(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
7 March 2013 (07.03.2013)

(10) International Publication Number
WO 2013/033528 A1

(51) International Patent Classification:
C12N 9/22 (2006.01) *C12P 19/34* (2006.01)

(21) International Application Number:
PCT/US2012/053330

(22) International Filing Date:
31 August 2012 (31.08.2012)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/530,273 1 September 2011 (01.09.2011) US
61/605,484 1 March 2012 (01.03.2012) US

(72) Inventors; and
(71) Applicants : **ONG, Jennifer** [US/US]; 159 Lafayette Street #4, Salem, MA 01970 (US). **EVANS, Jr., Thomas C.** [US/US]; 45 Fox Run Road, Topsfield, MA 01983 (US). **TANNER, Nathan** [US/US]; 2 Magnolia Way, Apartment 224, Peabody, MA 01960 (US).

(74) Agent: **STRIMPEL, HARRIET, M.**; New England Biolabs, Inc, 240 County Road, Ipswich, MA 01938 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report (Art. 21(3))*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*
— *with sequence listing part of description (Rule 5.2(a))*

(54) **Title**: COMPOSITIONS AND METHODS RELATING TO VARIANT DNA POLYMERASES AND SYNTHETIC DNA POLYMERASES

FIG. 5A

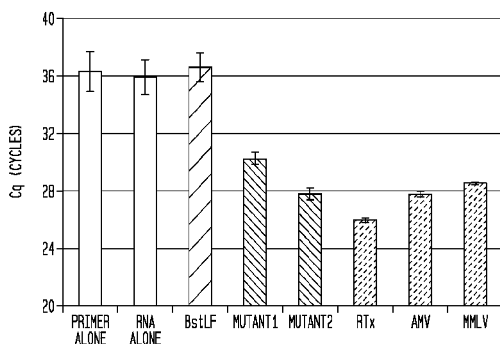
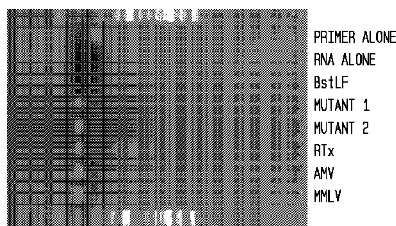


FIG. 5B



(57) **Abstract**: Compositions of novel polymerase variants and methods of identifying, making and using these novel polymerases are described. The variants have been shown to have advantageous properties such as increased thermostability, deoxyuridine nucleoside triphosphate tolerance, salt tolerance, reaction speed and/or increased reverse transcriptase properties. Uses for these improved enzymes have been demonstrated in isothermal amplification such as LAMP. Enhanced performance resulting from the use of these variants in amplification has been demonstrated both in reaction vessels and in dedicated automated amplification platforms.

WO 2013/033528 A1

Compositions and Methods Relating to Variant DNA Polymerases and Synthetic DNA Polymerases

5

FIELD OF THE INVENTION

A DNA polymerase from Geobacillus stearothermophilus has been described in Kong, et al., US Patent No. 5,814,506 (1998). This enzyme, which is a Bst DNA polymerase, belongs to DNA polymerase Family A and shares about 45% sequence identity with its better known relative Taq
10 DNA polymerase. Whereas Taq DNA polymerase is from a hyperthermophilic organism and is able to survive the high temperatures of the polymerase chain reaction, the Bst DNA polymerase reported in Kong, et al., is from a thermophilic organism, is optimally active between 60-70°C, but does not survive the high temperatures of PCR. The full
15 length (FL) Bst DNA polymerase is 876 amino acid residues and has 5'-3' endonuclease activity but not 3'-5' exonuclease activity. The large fragment (LF) of Bst DNA polymerase lacks both 5'-3' exonuclease activity and 3'-5' exonuclease activity and is only 587 amino acid residues with 289 amino acids being deleted from the N-terminal end.
20 The FL Bst DNA polymerase and the LF Bst DNA polymerase have been found to be useful for isothermal amplification techniques and DNA sequencing.

SUMMARY OF EMBODIMENTS OF THE INVENTION

25 Compositions and methods are described herein that relate to variants of DNA polymerases belonging to Family A DNA polymerases.

In embodiment 1, a variant Family A DNA polymerase comprises two or more amino acid sequence motifs selected from 3..EEK..5,
30 15..ADE..17, 65..SPQ..67, 86..RAI..88, 185..LTE..187, 186..TEL..188, 222..LKE..224, 306..VHP..308, 314…HTR..316, 555..LCK..557,

556..CKL..558 and 567..VEL..569, where the number preceding the amino acid in the motif corresponds to the location of that amino acid in the amino acid sequence of Figure 1, wherein the two or more motifs confer improved reaction speed in an amplification reaction and/or improved stability compared to the reaction speed and/or stability of any of SEQ ID NOs:1-23.

Other embodiments are defined in claims 2-49 appended hereto.

In embodiment 2, a variant polymerase of embodiment 1 has at least 75% but less than 100% identity to any of SEQ ID NOs:1-23.

In embodiment 3, a variant polymerase of embodiment 1 or 2 comprises at least three or four or five or six or seven or eight or nine or ten or eleven or twelve of the motifs.

In embodiment 4, a variant polymerase of any one of the preceding embodiments further comprises one or more mutations selected from the group of mutations consisting of (a)-(f) where the mutations in (a)-(f) are:

(a) A1E, G3(K, E or D), K5(L, A or V), E8(M or A), E9D, M10I, A13(D, E, T or V), I14D, V15A, V17(T, E or G), I18V;

(b) E20(M or D) M34(Q or L), E36D, I46F, L48(N or I), M57(L or I), P59(T or A), T61L, D65S, S66(F, E or P);

(c) Q67A, L69(V or K), A73E, M81V, A84R, V88(A or I), R99V, A102D, N113A, D117(T, S or A), A118D, G119(D or E), I121(A or V), V124K, E131H, S135(E or P), V144A, S147(P or A), L148(D or V), Q152(L or P);

(d) T153(A or V), Q170(R or E), M173(L or I), D175E, N178(E, K or R), Q183(L, E or R), L185F, T186(L or I), K187(E or D), Q190(L or M);

(e) A193(I or S), A194(L, S or T), N205(D or K), S216(L or E), R223(V, K or G), A224E, I225(Q or V), V247L, R307H, M316R; and

(f) A330T, D357L, D378N, D380E, I383A, Q387R, L390M; I400V, E406D, A410S, N411R, A433S, N437G, T439K, A452E, Q459(R or E), N463(V, E or D), L484D, D486E, V494L, T501M, Q530R, I552M, E557(K, Q or R) and T568(E or R).

In embodiment 5, a variant polymerase according to embodiment 4, comprising at least one mutated amino acid selected from each of groups (a)-(f).

In embodiment 6, a variant polymerase of embodiment 4 further comprises two, three, four, five, six, seven, eight, nine, ten, eleven, twelve, thirteen, fourteen, fifteen, sixteen, seventeen, eighteen, nineteen, or twenty mutant amino acids at the same position as the corresponding amino acids in (a)-(f) of SEQ ID NO:1.

In embodiment 7, a variant polymerase of any one of the preceding embodiments is described wherein said sequence motif(s) confer one or more improved properties selected from at least one of specific activity; reaction speed; thermostability; storage stability; dUTP tolerance and salt tolerance; increased performance in isothermal amplification; non-interference of pH during sequencing; improved strand displacement; altered processivity; altered ribonucleotide incorporation; altered modified nucleotide incorporation; and altered fidelity when compared to the corresponding parent polymerase.

In embodiment 8, a variant polymerase of any one of the preceding embodiments is described wherein a peptide is fused to one end of the variant polymerase directly or by means of a linker sequence.

5

In embodiment 9, an enzyme preparation comprises a variant polymerase according to any one of the preceding embodiments and a buffer.

10 In embodiment 10, an enzyme preparation according to embodiment 8 or 9 comprises a temperature dependent inhibitor of polymerase activity.

In embodiment 11, an enzyme preparation according to any of 15 embodiments 8 through 10, further comprises dNTPs.

In embodiment 12, a DNA encodes a variant polymerase as described in any of preceding embodiments.

20 In embodiment 13, a host cell comprises the DNA according to embodiment 12.

In embodiment 14, a process for preparing a variant of a parent Family A DNA polymerase having improved polymerase activity 25 compared with the parent polymerase, comprises synthesizing a polypeptide as defined in any one of embodiments 1-8; and characterizing the polymerase activity.

In embodiment 15, the process of embodiment 14 is described 30 wherein characterizing the polymerase activity, further comprises: determining in comparison with the parent polymerase, at last one of:

thermostability; stability in storage; tolerance to salt; performance in isothermal amplification; strand displacement; kinetics; processivity; fidelity; altered ribonucleotide incorporation; altered dUTP incorporation; and altered modified nucleotide incorporation.

5

In embodiment 16, a variant Family A DNA polymerase is obtainable by the process of embodiment 14 or embodiment 15.

In embodiment 17, a variant polymerase of any of embodiments 1
10 through 7 wherein the one or more motifs or one or more mutations selected from the group of mutations consisting of (a)-(f) have improved reverse transcriptase (Rtx) activity.

In embodiment 18, a method for reverse transcribing an RNA of
15 interest, comprises combining an RNA with a DNA polymerase variant or preparation thereof according to embodiments 1-11 to form a complementary DNA (cDNA).

In embodiment 19, a method according to embodiment 18 further
20 comprises amplifying the cDNA by means of the DNA polymerase variant or preparation thereof according to claims 1-11, to produce amplified DNA.

In embodiment 20, a method for amplifying DNA comprises
25 combining a target DNA with a DNA polymerase variant or preparation thereof according to embodiments 1-11, to produce amplified DNA.

In embodiment 21, a variant protein comprises: an amino acid sequence with at least 75% or 80% or 85% or 90% or 95% but less than
30 100% sequence identity to any of SEQ ID NOs:1-23, wherein the variant protein further comprises at least one mutated amino acid having a

position corresponding to SEQ ID NO:1 selected from the group of mutated amino acids consisting of (a)-(f) where the mutations in (a)-(f) are:

(a) A1E, G3(K, E or D), K5(L, A or V), E8(M or A), E9D, M10I, A13(D, E, T or V), I14D, V15A, V17(T, E or G), I18V;

(b) E20(M or D) M34(Q or L), E36D, I46F, L48(N or I), M57(L or I), P59(T or A), T61L, D65S, S66(F, E or P);

(c) Q67A, L69(V or K), A73E, M81V, A84R, V88(A or I), R99V, A102D, N113A, D117(T, S or A), A118D, G119(D or E), I121(A or V), V124K, E131H, S135(E or P), V144A, S147(P or A), L148(D or V), Q152(L or P);

(d) T153(A or V), Q170(R or E), M173(L or I), D175E, N178(E, K or R), Q183(L, E or R), L185F, T186(L or I), K187(E or D), Q190(L or M);

(e) A193(I or S), A194(L, S or T), N205(D or K), S216(L or E), R223(V, K or G), A224E, I225(Q or V), V247L, R307H, M316R; and

(f) A330T, D357L, D378N, D380E, I383A, Q387R, L390M; I400V, E406D, A410S, N411R, A433S, N437G, T439K, A452E, Q459(R or E), N463(V, E or D), L484D, D486E, V494L, T501M, Q530R, I552M, E557(K, Q or R) and T568(E or R).

