRLSQELAIPYEEAQAFIERYFQSFPKVRWIEKTLEEGRR
RGYVETLFGRRRYVPDLEARVKS VREAAERMAFNMPVQGTAA DLMKLAMVKLFPRLE

EMGARMILLQVHDEL VLEAPKERAEAVARLAKEVMEGVYPLAVPLEVEVGIGEDWLSA
KV。（加粗下划线为同源序列 SEQ ID NO: 12）

SEQ ID NO: 10 的序列为：

TVISYDNYVTILDEETLKAWIAKLEKAPVFAFDTETDSLDNISANLVGLSFAIEPGVA
AYIPVAHDYLDAPDQISRERALELLKPLLEDEKALKVGQNLKYDRGILANYGIELRGIAF
DTMLESYILNSVAGRHDMDSLAERWLKHKTTITFEEIAGKGKNQLTFNQIALDEAGRYAA
EDADVTQLHLKMWPDLQKHKGPLNVFENIEMPLVPVLSRIERNGVKIDPKVLHNHSEE
LTLRLAELEKKAYEIAAGEEFNLSSTKQLQTILFEKQGIK**PLKKTPGGAP**STSEEVLLEELAL
DYPLPKVILEYRGLAKLKSTYTDKLPLMINPKTGRVHTSYHQAVTATGRLSSTDPNLQNI
PVRNEEGRRIRQAFIAPEDYVIVSADYSQIELRIMAHLSRDKGLLTAFAGKDIHRATAAE
VFGLPLET VTSEQRRSAKAINFGLIYGMSAFGLARQLNIPRKEAQKYMDLYFERYPGVLK
YMERTRAQAKEQGYVETLDGRRLYLPDIKSSNGARRAAAERAANAPMQGTAADIKRA
MIAVDAWLQAEQPRVRMIMQVHDEL VFEVHRDDVDAVAKQIHQLMENCTRLDVPLL V
EVGSGENWDQAH。（加粗下划线为同源序列 SEQ ID NO: 13）

SEQ ID NO: 11 的序列为：

DINVKT VTDVTS DILVSPSAFVVEQIGDNYHEEPILGFSIVNETGAYFIPKDIAVESEV
FKEWVENDEQKKWVFDSKRAVVALRWQGIELKGAEFDTLAAYIINPGNSYDDVASVA
KDYGLHIVSSDES VYGKGAKRAVPSEDVLSEHLGRKALAIQSLREKL VQELENNDQLEL
FEELEMP LALILGEMESTGVKVDVDR LKRMGEELGAKLKEYEEKIHEIAGEPFNINSPKQ
LGVILFEKIGLP**VVKKTKTG**YSTSADVLEKLADKH DIVDYILQYRQIGKLQSTYIEGLLK
VTRPDSHKVHTRFNQALTQTGRLSSTDPNLQNIPIRLEEGRKIRQAFVPSEKDWLIFAADY
SQIELRVLAHISKDENLIEAFTNDMDIHTKTAMDV FHVAKDEVTSAMRRQAKAVNFGIV
YGISDYGLSQNLGITRKEAGAFIDRYLESFQGVKAYMEDSVQEAKQKGYVTTLMHRRR
YIPELTSRNFNIRSFAERTAMNTPIQGSAAADIKKAMIDMAAKLKEKQLKARLLLQVHDE
LIFEAPKEEIEILEKLVPEVMEHALALDVPLKVDFASGPSWYDAK。（加粗下划线为同源
序列 SEQ ID NO: 14）

在一些实施例中，所述 A 家族 DNA 聚合酶中的同源序列来源于拇指结构域序列。

在一些实施例中，所述 A 家族 DNA 聚合酶中的同源序列选自以下中的至少一种：

如 SEQ ID NO: 2 所示的 Bst DNA 聚合酶中的第 546 位至第 554 位氨基酸序列；

如 SEQ ID NO: 12 所示的 Taq DNA 聚合酶大片段中的第 199 位至第 209 位氨基酸序列；

如 SEQ ID NO: 13 所示的大肠杆菌 DNA 聚合酶大片段中的第 276 位至第 285 位氨基酸
序列；或

如 SEQ ID NO: 14 所示的大肠杆菌 Bsu 聚合酶大片段中的第 249 位至第 257 位氨基酸序列。

SEQ ID NO: 2 的序列为: VLKKTGTGY。

SEQ ID NO: 12 的序列为: AIGKTEKTGKR

SEQ ID NO: 13 的序列为: PLKKTPGGAP

SEQ ID NO: 14 的序列为: VVKKTGTGY

在一些实施例中, A 家族 DNA 聚合酶中的同源序列被替换为包含 SEQ ID NO: 3 所示氨基酸序列的序列、基本上由 SEQ ID NO: 3 所示氨基酸序列组成的序列、或由 SEQ ID NO: 3 所示氨基酸序列组成的序列。SEQ ID NO: 3 的序列为: PNREMKNQGSKKTTLGSTRRGIDNGRKLRLGRQF。

在一些实施例中, 重组蛋白的氨基酸序列选自 SEQ ID NO: 4、SEQ ID NO: 15、SEQ ID NO: 16 和 SEQ ID NO: 17 中的至少一种。

SEQ ID NO: 4 的序列为:

AEGEKPLEEMFAIVDVITEEMLADKAALVVEVMEENYHDAPIVGIALVNEHGRFF
MRPETALADSQFLAWLADETKKKSMFDAKRAVVALKWKGIELRGVAFDLLLLAAYLLN
PAQDAGDIAAVAKMKQYEA VRSDEAVYGGVGRSLPDEQTLAEHLVRKAAAIWALEQ
PFMDDLRNNEQDQLLTKLEQPLAAILAEMEFTGVNVDTKRLEQMGSELAEQLRAIEQRI
YELAGQEFNINSPKQLGVILFEKLQLPPNREMKNQGSKKTTLGSTRRGIDNGRKLRLGR
QFSTSAADVLEKLAPHHEIVENILHYRQLGKLQSTYIEGLLKVVRPDTGKVHTMFNQALT
QTGRLSSAEPNLQNIPIRLEEGRKIRQAFVPSEPDWLIFAADYSQIELRVLAHIADDDNLIE
AFQRDLDIHTKTAMDIFHVSEEEVTANMRRQAKAVNFGIVYGISDYGLAQNLNITRKEA
AEFIERYFASFPGVKQYMNIVQEAQKQGYVTLLHRRRYLPDITSRNFNVRSFAERTAM
NTPIQGSAAADIKKAMIDLAARLKEEQQLQARLLLQVHDELILEAPKEEIERLCELVPEVME
QAVTLRVPLKVDYHYGPTWYDAK。(加粗下划线为嵌合序列 SEQ ID NO: 3)

SEQ ID NO: 15 的序列为:

GAFVGFVLSRKEPMWADLLALAAARGGRVHRAPEPYKALRDLKEARGLLAKDLS
VLALREGLGLPPGDDPMLLAYLLDPSNTTPEGVARRYGGEWTEEAGERAALSERLFANL
WGRLEGEERLLWLYREVERPLSAVLAHVEATGVRLDVAYLRALSLEVAEEIARLEAEVF
RLAGHPFNLSRDQLERVLFDELGLPPNREMKNQGSKKTTLGSTRRGIDNGRKLRLGR
QFSTSAAVLEALREAHAVEKILQYRELTKLKSTYIDPLPELIHPRTGRLHTRFNQTATAT
GRLSSSDPNLQNIPIVRTPLGQRIRRAFIAEEGWLLVALDYSQIELRVLAHLSGDNELIRVF
QEGRDIHTETASWMFGVPREAVDPLMRRAAKTINFGVLYGMSAHRLSQELAIPEYEEAQA

FIERYFQSF PKVRAWIEKTLEEGRRRGYVETLFGRRRYVPDLEARVKS SVREAAERMAFN
MPVQGTAA DLMK LAMVKLFPRLEEMGARM LLQVHDEL VLEAPKERAEAVARLAKEV
MEGVYPLAVPLEVEVGIGEDWLSAKV。(加粗下划线为嵌合序列 SEQ ID NO: 3)

SEQ ID NO: 16 的序列为:

TVISYDNYVTILDEETLKAWIAKLEKAPVFAFDTETDSL DNISANLVGLSFAIEPGVA
AYIPVAHDYLDAPDQISRERALELLKPLLEDEKALKVGQNLKYDRGILANYGIELRGIAF
DTMLESYILNSVAGRHDMSLAERWLKHKTITFEEIAGKGKNQLTFNQIALDEAGRYAA
EDADVTLQLHLKMWPDLQKHKGPLNVFENIEMPLVPVLSRIERNGVKIDPKVLHNHSEE
LTLRLAELEKKAYEIAGEEFNLSSTKQLQTILFEKQGIK **PNREMKNQGSKKTLGSTRRG**
IDNGRKLRLGRQFSTSEEVLEELALDYPLPKVILEYRGLAKLKSTYTDKPLMINPKTGR
VHTSYHQA V TATGRLSSTD PNLQNIPVRNEEGRRIRQAFIAPEDYVIVSADYSQIELRIMA
HLSRDKGLLTAF AEGKDIHRATAAEVFGLPLET VTSEQRSAKAINFGLIYGMSAFGLAR
QLNIPRKEAQKYMDLYFERYPGVLKYMERTRAQAKEQGYVETLDGRRLYLPDIKSSNG
ARRAA AERA AINAPMQGTAADI KRAMIAVD A WLQAEQPRVRMIMQVHDEL VFEVHRD
D VDAVAKQIHQLMENCTRLDVPLLVEVGSGENWDQAH。(加粗下划线为嵌合序列 SEQ
ID NO: 3)

SEQ ID NO: 17 的序列为:

DINVKT VTDVTS DILVSPSAFVVEQIGDNYHEEPILGFSIVNETGAYFIPKDI AVESEV
FKEWVENDEQKKWVFDSKRAVVALRWQGI ELKGA EFDTL LAAYIINPGNSYDDVASVA
KDYGLHIVSSDES VYGKGAKRAVPSE DVLSEHLGRKALAIQSLREKL VQELENNDQLEL
FEELEMP LALILGEMESTGVKVDVDRLKRMGEELGAKLKEYEEKIHEIAGEPFNINSPKQ
LG VILFEKIGL **PNREMKNQGSKKTLGSTRRGIDNGRKLRLGRQF**STSADVLEKLADK
HDIVDYILQYRQIGKLQSTYIEGLLKVTRPDSHKVHTRFNQALTQTGRLSSTD PNLQNIPI
RLEEGRKIRQAFVPSEKDWLIFAADYSQIELRVLAHISKDENLIEAFTNDMDIHTKTAMD
VFHVAKDEVTSAMRRQAKAVNFGIVYGISDYGLSQNLGITRKEAGAFIDRYLESFQGVK
AYMEDSVQEAKQKGYVTTLMHRRRYIPELTSRNFNIRSF AERTAMNTPIQGSAA DIKKA
MIDMAAKLKEKQLKARLLLQVHDELIFEAPKEEIEILEKLVPEVMEHALALDVPLKVDFA
SGPSWYDAK。(加粗下划线为嵌合序列 SEQ ID NO: 3)

在一些实施例中, 重组蛋白还包含额外的保守突变、添加和缺失中的任一种。

在一些实施例中, 重组蛋白作为嵌合型 DNA 聚合酶的持续合成能力高于野生型的所述 A 家族 DNA 聚合酶。

根据本发明的又一个方面, 本发明提供了一种核酸, 该核酸编码如第一方面任一实施

方案所述的重组蛋白。

在一些实施例中，该核酸分子的核苷酸序列选自 SEQ ID NO: 5、SEQ ID NO: 18、SEQ ID NO: 19 和 SEQ ID NO: 20 中的至少一种。

SEQ ID NO: 5 的序列为：

GCCGAGGGTGAAAAACCTCTGGAAGAAATGGAATTTGCAATTGTGGATGTAATT
ACCGAAGAAATGCTGGCGGATAAAGCCGCTCTTGTGGTGGAAAGTCATGGAAGAAAA
CTATCATGATGCGCCGATTGTGGGCATTGCTCTGGTTAATGAACATGGCCGCTTCTTT
ATGCGTCCAGAAACAGCCCTGGCCGACTCACAAATTTCTGGCCTGGCTGGCGGACGAA
ACCAAAAAGAAATCAATGTTTGACGCGAAACGCGCTGTCGTCGCTCTGAAATGGAA
GGGGATCGAACTTCGTGGTGTGGCCTTCGATCTTCTGCTGGCAGCGTATTTGCTGAAT
CCCGCCAGGATGCCGGGGATATTGCAGCCGTGGCGAAAATGAAACAGTACGAAGC
GGTCCGCTCCGACGAGGCCGTGTACGGCAAGGGCGTCAAGCGTTCCTGCCGGATGA
ACAGACCCTGGCAGAACACCTGGTCAGAAAAGCCGCTGCCATCTGGGCACTGGAAC
AGCCGTTTATGGATGATCTGCGCAATAATGAGCAGGATCAGCTGCTGACAAAACCTGG
AACAGCCGCTGGCGGCAATTCTGGCCGAAATGGAATTTACGGGCGTAAATGTGGAT
ACCAAACGCCTGGAACAGATGGGCTCGGAACTGGCGGAACAGCTGCGTGCCATCGA
ACAGCGCATTTATGAACTGGCGGGCCAGGAATTTAATATCAATAGTCCTAAACAGCT
GGGTGTCATTCTGTTTGAAAACTGCAGTTGCCGCCGAACCGCGAAATGAAAAACCA
GGGCAGCAAAAAAACCTGGGCAGCACCCGCCGCGGCATTGATAACGGCCGCAAAC
TGCGCCTGGGCCGCCAGTTTAGTACCTCAGCGGATGTTCTGGAAAACTGGCTCCGC
ATCATGAGATCGTTGAAAACATTCTGCATTATCGTCAGCTGGGTAAACTGCAGAGCA
CCTATATAGAAGGGCTGCTGAAAGTCGTGCGCCCGGATACGGGTAAAGTCCATACCA
TGTTTAACCAGGCCCTGACCCAGACAGGGCGCCTGAGCAGTGCGGAGCCTAATTTAC
AGAATATACCGATTCTGTCTGGAAGAAGGTCGCAAAATTCGCCAGGCGTTTGTGCCAA
GCGAACCCGATTGGCTGATCTTTGCCGCAGATTACAGCCAGATTGAACTGCGCGTGC
TGGCGCATATTGCGGATGACGACAATTTAATTGAAGCGTTTCAGCGCGATTTAGATA
TTCACACCAAAACCGCGATGGATATTTTTTCATGTGAGCGAGGAAGAAGTTACGGCTA
ATATGCGTCGTCAGGCCAAAGCAGTGAATTTTCGGAATTGTTTACGGCATCAGTGATT
ATGGCCTGGCGCAGAATCTGAATATTACCCGTAAAGAGGCAGCAGAATTTATTGAAC
GTTATTTTGCAAGCTTTCCGGGTGTGAAACAGTATATGGAGAATATCGTTCAGGAAG
CCAAACAGAAAGGATACGTCACCACCCTGCTGCATCGTCGTCGTTACCTGCCGGATA
TTACCTCCCGCAATTTTAATGTGCGCAGTTTTGCAGAACGTACAGCGATGAATACAC

CTATTCAGGGCAGCGCGGCAGATATTATTA AAAAAGCGATGATCGATCTGGCAGCAC
GTCTGAAAGAAGAACAGCTGCAGGCACGTCTGCTGCTGCAGGTGCATGATGAACTG
ATTCTGGAAGCGCCGAAAGAAGAAATCGAACGCCTGTGCGAACTGGTCCCTGAAGT
GATGGAACAGGCGGTTACCCTGCGTGTACCGCTGAAAGTGGATTATCATTATGGCCC
GACCTGGTATGATGCTAAATAA。

SEQ ID NO: 18 的序列为:

GGGGCTTTCGTAGGATTTGTTCTATCAAGGAAGGAACCGATGTGGGCAGATTTG
CTGGCGCTGGCCGCGGCGAGAGGCGGTCCGCTACACCGTGCTCCGGAGCCGTATAA
GGCGCTGCGTGACTTGAAGGAGGCGAGAGGTTTGCTGGCGAAAGACCTGTCTGTTCT
GGCGCTGCGCGAGGGCCTTGGTTTACCGCCGGGTGACGATCCGATGCTGCTCGCTTA
CCTGCTGGACCCGAGCAATACCAACCCGGAAGGCGTCGCTAGGCGGTACGGCGGAG
AGTGGACCGAAGAGGCTGGTGAACGTGCGGCTCTGTCCGAACGCTTGTTTCGCCAATC
TGTGGGGTCGTCTCGAGGGCGAAGAACGCCTGTTATGGCTGTACCGCGAGGTTGAAA
GACCTCTGTCTGCTGTTCTTGCGCACGTGGAAGCTACCGGCGTGCGTTTAGACGTTGC
GTACCTGCGTGCTCTGTCCCTCGAGGTGGCGGAAGAAATTGCACGTTTGGAGGCAGA
AGTGTTCGCTCTGGCGGGTCATCCGTTTAACTTGAATAGCCGTGACCAGCTGGAGCG
CGTGCTGTTTCGATGAACTGGGCCTGCCACCGAACCGTGAAATGAAGAACCAGGGCTC
GAAAAAAACCCTGGGCAGCACGCGTCGTGGTATCGACAATGGTCGTAAACTGCGAT
TGGGCCGTCAGTTTAGCACGAGCGCCGCGGTGCTGGAAGCCCTGCGCGAAGCGCAT
GCAATTGTTGAGAAGATCCTGCAATATCGCGAGCTGACCAAACCTTAAGTCGACCTAT
ATCGATCCGTTACCGGAATTGATCCACCCGCGTACCGGTCGGCTGCATACCCGTTTTA
ACCAGACCGCGACGGCAACCGGCCGTTTATCCAGCTCTGATCCAAATTTGCAAAACA
TCCCGGTTTCGTACTCCGCTGGGTCAACGTATTTCGTCTGCATTTCATCGCGGAAGAAG
GTTGGCTGCTAGTCGCCCTGGACTACAGCCAAATTGAACTGCGTGTTCTGGCTCACCT
GAGTGGTGATGAGAACCTGATCCGTGTGTTTCAAGAAGGCCGTGACATCCATACCGA
AACTGCGAGCTGGATGTTTCGGTGTGCCACGTGAAGCGGTGGATCCGCTGATGCGTCG
TGCAGCTAAGACCATTAACCTTGGTGTGTTGTACGGCATGAGCGCACACCGCTTATC
GCAAGAGCTGGCCATTCCGTATGAAGAGGCGCAGGCGTTTATCGAGCGTTACTTCCA
GAGCTTTCGAAAGTGCGCGCGTGGATTGAGAAAACCCTCGAGGAAGGCCGTCGCC
GTGGTTATGTTGAGACGCTGTTTCGGTCGCCGCCGCTACGTTCCGGATCTGGAGGCGC
GCGTGAAAAGCGTTCGTGAAGCGGCAGAGCGTATGGCATTCAACATGCCGGTGCGAG
GGTACAGCGGCGGACTTAATGAAACTCGCTATGGTTAAGCTGTTCCCGCGTTTGGAG

GAGATGGGTGCACGTATGCTGCTACAGGTTACGACGAATTGGTTCTTGAGGCGCCT
AAAGAGCGCGCGGAAGCGGTGGCCCGCCTCGCCAAGGAGGTGATGGAAGGCGTCTA
TCCGCTGGCGGTTCCGCTCGAGGTAGAGGTGGCATCGGTGAAGATTGGCTATCCGC
GAAAGTTTAA。

SEQ ID NO: 19 的序列为:

ACTGTAATAAGTTATGATAATTACGTTACAATCCTGGACGAAGAAACCTTGAAG
GCGTGGATTGCGAAACTGGAAAAGGCGCCAGTGTTTCGCTTTTGATACCGAAACTGAC
TCGTTGGACAACATTAGCGCGAATCTAGTAGGCCTGTCTTTTGCAATTGAGCCGGGT
GTTGCCGCGTATATTCCGGTCGCGCATGACTACCTGGACGCGCCGGACCAGATCAGC
CGTGAACGTGCCCTGGAGTTGCTGAAGCCGCTGCTGGAGGATGAGAAGGCATTGAA
GGTGGGTCAGAACCTTAAGTACGATCGTGGCATCTTGGCCAACTACGGCATCGAGCT
CCGCGGTATCGCGTTCGACACAATGCTGGAAAGCTATATCCTGAACAGCGTCGCGGG
TCGCCATGATATGGACTCGCTGGCGGAACGTTGGCTGAAGCACAAAACCATCACCTT
CGAGGAAATTGCGGGCAAAGGCAAGAACCAGCTGACCTTTAATCAGATTGCACTGG
ACGAGGCAGGTCGTTATGCAGCAGAGGACGCTGATGTTACCCTGCAACTTCATCTGA
AGATGTGGCCGGATCTGCAGAAACACAAGGGTCCGCTGAACGTGTTTGAGAACATT
GAGATGCCGCTGGTGCCGGTCTTGTCCCGTATCGAGCGCAACGGTGTGAAAATCGAC
CCTAAAGTGCTGCATAATCATAGCGAAGAGCTGACTTTACGTCTCGCGGAGCTGGAG
AAAAAAGCTTATGAAATTGCTGGCGAGGAGTTCAACTTGTCCTCGACGAAACAGCTG
CAAACCATTTTGTGTTGAAAAGCAAGGTATCAAACCGAACAGGGAGATGAAAAACCA
GGGCAGCAAGAAGACCCTGGGCTCTACCAGACGTGGCATTGATAACGGCCGCAAAC
TGCGTCTGGGTCGCCAATTTAGCACGTCCGAAGAAGTGCTGGAGGAATTGGCGTTGG
ACTACCCGCTGCCGAAAGTGATCTTGGAATATCGTGGTCTGGCGAACTGAAGAGCA
CGTACACCGACAACTTCCCCTGATGATAAATCCGAAGACCGGTCGCGTGCACACCT
CCTACCATCAAGCGGTTACGGCTACCGGTCGTCTGAGCAGCACCGATCCGAATCTGC
AAAACATCCCGGTACGTAATGAAGAGGGTCGGCGCATTCGTCAGGCCTTCATCGCTC
CAGAGGACTACGTGATTGTTAGCGCTGATTATTCACAGATCGAATTGCGCATCATGG
CACACCTGAGTCGCGATAAGGGTCTGCTCACTGCGTTCGCCGAGGGCAAGGATATCC
ACCGTGCTACCGCGGCCGAAGTTTTCGGCTTGCCATTGGAGACAGTTACGTCCGAAC
AGCGTAGAAGCGCGAAAGCGATTAACCTTTGGCCTGATTTATGGTATGAGCGCGTTTG
GTCTGGCGCGTCAGTTGAATATCCCGCGTAAGGAGGCCCAAAAATACATGGATCTGT
ACTTCGAGCGTTATCCGGGCGTGCTGAAGTACATGGAACGCACCAGAGCACAAAGCT

AAAGAACAGGGTTATGTTGAAACCCTCGACGGTCGTCGACTCTACCTGCCGGACATC
AAAAGCTCCAATGGTGCACGTCGCGCGGCTGCCGAGCGCGCCGCCATTAATGCACCG
ATGCAAGGAACCGCAGCGGACATCATCAAGCGCGCTATGATCGCAGTGGATGCCTG
GCTGCAAGCGGAACAGCCGCGTGTTTCGTATGATTATGCAGGTCCACGACGAGTTGGT
GTTCGAGGTGCACCGTGATGATGTTGATGCGGTGCGGAAGCAAATTCATCAGCTGAT
GGAAACTGCACCCGTTTGGACGTTCCGCTCCTGGTTGAAGTTGGCTCTGGTGAGAA
CTGGGATCAGGCTCACTAA。

SEQ ID NO: 20 的序列为:

GACATAAATGTTAAAACTGTAACAGATGTCACGTCCGACATCCTGGTGAGCCCG
AGTGCGTTCGTCGTAGAGCAGATCGGTGATAATTATCATGAAGAACCGATTCTTGGC
TTTAGCATTGTTAATGAAACCGGTGCCTACTTCATCCCAAAAGACATTGCGGTGAG
AGCGAAGTTTTCAAAGAGTGGGTTGAAAACGACGAACAGAAAAAGTGGGTTTTCGA
CAGCAAACGCGCTGTGGTTGCGCTGCGCTGGCAGGGCATCGAGTTAAAGGGCGCAG
AATTTGACACCCTGCTGGCGGCGTACATTATCAACCCCGGTAATTCCTATGATGATGT
AGCCTCCGTGGCCAAGGACTACGGCCTGCATATTGTGAGCAGCGATGAGTCCGTTTA
CGGTAAAGGTGCGAAGCGCGCAGTGCCGTCAGAGGATGTCCTGTCTGAACACCTGG
GCCGAAAGGCGCTAGCGATTCAATCGCTGCGTGAAAAGCTGGTTCAAGAACTGGAG
AACACGATCAGCTGGAGTTGTTTGAAGAGTTGGAGATGCCGTTGGCGCTCATCTTG
GGCGAAATGGAATCTACTGGTGTTAAAGTGGACGTGGACCGTCTCAAGCGTATGGGT
GAAGAACTAGGAGCGAAGCTGAAGGAATATGAAGAGAAGATTCACGAGATCGCCG
GTGAGCCGTTTAACATCAACAGTCCGAAGCAGCTGGGCGTGATTCTGTTTGAGAAAA
TCGGTCTGCCACCGAATCGTGAAATGAAAAACCAGGGCTCCAAAAAGACCTTGGGC
AGCACGCGTCGTGGTATCGACAACGGTCGCAAGCTTCGCCTCGGTGCTCAATTTAGC
ACGTCTGCTGACGTCCTGGAGAAGTTGGCGGACAAACACGATATCGTGGATTACATT
TTGCAGTATCGTCAGATTGGCAAGCTCCAAAGCACCTATATCGAGGGCCTGCTTAAG
GTGACCAGACCTGATTCCCATAAAGTGCACACCCGTTTCAACCAGGCGCTGACCCAG
ACCGGTCGTCTGAGCTCAACCGATCCGAATCTGCAAAACATTCCGATCCGTTTGGAG
GAAGGTCGCAAAATCCGTCAGGCTTTCGTGCCGTCGGAGAAGGACTGGCTGATCTTT
GCGGCGGATTACTCGCAAATCGAACTGCGTGTTCTGGCGCATATCTCAAAGACGAG
AACCTGATTGAAGCTTTCACCAATGATATGGATATTCACACCAAAACCGCAATGGAC
GTGTTCCACGTTGCTAAAGACGAGGTTACGTCTGCAATGCGTCGCCAAGCGAAGGCG
GTGAACTTTGGTATTGTTTATGGCATCAGCGACTACGGCTTGAGCCAGAACCTGGGC

ATCACCCGTAAAGAGGCGGGTGCGTTCATTGATCGTTACCTGGAGAGCTTCCAGGGC
GTCAAAGCCTACATGGAGGACTCTGTGCAAGAAGCAAAGCAAAAAGGTTATGTTAC
TACGCTGATGCATAGACGTCGCTACATCCCGGAGTTGACCAGCCGTAATTTTAACAT
TCGCAGCTTCGCCGAACGTACCGCAATGAATACCCCGATTCAAGGTAGCGCGGCTGA
CATCATCAAAAAAGCCATGATCGATATGGCAGCGAACTAAAGGAGAAGCAACTGA
AGGCGCGTCTGCTGCTGCAGGTTACGACGAACTGATTTTCGAAGCTCCGAAAGAGG
AGATCGAAATTCTGGAGAAGTTAGTTCCGGAAGTGATGGAGCACGCACTCGCTTTAG
ATGTGCCACTGAAAGTTGATTTTGCGAGCGGTCCGAGCTGGTATGATGCAAAATAA。

根据本发明的又一个方面，本发明的实施方案提供了一种载体，该载体包含如第二方面任一实施方案所述的核酸。

根据本发明的又一个方面，本发明的实施方案提供了一种试剂盒，包含如第一方面任一实施方案所述的重组蛋白、第二方面任一实施方案所述的核酸、或第三方面任一实施方案所述的载体。

根据本发明的又一个方面，本发明的实施方案提供了一种制备第一方面任一实施方案所述的重组蛋白的方法，该方法包括：用在扩增反应中能够加强 DNA 聚合酶与目的 DNA 相互作用的氨基酸序列替换野生型的 A 家族 DNA 聚合酶中的同源序列获得的氨基酸序列，以获得所述重组蛋白。

在一些实施例中，所述方法包括：对来自所述野生型 A 家族 DNA 聚合酶的多个氨基酸序列进行序列比对，以确定嵌合序列；和用所述嵌合序列替换所述野生型 A 家族 DNA 聚合酶序列中的同源序列，以获得候选重组蛋白序列。

在一些实施例中，所述方法还包括：对所述候选重组蛋白序列进行计算机模拟分析，以确定重组蛋白序列，其中所示重组蛋白序列表现出减小的拇指结构域与手指结构域之间的空间构型。

在一些实施例中，对所述候选重组蛋白序列进行计算机模拟分析，以获得重组蛋白序列，包括：对所述候选重组蛋白序列进行结构预测，以获得重组蛋白的三级结构；对所述重组蛋白的三级结构进行模拟，以获得所述重组蛋白的分子对接结构；对所述分子对接结构进行基于配体的结构优化，以获得优化的重组蛋白结构；和对所述优化的重组蛋白结构进行动力学模拟，以确定重组蛋白序列。

在一些实施例中，所述方法还包括表达并纯化所述重组蛋白序列。

在一些具体实施例中，该方法具体包括：通过 A 家族聚合酶的序列数据库搜寻和同源序列对比，以确定候选嵌合序列用于替换野生型聚合酶中的同源序列；利用 Alphafold、Autodock 和 Gromacs 等工具对根据本发明实施例的具有嵌合序列的重组蛋白进行（1）结

构预测、(2) DNA 双链与重组蛋白的分子对接模拟和(3) 动力学模拟分析；利用原核表达系统对模拟分析确定的重组蛋白进行表达和纯化，并测定其活性；并且上机测序确认蛋白序列。

可以理解的是，该方法中的序列比对包括全局比对、局部比对、双重序列比对、多重序列比对等多种比对策略。

可以理解的是，A 家族 DNA 聚合酶的结构较为保守相似，通过所述制备重组蛋白的方法，能够应用于多种 A 家族 DNA 聚合酶的同源序列替换，并制备所述重组蛋白。

在本发明的实施例中，术语“分子对接”是指通过计算机模拟将小分子（配体）放置于大分子靶标（受体）的结合区域，使其通过调整物理化学参数计算配体与受体的结合力和结合方式，能够获得二者在其活性区域相结合时能量最低的构象。在本发明中，分子对接的配体与受体分别为嵌合聚合酶以及 DNA。模拟分子和蛋白质在原子水平上的相互作用，有助于发现对嵌合聚合酶进行进一步改造的策略。

在本发明的实施例中，术语“动力学模拟”是指用牛顿经典力学计算分子在空间中的轨迹，求解系统中的分子或原子间作用势能和系统外加约束共同作用的分子或原子的牛顿方程，模拟系统随时间推进的微观过程。蛋白质和配体的柔性使它们在相互结合时会寻求最吻合、能量最低的构象。通常情况下，对接完成后，通过打分函数得到的配体蛋白复合物准确度不是很高时，需要借助动力学模拟来计算结合能，验证蛋白与筛选出的小分子配体的有效结合情况。

在一些实施例中，所述 A 家族 DNA 聚合酶选自由大肠杆菌 DNA 聚合酶 I、T3 DNA 聚合酶、T5 DNA 聚合酶、T7 DNA 聚合酶、Taq DNA 聚合酶、Bsu DNA 聚合酶和 Bst DNA 聚合酶组成的组。