In embodiment 22, the variant may contain at least one amino acid corresponding to a mutated amino acid in SEQ ID NO:1 selected from each of groups (a) through (f).

In embodiment 23, the variant protein according to embodiment 21, further comprises at least one amino acid motif or at least two amino acid motifs selected from the group consisting of: from 3..EEK..5, 15..ADE..17, 65..SPQ..67, 86..RAI..88, 185..LTE..187,

186..TEL..188, 222..LKE..224, 306..VHP..308, 314...HTR..316, 555..LCK..557, 556..CKL..558 and 567..VEL..569.

In embodiment 24, a variant protein according to any of embodiments 21-23, further comprises two, three, four, five, six, seven, eight, nine, ten, eleven, twelve, thirteen, fourteen, fifteen, sixteen, seventeen, eighteen, nineteen, or twenty amino acids at the same positions as the corresponding mutant amino acids in (a)-(f) of SEQ ID NO:1.

In embodiment 25, the variant protein of embodiment 24 is described wherein the amino acid sequence is at least 80% identical to any one of SEQ ID NOs:1-23.

In embodiment 26, a variant protein according to embodiment 25 is described, wherein the amino acid sequence is at least 90% identical to any one of SEQ ID NOs:1-23.

In embodiment 27, a variant protein according to embodiment 26, is described wherein the amino acid sequence is at least 95% identical to any one of SEQ ID NOs:1-23.

In embodiment 28, a non-naturally occurring synthetic protein comprises: a fragment 1, a fragment 2, a fragment 3, a fragment 4, a fragment 5, a fragment 6, a fragment 7 and a fragment 8 wherein the fragments are covalently linked in numerical order, and wherein:

the fragment 1 is selected from Segment 1 having an amino acid sequence selected from the group consisting of SEQ ID NOs:24-39;

the fragment 2 is selected from Segment 2 having an amino acid sequence selected from the group consisting of SEQ ID NOs:40-56;

the fragment 3 is selected from Segment 3 having an amino acid sequence selected from the group consisting of SEQ ID NOs:57-72;

the fragment 4 is selected from Segment 4 having an amino acid sequence selected from the group consisting of SEQ ID NOs:73-87;

the fragment 5 selected from Segment 5 having an amino acid sequence selected from the group consisting of SEQ ID NOs: 88-99;

5 the fragment 6 selected from Segment 6 having an amino acid sequence selected from the group consisting of SEQ ID NOs:100-111;

the fragment 7 selected from Segment 7 having an amino acid sequence selected from the group consisting of SEQ ID NOs:112-125;

the fragment 8 selected from Segment 8 having an amino acid
10 sequence selected from the group consisting of SEQ ID NOs:126-138; and;

wherein the covalently linked fragments has an amino acid sequence that does not have 100% identity to SEQ ID NOs:1-23.

15 In embodiment 29, a synthetic protein according to embodiment 28 is described, wherein the amino acid sequence of the synthetic protein comprises at least one amino acid sequence motif selected from the group consisting of: 3..EEK..5, 15..ADE..17, 65..SPQ..67, 86..RAI..88, 185..LTE..187, 186..TEL..188, 222..LKE..224,
20 306..VHP..308, 314…HTR..316, 555..LCK..557, 556..CKL..558 and 567..VEL..569.

In embodiment 30, the synthetic protein of embodiment 28 is described, wherein the amino acid sequence comprises at least two or
25 three or four or five or six or seven or eight or nine or ten or eleven or twelve of the amino acid sequence motifs selected from the group consisting of: 3..EEK..5, 15..ADE..17, 65..SPQ..67, 86..RAI..88, 185..LTE..187, 186..TEL..188, 222..LKE..224, 306..VHP..308, 314…HTR..316, 555..LCK..557, 556..CKL..558 and 567..VEL..569.

30

In embodiment 31, a protein comprises at least 75% or 80% or 85% or 90% or 95% sequence identity with SEQ ID NOs:1 and further comprises one or more mutations (such as, for example, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, or 79 mutations) selected from the group consisting of A1E, G3(K, E or D), K5(L, A or V), E8(M or A), E9D, M10I, A13(D, E, T or V), I14D, V15A, V17(T, E or G), I18V, E20(M or D) M34(Q or L), E36D, I46F, L48(N or I), M57(L or I), P59(T or A), T61L, D65S, S66(F, E or P), Q67A, L69(V or K), A73E, M81V, A84R, V88(A or I), R99V, A102D, N113A, D117(T, S or A), A118D, G119(D or E), I121(A or V), V124K, E131H, S135(E or P), V144A, S147(P or A), L148(D or V), Q152(L or P), T153(A or V), Q170(R or E), M173(L or I), D175E, N178(E, K or R), Q183(L, E or R), L185F, T186(L or I), K187(E or D), Q190(L or M), : A193(I or S), A194(L, S or T), N205(D or K), S216(L or E), R223(V, K or G), A224E, I225(Q or V), V247L, R307H, M316R, A330T, D357L, D378N, D380E, I383A, Q387R, L390M, I400V, E406D, A410S, N411R, A433S, N437G, T439K, A452E, Q459(R or E), N463(V, E or D), L484D, D486E, V494L, T501M, Q530R, I552M, E557(K, Q or R) and T568(E or R); and optionally a sequence motif at a specified position in SEQ ID NO:1 selected from the group consisting of: 3..EEK..5, 15..ADE..17, 65..SPQ..67, 86..RAI..88, 185..LTE..187, 186..TEL..188, 222..LKE..224, 306..VHP..308, 314...HTR..316, 555..LCK..557, 556..CKL..558 and 567..VEL..569.

In embodiment 32, a variant protein or a synthetic protein according to any of embodiments of 21-31 is described, wherein a peptide is fused to one end of the variant protein. For example, the peptide may be fused to one end of the variant protein either directly or by means of a linker.

In embodiment 33, an enzyme preparation comprises a variant protein or a synthetic protein according to any of embodiments 21-31 and a buffer.

In embodiment 34, an enzyme preparation according to embodiment 33 further comprises a plurality of proteins.

In embodiment 35, an enzyme preparation according to embodiment 33 or 34 further comprises a reversible inhibitor of polymerase activity.

In embodiment 36, an enzyme preparation according to embodiment 33 or 34 further comprises dNTPs.

In embodiment 37, DNA encodes a variant protein or synthetic protein described in any of embodiments 21-36.

In embodiment 38, a host cell comprises the DNA according to embodiment 37.

In embodiment 39, a method for obtaining a variant of a parent protein has improved polymerase activity compared with the parent protein, comprises synthesizing a protein from any of embodiments 21-36; and characterizing the polymerase activity.

In embodiment 40, which is a method according to embodiment 39, characterizing the polymerase activity further comprises: determining in comparison with the parent protein, at least one of: thermostability; stability in storage; tolerance to salt; performance in isothermal amplification; strand displacement; kinetics; processivity; fidelity; altered ribonucleotide incorporation; altered dUTP incorporation; and altered modified nucleotide incorporation. Additionally, characterizing the polymerase activity includes detecting an increase in Rtx activity.

In embodiment 41, a method comprises:

(a) synthesizing a protein wherein the protein has an amino acid sequence which is capable of being generated from single selected protein fragments obtainable from 8 different segments described in Figure 2; and

(b) assaying the synthetic protein for polymerase activity.

In embodiment 42, a method according to embodiment 41 is provided, wherein the protein is synthesized by cloning a DNA sequence encoding the protein.

In embodiment 43, a method comprises:

(a) selecting a protein variant or synthetic protein according to any of claims embodiments 21-36 having an amino acid sequence; and

(b) expressing the protein variant or synthetic protein as a fusion protein with an additional peptide at an end of the amino acid sequence.

In embodiment 44, a method of isothermal amplification comprises:

(a) providing a preparation comprising a variant protein or synthetic protein according to any of claims 21-36;

(b) combining a target DNA with the preparation; and

(c) amplifying the target DNA at a temperature less than 90°C.

In embodiment 45, a method according to embodiment 44 is described, wherein the amplification reaction results in a quantitative measure of the amount of target DNA in the preparation.

In embodiment 46, a DNA polymerase having one or more improved properties for isothermal amplification compared with SEQ ID NO:1, where the one or more improved properties are selected from the group consisting of:

(a) an increased reaction speed where the increase is at least 10% and as much as 200%; 500% or 1000%;

(b) an increased temperature stability in the range of 50°C to 100°C, 50°C to 90°C, or 60°C to 90°C;

(c) an increased salt tolerance in the range of 10 mM – 1 M, or 20 mM – 200 mM or 500 mM monovalent salt;

(d) an increase in storage stability at 25°C, retaining at least 50% activity over 45 weeks, over 1 year or over 2 years;

(e) an enhanced dUTP tolerance of the range of an increase of 50% to 100% dUTP; and

(f) an increased reverse transcriptase activity by at least 2 fold;

wherein the DNA polymerase is a non-naturally occurring mutant of a wild type Bst DNA polymerase.

In embodiment 47, a DNA polymerase according to embodiment 46 is described having at least two or three or four or five or six of the improved properties.

In embodiment 48, a DNA polymerase according to embodiments 46 or 47 having at least 80% amino acid sequence identity but less than 100% amino acid sequence identity with any of SEQ ID NOs:1-23 and containing at least 12 artificially introduced single amino acid mutations that occur within a three amino acid motif that differs from a three amino acid motif in the corresponding site of a naturally occurring Bst polymerase.

In embodiment 49, a DNA polymerase according to embodiment 48 is described wherein at least one of the three amino acid motifs is selected from the group consisting of 3..EEK..5, 15..ADE..17, 65..SPQ..67, 86..RAI..88, 185..LTE..187, 186..TEL..188, 222..LKE..224, 306..VHP..308, 314...HTR..316, 555..LCK..557, 556..CKL..558 and 567..VEL..569.

In general in one aspect, the composition includes a variant protein, having an amino acid sequence with at least 75% or 80% or 85% or 90% or 95% but less than 100% identity to any of SEQ ID NOs:1-23. The variant protein may include at least one amino acid identified by a position in its amino acid sequence and an identity corresponding to any of the mutated amino acids in the corresponding position in SEQ ID NO:1 and listed in (a)-(f) as provided below, wherein the at least one amino acid is selected from the group consisting of:

(a) A1E, G3(K, E or D), K5(L, A or V), E8(M or A), E9D, M10I, A13(D, E, T or V), I14D, V15A, V17(T, E or G), I18V;

(b) E20(M or D) M34(Q or L), E36D, I46F, L48(N or I), M57(L or I), P59(T or A), T61L, D65S, S66(F, E or P);

(c) Q67A, L69(V or K), A73E, M81V, A84R, V88(A or I), R99V, A102D, N113A, D117(T, S or A), A118D, G119(D or E), I121(A or V), V124K, E131H, S135(E or P), V144A, S147(P or A), L148(D or V), Q152(L or P);

(d) T153(A or V), Q170(R or E), M173(L or I), D175E, N178(E, K or R), Q183(L, E or R), L185F, T186(L or I), K187(E or D), Q190(L or M);

(e) A193(I or S), A194(L, S or T), N205(D or K), S216(L or E), R223(V, K or G), A224E, I225(Q or V), V247L, R307H, M316R; and

(f) A330T, D357L, D378N, D380E, I383A, Q387R, L390M; I400V, E406D, A410S, N411R, A433S, N437G, T439K,

A452E, Q459(R or E), N463(V, E or D), L484D, D486E, V494L, T501M, Q530R, I552M, E557(K, Q or R) and T568(E or R).

5 In another aspect, the variant may contain at least one amino acid corresponding to a mutated amino acid in SEQ ID NO:1 and selected from each of groups (a) through (f).

In another aspect, the variant protein may include in addition to 10 the amino acids specified above, two, three, four, five, six, seven, eight, nine, ten, eleven, twelve, thirteen, fourteen, fifteen, sixteen, seventeen, eighteen, nineteen, or twenty amino acids at the same positions and identities as the corresponding mutant amino acids in (a)-(f) of SEQ ID NO:1.

15

In another aspect, the variant protein may include at least one or two or three or four or five or six or seven or eight or nine or ten or eleven or twelve amino acid sequence motifs selected from 3..EEK..5, 15..ADE..17, 65..SPQ..67, 86..RAI..88, 185..LTE..187, 186..TEL..188, 20 222..LKE..224, 306..VHP..308, 314…HTR..316, 555..LCK..557, 556..CKL..558 and 567..VEL..569, where the number preceding the amino acid in the motif corresponds to the location of that amino acid in the amino acid sequence as determined from Figure 1. The variant protein may include at least one or two or three or four or five or six or 25 seven or eight or nine or ten or eleven or twelve of these motifs in addition to one or more mutations in (a)-(f).

In another aspect, the variant protein has an amino acid sequence that is at least 80%, or at least 85% or at least 90% or at least 95% but 30 less than 100% identical to any one of SEQ ID NOs:1-23.

In another aspect, the variant protein of the sort described above has an amino acid sequence that is at least 80%, or at least 85% or at least 90% or at least 95% but less than 100% identical to any one of SEQ ID NOs:1-23.

5

In another aspect, a DNA polymerase is provided that comprises or consists of a plurality of peptide fragments selected from segments 1-8 covalently linked to form a single polypeptide that has less than 100% amino acid sequence identity with any of SEQ ID NOs:1-23.

10

In another aspect, a non-naturally occurring synthetic protein is provided that includes 8 fragments wherein the fragments include a Fragment 1 selected from Segment 1 and having an amino acid sequence selected from the group consisting of SEQ ID NOs:24-39; a Fragment 2

15 selected from Segment 2 and having an amino acid sequence selected from the group consisting of SEQ ID NOs:40-56, a Fragment 3 selected from Segment 3 and having an amino acid sequence selected from the group consisting of SEQ ID NOs:57-72, a Fragment 4 selected from Segment 4 and having an amino acid sequence selected from the group

20 consisting of SEQ ID NOs:73-87, a Fragment 5 selected from Segment 5 and having an amino acid sequence selected from the group consisting of SEQ ID NOs:88-99; a Fragment 6 selected from Segment 6 and having an amino acid sequence selected from the group consisting of SEQ ID NOs:100-111; a Fragment 7 selected from Segment 7 and having an

25 amino acid sequence selected from the group consisting of SEQ ID NOs:112-125; and a Fragment 8 selected from Segment 8 and having an amino acid sequence selected from the group consisting of SEQ ID NOs:126-138. Fragments 1-8 are covalently linked preferably in numerical order so as to form a single protein wherein the single protein

30 is not any of SEQ ID NOs:1-23.

In another aspect, the amino acid sequence of the synthetic protein comprises at least one or at least two or at least three or at least four or at least five or at least six or at least seven or at least eight or at least nine or at least ten or at least eleven amino acid sequence motifs selected from the group consisting of: 3..EEK..5, 15..ADE..17, 65..SPQ..67, 86..RAI..88, 185..LTE..187, 186..TEL..188, 222..LKE..224, 306..VHP..308, 314...HTR..316, 555..LCK..557, 556..CKL..558 and 567..VEL..569.

In another aspect, a non-naturally occurring protein is provided that comprises or consists of an amino acid sequence having at least 80% sequence identity with SEQ ID NO:1. The non-naturally occurring protein further comprises one or more mutations (such as, for example, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, or 79 mutations) selected from the group consisting of A1E, G3(K, E or D), K5(L, A or V), E8(M or A), E9D, M10I, A13(D, E, T or V), I14D, V15A, V17(T, E or G), I18V, E20(M or D) M34(Q or L), E36D, I46F, L48(N or I), M57(L or I), P59(T or A), T61L, D65S, S66(F, E or P), Q67A, L69(V or K), A73E, M81V, A84R, V88(A or I), R99V, A102D, N113A, D117(T, S or A), A118D, G119(D or E), I121(A or V), V124K, E131H, S135(E or P), V144A, S147(P or A), L148(D or V), Q152(L or P), T153(A or V), Q170(R or E), M173(L or I), D175E, N178(E, K or R), Q183(L, E or R), L185F, T186(L or I), K187(E or D), Q190(L or M), A193(I or S), A194(L, S or T), N205(D or K), S216(L or E), R223(V, K or G), A224E, I225(Q or V), V247L, R307H, M316R, A330T, D357L, D378N, D380E, I383A, Q387R, L390M, I400V, E406D, A410S, N411R, A433S, N437G, T439K, A452E, Q459(R or E), N463(V, E or D), L484D, D486E, V494L, T501M, Q530R, I552M, E557(K, Q or R) and T568(E or R); and optionally

a sequence motif at a specified position in SEQ ID NO:1 selected from the group consisting of: 3..EEK..5, 15..ADE..17, 65..SPQ..67, 86..RAI..88, 185..LTE..187, 186..TEL..188, 222..LKE..224, 306..VHP..308, 314…HTR..316, 555..LCK..557, 556..CKL..558 and
5 567..VEL..569.

In another aspect, a variant or synthetic protein as described herein may additionally comprise a peptide fused to the N-terminal end or the C-terminal end of the protein directly or via a linker.

10

In another aspect of the embodiments, an enzyme preparation is provided which contains a variant protein or a synthetic protein as described above and a buffer. The enzyme preparation may additionally contain a plurality of proteins described herein and/or a reversible
15 inhibitor of polymerase activity and/or dNTPs.

In another aspect of the embodiments, a polynucleotide is provided that encodes a variant protein or synthetic protein as described above. The polynucleotide may be expressed in a transformed host organism.

20

In general, methods are provided for synthesizing a variant or synthetic protein of the type described above having polymerase activity which in one aspect includes synthesizing a protein of the sort described above; and optionally determining whether the protein has a desired
25 property associated with polymerase activity, the polymerase activity being selected from the group consisting of increased thermostability; stability in storage; improved tolerance to salt; increased performance in isothermal amplification; does not alter the pH of a solution during sequencing; improved strand displacement; improved kinetics; altered
30 processivity; altered ribonucleotide incorporation, altered non-standard deoxyribonucleotide incorporation; altered dUTP incorporation; higher

fidelity; and increased Rtx activity as compared with the protein of any of SEQ ID NOs:1-23.

In another aspect, the method includes (a) synthesizing a protein wherein the protein has an amino acid sequence which is capable of being generated from single selected protein fragments obtainable from 8 different segments described in Figure 2; and (b) assaying the synthetic protein for polymerase activity and properties associated therewith. The protein may be synthesized by cloning a DNA sequence encoding the protein.

In another aspect, a method is provided that includes selecting a protein variant or synthetic DNA polymerase protein from those described above; and expressing the protein as a fusion protein with an additional peptide at one or both ends of the DNA polymerase amino acid sequence.

In another aspect, a method is provided for isothermal amplification that includes: (a) providing a preparation comprising of a variant protein or synthetic protein selected from those described above; (b) combining a target DNA with the preparation; and (c) amplifying the target DNA at a temperature less than 90°C to obtain an amplified target and optionally obtaining a quantitative measure of the amount of amplified DNA in the preparation.

In another aspect there is provided a DNA polymerase having one or more improved properties for isothermal amplification compared with SEQ ID NO:1, wherein the improved properties are selected from the group consisting of:

(a) an increased reaction speed in the range where the increase is at least 10% and as much as 20%; 500% or 1000%;

(b) an increased temperature stability in the range of 50°C to 100°C, 50°C to 90°C or 60°C to 90°C;

(c) an increased salt tolerance in the range of 10 mM – 1 M, or 20 mM – 200 mM or 500 mM monovalent salt;

5 (d) an increased storage stability at 25°C, retaining at least 50% activity over 45 weeks, over 1 year or over 2 years;

(e) an enhanced dUTP tolerance of the range of an increase of 50% to 100% dUTP; and

(f) an increased reverse transcriptase activity by at least 2

10 fold.

In the aforementioned aspect the DNA polymerase: (a) may have at least two or three or four or five or six of the improved properties; (b) may have at least 80% amino acid sequence identity but less than 100%

15 amino acid sequence identity with any of SEQ ID NOs:1-23 and containing at least 12 artificially introduced single amino acid mutations that occur in a three amino acid motif that differs from an amino acid in the corresponding site of a naturally occurring Bst polymerase; or (c) may be such that at least one of the three amino acid motifs is selected

20 from the group consisting of 3..EEK..5, 15..ADE..17, 65..SPQ..67, 86..RAI..88, 185..LTE..187, 186..TEL..188, 222..LKE..224, 306..VHP..308, 314…HTR..316, 555..LCK..557, 556..CKL..558 and 567..VEL..569.

25 **BRIEF DESCRIPTION OF THE FIGURES**

Figure 1A shows an alignment of 23 wild type Bst DNA polymerase (LF) sequences. Not shown is a methionine optionally added at the N-terminal end of each of SEQ ID NOs:1-23 to facilitate expression of the polymerase in a host cell.

30

Figure 1B shows sequence pair distances of the sequences in Figure 1A using the software program Lasergene® MegAlign™ (DNASTAR, Madison, WI).

5 Figure 2 shows a 115 fragments arrayed in 8 segments where a fragment selected from each segment joined in order to the neighboring fragment forms an intact synthetic protein having DNA reagent properties.