在一些实施例中，所述 A 家族 DNA 聚合酶选自具有 SEQ ID NO: 1 所示氨基酸序列的 Bst DNA 聚合酶、具有 SEQ ID NO: 9 所示的氨基酸序列的 Taq DNA 聚合酶、具有 SEQ ID NO: 10 所示氨基酸序列的大肠杆菌 DNA 聚合酶和具有 SEQ ID NO: 11 所示氨基酸序列的 Bsu DNA 聚合酶中的至少一种。

在一些实施例中，所述 A 家族 DNA 聚合酶中的同源序列来源于其拇指结构域。

在一些实施例中，所述 A 家族 DNA 聚合酶中的同源序列选自以下中的至少一种：

如 SEQ ID NO: 2 所示的 Bst DNA 聚合酶中的第 546 位至第 554 位氨基酸序列；

如 SEQ ID NO: 12 所示的 Taq DNA 聚合酶中的第 199 位至第 209 位氨基酸序列；

如 SEQ ID NO: 13 所示的大肠杆菌 DNA 聚合酶中的第 276 位至第 285 位氨基酸序列；

或

如 SEQ ID NO: 14 所示的大肠杆菌 Bsu 聚合酶中的第 249 位至第 257 位氨基酸序列。

在一些实施例中，所述 A 家族 DNA 聚合酶中的同源序列被替换为包含 SEQ ID NO: 3 所示氨基酸序列的序列、基本上由 SEQ ID NO: 3 所示氨基酸序列组成的序列、或由 SEQ ID NO: 3 所示氨基酸序列组成的序列。

在一些实施例中，所述重组蛋白的氨基酸序列选自 SEQ ID NO: 4、SEQ ID NO: 15、SEQ ID NO: 16 和 SEQ ID NO: 17 中的至少一种。

根据本发明的又一个方面，本发明的实施方案提供了一种扩增目的 DNA 的方法，使用前述第一方面任一实施方案所述的重组蛋白对目的 DNA 进行扩增。

在一些实施例中，所述扩增是等温扩增，所述等温扩增选自滚环扩增 RCA、多重置换扩增 MDA、重组酶聚合酶扩增反应 RPA、链置换扩增 SDA、环介导等温扩增 LAMP。

在一些实施例中，所述扩增用于构建 DNA 文库或测序。

实施例

实施例 1

本实施例基于 A 家族 DNA 聚合酶中 Bst 聚合酶、Taq DNA 聚合酶、大肠杆菌 DNA 聚合酶和 Bsu DNA 聚合酶的序列，进行数据库搜寻和 MSA 同源序列对比，以确定候选嵌合序列用于替换野生型聚合酶中的同源序列；利用 Alphafold、Autodock 和 Gromacs 等工具对本发明实施例中的具有嵌合序列的重组蛋白进行结构预测，DNA 双链与重组蛋白的分子对接和动力学模拟分析等计算模拟。

1.1 确定嵌合序列

利用 Protein-blast 工具在 NCBI 蛋白质序列库中比对 A 家族聚合酶的同源序列，确定出了嵌合序列 (SEQ ID NO: 3) 与 Bst DNA 聚合酶拇指结构域中氨基酸第 546-554 位 (SEQ ID NO: 2)、Taq DNA 聚合酶大片段拇指结构域中氨基酸第 199-209 位 (SEQ ID NO: 12)、大肠杆菌 DNA 聚合酶大片段拇指结构域中氨基酸第 276-285 位 (SEQ ID NO: 13) 和 Bsu DNA 聚合酶大片段拇指结构域中氨基酸第 249-257 位 (SEQ ID NO: 14) 同源，比对结果如图 2 所示。该嵌合序列来自人类 DNA 聚合 θ (Human pol theta, POL θ)，并且 SEQ ID NO: 3 比 SEQ ID NO: 2 长约 24 个氨基酸、比 SEQ ID NO: 12 长约 XX 个氨基酸、比 SEQ ID NO: 13 长约 XX 个氨基酸、比 SEQ ID NO: 14 长约 XX 个氨基酸。

1.2 用嵌合序列替换野生型 A 家族 DNA 聚合酶中的同源序列

1.2.1 根据上述同源序列比对，将野生型 Bst DNA 聚合酶大片段 (Bst-LF WT) 中的同源序列 SEQ ID NO: 2 替换为嵌合序列 SEQ ID NO: 3，从而得到嵌合型 Bst DNA 聚合酶，具有 SEQ ID NO: 4 所述的序列。

野生型 Bst DNA 聚合酶大片段 (Bst-LF WT) 的序列 SEQ ID NO: 1:

>Bst-LF WT [Geobacillus stearothermophilus] 氨基酸序列 (293-831)

AEGEKPLEEMEF AIVDVITEEMLADKAALVVEVMEENYHDAPIVGIALVNEHGRFF
MRPETALADSQFLAWLADETKKKSMFDAKRAVVALKWKGIELRGVAFDLLLLAAYLLN
PAQDAGDIAAVAKMKQYEAVRSDEAVYGGVVKRSLPDEQTLAEHLVRKAAAIWALEQ
PFMDDLRNNEQDQLLTKLEQPLAAILAEMEFTGVNVDTKRLEQMGSELAEQLRAIEQRI
YELAGQEFNINSPKQLGVILFEKLQLP**VLKKTGTGYSTS**ADVLEKLAPHHEIVENILHYR
QLGKLQSTYIEGLLKVVRPDTGKVHTMFNQALTQTGRLSSAEPNLQNIPIRLEEGRKIRQ
AFVPSEPDWLIFAADYSQIELRVLAHIADDDNLIEAFQRDLDIHTKTAMDIFHVSEEEVTA
NMRRQAKAVNFGIVYGISDYGLAQNLNITRKEAAEFIERYFASFPGVKQYMENIVQEAK
QKGYVTLLHRRRYLPDITSRNFNVRSAERTAMNTPIQGSAAIIKKAMIDLAARLKEE
QLQARLLLQVHDELILEAPKEEIERLCELVPEVMEQAVTLRVPLKVDYHYGPTWYDAK。

(加粗下划线为同源序列 SEQ ID NO: 2)

嵌合型 Bst DNA 聚合酶 (Bst-HS-1) 的氨基酸序列 (SEQ ID NO: 4):

AEGEKPLEEMEF AIVDVITEEMLADKAALVVEVMEENYHDAPIVGIALVNEHGRFF
MRPETALADSQFLAWLADETKKKSMFDAKRAVVALKWKGIELRGVAFDLLLLAAYLLN
PAQDAGDIAAVAKMKQYEAVRSDEAVYGGVVKRSLPDEQTLAEHLVRKAAAIWALEQ
PFMDDLRNNEQDQLLTKLEQPLAAILAEMEFTGVNVDTKRLEQMGSELAEQLRAIEQRI
YELAGQEFNINSPKQLGVILFEKLQLP**PNREMKNQGSKKTGLGSTRRGIDNGRKLRLGR**
QFSTSADVLEKLAPHHEIVENILHYRQLGKLQSTYIEGLLKVVRPDTGKVHTMFNQALT
QTGRLSSAEPNLQNIPIRLEEGRKIRQAFVPSEPDWLIFAADYSQIELRVLAHIADDDNLIE
AFQRDLDIHTKTAMDIFHVSEEEVTANMRRQAKAVNFGIVYGISDYGLAQNLNITRKEA
AEFIERYFASFPGVKQYMENIVQEAKQKGYVTLLHRRRYLPDITSRNFNVRSAERTAM
NTPIQGSAAIIKKAMIDLAARLKEEQQLQARLLLQVHDELILEAPKEEIERLCELVPEVME
QAVTLRVPLKVDYHYGPTWYDAK。(加粗下划线为嵌合序列 SEQ ID NO: 3)

1.2.2 根据上述同源序列比对,将野生型 Taq DNA 聚合酶大片段 (Taq-LF WT) 中的同源序列 SEQ ID NO: 12 替换为嵌合序列 SEQ ID NO: 3,从而得到嵌合型 Taq DNA 聚合酶,具有 SEQ ID NO: 15 所述的序列。

野生型 Taq DNA 聚合酶大片段 (Taq-LF WT) 的序列 SEQ ID NO: 9:

>Taq-LF WT [Thermus aquaticus] 氨基酸序列 (1-529)

GAFVGFVLSRKEPMWADLLALAAARGGRVHRAPEPYKALRDLKEARGLLAKDLS
VLALREGLGLPPGDDPMLLAYLLDPSNTTPEGVARRYGGEWTEEAGERAAALSERLFANL

WGRLEGEERLLWLYREVERPLSAVLAHVEATGVRLDVAYLRALSLEVAEEIARLEAEVF
 RLAGHPFNLNSRDQLERVLFDELGLP**AIGKTEKTGKR**STSAAVLEALREAHAIVEKILQ
 YRELTKLKSTYIDPLPELIHPRTGRLHTRFNQTATATGRLSSSDPNLQNIPTVPLGQRIRRA
 AFIAEEGWLLVALDYSQIELRVLAHLSGDNLRVVFQEGRDIHTETASWMFGVPREAVDP
 LMRRAAKTINFGVLYGMSAHRLSQELAIPYEEAQAFIERYFQSFPKVRAWIEKTLEEGRR
 RGYVETLFGRRRYVPDLEARVKS VREAAERMAFNMPVQGTAAADLMKLAMVKLFPRLE
 EMGARMMLQVHDEL VLEAPKERA EAVARLAKEVM EGVYPLAVPLEVEVGIGEDWLSA
 KV。（加粗下划线为同源序列 SEQ ID NO: 12）

嵌合型 Taq DNA 聚合酶（Taq-HS-1）的氨基酸序列（SEQ ID NO: 15）：

GAFVGVFLSRKEPMWADLLALAAARGGRVHRAPEPYKALRDLKEARGLLAKDLS
 VLALREGLGLPPGDDPMLLAYLLDPSNTTPEGVARRYGGEWTEEAGERAAALSERLFANL
 WGRLEGEERLLWLYREVERPLSAVLAHVEATGVRLDVAYLRALSLEVAEEIARLEAEVF
 RLAGHPFNLNSRDQLERVLFDELGLP**PNREMKNQGSKKT**LGSTRRGIDNGRKLRLGR
QFSTSAAVLEALREAHAIVEKILQYRELTKLKSTYIDPLPELIHPRTGRLHTRFNQTATAT
 GRLSSSDPNLQNIPTVPLGQRIRRAFI AEEGWLLVALDYSQIELRVLAHLSGDNLRVVF
 QEGRDIHTETASWMFGVPREAVDPLMRRAAKTINFGVLYGMSAHRLSQELAIPYEEAQAF
 IERYFQSFPKVRAWIEKTLEEGRRRGYVETLFGRRRYVPDLEARVKS VREAAERMAFN
 MPVQGTAAADLMKLAMVKLFPRLEEMGARMMLQVHDEL VLEAPKERA EAVARLAKEV
 MEGVYPLAVPLEVEVGIGEDWLSAKV。（加粗下划线为嵌合序列 SEQ ID NO: 3）

1.2.3 根据上述同源序列比对，将野生型大肠杆菌 DNA 聚合酶大片段（Ecoli-LF WT）中的同源序列 SEQ ID NO: 13 替换为嵌合序列 SEQ ID NO: 3，从而得到嵌合型大肠杆菌 DNA 聚合酶，具有 SEQ ID NO: 16 所述的序列。

野生型大肠杆菌 DNA 聚合酶大片段（Ecoli -LF WT）的序列 SEQ ID NO: 10：

> Ecoli -LF WT [Escherichia coli] 氨基酸序列（1-606）

TVISYDNYVTILDEETLKAWIAKLEKAPVFAFDTETDSDLNISANLVGLSFAIEPGVA
 AYIPVAHDYLDAPDQISRERALELLKPLLEDEKALKVGQNLKYDRGILANYGIELRGIAF
 DTMLESYILNSVAGRHDMDSLAERWLKHKTTTFEEIAGKGKNQLTFNQIALDEAGRYAA
 EDADVTLQLHLKMWPDLQKHKGPLNVFENIEMPLVPVLSRIERNGVKIDPKVLHNHSEE
 LTLRLAELEKKAYEIIAGEEFNLSSTKQLQTILFEKQGIK**PLKKT**PGGAPSTSEEVL EELAL
 DYPLPKVILEYRGLAKLKSTYTDKLPLMINPKTGRVHTSYHQAVTATGRLSSTDPNLQNI
 PVRNEEGRRIRQAFIAPEDYVIVSADYSQIELRIMAHLSRDKGLLTAF AEGKDIHRATAAE

VFGLPLETVTSEQRSSAKAINFGLIYGMSAFGLARQLNIPRKEAQKYMDLYFERYPGVLK
YMERTRAQAKEQGYVETLDGRRLYLPDIKSSNGARRAAAERA AINAPMQGTAADIHKRA
MIAVDAWLQAEQPRVRMIMQVHDEL VFEVHRDDVDAVAKQIHQLMENCTRLDVPLL V
EVGSGENWDQAH。(加粗下划线为同源序列 SEQ ID NO: 13)

嵌合型大肠杆菌 DNA 聚合酶 (Ecoli-HS-1) 的氨基酸序列 (SEQ ID NO: 16):

TVISYDNYVTILDEETLKAWIAKLEKAPVFAFD TETDSL DNISANLVGLSFAIEPGVA
AYIPVAHDYLDAPDQISRERALELLKPLLEDEKALKVGQNLKYDRGILANYGIELRGIAF
DTMLESYILNSVAGRHDMDSLAERWLKHK TITFEEIAGKGKNQLTFNQIALDEAGRYAA
EDADVTLQLHLKMWPDLQKHKGPLNVFENIEMPLVPVLSRIERNGVKIDPKVLHNHSEE
LTLRLAELEKKAYEIA GEEFNLSSTKQLQTILFEKQGIK **PNREMKNQGSKKT LGSTRRG**
IDNGRKLRLGRQFSTSEEVL EELALDYPLPKVILEYRGLAKLKSTYTDKLPLMINPKTGR
VHTSYHQA V TATGRLSSTD PNLQNIPVRNEEGRRIRQAFIAPEDYVIVSADYSQIELRIMA
HLSRDKGLLTAF AEGKDIHRATAAEVFGLPLETVTSEQRSSAKAINFGLIYGMSAFGLAR
QLNIPRKEAQKYMDLYFERYPGVLK YMERTRAQAKEQGYVETLDGRRLYLPDIKSSNG
ARRAAAERA AINAPMQGTAADIHKRAMIAVDAWLQAEQPRVRMIMQVHDEL VFEVHRD
DVDAVAKQIHQLMENCTRLDVPLL VEVGSGENWDQAH。(加粗下划线为嵌合序列 SEQ
ID NO: 3)

1.2.4 根据上述同源序列比对, 将野生型 Bsu DNA 聚合酶大片段 (Bsu-LF WT) 中的
同源序列 SEQ ID NO: 14 替换为嵌合序列 SEQ ID NO: 3, 从而得到嵌合型 Bsu DNA 聚合酶,
具有 SEQ ID NO: 17 所述的序列。

野生型 Bsu DNA 聚合酶大片段 (Bsu -LF WT) 的序列 SEQ ID NO: 11:

> Bsu-LF WT [Bacillus subtilis] 氨基酸序列 (1-603)

DINVKT VTDVTS DILVSPSAFVVEQIGDNYHEEPILGFSIVNETGAYFIPKDIAVESEV
FKEWVENDEQKKWVFDSKRAVVALRWQGIELKGAEFDTLAAYIINPGNSYDDVASVA
KDYGLHIVSSDES VYGKGAKRAVPSEDVLSEHLGRKALAIQSLREKL VQELENNDQLEL
FEELEMP LALILGEMESTGVKVDVDRLKRMGEELGAKLKEYEEKIHEIAGEPFNINSPKQ
LGVILFEKIGLP **VVKKT KTGY**STSADVLEKLADKH DIVDYILQYRQIGKLQSTYIEGLLK
VTRPD SHK V HTRFNQALTQTGRLSSTD PNLQNIPIRLEEGRKIRQAFVPSEKDWLIFAADY
SQIELRVL AHISKDENLIEAFTNDMDIHTKTAMDVFHVAKDEVTSAMRRQAKAVNFGIV
YGISDYGLSQNLGITRKEAGAFIDRYLESFQGVKAYMEDSVQEAKQKGYVTTLMHRRR
YIPELTSRNFNIRSFAERTAMNTPIQGSAA DIKKAMIDMAAKLKEKQLKARLLLQVHDE

LIFEAPKEEIEILEKLVPVMEHALALDVPLKVDFASGPSWYDAK。（加粗下划线为同源序列 SEQ ID NO: 14）

嵌合型 Bsu DNA 聚合酶（Bsu-HS-1）的氨基酸序列（SEQ ID NO: 17）:

DINVKTVDVTSDILVSPSAFVVEQIGDNYHEEPILGFSIVNETGAYFIPKDIAVESEV
FKEWVENDEQKKWVFDSKRAVVALRWQGIELKGAEFDTLLAAYIINPGNSYDDVASVA
KDYGLHIVSSDESUYGKGAKRAVPSEDVLSEHLGRKALAIQSLREKLVQELENNDQLEL
FEELEMPALILGEMESTGVKVDVDRCLKRMGEELGAKLKEYEEKIHEIAGEPFNINSPKQ
LGVILFEKIGLPPNREMKNQGSKKTLGSTRRGIDNGRKLRLGRQFSTSADVLEKLADK
HDIVDYILQYRQIGKLQSTYIEGLLKVTRPDSHKVHTRFNQALTQTGRLSSTDPNLQNIPI
RLEEGRKIRQAFVPSEKDWLIFAADYSQIELRVLAHISKDENLIEAFTNDMDIHTKTAMD
VFHVAKDEVTSAMRRQAKAVNFGIVYGISDYGLSQNLGITRKEAGAFIDRYLESFQGVK
AYMEDSVQEAKQKGYVTTLMHRRRYIPELTSRNFNIRSF AERTAMNTPIQGSAAIIKKA
MIDMAAKLKEKQLKARLLLQVHDELIFEAPKEEIEILEKLVPVMEHALALDVPLKVDFAS
GPSWYDAK。（加粗下划线为嵌合序列 SEQ ID NO: 3）

1.3 结构预测

利用 Alphafold2 对嵌合型 A 家族 DNA 聚合酶 Bst-HS-1、Taq-HS-1、Ecoli-HS-1 和 Bsu-HS-1 的序列进行重组蛋白三级结构预测。

1.4 分子对接模拟

基于所预测的结构，利用 Autodock 工具在 dNTP 的存在下对 DNA 双链与嵌合型 A 家族 DNA 聚合酶 Bst-HS-1、Taq-HS-1、Ecoli-HS-1 和 Bsu-HS-1 的进行分子对接模拟，确认改造后具有延长的拇指结构域的嵌合型 DNA 聚合酶与目的 DNA 的结合强度。如图 3 所示，分子对接模拟显示了嵌合序列不仅缩小了拇指结构域与手指结构域之间的空间，从而使模板链保持在拇指结构域和手指结构域之间，而且能够与新合成的双链 DNA 持续保持紧密结合，从而稳定了聚合酶-DNA 结构。

1.5 配体结构优化

用 Gromacs 工具对接后的嵌合型 A 家族 DNA 聚合酶 Bst-HS-1、Taq-HS-1、Ecoli-HS-1 和 Bsu-HS-1 的与 DNA 分子的复合物进行配体结构优化。

1.6 动力学模拟

将经过优化的结构进行 25ns 动力学模拟，模拟温度为 330K。如图 4 所示，嵌合型 Bst DNA 聚合酶中手指结构域中的第 515 至 517 位氨基酸序列 NFN 与拇指结构域中的 K268 可持续保持在 4-6 埃的距离，使 DNA 模板链能被夹在二者形成的口袋结构中，同时延长的区域也能和新合成的双链 DNA 持续保持紧密结合。

此外，实验结构表明，还可以对嵌合序列中的数个位点进行额外突变，以加强嵌合体的热稳定性及聚合活性，并能进一步提高嵌合型聚合酶对 DNA 双链的亲合力及持续合成能力。

对于所有 A 家族聚合酶，如 Bst、Taq、Ecoli、BSU 等，都可以在相应的拇指区域插入该嵌合片段，即可起到与 Bst 上改造结构类似作用，加强与 DNA 双链的亲合力及持续合成能力。

下列数据为 4 种 A 家族聚合酶与其嵌合突变体和 DNA 双链结合能模拟计算的对比，可见所有聚合酶在加入嵌合片段后都能加强与 DNA 双链的亲合力：

1. Bst-LF WT: -164.2 kJ/mol
Bst-HS-1: -272.0 kJ/mol
 $\Delta\Delta G = -107.8$ kJ/mol
2. Taq-LF WT: -150.2 kJ/mol
Taq-HS-1: -277.6 kJ/mol
 $\Delta\Delta G = -127.4$ kJ/mol
3. Ecoli-LF WT: -186.3 kJ/mol
Ecoli-HS-1: -327.9 kJ/mol
 $\Delta\Delta G = -141.6$ kJ/mol
4. BSU-LF WT: -178.4 kJ/mol
BSU-HS-1: -283.3 kJ/mol
 $\Delta\Delta G = -104.9$ kJ/mol

实施例 2

本实施例基于合理设计的嵌合型聚合酶 Bst-HS-1，利用原核表达系统对其进行蛋白的表达和纯化，以获得用于实际应用的 Bst-HS 嵌合型 DNA 聚合酶。

2.1 制备嵌合型 DNA 聚合酶质粒

针对实施例 1 确定的嵌合型 DNA 聚合酶序列（即 SEQ ID NO: 4）设计用于引入嵌入序列（SEQ ID NO: 3）的上游引物和下游引物，以包含编码野生型 Bst DNA 聚合酶（SEQ ID NO: 1）的核酸序列（SEQ ID NO: 6）质粒为模板，使用诺唯赞点突变试剂盒（C112-01）

进行点突变 PCR, 以获得编码 DNA 聚合酶序列(即 SEQ ID NO: 4)的核酸序列(SEQ ID NO: 5)。

编码野生型 Bst DNA 聚合酶的核酸序列 (SEQ ID NO: 6) 的序列为:

GCCGAGGGTGAAAAACCTCTGGAAGAAATGGAATTTGCAATTGTGGATGTAATTACC
GAAGAAATGCTGGCGGATAAAGCCGCTCTTGTGGTGGAAAGTCATGGAAGAAAATA
TCATGATGCGCCGATTGTGGGCATTGCTCTGGTTAATGAACATGGCCGCTTCTTTATG
CGTCCAGAAACAGCCCTGGCCGACTCACAATTTCTGGCCTGGCTGGCGGACGAAACC
AAAAAGAAATCAATGTTTGACGCGAAACGCGCTGTCGTCGCTCTGAAATGGAAGGG
GATCGAACTTCGTGGTGTGGCCTTCGATCTTCTGCTGGCAGCGTATTTGCTGAATCCC
GCCCAGGATGCCGGGGATATTGCAGCCGTGGCGAAAATGAAACAGTACGAAGCGGT
CCGCTCCGACGAGGCCGTGTACGGCAAGGGCGTCAAGCGTTCCTGCCGGATGAAC
AGACCCTGGCAGAACACCTGGTCAGAAAAGCCGCTGCCATCTGGGCACTGGAACAG
CCGTTTATGGATGATCTGCGCAATAATGAGCAGGATCAGCTGCTGACAAAACCTGGAA
CAGCCGCTGGCGGCAATTCTGGCCGAAATGGAATTTACGGGCGTAAATGTGGATACC
AAACGCCTGGAACAGATGGGCTCGGAACTGGCGGAACAGCTGCGTGCCATCGAACA
GCGCATTTATGAACTGGCGGGCCAGGAATTTAATATCAATAGTCCTAAACAGCTGGG
TGTCATTCTGTTTGAAAAACTGCAGTTGCCGGTGCTGAAAAAGACGAAAACCGGTTA
TAGTACCTCAGCGGATGTTCTGGAAAAACTGGCTCCGCATCATGAGATCGTTGAAAA
CATTCTGCATTATCGTCAGCTGGGTAAACTGCAGAGCACCTATATAGAAGGGCTGCT
GAAAGTCGTGCGCCCGGATACGGGTAAAGTCCATACCATGTTTAACCAGGCCCTGAC
CCAGACAGGGCGCCTGAGCAGTGCGGAGCCTAATTTACAGAATATACCGATTCTGTCT
GGAAGAAGGTCGCAAAATTCGCCAGGCGTTTGTGCCAAGCGAACCCGATTGGCTGA
TCTTTGCCGCAGATTACAGCCAGATTGAACTGCGCGTGCTGGCGCATATTGCGGATG
ACGACAATTTAATTGAAGCGTTTCAGCGCGATTTAGATATTCACACCAAAACCGCGA
TGGATATTTTTCATGTGAGCGAGGAAGAAGTTACGGCTAATATGCGTCGTCAGGCCA
AAGCAGTGAATTTTCGGAATTGTTTACGGCATCAGTGATTATGGCCTGGCGCAGAATC
TGAATATTACCCGTAAAGAGGCAGCAGAATTTATTGAACGTTATTTTGCAAGCTTTC
CGGGTGTGAAACAGTATATGGAGAATATCGTTCAGGAAGCCAAACAGAAAGGATAC
GTCACCACCCTGCTGCATCGTCGTCGTTACCTGCCGGATATTACCTCCCGCAATTTTA
ATGTGCGCAGTTTTGCAGAACGTACAGCGATGAATACACCTATTCAGGGCAGCGCGG
CAGATATTATTAAGGAGCGATGATCGATCTGGCAGCACGTCTGAAAGAAGAACAG
CTGCAGGCACGTCTGCTGCTGCAGGTGCATGATGAACTGATTCTGGAAGCGCCGAAA

GAAGAAATCGAACGCCTGTGCGAACTGGTCCCTGAAGTGATGGAACAGGCGGTTAC
CCTGCGTGTACCGCTGAAAGTGGATTATCATTATGGCCCGACCTGGTATGATGCTAA
ATAA。

2.1.1 按照如表 1 所示的 PCR 反应体系和如表 2 所示的 PCR 反应条件进行上述点突变 PCR，以获得包含扩增产物的反应溶液。

表 1 PCR 反应体系

名称	体积
2xMix	10.4 μl
dNTPs（10mM）	0.4 μl
上游引物（SEQ ID NO: 7）+下游引物（SEQ ID NO: 8）（10 μM）	0.8 μl
水	8.4 μl
总体积	20 μl

上游引物（SEQ ID NO: 7）：TACTTTCAGGGCGCCGAGGGTGAAAAACCTCTG

下游引物（SEQ ID NO: 8）：GGGTTAACCTTATTAGCATCATACCAGGTCGGG

表 2 PCR 反应条件

循环数	温度	时间
-	95℃	30 s
25	95℃	10 s
	60℃	30 s
	72℃	1 min
	72℃	10 min
-	4℃	∞

2.1.2 向体积为 20 μl 的包含扩增产物的反应溶液添加 0.4 μl dpn1 酶，在 37℃ 的温度对扩增产物消化 3 小时，以获得包含消化产物的溶液。

2.1.3 取体积为 5 μl 的包含消化产物的溶液对 100 μl 感受态细胞进行转化，并将转化后的感受态细胞涂板于固体 LB 培养基，在 4℃ 过夜孵育。

2.1.4 挑取单克隆菌落，在 37℃ 对其摇菌后送样测序，以获得测序结果。

2.1.5 将测序结果进行比对，以确定含有正确突变的质粒。

2.2 Bst-HS 嵌合型突变体的诱导表达

2.2.1 利用 Bst-HS 嵌合突变体质粒对 BL21 感受态细胞进行转化，以获得转化有

Bst-HS 嵌合突变体的细胞。

2.2.2 向 50 ml 离心管加入转化有 Bst-HS 嵌合突变体的细胞和 20 mL 新鲜培养基，并在 37℃ 的温度以 220 rpm 振摇培养过夜，以获得第一菌液。

2.2.3 将 20 ml 第一菌液转移到含 1 L 培养基的 5 L 锥形瓶中，在 37℃ 以 220 rpm 振摇培养，直到 OD 值为 0.8-1 获得第二菌液。

2.2.4 向第二菌液加入 IPTG 诱导剂达到 IPTG 的终浓度为 0.2 mM，并在 16℃ 的温度进行孵育，持续 16 小时，以获得第三菌液，该第三菌液包含表达 Bst-HS 嵌合突变体的菌体。

2.2.5 将第三菌液以 12000 rpm 离心 10 min，收集作为沉淀物的菌体。

2.3 Bst-HS 嵌合突变体蛋白纯化

2.3.1 将沉淀物混悬于 40 ml A 溶液，并在 700bar 的压力下以高压破碎机对菌体进行破碎，持续 5 分钟，以获得第四菌液。

A 溶液 (pH7.4)

试剂名称	相对分子量	终浓度	用量 (g/L)	5L 用量 (g)
K ₂ HPO ₄ ·3H ₂ O	228.22	50 mM	9.15	45.75
KH ₂ PO ₄	136.09	50 mM	1.35	6.75
NaCl	58.44	500 mM	29.22	146.1
Imidazole	68.08	20 mM	1.36	6.8

2.3.2 在 60℃ 温度的水浴中对第四菌液进行热处理，持续 15 分钟，以获得终溶液。将终溶液以 13000 rpm 离心 30 分钟，然后用 0.22 μm 滤膜对上清液进行过滤并收集滤液。

2.3.3 将滤液加载到镍柱上进行纯化，其中用 50%B 溶液进行洗脱，以获得目的蛋白。

B 溶液 (pH7.4)

试剂名称	相对分子量	终浓度	用量 (g/L)
KH ₂ PO ₄	136.09	50 mM	1.35
NaCl	58.44	500 mM	29.22
Imidazole	68.08	500 mM	34.4

2.3.4 通过透析的方式，将包含目的蛋白的 B 溶液置换为透析液，以获得纯化后的目的蛋白。

透析液（pH7.6）

试剂名称	相 对 分子量	终浓度	用 量 (g/L)	5L 用量 (g)
Tris	121.14	20 mM	2.42	12.1
KCl	75.44	200 mM	14.91	74.55
EDTA-2Na	327.24	0.2 mM	0.0844	0.422

2.3.5 将纯化后的目的蛋白分装储存，以用于后续的活性测定。

实施例 3

本实施例中，对纯化后的嵌合型 DNA 聚合酶 Bst-HS 与野生型 DNA 聚合酶大片段 Bst-LF-WT 进行聚合活性测试，对比二者的持续合成能力。具有 His6 标签的野生型 Bst DNA 聚合酶溶液及实施例 2 制备的具有 His6 标签的嵌合型 DNA 聚合酶 Bst DNA 溶液为待测酶溶液。

3.1 嵌合型 DNA 聚合酶在 RCA 扩增中的合成能力

3.1.1 制备文库

将 40 μl 的初始单链环状 DNA（自合成 Ecoli V3 单链环状文库，1.82 ng/μl）与 1.06 μl 的文库互补引物（5 μM）以 1:1 的摩尔比混合，并按照表 2 所述的条件进行退火，以获得文库+引物的混合物（浓度约为 1.77 ng/μl，）。

3.1.2 RCA 反应

将文库+引物的混合物配制在如下表 3 的 RCA 反应体系中，并在 65℃ 的温度孵育 1 小时。使用 0.1 μl 的 0.5M EDTA 终止扩增反应，以获得 RCA 反应产物。

表 3 RCA 反应体系

试剂名称	体积
文库+引物的混合物	1.7 μl（3ng）
水	3.8 μl
dNTPs(2mM)	3 μl
待测酶溶液	0.5 μl
10X 缓冲液	1 μl
总体积	10 μl