10 Figure 3A and 3B show melt peaks for a parent Bst DNA polymerase FL or LF and variant DNA polymerases.

Figure 3A shows the melt peaks for the variant DNA polymerase (FL) which has a melting temperature (Tm) = 73.5°C (Δ) and the parent
15 Bst DNA polymerase (FL) has a Tm = 68°C (O).

Figure 3B shows the melt peaks for a parent Bst DNA polymerase LF (O) which has a Tm = 65°C while the variant DNA polymerase (Δ) has a Tm = 70°C.

20

The reactions were performed in 1x Detergent-free ThermoPol™ Buffer (New England Biolabs, Ipswich, MA) and 1x SYPRO® Orange (Life Technologies, Carlsbad, CA).

25 Figures 4A-E show how the properties of a variant DNA polymerase can be screened for significant beneficial properties using an isothermal amplification protocol (Notomi, et al., *Nucleic Acids Res*earch, 28:E63 (2000)) and lambda DNA.

Figure 4A shows an analysis of reaction speed. The variant DNA polymerase shows faster DNA amplification than the parent Bst DNA polymerase.

5 Figure 4B shows the results of an assay to determine salt tolerance. The time in which the amplification reaction took to reach a threshold level of product was graphed against increasing KCl concentration in the reaction. The variant DNA polymerase was more tolerant to changes in salt concentration than the parent Bst DNA 10 polymerase.

Figure 4C shows the results of an assay to determine an increase in thermostability of a variant DNA polymerase by at least 3°C compared with the parent Bst DNA polymerase. The time in which the amplification 15 reaction took to reach a threshold level of product was graphed against increasing reaction temperature. The variant DNA polymerase was able to amplify DNA at a higher temperature than the parent Bst DNA polymerase.

20 Figure 4D shows the results of an assay for storage stability in which a variant polymerase remains stable for at least 28 weeks at room temperature (22°C) versus about 13 weeks for the parent Bst DNA polymerase (8000U/ml for each enzyme was used).

25 Figure 4E shows the results of an assay for dUTP tolerance in which a parent Bst DNA polymerase is significantly inhibited by increasing amounts of dUTP while the variant DNA polymerase activity is relatively stable as dUTP levels increase (1.4 mM dUTP corresponds to complete substitution of dTTP with dUTP). The ability to incorporate dUTP without 30 inhibition of the polymerase is a useful feature of a DNA polymerase for various applications including strand modification and differentiation.

Thermophilic archaeal DNA polymerases do not amplify DNA effectively in the presence of dUTP. Taq DNA polymerase can incorporate dUTP into substrate but Taq DNA polymerase is not suitable for isothermal amplification because it is not capable of the requisite amount of strand displacement.

Figure 5A and 5B shows that the DNA polymerase mutants described herein with improved polymerase activity also have improved reverse transcriptase activity.

Figure 5A shows the results of determining Rtx activity using RT-qPCR. The lower the value of cycles (Cq) the greater the activity of the Rtx. From left to right, the bar chart shows Primer alone, RNA alone, Bst polymerase large fragment (BstLF), 2 mutants of the DNA polymerase described herein, Rtx, Avian Myeloblastosis Virus Reverse Transcriptase (AMV) and Moloney Murine Leukemia Virus Reverse Transcriptase (MMLV).

Figure 5B shows gel electrophoresis of amplified DNA resulting from an RNA template and BstLF DNA polymerase or mutants.
The lanes are labeled left to right as follows: primer alone, RNA alone, BstLF, Mutant 1 and 2, Rtx, AMV and MMLV.

DETAILED DESCRIPTION OF THE EMBODIMENTS

As used herein, the term “synthetic” with respect to proteins or peptides refers to a non-naturally occurring amino acid sequence that is generated either by expression of a gene encoding the non-naturally occurring amino acid sequence or is generated by chemical synthesis. The gene encoding the non-naturally occurring amino acid sequence may be generated, for example, by mutagenesis of a naturally occurring gene sequence or by total chemical synthesis.

A "variant" protein refers to a protein that differs from a parent protein by at least one amino acid that is the product of a mutation. A variant polymerase is intended to include a "synthetic" protein and vice versa as the context permits. The examples utilize a variant DNA polymerase but it will be understood to a person of ordinary skill in the art that the assays described in the examples are applicable to analyzing synthetic proteins also.

"Non-naturally occurring" refers to a sequence or protein that at the date in which the embodiments of the invention are presented herein, no naturally occurring amino acid sequence corresponding to the alleged non-naturally occurring amino acid has been described in the publically available databases.

"Isothermal amplification" refers to a DNA amplification protocol that is conducted at a temperature below 90°C after an initial denaturation step, where an initial denaturation step is required.

The term "stability" as used in the claims includes thermostability and storage stability as illustrated in Figure 4 and in the examples.

We have developed a set of variant proteins that are mutants of a highly conserved family of DNA polymerases belonging to Family A DNA polymerases. One or more of the amino acid mutations and/or amino acid motifs described herein are capable of enhancing the properties of these polymerases such as those properties determined by the assays described in the examples.

The Family A DNA polymerases are highly conserved so that it will be readily appreciated that with the teaching of the present

embodiments, a person of ordinary skill in the art could select a naturally occurring DNA polymerase sequence (such as from GenBank) having at least 80% sequence identity with SEQ ID NOs:1-23 and introduce one or more of the specified mutations and/or motifs described herein to obtain polymerases with improved properties such as the type described in the examples.

In one embodiment, the DNA polymerase mutant proteins comprise or consist of an amino acid sequence that has at least 75% amino acid sequence identity, at least 80% amino acid sequence identity, or at least 85% amino acid sequence identify and as much as 90% amino acid sequence identity or 95% amino acid sequence identity to the parent DNA polymerase provided in the sequences described in SEQ ID NOs:1-23 wherein the amino acid sequence is less than 100% identical to the amino acid sequence of any of SEQ ID NOs:1-23.

Percentage sequence identity may be calculated by any method known in the art such as for example, using the BLOSUM62 matrix and the methods described in Henikoff, et al., *PNAS,* 89 (22):10915–10919 (1992)).

The at least one amino acid mutation in the variants is identified using the numbering scheme described in Figure 1 with a reference amino acid as it occurs in SEQ ID NO:1 replaced by a desired amino at the specified position.

Accordingly, a parent polymerases having amino acid sequences with at least 75%, 80%, 85%, 90%, or 95% sequence identity to any of SEQ ID NOs:1-23 may be altered by at least one mutation selected from the group consisting of: A1E, G3(K, E or D), K5(L, A or V), E8(M or A), E9D, M10I, A13(D, E, T or V), I14D, V15A, V17(T, E or G), I18V, E20(M

or D) M34(Q or L), E36D, I46F, L48(N or I), M57(L or I), P59(T or A), T61L, D65S, S66(F, E or P), Q67A, L69(V or K), A73E, M81V, A84R, V88(A or I), R99V, A102D, N113A, D117(T, S or A), A118D, G119(D or E), I121(A or V), V124K, E131H, S135(E or P), V144A, S147(P or A), L148(D or V), Q152(L or P), T153(A or V), Q170(R or E), M173(L or I), D175E, N178(E, K or R), Q183(L, E or R), L185F, T186(L or I), K187(E or D), Q190(L or M), A193(I or S), A194(L, S or T), N205(D or K), S216(L or E), R223(V, K or G), A224E, I225(Q or V), V247L, R307H, M316R, A330T, D357L, D378N, D380E, I383A, Q387R, L390M, I400V, E406D, A410S, N411R, A433S, N437G, T439K, A452E, Q459(R or E), N463(V, E or D), L484D, D486E, V494L, T501M, Q530R, I552M, E557(K, Q or R) and T568(E or R).

The variant may optionally include one or more motifs selected from the group consisting of: 3..EEK..5, 15..ADE..17, 65..SPQ..67, 86..RAI..88, 185..LTE..187, 186..TEL..188, 222..LKE..224, 306..VHP..308, 314…HTR..316, 555..LCK..557, 556..CKL..558 and 567..VEL..569.

The DNA polymerase protein variants described above may be screened using at least one method described in Examples 1-6 so as to identify those variants having at least one of the functional properties that are at least typical of a Family A DNA polymerase, such as, Bst DNA Polymerase with an amino acid sequence corresponding to SEQ ID NO:1. The DNA polymerase may additionally have improved properties as compared with the wild type Family A DNA polymerases such as those including one of specific activity, reaction speed, thermostability, storage stability, dUTP tolerance, salt tolerance and reverse transcriptase activity.

In another embodiment, a synthetic protein is described that contains sequences from single fragments selected from each of 8 segments assembled in order of the 8 numbered segments (see Figure 2). The synthetic protein may be synthesized either as a single DNA or protein sequence or as a set of polynucleotides or peptides that are ligated together using techniques known in the art (see for example Gibson Assembly™ Master Mix (New England Biolabs, Ipswich, MA), US Patent No. 7,435,572 or US Patent No. 6,849,428):

a Fragment 1 selected from Segment 1 having an amino acid sequence selected from the group consisting of SEQ ID NOs:24-39;

a Fragment 2 selected from Segment 2 having an amino acid sequence selected from the group consisting of SEQ ID NOs:40-56;

a Fragment 3 selected from Segment 3 having an amino acid sequence selected from the group consisting of SEQ ID NOs:57-72;

a Fragment 4 selected from Segment 4 having an amino acid sequence selected from the group consisting of SEQ ID NOs:73-87;

a Fragment 5 selected from Segment 5 having an amino acid sequence selected from the group consisting of SEQ ID NOs:88-99;

a Fragment 6 selected from Segment 6 having an amino acid sequence selected from the group consisting of SEQ ID NOs:100-111;

a Fragment 7 selected from Segment 7 having an amino acid sequence selected from the group consisting of SEQ ID NOs:112-125;

a Fragment 8 selected from Segment 8 having an amino acid sequence selected from the group consisting of SEQ ID NOs:126-138.

A proviso for creating a synthetic protein is that the synthetic protein has a sequence that differs from any SEQ ID NOs:1-23.

Preferably, a synthetic protein comprising segments 1-8 has at least one, two, three, four, five, six, seven, eight, nine or 10 sequence motifs selected from 3..EEK..5, 15..ADE..17, 65..SPQ..67, 86..RAI..88,

185..LTE..187, 186..TEL..188, 222..LKE..224, 306..VHP..308, 314...HTR..316, 555..LCK..557, 556..CKL..558 and 567..VEL..569.