3.1.3 RCA 产物浓度检测

使用 qubit ssDNA 检测试剂盒，按说明书操作，使用 Qubit fluorometer 3.0 检测 RCA 反应产物浓度。结果如下：

Bst-LF-WT: 33.6 ng/uL

Bst-HS-1: 39.9 ng/uL

由此可见，在 RCA 等温扩增中，重组蛋白（即嵌合型 A 家族 DNA 聚合酶）的合成能力高于野生型 A 家族 DNA 聚合酶。

3.2 嵌合型 DNA 聚合酶在 MDA 扩增中的合成能力

3.2.1 DNB 的制备

3.2.1.1 将单链环状 DNA（自合成 Ecoli V3 单链环状文库，浓度为 c）与 DNB 制备缓冲液、和 TE 缓冲液混合在 PCR 管中，并将 PCR 管置于冰盒上约 0.5 小时。使用漩涡振荡器对 PCR 管震荡 5 秒以使混合物混合均匀，短暂离心并置于冰上备用。将 PCR 管置于 PCR 仪中进行如下程序：95℃持续 1min，65℃持续 1min，40℃持续 1min，热盖温度为 102℃，4℃保持。

组分	量 (μl)
单链环状 DNA	3/c
TE 缓冲液	10-3/c
DNB 制备缓冲液	10
总体积	20

3.2.1.2 反应后，将 PCR 管置转移至冰上，并向 PCR 管中加入如下表所示体积的 DNB 聚合酶混合液 I 和 DNB 聚合酶混合液 II。将 PCR 管再次置于 PCR 仪中，进行如下程序：30℃持续 30 分钟，热盖温度 35℃，4℃保持。

步骤 3.2.1.1 获得产物	20μl
DNB 聚合酶混合液 I	20μl
DNB 聚合酶混合液 II	2μl
总体积	42μl

3.2.1.3 向 PCR 管加入 1 微升 DNB 终止缓冲液，震荡混匀以获得 DNB。

3.2.2 MDA 反应

3.2.2.1 取 3.2.1 制备的 DNB，按下述反应体系在 PCR 管中混匀，进行如下程序：95℃

持续 1 分钟，65℃持续 1 分钟，40℃持续 1 分钟，热盖温度为 102℃，4℃保持。

组分	量 (μl)
DNB	20
MP1	6

3.2.2.2 反应后，将 PCR 管转移至冰上，并向 PCR 管中加入下表所示成分。混匀后，将 PCR 管再次置于 PCR 仪中，进行如下程序：65℃持续 30 分钟；热盖温度 70℃，4℃保持。

组分	量 (μl)
3.2.2.1 产物	26
10×缓冲液	4
dNTPs (25 mM)	1
待测酶溶液	2
DEPC	补齐至 40

3.2.3MDA 产物浓度检测

采用 Qubit dsDNA Assay Kit 按说明书操作，使用 Qubit fluorometer 3.0 检测 MDA 产物浓度。

检测结果如下：

Bst-LF-WT: 29.6 ng/uL

Bst-HS-1: 34.8 ng/uL

由此可见，在 MDA 等温扩增中，重组蛋白（即嵌合型 A 家族 DNA 聚合酶）的合成能力高于野生型 A 家族 DNA 聚合酶。

在本发明中，术语“第一”、“第二”仅用于描述目的，而不能理解为指示或暗示相对重要性或者隐含指明所指示的技术特征的数量。由此，限定有“第一”、“第二”的特征可以明示或者隐含地包括至少一个该特征。在本发明的描述中，“多个”的含义是至少两个，例如两个，三个等，除非另有明确具体的限定。

在本发明中，术语“一个实施例”、“一些实施例”、“示例”、“具体示例”、或“一些示例”等意指结合该实施例或示例描述的具体特征、结构、材料或者特点包含于本发明的至少一个实施例或示例中。在本说明书中，对上述术语的示意性表述不必须针对的是相同的实施例或示例。而且，描述的具体特征、结构、材料或者特点可以在任一个或多个实施例或示

例中以合适的方式结合。此外，在不相互矛盾的情况下，本领域的技术人员可以将本说明书中描述的不同实施例或示例以及不同实施例或示例的特征进行结合和组合。

尽管上面已经示出和描述了本发明的实施例，可以理解的是，上述实施例是示例性的，不能理解为对本发明的限制，本领域的普通技术人员在本发明的范围内可以对上述实施例进行变化、修改、替换和变型。

项目名称: FS221389C1P

状态: 已生成

创建日期: 2022-12-30

Last modified: 2022-12-30

基本信息

当前申请

申请号: before

知识产权局:

申请号:

申请日:

申请人档案名: FS221389C1P

发明名称

语言	发明人姓名
zh	重组蛋白及其应用

申请人和发明人:

申请人姓名或名称: 深圳华大智造科技股份有限公司

语言: zh

拉丁名称: MGI Tech Co., Ltd.

居住地址:

通信地址:

序列

序列 1: "SEQ ID NO: 1"

长度	分子类型	生物体	含有 DNA 和 RNA 片段	跳过的序列
587	AA	Bacillus stearothermophilus	否	否

特征

特征键	特征位置	限定符
source	1..587	mol_type = protein organism = Bacillus stearothermophilus

残基

AEGEKPLEEM EFAIVDVITE EMLADKAALV VEVMEENYHD APIVGIALVN EHGRFFMRPE 60

TALADSQFLA WLADETKKKS MFDRAKRAVVA LKWKGIELRG VAFDLLLLAAY LLNPAQDAGD 120
IAAVAKMKQY EAVRSDEAVY GKGVKRS LPD EQTLAEHLVR KAAAIWALEQ PFMDDLRNNE 180
QDQLLTKEQ PLAAILAEME FTGVNVDTKR LEQMGSELAQ QLRAIEQRIY ELAQGEFNIN 240
SPKQLGVILF EKLQLPVLKK TKTGYSTSD VLEKLAPHHE IVENILHYRQ LGKLQSTYIE 300
GLLKVVVRPDT GKVHTMFNQA LTQTGRLLSA EPNLQNIPIR LEEGRKIRQA FVPSEPDWLI 360
FAADYSQIEL RVLAIADDD NLEAFQDRL DIHTKTAMDI FHVSEFEVTA NMRRQAKAVN 420
FGIVYGISDY GLAQNLNITR KEAAEFIERY FASFPQVKQY MENIVQEAQK KGYVTLLHR 480
RRYLPDITSR NFNVSFAER TAMNPIQGS AADIKKAMI DLAARLKEEQ LQARLLQVH 540
DELILEAPKE EIERLCELVP EVMEQAVTLR VPLKVDYHYG PTWYDAK 587

序列 2: "SEQ ID NO: 2"

长度	分子类型	生物体	含有 DNA 和 RNA 片段	跳过的序列
9	AA	Bacillus stearothermophilus	否	否

特征

特征键	特征位置	限定符
source	1..9	mol_type = protein organism = Bacillus stearothermophilus

残基

VLKKTGTGY 9

序列 3: "SEQ ID NO: 3"

长度	分子类型	生物体	含有 DNA 和 RNA 片段	跳过的序列
33	AA	Homo sapiens	否	否

特征

特征键	特征位置	限定符
source	1..33	mol_type = protein organism = Homo sapiens

残基

PNREMKNQGS KKTIGSTRRG IDNGRKRLRG RQF 33

序列 4: "SEQ ID NO: 4"

长度	分子类型	生物体	含有 DNA 和 RNA 片段	跳过的序列
611	AA	Artificial sequence	否	否

特征

特征键	特征位置	限定符
source	1..611	mol_type = protein organism = Artificial sequence

残基

AEEGKPLEEM	EFAIVDVITE	EMLADKAALV	VEVMEENYHD	APIVGIALVN	EHGRFFMRPE	60
TALADSQFLA	WLADETKKKS	MFDAKRAVVA	LKWKGIELRG	VAFDLLLLAY	LLNPAQDAGD	120
IAAVAKMKQY	EAVRSDEAVY	GKGVKRSLPD	EQTIAEHLVR	KAAAIWALEQ	PFMDDLNNNE	180
QDQLLTKELE	PLAAILAEME	FTGVNVDTKR	LEQMGSELAE	QLRAIEQRIY	ELAQQEFNIN	240
SPKQLGVILF	EKLQLPPNRE	MKNQGSKKTLL	GSTRRGIDNG	RKRLRGQFIS	TSADVLEKLA	300
PHHEIVEENIL	HYRQLGKLQS	TYIEGLLKVV	RPDTGKVHTM	FNQALTQTGR	LSSAEPNLQN	360
IPIRLEEGRK	IRQAFVPSEP	DWLIFAADYS	QIELRVLAHI	ADDDNLIEAF	QRDLDIHTKT	420
AMDIFHVSEE	EVTANMRQA	KAVNFGIVYG	ISDYGLAQLN	NITRKEAAEF	IERYFASFPG	480
VKQYMENIVQ	EAKQKGYVTT	LLHRRRYLPD	ITSRNFNVRS	FAERTAMNTP	IQGSAADIHK	540
KAMIDLAAARL	KEEQQLQARLL	LQVHDELILE	APKEEIERLC	ELVPEVMEQA	VTLRVPLKVD	600
YHYGPTWYDA	K					611

序列 5: "SEQ ID NO: 5"

长度	分子类型	生物体	含有 DNA 和 RNA 片段	跳过的序列
1836	DNA	Artificial sequence	否	否

特征

特征键	特征位置	限定符
source	1..1836	mol_type = other DNA organism = Artificial sequence

残基

gccgaggggtg	aaaaacctct	ggaagaaatg	gaatttgcaa	ttgtggatgt	aattaccgaa	60
gaaatgctgg	cggataaaagc	cgctcttggtg	gtggaagtca	tggaaagaaa	ctatcatgat	120
gcgcgattgg	tgggcattgc	tctggttaaat	gaacatggcc	gcttctttat	gcgtccagaa	180
acagccctgg	ccgactcaca	atctctggcc	tggctggcgg	acgaaaccaa	aaagaaatca	240
atgtttgacg	cgaaacgcgc	tgtctcgct	ctgaaatgga	aggggatcga	acttcgtggt	300
gtggccttgc	atcttctgct	ggcagcgtat	ttgctgaatc	ccgcccagga	tgcgggggat	360
attgcagccg	tggcgaaaat	gaaacagtac	gaagcggctc	gctccgacga	ggcctgttac	420
ggcaaggggcg	tcaagcgttc	cctgcccggat	gaacagaccc	tggcagaaca	cctgggtcaga	480
aaagccgctg	ccatctgggc	actggaacag	cggtttatgg	atgatctgcg	caataatgag	540
caggatcagc	tgcctgacaaa	actggaacag	ccgctggcgg	caattctggc	cgaatggaa	600
tttacggggcg	taaaatgtgga	taccaaaacgc	ctggaacaga	tgggctcga	actggcggaa	660
cagctgcgtg	ccatcgaaca	gcgcatttat	gaactggcgg	gccagggaatt	taatatcaat	720
agtcctaaac	agctgggtgt	cattctgttt	gaaaaactgc	agttgcccgc	gaaccgcgaa	780
atgaaaaaac	agggcagcaa	aaaaacccctg	ggcagcacc	gccgcggcat	tgataacggc	840
cgcaaacctgc	gcctgggccc	ccagtttagt	acctcagcgg	atgttctgga	aaaactggct	900
ccgcatcatg	agatcgttga	aaacattctg	cattatcgtc	agctgggtaa	actgcagagc	960
acctatatag	aagggtgct	gaaaagtcgtg	cgcccgata	cggttaaagt	ccataccatg	1020
tttaaccagg	ccctgaccca	gacagggcgc	ctgagcagt	cggagcctaa	tttacagaa	1080
ataccgattc	gtctggaaga	aggtcgcaaa	attcgccagg	cgtttctggt	aagcgaaccc	1140
gattggtcta	tctttgcgcg	agattacagc	cagattgaac	tgcgcgtgct	ggcgcattat	1200
gcggatgacg	acaatttaaat	tgaagcgttt	cagcgcgatt	tagatattca	caccaaaacc	1260
gcgatggata	tttttcatgt	gagcgaggaa	gaagttaacg	ctaatatcg	tgcgcagccc	1320
aaagcagtg	atttcggaat	tgtttacggc	atcagtgtat	atggcctggc	gcagaatctg	1380
aatattacc	gtaaagaggg	agcagaattt	attgaaactg	atgttgcaag	ctttccgggt	1440
gtgaaacagt	atatggagaa	tatcgttcag	gaagccaaac	agaaaaggata	cgtcacccac	1500
ctgctgcatt	gtcgtcgtta	cctgcccggat	attacccccc	gcaattttta	tgtgcgcagt	1560
tttgacagac	gtacagcgat	gaatacacct	attcagggca	gcgcggcaga	tattattaaa	1620
aaagcagtg	tgcattctgg	agcagcgtct	aaagaagaa	agctgcaggc	acgtctgctg	1680
ctgcaggtgc	atgatgaact	gattctggaa	gcgcgcgaa	aaagaatcga	acgcctgtgc	1740
gaactggctc	ctgaagtgtat	ggaacagggc	gttaccctgc	gtgtaccgct	gaaagtggat	1800
tatcatttatg	gcccgcacctg	gtatgatgct	aaataa			1836

序列 6: "SEQ ID NO: 6"

长度	分子类型	生物体	含有 DNA 和 RNA 片段	跳过的序列
1764	DNA	Bacillus stearothermophilus	否	否

特征

特征键	特征位置	限定符
source	1..1764	mol_type = genomic DNA organism = Bacillus stearothermophilus

残基

gccgaggggtg	aaaaacctct	ggaagaaatg	gaatttgcaa	ttgtggatgt	aattaccgaa	60
-------------	------------	------------	------------	------------	------------	----

序列 7: "SEQ ID NO: 7"

特征

残基

序列 8: "SEQ ID NO: 8"

特征

残基

序列 9: "SEQ ID NO: 9"

长度	分子类型	生物体	含有 DNA 和 RNA 片段	跳过的序列
----	------	-----	-----------------	-------

长度	分子类型	生物体	含有 DNA 和 RNA 片段	跳过的序列
529	AA	Thermus aquaticus	否	否

特征

特征键	特征位置	限定符
source	1..529	mol_type = protein organism = Thermus aquaticus

残基

```
GAFVGFVLSR  KEPMWADLLA  LAAARGGRVH  RAPEPYKALR  DLKEARGLLA  KDLSVLALRE  60
GLGLPPGDDP  MLLAYLLDPS  NTTPEGVARR  YGGEWTEEAG  ERAALSERLE  ANLWGRLEGE  120
ERLLWLRYEV  ERPLSAVLAH  VEATGVRLDV  AYLRALSLEV  AEEIARLEAE  VFRLAGHPFN  180
LNSRDQLERV  LFDELGLPAI  GKTEKTGKRS  TSAAVLEALR  EAHAVEKIL  QYRELTKLKS  240
TYIDPLPELI  HPRTGRLHTR  FNQTATATGR  LSSSDPNLQN  IPVRTPLGQR  IRRAFIAEEG  300
WLLVALDYSQ  IELRVLAHLS  GDENLIRVFO  EGRDIHTETA  SWMFQVPREA  VDPLMRRAAK  360
TINFGVLYGM  SAHRLSQELA  IPYEEAQAFI  ERYFQSFPKV  RAWIEKTEE  GRRRGYVETL  420
FGRRRYVPDL  EARVKSVREA  AERMAFNMPV  QGTAADLMKL  AMVKLFPRLE  EMGARMLLQV  480
HDELVLAPK  ERAEAVARLA  KEVMEGVYPL  AVPLEVEVGI  GEDWLSAKV  529
```

序列 10: "SEQ ID NO: 10"

长度	分子类型	生物体	含有 DNA 和 RNA 片段	跳过的序列
606	AA	Escherichia coli	否	否

特征

特征键	特征位置	限定符
source	1..606	mol_type = protein organism = Escherichia coli

残基

```
TVISYDNYVT  ILDEETLKAW  IAKLEKAPVF  AFDTETDSL  NISANLVGLS  FAIEPGVAAY  60
IPVAHDYLD  PDQISRERAL  ELLKPLEDE  KALKVGQNL  YDRGILANYG  IELRGIAFD  120
MLESYILNSV  AGRHDMDSL  ERWLKHKTI  FEEIAGKGN  QLTFNQIALD  EAGRYAAED  180
DVTQLHLKM  WPDQLQHKCP  LNVFENIEMP  LVPVLSRIE  NGVKIDPKVL  HNHSEELTL  240
LAELEKKAY  IAGEEFNLSS  TKQLQTLFE  KQGIKPLKT  PGGAPSTSEE  VLEELALDY  300
LPKVILEYRG  LAKLKSTYTD  KLPLMINPKT  GRVHTSYHQ  VTATGRLSST  DPNLQNI  360
NEEGRRIRQA  FIAPEDYVIV  SADYSQIELR  IMAHLSRDK  LLTAFAEKGD  IHRATAAEV  420
GLPLETVTSE  QRRSAKAINF  GLIYCMSAF  LARQLNIPRK  EAQKYMPLY  ERYPGVLKYM  480
ERTRAQAKEQ  GYVETLDGR  LYLPDIKSSN  GARRAAERA  AINAPMOGTA  ADIIKRAMIA  540
VDAWLQAEQ  RVRMIMQVHD  ELVFEVHRDD  VDAVAKQIH  LMENCTRLDV  PLLVEVGSGE  600
NWDQAH  606
```

序列 11: "SEQ ID NO: 11"

长度	分子类型	生物体	含有 DNA 和 RNA 片段	跳过的序列
579	AA	Bacillus subtilis	否	否

特征

特征键	特征位置	限定符
source	1..579	mol_type = protein organism = Bacillus subtilis

残基

DINVKTVTDV	TSDILVSPSA	FVVEQIGDNY	HEEPILGFSI	VNETGAYFIP	KDIAVESEVF	60
KEWVENDEQK	KWVFDSCRRAV	VALRWQGIEL	KGAEFDTLA	AYIINPGNSY	DDVASVAKDY	120
GLHIVSSDES	VYGGAKRAV	PSEDVLSEHL	GRKALAIQSL	REKLVQELEN	NDQLELFEEL	180
EMPLALILGE	MESTGVKVDV	DRLKRMGEEL	GAKLKEYEEK	IHEIAGEPFN	INSPKQLGVI	240
LEFKIGLPVV	KKTKTGYSTS	ADVLEKLADK	HDIVDYILOQ	RQIGKLQSTY	IEGLLKVTRP	300
DSHKVHTRFN	QALTQTGRLS	STDPNLQNI	IRLEEGRKIR	QAFVPSEKDW	LIFAADYSQI	360
ELRVLAHISK	DENLIEAFTN	DMDIHTKTAM	DVFHVAKDEV	TSAMRRQAKA	VNFGIVYGIS	420
DYGLSQNLGI	TRKEAGAFID	RYLESEFQGVK	AYMEDSVQEA	KQKGYVTTL	HRRRYIPELT	480
SRNFNIRSFA	ERTAMNTPIQ	GSAADIKKKA	MIDMAAKLKE	KQLKARLLLO	VHDELIFEAP	540
KEEIEILEKL	VPEVMEHALA	LDVPLKVDFA	SGPSWYDAK			579

序列 12: "SEQ ID NO: 12"

长度	分子类型	生物体	含有 DNA 和 RNA 片段	跳过的序列
11	AA	Thermus aquaticus	否	否

特征

特征键	特征位置	限定符
source	1..11	mol_type = protein organism = Thermus aquaticus

残基

AIGKTEKTKG R 11

序列 13: "SEQ ID NO: 13"

长度	分子类型	生物体	含有 DNA 和 RNA 片段	跳过的序列
10	AA	Escherichia coli	否	否

特征

特征键	特征位置	限定符
source	1..10	mol_type = protein organism = Escherichia coli

残基

PLKKTPGGAP 10

序列 14: "SEQ ID NO: 14"

长度	分子类型	生物体	含有 DNA 和 RNA 片段	跳过的序列
9	AA	Bacillus subtilis	否	否

特征

特征键	特征位置	限定符
source	1..9	mol_type = protein organism = Bacillus subtilis

残基

VVKKTKTGY

9

序列 15: "SEQ ID NO: 15"

长度	分子类型	生物体	含有 DNA 和 RNA 片段	跳过的序列
551	AA	Artificial sequence	否	否

特征

特征键	特征位置	限定符
source	1..551	mol_type = protein organism = Artificial sequence

残基

GAFVGFVLSR KEPMWADLLA LAAARGGRVH RAPEPYKALR DLKEARGLLA KDLSVLALRE 60
GLGLPPGDDP MLLAYLLDES NTTPEGVARR YGGEWTEEAG ERAALSERLF ANLWGRLEGE 120
ERLLWLRYREV ERPLSAVLAH VEATGVRLDV AYLRAALSLEV AEEIARLEAE VFRLAGHFFN 180
LNSRDQLERV LFDLGLPPN REMKNQGSKK TLGSTRRGID NGRKLRLLGRQ FSTSAAVLEA 240
LREAHAVEK ILQYRELTKL KSTYIDPLPE LIHERTGRLH TRFNQTATAT GRLLLLSSDPNL 300
QNIPVTRPLG QRIRRAFIAE EGWLLVALDY SQIELRVLAH LSGDENLIRV FQEGRDIHTE 360
TASWMFGVPR EAVDPLMRRA AKTINFGVLY GMSAHLRSQE LAIPYEEAAQ FIERYFQSFP 420
KVRAWIEKTL EEGRRRGYVE TLFGRRRYVP DLEARVKSVR FFAERMAFNM PVQGTAAADLM 480
KLAMVKLFPR LEEMGARMML QVHDELVLVEA PKERAEAVAR LAKEVMEGVY PLAVPLEVEV 540
GIGEDWLSAK V 551

序列 16: "SEQ ID NO: 16 "

长度	分子类型	生物体	含有 DNA 和 RNA 片段	跳过的序列
629	AA	Artificial sequence	否	否

特征

特征键	特征位置	限定符
source	1..629	mol_type = protein organism = Artificial sequence

残基

TVISYDNYVT ILDEETLKAW IAKLEKAPVF AFDTETDSDL NISANLVGLS FAIEPGVAAY 60
IPVAHDYLD PDQISRERAL ELLKPLLEDE KALKVGGNLK YDRGILANYG IELRGIAFD 120
MLESYILNSV AGRHDMDSLA ERWLKHKTTT FEEIAGKGN QLTFNQIALD EAGRYAAEDA 180
DVTQLQLHLM WPDQLQHKCP LNVFENIEMP LVPVLSRIER NCVKIDPKVL HNHSEELTLR 240
LAELEKKAYE IAGEEFNLSS TYQLQTLFE KQGIKPNREM KNQGSKKTLG STRRGIDNGR 300
KLRLGRQFST SEEVLEELAL DYPLPKVILE YRGLAKLKST YTDKLPMLIN PKTGRVHTSY 360
HQAVTATGRL SSTDPNLQNI PVRNEEGRRI RQAFIAPEDY VIVSADYSQI ELRIMAHLSR 420
DRGLLTAFAE GKDIHRATAE EVFGLPLETV TSEQRRSAKA INFGLIYGMS AFGLARQLNI 480
PRKEAQKMD LYFERYPGVL KYMERTRAQA KEQGYVETLD GRRLYLPDIK SSNGARRAAA 540
ERAAINAPMQ GTAADIKRA MIAVDLWLQA EQPRVRMIMQ VHDELVFEVH RDDVDVAKQ 600
IHQLMENCTR LDVPLLVEVG SGENWDQAH 629

序列 17: "SEQ ID NO: 17"

长度	分子类型	生物体	含有 DNA 和 RNA 片段	跳过的序列
603	AA	Artificial sequence	否	否

特征

特征键	特征位置	限定符
source	1..603	mol_type = protein organism = Artificial sequence

残基

DINVKTVDV	TSDILVSPSA	FVVEQIGDNY	HEEPILGFSI	VNETGAYFIP	KDIAVESEVF	60
KEWVENDEQK	KWVFDKRAV	VALRWQGIEL	KGAFFDTLLA	AYIINPGNSY	DDVASVAKDY	120
GLHIVSSDES	VYKGAKRAV	PSEDVLSEHL	GRKALAIQSL	REKLVOELEN	NDQLELFEEL	180
EMPLALILGE	MESTGVKVDV	DRLKRMGEEL	GARKLEYEEK	IHEIAGEPFN	INSPKQLGVI	240
LFEKIGLPPN	REMNQGSKK	TLGSTRRGID	NGRKLRLGRQ	FSTSADVLEK	LADKHDIVDY	300
ILQYRQIGKL	QSTYIEGLLK	VTRPDSHKVH	TRFNQALTQT	GRLSSTDPNL	QNIPIRLEEG	360
RRIRQAFVPS	EKDWLIFAAD	YSQIELRVLA	HISKDENLIE	AFTNDMDIHT	KTAMDVFHVA	420
KDEVTSAMRR	OAKAVNFRGIV	YGISDYGLSQ	NLGITRKEAG	AFIDRYLESE	QGVKAYMEDS	480
VQFAKQKGYV	TTLMHRRYI	PELTSRNFNI	RSFAERTAMN	TPIQGSAADI	IKKAMIDMAA	540
KLKEKQLKAR	LLLQVHDELI	FEAPKEEIEI	LEKLVPEVME	HALALDVPLK	VDFASGSPSWY	600
DAK						603

序列 18: "SEQ ID NO: 18"

长度	分子类型	生物体	含有 DNA 和 RNA 片段	跳过的序列
1656	DNA	Artificial sequence	否	否

特征

特征键	特征位置	限定符
source	1..1656	mol_type = other DNA organism = Artificial sequence

残基

ggggcctttcg	taggatttgt	totatcaagg	aaggaaccga	tgtgggcaga	tttgctggcg	60
ctggccgcggg	cgagaggcgg	togcgtacac	cgtgctccgg	agccgtataa	ggcgctgcgt	120
gactttgaagg	aggcgagagg	tttgctggcg	aaagacctgt	ctgttctggc	gctgcgcgag	180
ggccttggtt	taccgcgggg	tgaagatccg	atgctgctcg	cttacctgct	ggaccgcgag	240
aataaccaccc	cgggaaggcgt	cgctaggcgg	tacggcgagg	agtggaccga	agaggctggt	300
gaacgtgcgg	ctctgtccga	acgcttggtc	gccaatctgt	ggggctgctc	cgagggcgaa	360
gaacgcctgt	tatggctgta	cgcgcagggt	gaaagacctc	tgtctgctgt	tcttgcgcac	420
gtggaaagcta	cggcgctgcg	tttagacggt	gcgtacctgc	gtgctctgtc	cctcgagggtg	480
gcggaaagaaa	ttgcacgttt	ggaggcgagaa	gtgttcgctc	tggcggggtca	tccggtttaac	540
ttgaatagcc	gtgaccagct	ggagcgcgtg	ctgttcgatg	aaactgggct	gccaccgaac	600
cgtgaaatga	agaaccaggg	ctcgaaaaaa	accctgggca	gcacgcgtcg	tggtatcgac	660
aatggtcgta	aaactgcgatt	gggcgcgtcag	tttagcacga	gcgcgcgggt	gctggaaagcc	720
ctgcgcgaag	cgcatgcaat	tgttgagaag	atcctgcaat	atcgcgagct	gaccaaaactt	780
aagtgcacct	atatacgatcc	gttaccggaa	ttgateccacc	cgcgtaccgg	tccggtgcatt	840
acccggtttta	accagaccgc	gaacggcaacc	ggcggtttat	ccagctctga	tccaaaatttg	900
caaaaacatcc	cggttcgtac	tccgctgggt	caacgtattc	gtcgtgcatt	catcgcggaat	960
gaaggttggc	tgctagtcgc	cctggactac	agccaaaattg	aaactgctgt	tctggctcac	1020
ctgagtggtg	atgagaacct	gategctgtg	tttcaagaag	ggcgtgacat	ccataccgaa	1080
actgcgagct	ggatgtttcgg	tgtgcccacgt	gaagcgggtgg	atccgctgat	gcgtcgtgca	1140
gctaaagacca	ttaacttttg	tgtgttgtac	ggcatgagcg	cacacccgctt	atcgcaagag	1200
ctggccattc	cgtatgaaga	ggcgcaaggcg	tttatcgagc	gttaacttcca	gagctttccg	1260
aaagtgcgcg	cgtggattga	gaaaaaccctc	gaggaaggcc	gtcgcggtgg	ttatgttgag	1320
acgctgttcg	ctcgccgcgg	ctacgttccg	gatctggagg	cgcgcgtgaa	aagcgttcgt	1380
gaagcggcag	agcgtatggc	attcaacatg	ccggtgcagg	gtacagcggc	ggacttaatg	1440
aaactcgcta	tgggttaagct	gttcccgcg	ttggaggaga	tgggtgcacg	tatgctgcta	1500
caggttcacg	acgaattggt	tcttgaggcg	ctttaaagagc	ggcggggaagc	ggtggccccgc	1560
ctcgccaagg	aggtgatgga	aggcgtctat	ccgctggcgg	ttccgctcga	ggtagagggtc	1620
ggcatcggtg	aagattggct	atccgcgaaa	gttttaa			1656

序列 19: "SEQ ID NO: 19"

长度	分子类型	生物体	含有 DNA 和 RNA 片段	跳过的序列
1890	DNA	Artificial sequence	否	否

特征

特征键	特征位置	限定符
-----	------	-----

特征键	特征位置	限定符
source	1..1890	mol_type = other DNA organism = Artificial sequence

残基

```
actgtgaataa gttatgataa ttacgttaca atcctggacg aagaaacctt gaaggcgtgg 60
attgcgaaac tggaaaaggc gccagtgttc gcttttgata ccgaaacctg ctcgttggac 120
aacattagcg cgaatctagt aggcctgtct ttggcaattg agcccggtgt tgcgcgtat 180
attccggtcg cgaatgacta cctggacgag ccggaccaga tcagccgtga acgtgccttg 240
gaattgctga agccgctgct ggaggatgag aaggcattga aggtgggtga gaaccttaag 300
tacgatcgtg gcatcttgcc caactacggc atcgagctcc ggggtatcgc gttcgacaca 360
atgctggaaa tctatatcct gaacacgctc gccggtcgcc atgatatgga ctgctggcg 420
gaacgttggc tgaagcacaa ttcagagaaa ttgcgggcaa aggcgaagaa 480
cagctgaact ttaatcagat tgcaactggc gaggcaggtc gttatgcagc aaggaagcgt 540
gatgttcccc tgcacctca tctgaagatg tggccggtac tgcagaaaac caagggtccg 600
ctggaacgtg ttgagaacat tgagatgccg ctggtgccgg tcttgtcccg tatcgagcgc 660
aacggtgtga aaatcgaccc taaagtgtct cataatcata gccgaagagt gactttacgt 720
ctcgcgagag tggagaaaaa agcttatgaa attgctggcg aggagttcaa cttgtcctcg 780
acgaaacagc tgcaaaacct tttgtttgaa aagcaagtta tcaaacccga caggagatg 840
aaaaaccagg gcagcaagaa gacctgggc tctaccagac gtggcattga taacggccgc 900
aaactcgctc tgggtcgcca atttagcacg tccgaagaag tgctggagga attggcgttg 960
gactaccgcg tgcgaaaagt gatcttgaa tatcgtggtc tggcgaaact gaagagcacg 1020
tacaccgaca aaatccctct cgtgataaat ccgaagaccg gtcgctgca cactctctac 1080
catcaagcgg ttacggctac cggctgctctg acgagcaccg atccgaatct gcaaaacatc 1140
ccggtacgta atgaagaggg tccggcgcatt cgtcaggcct tcatcgctcc agaggactac 1200
gtgattgtta gcgctgatta ttcaacagatc gaattgcgca tcatggcaca cctgagtgcg 1260
gataaagggtc tgctcactgc gttcgcgcag ggcaaggata tccaccgctc taccgcggcc 1320
gaagtgttcg gcttgcaatt ggagacagtt acgtccgaac agcgtagaag cgcgaaagcg 1380
attaaacttg gcctgattta tggatgagc gcgtttggtc tggcgcgta gttgaatatc 1440
ccgcgttaag aggcaccaaa atacatggat ctgtacttgc agcgttatcc gggcgtgctg 1500
aagtacatgg aaacgcaccg agcacaaact aaagaacagg gttatgttga aacctcgac 1560
ggtcgtcgac tctacctgcc ggacatcaaa agctccaatg gtgcacgtcg cgcggtgcc 1620
gagcgcgcgc ccattaatgc accgatgcaa ggaaccgcag cggacatcat caagcgcgt 1680
atgatcgacg tggatgccg gctgcaagcg gaacagccgc gttgtcgtat gatfatgcag 1740
gtccacgacg agttggtgtt cgaggtgccac cgtgatgatg ttgatgcggt cgcgaagcaa 1800
attcatcagc tgatggaaaa ctgcaccctt ttggacgttc cgctcctggt tgaagttggc 1860
tctggtgaga actgggatca ggctcactaa
```