The synthetic proteins described herein and characterized by a non-natural amino acid sequence generally retain DNA binding properties making these synthetic proteins useful for example as DNA detection reagents. The variants may be screened using at least one method described in Examples 1-6, or by other screening methods common used in the art, so as to identify those variants having at least one of the functional properties that are at least typical of a Family A DNA polymerase and/or have one or more improved properties selected from at least one of specific activity; reaction speed; thermostability; storage stability; dUTP tolerance and salt tolerance; increased performance in isothermal amplification; non-interference of pH during sequencing; improved strand displacement; altered processivity; altered ribonucleotide incorporation; altered modified nucleotide incorporation; and altered fidelity when compared to the corresponding parent polymerase. The improved properties of these mutant enzymes have been demonstrated to enhance the performance of sequencing platforms (for example, the Ion Torrent™ sequencer (Life Technologies, Carlsbad, CA)). The improved properties of these mutant enzymes enhance their use in isothermal amplification for diagnostic applications.

The DNA polymerase variants and synthetic proteins described herein may be expressed in suitable non-native host cells such as *E. coli* according to standard methods known in the art. To facilitate expression, the variant DNA polymerase may additionally have a methionine in front of the first amino acid at the N-terminal end. Host cells may be transformed with DNA encoding the variant optionally contained in a suitable expression vector (see New England Biolabs catalog 2019-10 or 2011-12 for expression vectors known in the art for

this purpose). Transformation is achieved using methods well known in the art.

The DNA polymerase variants and synthetic proteins characterized herein may further be modified by additions and/or deletions of peptides at their N-terminal and/or C-terminal ends. For example, fusion of a peptide to a synthetic protein may include fusion of one or more of a DNA binding domain (such as Sso7d from archaea), an exonuclease domain (such as amino acids 1-289 of Bst DNA polymerase), a peptide lacking exonuclease activity (for example, a mutated exonuclease domain similar to amino acids 1-289 of Bst DNA polymerase), an affinity binding domain such as a Histidine tag, chitin binding domain, or intein, and a solubility tag such as maltose binding domain (MBP). The addition of a peptide fused to an end of the amino acid sequence of the DNA polymerase may be used to enhance one or more of the functional features described in Examples 1-6. Aptamers may be fused to one end of the mutant DNA polymerase.

The variants may be stored in a storage or reaction buffer that includes a detergent such as a non-ionic detergent, a zwitterionic detergent, an anionic detergent or a cationic detergent. The storage or reaction buffer may further include one or more of: a polynucleotide, for example, an aptamer for facilitating a hot start; polynucleotide primers, dNTPs, target polynucleotides; additional polymerases including additional DNA polymerases; RNA polymerases and/or reverse transcriptases; crowding agents such as polyethylene glycol; and/or other molecules known in the art for enhancing the activity of the DNA polymerase variants.

The DNA polymerase variant and synthetic proteins may be used for DNA synthesis, DNA repair, cloning and sequencing (see for example

US Patent No. 7,700,283 and US Application Publication No. US 2011/0201056) and such as illustrated in the examples and also for temperature dependent amplification methods. Examples of isothermal amplification methods in addition to loop-mediated isothermal amplification (LAMP) used in the present examples include helicase dependent amplification (HDA) (see for example US Patent No. 7,829,284, US Patent No. 7,662,594, and US Patent No. 7,282,328); strand displacement amplification (SDA); nicking enzyme amplification reaction; recombinase polymerase amplification; padlock amplification; rolling circle amplification; and multiple displacement amplification (see for example Gill, et al., *Nucleosides, Nucleotides and Nucleic Acids*, 27:224-243 (2008)). The variant and synthetic DNA polymerases described herein may also be used in sample preparation for sequencing by synthesis techniques known in the art. The variant and/or synthetic polymerases may also be used in quantitative amplification techniques known in the art that may be performed at a temperature at which the variant or synthetic protein effectively polymerizes nucleotides.

EXAMPLES

The examples below illustrate assays and properties of Bst DNA polymerase variants described above.

Example 1: Assay for determining the properties of a variant DNA polymerase

(a) Loop-mediated isothermal amplification (LAMP)

The properties of a variant polymerase can be determined using an isothermal amplification procedure such as a LAMP protocol (Nagamine, et al., *Mol. Cell. Probes*, 16:223-229(2002); Notomi, et al., *Nucleic Acids Research*, 28:E63 (2000)).

The LAMP reaction used bacteriophage λ genomic DNA (New England Biolabs, Ipswich, MA) as the template. The LAMP primers used

here were:
FIP (5'-CAGCCAGCCGCAGCACGTTCGCTCATAGGAGATATGGTAGAGCCGC-3') (SEQ ID NO:139),
BIP (5'GAGAGAATTTGTACCACCTCCCACCGGGCACATAGCAGTCCTAGGGAC
5 AGT-3') (SEQ ID NO:140),
F3 (5'-GGCTTGGCTCTGCTAACACGTT-3') (SEQ ID NO:141),
B3 (5'-GGACGTTTGTAATGTCCGCTCC-3') (SEQ ID NO:142),
LoopF (5'-CTGCATACGACGTGTCT-3')(SEQ ID NO:143),
LoopB (5'-ACCATCTATGACTGTACGCC-3') (SEQ ID NO:144).
10

The LAMP reaction used 0.4 U–0.2 U variant Polymerase/μL, 1.6 μM FIP/BIP, 0.2 μM F3/B3, 0.4 μM LoopF/LoopB, and 5 ng lambda DNA in a buffer containing 1x ThermoPol Detergent-free, 0.1% Tween 20, 6-8mM $MgSO_4$ and 1.4μMdNTP. The reaction was followed by monitoring
15 turbidity in real time using the Loopamp® Realtime Turbidimeter LA-320c (SA Scientific, San Antonio, TX) or with a CFX96™ Real-Time fluorimeter (Bio-Rad, Hercules, CA). The reaction conditions were varied to determine the optimum range that the variant DNA polymerase could perform LAMP. This was compared with the parent Bst DNA polymerase.
20 The parent Bst DNA polymerase was typically used at 65°C in these LAMP reaction conditions. However, the temperature was varied to determine the optimum temperature for a particular variant. Different salt conditions and rates of reaction were tested and variants identified which were 10%-50% faster than the parent polymerase and had an increased
25 salt tolerance to as much as 200mM KCl.
The results are shown in Figure 4.

(b) DNA polymerase activity assay using modified nucleotides in a comparison of the activity of a fusion variant protein with
30 exonuclease activity, with full length parent Bst polymerase.

This assay was used to determine the activity of the variant

polymerase having exonuclease activity as a result of an additional 289 amino acid sequence at the N-terminal end that has been described in detail for parent DNA Bst polymerase. The activity was measured by incorporation of a radioactive ^{3}H-dTTP in a DNA substrate using various concentrations of a variant polymerase. A DNA polymerase reaction cocktail (40 µl) was prepared by mixing 30 nM single-stranded M13mp18, 82 nM primer #1224 (5'-CGCCAGGGTTTTCCCAGTCACGAC-3') (SEQ ID NO:145), 200 µM dATP, 200 µM dCTP, 200 µM dGTP, and 100 or 200 µM dTTP including 0.6 to 0.8 µCi [3H]-dTTP. The DNA polymerase reaction cocktail was mixed with DNA polymerase (2.2 to 8.7 ng for the parent Bst DNA polymerase (FL), 0.27 to 1 ng for the fusion variant, or 2.5 to 20 ng for the parent BstLF), or water for the no enzyme control, and incubated at 65°C for 5 minutes. Reactions were halted and precipitated by acid precipitation as follows. A 30 µl aliquot of each reaction was spotted onto 3 mm Whatman discs and immediately submerged into cold 10% Trichloroacetic acid (TCA) in 1 L beaker in an ice bucket. A total counts control was spotted as described but not washed. Filters were washed three times with cold 10% TCA for 10 minutes with vigorous shaking and twice with room temperature 95% isopropanol for 5 minutes. Filters were dried under a heat lamp for 10 minutes and counted using a scintillation counter. The pmoles of dNTPs incorporated were calculated for each sample from the fraction of radioactive counts incorporated, multiplied by the total amount of dNTPs and the volume of the reaction.

A tenfold increase in specific activity of the fusion variant polymerase was found compared with the parent FL Bst polymerase where the fusion variant DNA polymerase was present in the mixture at 506,000 U/mg while the parent Bst DNA polymerase was present at 48,000 U/mg. (1 unit = incorporation of 10 nmol dNTP in 30 minutes at 65°C).

A 15% increase in activity of the variant polymerase compared with the parent Bst large fragment DNA polymerases was observed in which the variant DNA polymerase was present in the mixture at 370,000 U/mg and the parent BstLF was present at 260,000 U/mg.

Example 2: Variant DNA polymerase thermostability

The thermostability of the variant DNA polymerase was assessed by incubating the polymerase at differing temperatures followed by performing either one or both of the DNA polymerase assay described in Example 1. The results are shown in Figure 4C.

Example 3: Inhibitor resistance of the variant DNA polymerase

The resistance of a variant DNA polymerase to inhibitors such as blood is determined by adding increasing concentrations of the inhibitor into the DNA polymerase assay and determining the change, if any, in the apparent specific activity of the protein. The DNA polymerase assay was performed as described in Example 1 at 65°C.

Another inhibitor of DNA polymerase is dUTP which is used to prevent carryover contamination in isothermal amplification by replacing dTTP. In this case it is desirable for the polymerase to be insensitive to dUTP inhibition so as to utilize dUTP as a substrate for LAMP. Figure 4E shows that the mutant polymerase can efficiently utilize dUTP while the wild type Bst polymerase is inhibited by substituting dTTP with dUTP in the amplification reaction.

Example 4: Increased resistance to high salt concentration

The resistance of a variant DNA polymerase to increased salt concentration was determined by adding increasing concentrations of salt (for example, KCl or NaCl) to the DNA polymerase assay described in Example 1 and determining the activity of the protein at 65°C and

comparing its activity to parent Bst DNA polymerase (see Figure 4B).

Example 5: Increased stability in storage

The stability of a variant DNA polymerase during storage was
5 determined by incubating the enzyme in storage buffer (10 mMTris-HCl pH 7.5, 50 mM KCl, 1 mM Dithiothreitol, 0.1 mM EDTA, 50% Glycerol, 0.1% Triton X-100) at a temperature ranging from 4°C to 65°C for a time period ranging from 1 day to 28 weeks, and assaying DNA polymerase activity remaining after storage using the LAMP method
10 described in Example 1. The remaining activity was compared to a sample stored at -20°C for the same amount of time. The stability of the variant was then compared to the stability of parent Bst DNA polymerase (See Figure 4D). When this period was extended to 60 weeks, no detectable loss of activity of the mutants was observed even in the
15 absence of glycerol.

Example 6: Assay for determining the melting temperature of a variant polymerase for comparison with a parent DNA polymerase using a SYPRO Orange assay

20 The assay was performed as follows: Each 50 µl reaction contains 1x ThermoPol Buffer, detergent-free (20 mM Tris-HCl pH 8.8, 10 mM $(NH_4)_2SO_4$, 10 mM KCl, 2 mM MgSO4, 1x SYPRO Orange protein gel stain, and DNA polymerase concentrations ranging from 2.2 to 17.5 µg (parent BstLF mutant) or 0.6 to 4.8 µg (parent Bst FL mutant). The
25 reactions were placed in a CFX96 Real-Time System. The temperature was raised 1°C per second from 20 to 100°C, and the fluorescence (in the FRET channel) was read at each temperature. Here, the Tm is the inflection point of the sigmodial curve of fluorescence plotted against temperature. The inverted first derivative of the fluorescence emission in
30 Figure 3A and 3B is shown in relation to temperature, where the location of the minima corresponded to the value of the Tm (see Figure 3).

Example 7: Whole genome amplification using a variant Bst DNA polymerase

The variant DNA polymerase can be tested for suitability in whole genome amplification using the methods termed hyperbranched strand displacement amplification (Lage, et al., *Genome Research*, 13 (2):294-307 (2003)) or multiple-strand displacement amplification (Aviel-Ronen, et al., *BMC Genomics*, 7:312 (2006)).

Example 8: DNA sequencing on a semiconductor device using a variant DNA polymerase

The variant DNA polymerase can be tested for its suitability in DNA sequencing, for example, as described in Rothberg, et al., *Nature*, 475(7356):348-352(2011), an integrated semiconductor device enabling non-optical genome sequencing.

Example 9: Solid-phase DNA amplification using a variant polymerase

Variant DNA polymerase can be tested for its suitability in solid-phase DNA amplification, for example as described in (Adessi, et al., *Nucleic Acids Research,* 28:E87 (2000), which describes a method for the amplification of target sequences with surface bound oligonucleotides.

Example 10: Enhanced reverse transcriptase activity

The reverse activity of the mutant Bst DNA polymerase was determined using a two-step RT-qPCR assay (Sambrook, et al., *Molecular Cloning – A Laboratory Manual*, 3rd ed., Cold Harbor Laboratory Press (2001)). The first step was for cDNA synthesis using the mutant enzymes and various traditional reverse transcriptases. The second measures the amount of synthesized cDNA by qPCR. The RT step was performed using 6uM Hexamer (Random Primer Mix, New England Biolabs, Ipswich, MA) as primers in Isothermal Amplification Buffer (New

England Biolabs, Ipswich, MA) supplemented with 6mM Mg and 200uM dNTP with 0.1ug Jurkat Total RNA (Life Technologies, Carlsbad, CA) and incubated at 65°C for 20 minutes. 1ul of the RT product was added to qPCR reaction for GAPDH gene with 200nM of forward (5'-AGAACGGGAAGCTTGTCATC) (SEQ ID NO:146) and reverse primer (5'-CGAACATGGGGGCATCAG) (SEQ ID NO:147), 200uM dNTP, 1.25 unit of Taq DNA polymerase in 25ul of 1x Standard Taq Buffer (New England Biolabs, Ipswich, MA) containing 2 uM of dsDNA-binding fluorescent dye SYTO®9 (Life Technologies, Carlsbad, CA). The PCR cycles were: 95°C for 1 minute, then 50 cycles at 95°C for 10 seconds, 61°C for 15 seconds and 68°C for 30 seconds, and a final step of 68°C for 5 minutes. The PCR was performed on a CFX96 Real-Time PCR machine and the Cq value was obtained as an indication of the amount of specific cDNA being synthesized (Figure 5A). Mutant 1 and mutant 2 (4th and 5th bar from left in bar chart) make abundant cDNA as indicated by having Cq values similar to that of traditional RTs (6TH, 7th, and 8th bar from left) in qPCR. Wild type BstLF (3rd bar from the left) is the same as controls (1st and 2nd bar from left) without RT. After completion of the PCR reaction, 10ul of PCR product was analyzed by electrophoresis in a 1.5% agarose gel (Figure 5B) to verify the size of the PCR product. The lanes from left to right are primer alone, RNA alone, BstLF, mutant 1, mutant 2, Rtx, AMV and MMLV. Mutant 1, mutant 2 and all RTs (Rtx, AMV and MMLV) lanes gave a band of expected size (207 base pairs) but no specific band with wild type BstLF or controls. These results demonstrate that mutant 1 and mutant 2 has much improved Rtx activity compared to wild type BstLF.

All references cited herein, as well as U.S. provisional application serial number 61/530,273 filed September 1, 2011 and U.S. provisional application serial number 61/605,484 filed March 1, 2012, are herein incorporated by reference.

What is claimed is:

1. A variant Family A DNA polymerase comprising two or more amino acid sequence motifs selected from 3..EEK..5, 15..ADE..17, 65..SPQ..67, 86..RAI..88, 185..LTE..187, 186..TEL..188, 222..LKE..224, 306..VHP..308, 314...HTR..316, 555..LCK..557, 556..CKL..558 and 567..VEL..569, where the number preceding the amino acid in the motif corresponds to the location of that amino acid in the amino acid sequence of Figure 1, wherein the two or more motifs confer improved reaction speed in an amplification reaction and/or improved stability compared to the reaction speed and/or stability of any of SEQ ID NOs:1-23.

2. A variant polymerase of claim 1 which has at least 75% but less than 100% identity to any of SEQ ID NOs:1-23.

3. A variant polymerase of claim 1 or claim 2 comprising at least three or four or five or six or seven or eight or nine or ten or eleven or twelve of said motifs.

4. A variant polymerase of any one of the preceding claims further comprising one or more mutations selected from the group of mutations consisting of (a)-(f) where the mutations in (a)-(f) are:

(a) A1E, G3(K, E or D), K5(L, A or V), E8(M or A), E9D, M10I, A13(D, E, T or V), I14D, V15A, V17(T, E or G), I18V;

(b) E20(M or D) M34(Q or L), E36D, I46F, L48(N or I), M57(L or I), P59(T or A), T61L, D65S, S66(F, E or P);

(c) Q67A, L69(V or K), A73E, M81V, A84R, V88(A or I), R99V, A102D, N113A, D117(T, S or A), A118D, G119(D or E), I121(A or V), V124K, E131H, S135(E or P), V144A, S147(P or A), L148(D or V), Q152(L or P);

(d) T153(A or V), Q170(R or E), M173(L or I), D175E, N178(E, K or R), Q183(L, E or R), L185F, T186(L or I), K187(E or D), Q190(L or M);

(e) A193(I or S), A194(L, S or T), N205(D or K), S216(L or E), R223(V, K or G), A224E, I225(Q or V), V247L, R307H, M316R; and

(f) A330T, D357L, D378N, D380E, I383A, Q387R, L390M; I400V, E406D, A410S, N411R, A433S, N437G, T439K, A452E, Q459(R or E), N463(V, E or D), L484D, D486E, V494L, T501M, Q530R, I552M, E557(K, Q or R) and T568(E or R).

5. A variant polymerase according to claim 4, comprising at least one mutated amino acid selected from each of groups (a)-(f).

6. A variant polymerase of claim 4 which further comprises two, three, four, five, six, seven, eight, nine, ten, eleven, twelve, thirteen, fourteen, fifteen, sixteen, seventeen, eighteen, nineteen, or twenty mutant amino acids at the same position as the corresponding amino acids in (a)-(f) of SEQ ID NO:1.

7. A variant polymerase of any one of the preceding claims wherein said sequence motif(s) confer one or more improved properties selected from at least one of specific activity; reaction speed; thermostability; storage stability; dUTP tolerance and salt tolerance; increased performance in isothermal amplification; non-interference of pH during sequencing; improved strand displacement; altered processivity; altered ribonucleotide incorporation; altered modified nucleotide incorporation; and altered fidelity when compared to the corresponding parent polymerase.

8. A variant polymerase of any one of the preceding claims wherein a peptide is fused to one end of the variant polymerase directly or by means of a linker sequence.

5 9. An enzyme preparation comprising a variant polymerase according to any one of the preceding claims and a buffer.

10. An enzyme preparation according to claim 9 comprises a temperature dependent inhibitor of polymerase activity.

10

11. An enzyme preparation according to claim 9 or claim 10, further comprising dNTPs.

12. A DNA encoding a variant polymerase as described in any of
15 claims 1-8.

13. A host cell comprising the DNA according to claim 12.

14. A process for preparing a variant of a parent Family A DNA
20 polymerase having improved polymerase activity compared with the parent polymerase; comprising synthesizing a polypeptide as defined in any one of claims 1-8; and characterizing the polymerase activity.

25 15. The process of claim 14, wherein characterizing the polymerase activity, further comprises: determining in comparison with the parent polymerase, at last one of: thermostability; stability in storage; tolerance to salt; performance in isothermal amplification; strand displacement; kinetics; processivity; fidelity;
30 altered ribonucleotide incorporation; altered deoxyuracil incorporation; and altered modified nucleotide incorporation.

16. A variant Family A DNA polymerase obtainable by the process of claim 14 or claim 15.

17. A variant polymerase of any of claims 1-7 wherein the one or more motifs or one or more mutations selected from the group of mutations consisting of (a)-(f) have improved reverse transcriptase activity.

18. A method for reverse transcribing an RNA of interest, comprising combining an RNA with a DNA polymerase variant or preparation thereof according to any of claims 1-11 to form a cDNA.

19. A method according to claim 18 further comprising amplifying the cDNA by means of the DNA polymerase variant or preparation thereof according to claims 1-11, to produce amplified DNA.

20. A method for amplifying DNA, comprising combining a target DNA with a DNA polymerase variant or preparation thereof according to claims 1-11, to produce amplified DNA.

21. A variant protein, comprising: an amino acid sequence with at least 75% or 80% or 85% or 90% or 95% but less than 100% sequence identity to any of SEQ ID NOs:1-23, wherein the variant protein further comprises at least one mutated amino acid having a position corresponding to SEQ ID NO:1 selected from the group of mutated amino acids consisting of (a)-(f) where the mutations in (a)-(f) are:

(a) A1E, G3(K, E or D), K5(L, A or V), E8(M or A), E9D, M10I, A13(D, E, T or V), I14D, V15A, V17(T, E or G), I18V;

(b) E20(M or D) M34(Q or L), E36D, I46F, L48(N or I), M57(L or I), P59(T or A), T61L, D65S, S66(F, E or P);

(c) Q67A, L69(V or K), A73E, M81V, A84R, V88(A or I), R99V, A102D, N113A, D117(T, S or A), A118D, G119(D or E), I121(A or V), V124K, E131H, S135(E or P), V144A, S147(P or A), L148(D or V), Q152(L or P);

5 (d) T153(A or V), Q170(R or E), M173(L or I), D175E, N178(E, K or R), Q183(L, E or R), L185F, T186(L or I), K187(E or D), Q190(L or M);

(e) A193(I or S), A194(L, S or T), N205(D or K), S216(L or E), R223(V, K or G), A224E, I225(Q or V), V247L, R307H,

10 M316R; and

(f) A330T, D357L, D378N, D380E, I383A, Q387R, L390M; I400V, E406D, A410S, N411R, A433S, N437G, T439K, A452E, Q459(R or E), N463(V, E or D), L484D, D486E, V494L, T501M, Q530R, I552M, E557(K, Q or R) and T568(E

15 or R).