序列 20: "SEQ ID NO: 20"

长度	分子类型	生物体	含有 DNA 和 RNA 片段	跳过的序列
1812	DNA	Artificial sequence	否	否

特征

特征键	特征位置	限定符
source	1..1812	mol_type = other DNA organism = Artificial sequence

残基

```
gacataaatg ttaaaaactg aacagatgtc acgtccgaca tcctggtgag cccgagtgcg 60
ttcgtcgtag agcagatogg tgataaattat catgaagaa cgttcttgg cttfagcatt 120
gttaaatgaaa ccggtgacct cttcatccca aaagacattg cggtagagag cgaagttttc 180
aaagagtggg ttgaaaacga cgaacagaaa aagtggtttf tgcacagcaa acgcgtgtg 240
gttgccgtcg gctggccagg catcgagttt aaggcccgag aatttgacac cctgctggcg 300
gcgtacattt tcaaccctgg taattcctat gatgatgtag cctccgtggc caaggactac 360
ggcctgcata ttgtgagcag cgatgagtcg gtttacgtga aagtgcgaa gcgcgcagtg 420
cgcgcagagg atgtcctgtc tgaacacctg gccgaaaagg cgtacgcgat tcaatcgctg 480
cgtgaaaagg tggttcaaga actggagaa caccgatcagc tggagttggt tgaagagttg 540
gagatgccgt tggcgctcat cttggcgcaa atggaatcta ctggtgttaa agtggaagtg 600
gaccgtctca agcgttatgg tgaagaaact ggaagcaagc tgaaggaata tgaagagaag 660
attcacgaga tccgcgtgga gccgtttaac atcaacagtc cgaagcagct gggcgtgatf 720
ctgtttgaga aaatcggtct gccacogaat cgtgaaatga aaaaaccagg ctccaaaagg 780
accttgggca gcaacgcgtc fggatcgac aacggtcgca agcttcgct cgtcgtcaa 840
tttagcagct ctgctgacgt cctggagaa ttggcggaca aacacgatat cgtggattac 900
attttgcagt atcgtcagat tggcaagctc caaagcacct atatcgaggg cctgcttaag 960
gtgaccagac ctgattccca taaagtgcac acccgtttca accagggcgt gaccagacc 1020
ggctcgtcga gctcaacoga tccgaactcg caaaacattc cgtccggtt ggaggaaggt 1080
cgcaaaatcc gtcaggcttt cgtgccgtcg gagaaggact ggctgatctt tgcggcggt 1140
tactcgcaaa tgcagctcgc tgttctggcg catatctcca aagacgagaa cctgattgaa 1200
gctttcacca atgatatgga tattcacacc aaaaccgcaa tggacgtgtt ccacgtgct 1260
aaagacgagg ttacgtctgc aatgcgtcgc caagcgagg cggtgaaact tggattgtt 1320
tatggcatca ggcactacgg cttgagccag aacctgggca tcaaccgtta agaggcgggt 1380
cggttcattg atcgttacct ggaagcgttc cagggcgtca aagcctacat ggaggactct 1440
gtgcagaagg caaagcaaaa aggttatgtt actacgctga tgcataagac tgcctacatc 1500
ccgaggttga ccagcgttaa ttttaacatt cgcagcttcg ccgaacgtta cgcctaatga 1560
accccgattc aaggtagcgc ggctgacatc atcaaaaaag coactgatga tatggcagcg 1620
aaactaaagg agaagcaact ctgctgctgc aggttcacga cgaactgatf 1680
ttcgaaagctc cgaagagga gatcgaaatt ctggagaagt tagttccgga agtgatggag 1740
cacgcaactc ctttagatgt gccactgaaa gttgatattf cgagcgggtc gagctgggat 1800
gatgcaaaat aa
```

权利要求书

1. 一种重组蛋白，其特征在于，所述重组蛋白的氨基酸序列是用在扩增反应中能够加强 DNA 聚合酶与目的 DNA 相互作用的氨基酸序列替换野生型的 A 家族 DNA 聚合酶中的同源序列获得的氨基酸序列。

2. 根据权利要求 1 所述的重组蛋白，其特征在于，所述 A 家族 DNA 聚合酶选自由大肠杆菌 DNA 聚合酶 I、T3 DNA 聚合酶、T5 DNA 聚合酶、T7 DNA 聚合酶、Taq DNA 聚合酶、Bsu DNA 聚合酶和 Bst DNA 聚合酶组成的组。

3. 根据权利要求 1 或 2 所述的重组蛋白，其特征在于，所述 A 家族 DNA 聚合酶选自具有 SEQ ID NO: 1 所示氨基酸序列的 Bst DNA 聚合酶、具有 SEQ ID NO: 9 所示氨基酸序列的 Taq DNA 聚合酶、具有 SEQ ID NO: 10 所示氨基酸序列的大肠杆菌 DNA 聚合酶和具有 SEQ ID NO: 11 所示氨基酸序列的 Bsu DNA 聚合酶中的至少一种。

4. 根据权利要求 1 至 3 中任一项所述的重组蛋白，其特征在于，所述 A 家族 DNA 聚合酶中的同源序列来源于其拇指结构域。

5. 根据权利要求 1 至 4 中任一项所述的重组蛋白，其特征在于，所述 A 家族 DNA 聚合酶中的同源序列选自以下中的至少一种：

如 SEQ ID NO: 2 所示的 Bst DNA 聚合酶中的第 546 位至第 554 位氨基酸序列；

如 SEQ ID NO: 12 所示的 Taq DNA 聚合酶中的第 199 位至第 209 位氨基酸序列；

如 SEQ ID NO: 13 所示的大肠杆菌 DNA 聚合酶中的第 276 位至第 285 位氨基酸序列；

或

如 SEQ ID NO: 14 所示的大肠杆菌 Bsu 聚合酶中的第 249 位至第 257 位氨基酸序列。

6. 根据权利要求 1 至 5 中任一项所述的重组蛋白，其特征在于，所述 A 家族 DNA 聚合酶中的同源序列被替换为包含 SEQ ID NO: 3 所示的氨基酸序列的序列、基本上由 SEQ ID NO: 3 所示氨基酸序列组成的序列、或由 SEQ ID NO: 3 所示氨基酸序列组成的序列。

7. 根据权利要求 1 至 6 中任一项所述的重组蛋白，其特征在于，所述重组蛋白的氨基酸序列选自 SEQ ID NO: 4、SEQ ID NO: 15、SEQ ID NO: 16 和 SEQ ID NO: 17 中的至少一种。

8. 根据权利要求 1 至 7 中任一项所述的重组蛋白，其特征在于，所述重组蛋白还包含额外的保守突变、添加和缺失中的任一种。

9. 根据权利要求 1 至 8 中任一项所述的重组蛋白，其特征在于，所述重组蛋白作为嵌合型 DNA 聚合酶的持续合成能力高于野生型的所述 A 家族 DNA 聚合酶。

10. 一种核酸，其特征在于，所述核酸编码如权利要求 1 至 9 中任一项所述的重组蛋

白。

11. 根据权利要求 10 所述的核酸，其特征在于，所述核酸分子的核苷酸序列选自 SEQ ID NO: 5、SEQ ID NO: 18、SEQ ID NO: 19 和 SEQ ID NO: 20 中的至少一种。

12. 一种载体，其特征在于，所述载体包含如权利要求 11 所述的核酸。

13. 一种试剂盒，包含如权利要求 1 至 9 中任一项所述的重组蛋白、如权利要求 10 或 11 所述的核酸、或如权利要求 12 所述的载体。

14. 一种制备权利要求 1 至 9 中任一项所述的重组蛋白的方法，其特征在于，所述方法包括：用在扩增反应中能够加强 DNA 聚合酶与目的 DNA 相互作用的氨基酸序列替换野生型的 A 家族 DNA 聚合酶中的同源序列获得的氨基酸序列，以获得所述重组蛋白。

15. 根据权利要求 14 所述的制备重组蛋白的方法，其特征在于，所述方法包括：

对来自所述野生型 A 家族 DNA 聚合酶的多个氨基酸序列进行序列比对，以确定嵌合序列；和

用所述嵌合序列替换所述野生型 A 家族 DNA 聚合酶序列中的同源序列，以获得候选重组蛋白序列。

16. 根据权利要求 15 所述的方法，其特征在于，所述方法还包括：

对所述候选重组蛋白序列进行计算机模拟分析，以确定重组蛋白序列，

其中所示重组蛋白序列表现出减小的拇指结构域与手指结构域之间的空间构型。

17. 根据权利要求 16 所述的方法，其特征在于，对所述候选重组蛋白序列进行计算机模拟分析，以获得重组蛋白序列，包括：

对所述候选重组蛋白序列进行结构预测，以获得重组蛋白的三级结构；

对所述重组蛋白的三级结构进行模拟，以获得所述重组蛋白的分子对接结构；

对所述分子对接结构进行基于配体的结构优化，以获得优化的重组蛋白结构；和

对所述优化的重组蛋白结构进行动力学模拟，以确定重组蛋白序列。

18. 根据权利要求 16 或 17 所示的方法，其特征在于，还包括：表达并纯化所述重组蛋白序列。

19. 根据权利要求 14 至 18 中任一项所述的方法，其特征在于，所述 A 家族 DNA 聚合酶选自由大肠杆菌 DNA 聚合酶 I、T3 DNA 聚合酶、T5 DNA 聚合酶、T7 DNA 聚合酶、Taq DNA 聚合酶、Bsu DNA 聚合酶和 Bst DNA 聚合酶组成的组。

20. 根据权利要求 14 至 19 中任一项所述的方法，其特征在于，所述 A 家族 DNA 聚合酶选自具有 SEQ ID NO: 1 所示氨基酸序列的 Bst DNA 聚合酶、具有 SEQ ID NO: 9 所示的氨基酸序列的 Taq DNA 聚合酶、具有 SEQ ID NO: 10 所示氨基酸序列的大肠杆菌 DNA 聚合酶和具有 SEQ ID NO: 11 所示氨基酸序列的 Bsu DNA 聚合酶中的至少一种。

21. 根据权利要求 14 至 20 中任一项所述的方法，其特征在于，所述 A 家族 DNA 聚合酶中的同源序列来源于其拇指结构域。

22. 根据权利要求 14 至 21 中任一项所述的方法，其特征在于，所述 A 家族 DNA 聚合酶中的同源序列选自以下中的至少一种：

如 SEQ ID NO: 2 所示的 Bst DNA 聚合酶中的第 546 位至第 554 位氨基酸序列；

如 SEQ ID NO: 12 所示的 Taq DNA 聚合酶中的第 199 位至第 209 位氨基酸序列；

如 SEQ ID NO: 13 所示的大肠杆菌 DNA 聚合酶中的第 276 位至第 285 位氨基酸序列；

或

如 SEQ ID NO: 14 所示的大肠杆菌 Bsu 聚合酶中的第 249 位至第 257 位氨基酸序列。

23. 根据权利要求 14 至 22 中任一项所述的方法，其特征在于，所述 A 家族 DNA 聚合酶中的同源序列被替换为包含 SEQ ID NO: 3 所示氨基酸序列的序列、基本上由 SEQ ID NO: 3 所示氨基酸序列组成的序列、或由 SEQ ID NO: 3 所示氨基酸序列组成的序列。

24. 根据权利要求 14 至 23 中任一项所述的方法，其特征在于，所述重组蛋白的氨基酸序列选自 SEQ ID NO: 4、SEQ ID NO: 15、SEQ ID NO: 16 和 SEQ ID NO: 17 中的至少一种。

25. 一种扩增目的 DNA 的方法，其特征在于，使用权利要求 1 至 9 中任一项所述的重组蛋白对所述目的 DNA 进行扩增。

26. 根据权利要求 25 所述的方法，其特征在于，所述扩增是等温扩增，所述等温扩增选自滚环扩增 RCA、多重置换扩增 MDA、重组酶聚合酶扩增反应 RPA、链置换扩增 SDA 或环介导等温扩增 LAMP。

27. 根据权利要求 25 或 26 所述的方法，其特征在于，所述扩增用于构建 DNA 文库或测序。

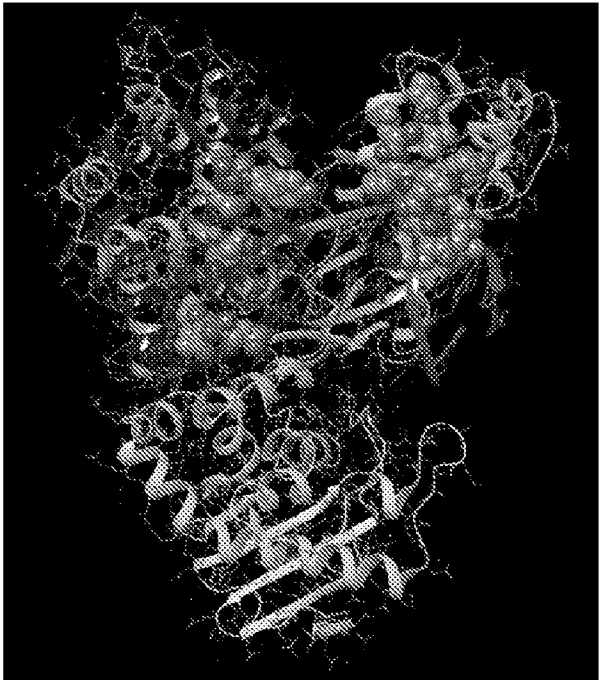
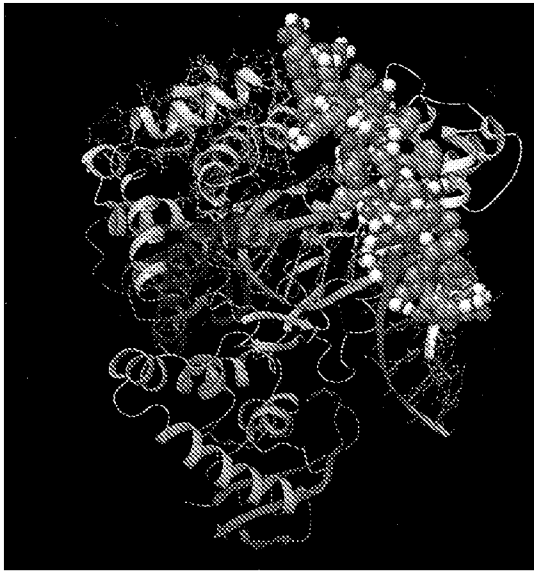