22. The variant protein according to claim 21, further comprising at least one amino acid mutation selected from each of groups (a)-(f).

20

23. The variant protein according to claim 21 or claim 22, further comprising at least one amino acid motif or at least two amino acid motifs selected from the group consisting of: 3..EEK..5, 15..ADE..17, 65..SPQ..67, 86..RAI..88, 185..LTE..187,

25 186..TEL..188, 222..LKE..224, 306..VHP..308, 314...HTR..316, 555..LCK..557, 556..CKL..558 and 567..VEL..569.

24. A variant protein according to claim 21 -23, further comprising two, three, four, five, six, seven, eight, nine, ten,

30 eleven, twelve, thirteen, fourteen, fifteen, sixteen, seventeen, eighteen, nineteen, or twenty amino acids at the same position as the corresponding mutant amino acids in (a)-(f) of SEQ ID NO:1.

25. The variant protein of any of claims 21-24, wherein the amino acid sequence is at least 80% identical to any one of SEQ ID NOs:1-23.

5

26. A variant protein according to claim 25, wherein the amino acid sequence is at least 90% identical to any one of SEQ ID NOs:1-23.

10 27. A variant protein according to claim 26, wherein the amino acid sequence is at least 95% identical to any one of SEQ ID NOs:1-23.

28. A non-naturally occurring synthetic protein, comprising: a
15 fragment 1, a fragment 2, a fragment 3, a fragment 4, a fragment 5, a fragment 6, a fragment 7 and a fragment 8 wherein the fragments are covalently linked, and wherein:

the fragment 1 is selected from Segment 1 having an amino acid sequence selected from the group consisting of SEQ ID
20 NOs:24-39;

the fragment 2 is selected from Segment 2 having an amino acid sequence selected from the group consisting of SEQ ID NOs:40-56;

the fragment 3 is selected from Segment 3 having an amino
25 acid sequence selected from the group consisting of SEQ ID NOs:57-72;

the fragment 4 is selected from Segment 4 having an amino acid sequence selected from the group consisting of SEQ ID NOs:73-87;

30 the fragment 5 selected from Segment 5 having an amino acid sequence selected from the group consisting of SEQ ID NOs: 88-99;

the fragment 6 selected from Segment 6 having an amino acid sequence selected from the group consisting of SEQ ID NOs:100-111;

5 the fragment 7 selected from Segment 7 having an amino acid sequence selected from the group consisting of SEQ ID NOs:112-125;

the fragment 8 selected from Segment 8 having an amino acid sequence selected from the group consisting of SEQ ID NOs:126-138; and

10 wherein the covalently linked fragments having an amino acid sequence that does not have 100% identity to SEQ ID NOs:1-23.

29. A synthetic protein according to claim 28, wherein the amino acid sequence of the synthetic protein comprises at least one

15 amino acid sequence motif selected from the group consisting of: 3..EEK..5, 15..ADE..17, 65..SPQ..67, 86..RAI..88, 185..LTE..187, 186..TEL..188, 222..LKE..224, 306..VHP..308, 314…HTR..316, 555..LCK..557, 556..CKL..558 and 567..VEL..569.

20 30. The synthetic protein of claim 28, wherein the amino acid sequence comprises at least two or three or four or five or six or seven or eight or nine or ten or eleven or twelve of the amino acid sequence motifs selected from the group consisting of: 3..EEK..5, 15..ADE..17, 65..SPQ..67, 86..RAI..88, 185..LTE..187,

25 186..TEL..188, 222..LKE..224, 306..VHP..308, 314…HTR..316, 555..LCK..557, 556..CKL..558 and 567..VEL..569.

31. A protein comprising at least 80% sequence identity with SEQ ID NOs:1 and further comprising one or more mutations (such as,

30 for example, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34,

35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, or 79 mutations) selected from the group consisting of: A1E, G3(K, E or D), K5(L, A or V), E8(M or A), E9D, M10I, A13(D, E, T or V), I14D, V15A, V17(T, E or G), I18V, E20(M or D) M34(Q or L), E36D, I46F, L48(N or I), M57(L or I), P59(T or A), T61L, D65S, S66(F, E or P), Q67A, L69(V or K), A73E, M81V, A84R, V88(A or I), R99V, A102D, N113A, D117(T, S or A), A118D, G119(D or E), I121(A or V), V124K, E131H, S135(E or P), V144A, S147(P or A), L148(D or V), Q152(L or P), T153(A or V), Q170(R or E), M173(L or I), D175E, N178(E, K or R), Q183(L, E or R), L185F, T186(L or I), K187(E or D), Q190(L or M), A193(I or S), A194(L, S or T), N205(D or K), S216(L or E), R223(V, K or G), A224E, I225(Q or V), V247L, R307H, M316R, A330T, D357L, D378N, D380E, I383A, Q387R, L390M, I400V, E406D, A410S, N411R, A433S, N437G, T439K, A452E, Q459(R or E), N463(V, E or D), L484D, D486E, V494L, T501M, Q530R, I552M, E557(K, Q or R) and T568(E or R); and optionally a sequence motif at a specified position in SEQ ID NO:1 selected from the group consisting of: 3..EEK..5, 15..ADE..17, 65..SPQ..67, 86..RAI..88, 185..LTE..187, 186..TEL..188, 222..LKE..224, 306..VHP..308, 314…HTR..316, 555..LCK..557, 556..CKL..558 and 567..VEL..569.

32. A variant protein or a synthetic protein according to any of claims 21-31, wherein a peptide is fused to one end of the variant protein either directly or by means of a linker.

33. An enzyme preparation comprising a variant protein or a synthetic protein according to any of claims 21-31 and a buffer.

34. An enzyme preparation according to claim 33, further comprising a plurality of proteins.

35. An enzyme preparation according to claim 33 or claim 34, further comprising a reversible inhibitor of polymerase activity.

36. An enzyme preparation according to claim 33 or claim 34, further comprising dNTPs.

37. A DNA encoding a variant protein or synthetic protein described in any of claims 21-36.

38. A host cell comprising the DNA according to claim 37.

39. A method for obtaining a variant of a parent protein having improved polymerase activity compared with the parent protein; comprising synthesizing a protein from any of claims 21-36; and characterizing the polymerase activity.

40. A method according to claim 39, wherein characterizing the polymerase activity, further comprises: determining in comparison with the parent protein, at last one of: thermostability; stability in storage; tolerance to salt; performance in isothermal amplification; strand displacement; kinetics; processivity; fidelity; altered ribonucleotide incorporation; altered dUTP incorporation; and altered modified nucleotide incorporation.

41. A method, comprising:

(a) synthesizing a protein wherein the protein has an amino acid sequence which is capable of being generated from single selected protein fragments obtainable from 8 different segments described in Figure 2; and

(b) assaying the synthetic protein for polymerase activity.

42. A method according to claim 41, wherein the protein is synthesized by cloning a DNA sequence encoding the protein.

43. A method, comprising:
    (a) selecting a protein variant or synthetic protein according to any of claims 21-36 having an amino acid sequence; and
    (b) expressing the protein variant or synthetic protein as a fusion protein with an additional peptide at an end of the amino acid sequence.

44. A method of isothermal amplification, comprising:
    (a) providing a preparation comprising a variant protein or synthetic protein according to any of claims 21-32;
    (b) combining a target DNA with the preparation; and
    (c) amplifying the target DNA at a temperature less than 90°C.

45. A method according to claim 44, wherein the amplification reaction results in a quantitative measure of the amount of target DNA in the preparation.

46. A DNA polymerase having one or more improved properties for isothermal amplification compared with SEQ ID NO:1, selected from the group consisting of:
    (a) an increased reaction speed in the range where the increase is at least 10% and as much as 20%; 500% or 1000%;
    (b) an increased temperature stability in the range of 50°C to 100°C, 50°C to 90°C or 60°C to 90°C;
    (c) an increased salt tolerance in the range of 10 mM – 1 M, or 20 mM – 200 mM or 500 mM monovalent salt;
    (d) an increased storage stability at 25°C, retaining at least 50% activity over 45 weeks, over 1 year or over 2 years;

(e) an enhanced dUTP tolerance of the range of an increase of 50% to 100% dUTP; and

(f) an increased reverse transcriptase activity by at least 2 fold.

5 47. A DNA polymerase according to claim 46, having at least two or three or four or five or six of the improved properties.

48. A DNA polymerase according to claims 46 or 47 having at least 80% amino acid sequence identity but less than 100% amino

10 acid sequence identity with any of SEQ ID NOs:1-23 and containing at least 12 artificially introduced single amino acid mutations that occur in a three amino acid motif that differs from an amino acid in the corresponding site of a naturally occurring Bst polymerase.

15 49. A DNA polymerase according to claim 48, wherein at least one of the three amino acid motifs is selected from the group consisting of: 3..EEK..5, 15..ADE..17, 65..SPQ..67, 86..RAI..88, 185..LTE..187, 186..TEL..188, 222..LKE..224, 306..VHP..308, 314…HTR..316, 555..LCK..557, 556..CKL..558 and 567..VEL..569.

20

50. A DNA polymerase according to any of claims 1-45 having one or more improved properties for isothermal amplification compared with SEQ ID NO:1, selected from the group consisting of:

(a) an increased reaction speed in the range where the increase

25 is at least 10% and as much as 20%; 500% or 1000%;

(b) an increased temperature stability in the range of 50°C to 100°C, 50°C to 90°C or 60°C to 90°C;

(c) an increased salt tolerance in the range of 10 mM – 1 M, or 20 mM – 200 mM or 500 mM monovalent salt;

30 (d) an increased storage stability at 25°C, retaining at least 50% activity over 45 weeks, over 1 year or over 2 years;

(e) an enhanced dUTP tolerance of the range of an increase of 50% to 100% dUTP; and

(f) an increased reverse transcriptase activity by at least 2 fold.