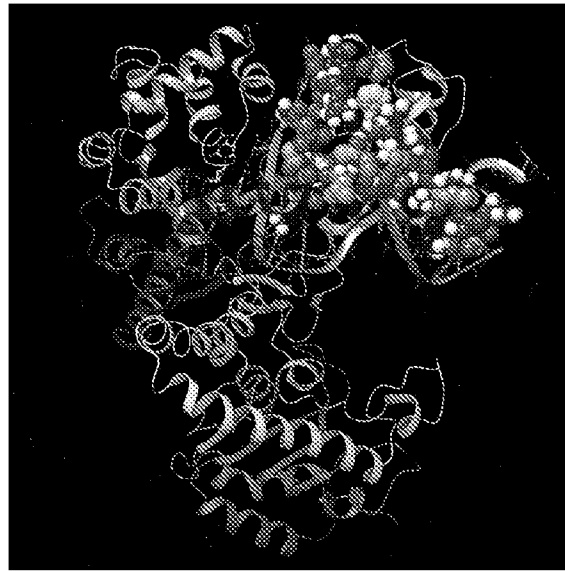
图 1

		-1	306R	316A	320N	336A	346L	356G	366K	375A	386E	
Bst_LF_W71-567	1											100
Bst_HS_U1-611	1											100
		396L	406G	416M	426E	436A	446H	456E	466R	475E	486A	
Bst_LF_W71-567	101											200
Bst_HS_U1-611	101											200
		496D	506E	516G	526F	536V	546V	552-	552-	552T	562E	
Bst_LF_W71-567	201											276
Bst_HS_U1-611	201											300
		572E	582K	592L	602V	612G	622N	632E	642P	652A	662L	
Bst_LF_W71-567	377											376
Bst_HS_U1-611	301											400
		672M	682H	692V	702R	712I	722A	732A	742S	752N	762Y	
Bst_LF_W71-567	377											476
Bst_HS_U1-611	401											500
		772Y	782N	792M	802D	812N	822A	832L	842E	852M	862L	
Bst_LF_W71-567	477											576
Bst_HS_U1-611	501											600
		872W										
Bst_LF_W71-567	677											667
Bst_HS_U1-611	601											611

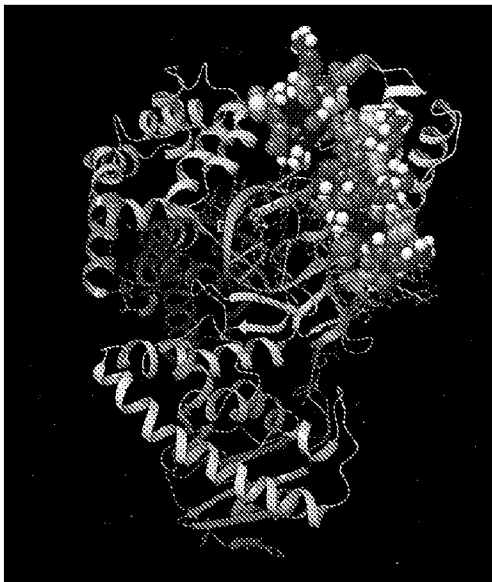
图 2



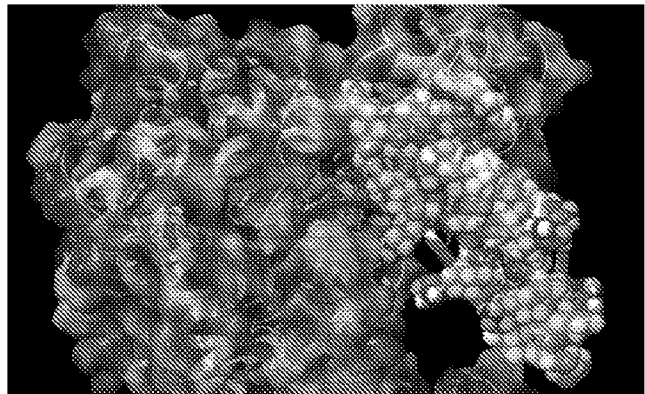
a



b



c



d

图 3

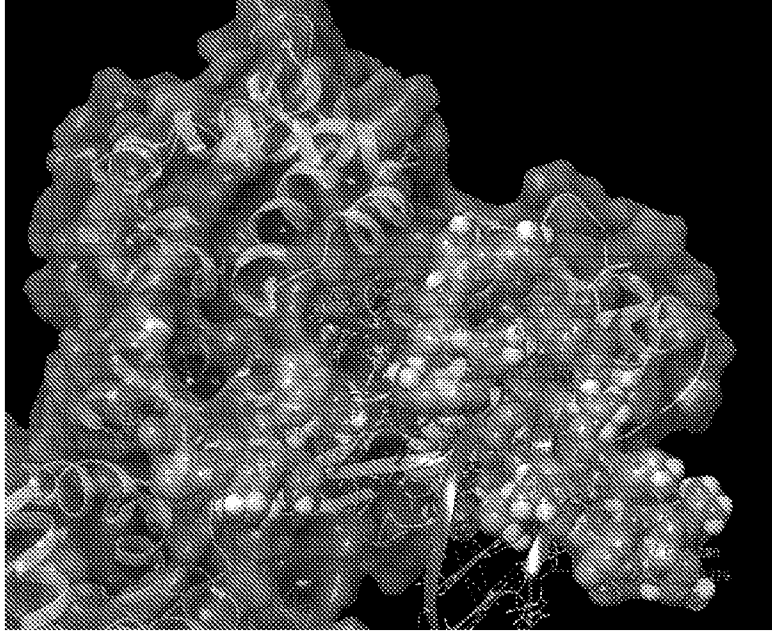


图 4

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2022/144377

A. CLASSIFICATION OF SUBJECT MATTER

C12N9/12(2006.01)i; C12N15/52(2006.01)i; C12N15/10(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNXTX, WPABSC, ENTXTX, ENTXT, USTXT, WOTXT, CJFD, VEN, CNKI, 万方, WANFANG, 百度, BAIDU, PUBMED, ISI Web of Knowledge, NCBI GenBank, EBI-EMBL, 读秀, DUXIU: SEQ ID Nos: 1-20, 重组蛋白, 嵌合, 聚合酶, DNA聚合酶, DNAP, 融合, 替换, 同源, 对比, 比对, Bst, Bsu, phi29, phi29, 拇指结构, TPR2, 结合, 合成, 酶活, recombinant protein, chimera, polymerase, DNA polymerase, fusion, replace, homology, alignment, thumb structure, bind, synthetic, enzymatic activity

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN 102648275 A (CONSEJO SUPERIOR INVESTIGACION) 22 August 2012 (2012-08-22) description, paragraph 12	5, 8-10, 13-22, 25-27
X	PAVLOV, A. R. et al. "Helix-hairpin-helix motifs confer salt resistance and processivity on chimeric DNA polymerases" PNAS, Vol. 99, No. 21, 15 October 2002 (2002-10-15), 13510-13515 pages 13510-13515, and figure 1	1-4, 8-10, 13-14, 19-21, 25-27
Y	PAVLOV, A. R. et al. "Helix-hairpin-helix motifs confer salt resistance and processivity on chimeric DNA polymerases" PNAS, Vol. 99, No. 21, 15 October 2002 (2002-10-15), 13510-13515 pages 13510-13515, and figure 1	5, 8-10, 13-22, 25-27
Y	CN 103881989 A (BIOTECHNOLOGY RESEARCH INSTITUTE, CHINESE ACADEMY OF AGRICULTURAL SCIENCES) 25 June 2014 (2014-06-25) claims 1-4, description, paragraph 41, and figures 1-2	5, 8-10, 13-22, 25-27



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"D" document cited by the applicant in the international application

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

13 September 2023

Date of mailing of the international search report

19 September 2023

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
China No. 6, Xitucheng Road, Jimenqiao, Haidian District,
Beijing 100088

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2022/144377

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	VEGA, M. D. et al. "Improvement of ϕ 29 DNA polymerase amplification performance by fusion of DNA binding motifs" <i>PNAS</i> , Vol. 107, No. 38, 21 September 2010 (2010-09-21), 16506-16511 16506-16511	1-27
A	GAO, Yaping et al. "Chimeric Phi29 DNA polymerase with helix-hairpin-helix motifs shows enhanced salt tolerance and replication performance" <i>Microbial Biotechnology</i> , 31 December 2021 (2021-12-31), 1642-1656 1642-1656	1-27
A	CN 113755465 A (WUHAN ABCLONAL CO., LTD.) 07 December 2021 (2021-12-07) entire document	1-27

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2022/144377

Box No. I **Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)**

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:
 - a. ☒ forming part of the international application as filed.
 - b. ☐ furnished subsequent to the international filing date for the purposes of international search (Rule 13*ter*.1(a)),
☐ accompanied by a statement to the effect that the sequence listing does not go beyond the disclosure in the international application as filed.
2. ☐ With regard to any nucleotide and/or amino acid sequence disclosed in the international application, this report has been established to the extent that a meaningful search could be carried out without a WIPO Standard ST.26 compliant sequence listing.
3. Additional comments:

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2022/144377

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	102648275	A	22 August 2012	PL	2450436	T3	31 March 2014
				DK	2450436	T3	13 January 2014
				WO	2011000997	A1	06 January 2011
				CL	2011003312	A1	01 June 2012
				JP	2012531220	A	10 December 2012
				JP	5778671	B2	16 September 2015
				CO	6480970	A2	16 July 2012
				ES	2359058	A1	18 May 2011
				ES	2359058	B1	27 March 2012
				HK	1173188	A1	10 May 2013
				US	2013236886	A1	12 September 2013
				EP	2450436	A1	09 May 2012
				EP	2450436	A4	28 November 2012
				EP	2450436	B1	02 October 2013
				NZ	597884	A	29 November 2013
				HRP	20131163	T1	14 February 2014
				KR	20120052954	A	24 May 2012
				KR	101751554	B1	28 June 2017
				BR	112012000026	A2	08 September 2015
				US	2012190014	A1	26 July 2012
				US	8404808	B2	26 March 2013
				IL	217324	A0	29 February 2012
				CA	2766725	A1	06 January 2011
				CA	2766725	C	11 October 2016
				PT	2450436	E	27 December 2013
				CY	1114747	T1	14 December 2016
				AU	2010267954	A1	16 February 2012
				AU	2010267954	B2	11 December 2014
				MX	2012000133	A	30 April 2012
				ES	2438786	T3	20 January 2014
				SG	177410	A1	28 February 2012
				SI	2450436	T1	28 February 2014
				SMT	201300153	B	07 March 2014
CN	103881989	A	25 June 2014	WO	2015131302	A1	11 September 2015
				US	2018298357	A1	18 October 2018
				US	10301606	B2	28 May 2019
				US	2016369249	A1	22 December 2016
				US	10047349	B2	14 August 2018
CN	113755465	A	07 December 2021	None			

国际检索报告

国际申请号

PCT/CN2022/144377

A. 主题的分类

C12N9/12(2006.01)i; C12N15/52(2006.01)i; C12N15/10(2006.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

C12N

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

CNXTX, WPABSC, ENTXTX, ENTXT, USTXT, WOTXT, CJFD, VEN, CNKI, 万方, 百度, PUBMED, ISI Web of Knowledge, NCBI GenBank, EBI-EMBL, 读秀:SEQ ID Nos: 1-20, 重组蛋白, 嵌合, 聚合酶, DNA聚合酶, DNAP, 融合, 替换, 同源, 对比, 比对, Bst, Bsu, phi29, ϕ 29, 拇指结构, TPR2, 结合, 合成, 酶活, recombinant protein, chimer+, polymerase, DNA polymerase, fusion, replace+, homology, alignment, thumb structure, bind+, synthetic, enzymatic activity

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
Y	CN 102648275 A (康斯乔最高科学研究公司) 2012年8月22日 (2012 - 08 - 22) 说明书第12段	5, 8-10, 13-22, 25-27
X	PAVLOV, A. R. 等. "Helix-hairpin-helix motifs confer salt resistance and processivity on chimeric DNA polymerases" PNAS, 第99卷, 第21期, 2002年10月15日 (2002 - 10 - 15), 13510-13515 第13510-13515页, 图1	1-4, 8-10, 13-14, 19-21, 25-27
Y	PAVLOV, A. R. 等. "Helix-hairpin-helix motifs confer salt resistance and processivity on chimeric DNA polymerases" PNAS, 第99卷, 第21期, 2002年10月15日 (2002 - 10 - 15), 13510-13515 第13510-13515页, 图1	5, 8-10, 13-22, 25-27
Y	CN 103881989 A (中国农业科学院生物技术研究所) 2014年6月25日 (2014 - 06 - 25) 权利要求1-4, 说明书第41段, 图1-2	5, 8-10, 13-22, 25-27

☒ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“D” 申请人在国际申请中引证的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

2023年9月13日

国际检索报告邮寄日期

2023年9月19日

ISA/CN的名称和邮寄地址

中国国家知识产权局

中国北京市海淀区蓟门桥西土城路6号 100088

授权官员

沈晶晶

电话号码 (+86) 010-53961944

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
A	VEGA, M. D. 等. "Improvement of ϕ 29 DNA polymerase amplification performance by fusion of DNA binding motifs" PNAS, 第107卷, 第38期, 2010年9月21日 (2010 - 09 - 21), 16506-16511 16506-16511	1-27
A	GAO, Y. P. 等. "Chimeric Phi29 DNA polymerase with helix-hairpin-helix motifs shows enhanced salt tolerance and replication performance" Microbial Biotechnology, 2021年12月31日 (2021 - 12 - 31), 1642-1656 1642-1656	1-27
A	CN 113755465 A (武汉爱博泰克生物科技有限公司) 2021年12月7日 (2021 - 12 - 07) 全文	1-27

第I栏

核苷酸和/或氨基酸序列(续第1页第1.c项)

1. 关于国际申请中所公开的任何核苷酸和/或氨基酸序列, 国际检索是基于下列序列表进行的:

- a. ☒ 作为国际申请的一部分提交的;
- b. ☐ 为国际检索的目的在国际申请日之后提交(细则13之三.1(a)),
☐ 附有说明序列表不超出所提交国际申请公开范围的声明。

2. ☐ 本报告是在没有收到符合WIPO ST. 26标准的序列表的情况下, 考虑了国际申请中披露的任何核苷酸和/或氨基酸序列, 在可进行有意义检索的范围内做出的。

3. 补充意见:

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2022/144377

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	102648275	A	2012年8月22日	PL	2450436	T3	2014年3月31日
				DK	2450436	T3	2014年1月13日
				WO	2011000997	A1	2011年1月6日
				CL	2011003312	A1	2012年6月1日
				JP	2012531220	A	2012年12月10日
				JP	5778671	B2	2015年9月16日
				CO	6480970	A2	2012年7月16日
				ES	2359058	A1	2011年5月18日
				ES	2359058	B1	2012年3月27日
				HK	1173188	A1	2013年5月10日
				US	2013236886	A1	2013年9月12日
				EP	2450436	A1	2012年5月9日
				EP	2450436	A4	2012年11月28日
				EP	2450436	B1	2013年10月2日
				NZ	597884	A	2013年11月29日
				HRP	20131163	T1	2014年2月14日
				KR	20120052954	A	2012年5月24日
				KR	101751554	B1	2017年6月28日
				BR	112012000026	A2	2015年9月8日
				US	2012190014	A1	2012年7月26日
				US	8404808	B2	2013年3月26日
				IL	217324	A0	2012年2月29日
				CA	2766725	A1	2011年1月6日
				CA	2766725	C	2016年10月11日
				PT	2450436	E	2013年12月27日
				CY	1114747	T1	2016年12月14日
				AU	2010267954	A1	2012年2月16日
				AU	2010267954	B2	2014年12月11日
				MX	2012000133	A	2012年4月30日
				ES	2438786	T3	2014年1月20日
				SG	177410	A1	2012年2月28日
				SI	2450436	T1	2014年2月28日
				SMT	201300153	B	2014年3月7日
CN	103881989	A	2014年6月25日	WO	2015131302	A1	2015年9月11日
				US	2018298357	A1	2018年10月18日
				US	10301606	B2	2019年5月28日
				US	2016369249	A1	2016年12月22日
				US	10047349	B2	2018年8月14日
CN	113755465	A	2021年12月7日	无			