5

FIG. 1A-1

```
                ---------+---------+---------+---------+---------+---------+
                         10        20        30        40        50        60
                ---------+---------+---------+---------+---------+---------+
SEQ ID NO. 1    AEGEKPLEEMEFAIVDVITEEMLADKAALVVEVMEENYHDAPIVGIALVNEHGRFFMRPE      60
SEQ ID NO. 2    AEGRKPLEEMEFAIVDVITEEMLADKAALVVEVMEENYHDAPIVGIALVNEHGRFFMRPE      60
SEQ ID NO. 3    AEDETPLMEMEFVIADGITDEMLADKAALVVEVMEENYHDAPIVGIALVNEHGRFFLRAE      60
SEQ ID NO. 4    AEDETPLMEMEFVIADGITDEMLADKAALVVEVMEENYHDAPIVGIALVNEHGRFFLRAE      60
SEQ ID NO. 5    AEDETPLMEMEFVAADGITDEMLADKAALVVEVMEENYHDAPIVGIALVNEHGRFFLRAE      60
SEQ ID NO. 6    AEDETPLMEMEFVAADGITDEMLADKAALVVEVMEENYHDAPIVGIALVNEHGRFFLRAE      60
SEQ ID NO. 7    AEDETPLMEMEFVAADGITDEMLADKAALVVEVMEENYHDAPIVGIALVNEHGRFFLRAE      60
SEQ ID NO. 8    AEDETPLMEMEFVAADGITDEMLADKAALVVEVMEENYHDAPIVGIALVNEHGRFSLRAE      60
SEQ ID NO. 9    AEDETPLMEMEFVAADGITDEMLADKAALVVEVMEENYHDAPIVGIALVNEHGRFSLRAE      60
SEQ ID NO. 10   DEGEKPLAGMDFAIADSVTDEMLADKAALVVEVVGDNYHHAPIVGIALANERGRFFLRPE      60
SEQ ID NO. 11   DEGEKPLAGMDFAIADSVTDEMLADKAALVVEVVGDNYHHAPIVGIALANERGRFFLRPE      60
SEQ ID NO. 12   SEEEKPLAKMAFTLADRVTEEMLADKAALVVEVVEENYHDAPIVGIAVVNEHGRFFLRPE      60
SEQ ID NO. 13   SEEEKPLAKMAFTLADRVTEEMLADKAALVVEVVEENYHDAPIVGIAVVNEHGRFFLRPE      60
SEQ ID NO. 14   SEEEKPLAKMAFTLADRVTEEMLADKAALVVEVVEENYHDAPIVGIAVVNEHGRFFLRPE      60
SEQ ID NO. 15   SEEEKPLAKMAFTLADRVTEEMLADKAALVVEVVEENYHDAPIVGIAVVNEHGRFFLRPE      60
SEQ ID NO. 16   SEEEKPLAKMAFTLADRVTEEMLADKAALVVEVVEENYHDAPIVGIAVVNEHGRFFLRPE      60
SEQ ID NO. 17   SEEEKPLAKMAFTLADRVTEEMLADKAALVVEVVEENYHDAPIVGIAVVNEHGRFFLRPE      60
SEQ ID NO. 18   DEGEKPLAGMDFAIADSVTDEMLADKAALVVEVVGDNYHHAPIVGIALANERGRFFLRPE      60
SEQ ID NO. 19   EETEVALSDIDYTIVDEVTESILSDEAALVVEVLESNYHKAPILGFAIANEHGNFFIRTD      60
SEQ ID NO. 20   KKEEAVLPNIDYTIVEEVSESVLADRSALVVEVLESNYHKAPILGFALANEHGHFFIRTD      60
SEQ ID NO. 21   EKTAVALPKIRYEIVDEVTEAILSDEAALVVEVLESNYHKAPILGFAIANEHGNFFIRTD      60
SEQ ID NO. 22   EKTAVALPEIGYKIVDEVTEAILSDEAALVVEVLESNYHKAPILGFAIANEHGNFFIRTD      60
SEQ ID NO. 23   SQEQLSLTDISFVTVQTIDEHMLTKEGALVVEVLDENYHQAPIVGFALVNERGHFFIPTD      60

                ---------+---------+---------+---------  +---------+--------
                         70        80        90          100       110
                ---------+---------+---------+---------  +---------+--------
SEQ ID NO. 1    TALADSQFLAWLADETKKKSMFDAKRAVVALKWKGIELR--GVAFDLLLAAYLLNPAQDA     118
SEQ ID NO. 2    TALADSQFLAWLADETKKKSMFDAKRAVVALKWKGIDVR--GVAFDLLLAAYLLNPAQDA     118
SEQ ID NO. 3    MALADSQFLAWLADETKKKSMFDAKRAAVALKWKGIELR--GVAFDLLLAAYLLNPAQDA     118
SEQ ID NO. 4    MALADSQFLAWLADETKKKSMFDAKRAAVALKWKGIELR--GVAFDLLLAAYLLNPAQDA     118
SEQ ID NO. 5    MALADPQFVAWLADETKKKSMFDAKRASVALKWKGIELR--GVAFDLLLAAYLLNPAQDA     118
SEQ ID NO. 6    MALADPQFVAWLADETKKKSMFDAKRASVALKWKGIELR--GVAFDLLLAAYLLNPAQDA     118
SEQ ID NO. 7    MALADPQFVAWLADETKKKSMFDAKRASVALKWKGIELR--GVAFDLLLAAYLLNPAQDA     118
SEQ ID NO. 8    MALADPQFVAWLADETKKKSMFDAKRASVALKWKGIELR--GVAFDLLLAAYLLNPAQDA     118
SEQ ID NO. 9    MALADPQFVAWLADETKKKSMFDAKRASVALKWKGIELR--GVAFDLLLAAYLLNPAQDA     118
SEQ ID NO. 10   TALADPKFLAWLGDETKKKTMFDSKRAAVALKWKGIELR--GVVFDLLLAAYLLDPAQAA     118
SEQ ID NO. 11   TALADPKFLAWLGDETKKKTMFDSKRAAVALKWKGIELR--GVVFDLLLAAYLLDPAQAA     118
SEQ ID NO. 12   TALADPQFVAWLGDETKKKSMFDSKRAAVALKWKGIELC--GVSFDLLLAAYLLDPAQGV     118
SEQ ID NO. 13   TALADPQFVAWLGDETKKKSMFDSKRAAVALKWKGIELC--GVSFDLLLAAYLLDPAQGV     118
SEQ ID NO. 14   TALADPQFVAWLGDETKKKSMFDSKRAAVALKWKGIELC--GVSFDLLLAAYLLDPAQGV     118
SEQ ID NO. 15   TALADPQFVAWLGDETKKKSMFDSKRAAVALKWKGIELC--GVSFDLLLAAYLLDPAQGV     118
SEQ ID NO. 16   TALADPQFVAWLGDETKKKSMFDSKRAAVALKWKGIELC--GVSFDLLLAAYLLDPAQGV     118
SEQ ID NO. 17   TALADPQFVAWLGDETKKKSMFDSKRAAVALKWKGIELC--GVSFDLLLAAYLLDPAQGV     118
SEQ ID NO. 18   TAVADPKFLAWLGDETKKKTMFDSKRAAVALNGKGIELAGVGVVFDLLLAAYLLDPAQAA     120
SEQ ID NO. 19   TALSSSLFTQWLEDESKKKSVFDGKRAIVSLKWQGIQLR--GVQFDLLIASYLLNPSQST     118
SEQ ID NO. 20   TALSSPLFTAWLEDETKKKSVFDGKRAIVALKWKGIELR--GIEFDLLIASYLLNPSQST     118
SEQ ID NO. 21   TALSSSLFTTWLEDESKKKSVFDGKRAIVSLKWQGVHLR--GIQFDLLIASYLLNPSQST     118
SEQ ID NO. 22   TALSSSLFTTWLEDESKKKSVFDGKRAIVSLKWQGVHLR--GIQFDLLIASYLLNPSQST     118
SEQ ID NO. 23   IALASSRFKRWLEDEQCKKSVFDAKRAIVALKWNGIELK--GVDFDLLLAAYLLNPTDAN     118
```

2/18

FIG. 1A-2

```
                 -+---------+---------+---------+---------+---------+--------
                  120       130       140       150       160       170
                 -+---------+---------+---------+---------+---------+--------
SEQ ID NO. 1     GDIAAVAKMKQYEAVRSDEAVYGKGVKRSLPDEQTLAEHLVRKAAAIWALEQPFMDDLRN      178
SEQ ID NO. 2     GDIAAVAKMKQYEAVRSDEAVYGKGVKRSLPDEQTLAEHLVRKAAAIWALEQPFMDDLRN      178
SEQ ID NO. 3     GDVAAVAKMKQYEAVRPDEAVYGKGAKRSLPDEPTLAEHLVRKAAAIWALERPFLDELRS      178
SEQ ID NO. 4     GDVAAVAKMKQYEAVRPDEAVYGKGAKRSLPDEPTLAEHLVRKAAAIWALERPSLDELRS      178
SEQ ID NO. 5     GDVAAVAKMKQYEAVRPDEAVYGKGAKRSLPDEPTLAEHLVRKAAAIWALERPFLDELRS      178
SEQ ID NO. 6     GDVAAVAKMKQYEAVRPDEAVYGKGAKRSLPDEPTLAEHLVRKAAAIWALERPFLDELRS      178
SEQ ID NO. 7     GDVAAVAKMKQYEAVRPDEAVYGKGAKRSLPDEPTLAEHLVRKAAAIWALERPFLDELRS      178
SEQ ID NO. 8     GDVAAVAKMKQYEAVRPDEAVYGKGAKRSLPDEPTLAEHLVRKAAAIWALERPFLDELRS      178
SEQ ID NO. 9     GDVAAVAKMKQYEAVRPDEAVYGKGAKRSLPDEPTLAEHLVRKAAAIWALERPFLDELRS      178
SEQ ID NO. 10    GDVAAVAKMHQYEAVRSDEAVYGKGAKRTVPDEPTLAEHLVRKAAAIWALEEPLMDELRR      178
SEQ ID NO. 11    GDVAAVAKMHQYEAVRSDEAVYGKGAKRTVPDEPTLAEHLARKAAAIWALEEPLMDELRR      178
SEQ ID NO. 12    DDVAAAAKMKQYEAVRSDEAVYGKGAKRAVPDEPVLAEHLVRKAAAIWALERPFLDELRR      178
SEQ ID NO. 13    DDVAAAAKMKQYEAVRSDEAVYGKGAKRAVPDEPVLAEHLVRKAAAIWALERPFLDELRR      178
SEQ ID NO. 14    DDVAAAAKMKQYEAVRPDEAVYGKGAKRAVPDEPVLAEHLVRKAAAIWALERPFLDELRR      178
SEQ ID NO. 15    DDVAAAAKMKQYEAVRPDEAVYGKGAKRAVPDEPVLAEHLVRKAAAIWALERPFLDELRR      178
SEQ ID NO. 16    DDVAAAAKMKQYEAVRPDEAVYGKGAKRAVPDEPVLAEHLVRKAAAIWELERPFLDELRR      178
SEQ ID NO. 17    DDVAAAAKMKQYEAVRPDEAVYGKGAKRAVPDEPVLAEHLVRKAAAIWALERPFLDELRR      178
SEQ ID NO. 18    GDVAAVAKMHQYEAVRSDEAVYGKGAKRTVPDEPTLAEQLVRKAAAIWALEEPLMDELRR      180
SEQ ID NO. 19    EDVASIAKTKQYTDVQSDEAIYGKGAKQKIPDEPVLAEHLVRKAAAIRALEQDFICDLQE      178
SEQ ID NO. 20    EDVASVAKTKQYMAVQSDEAVYGKGAKQTVPEEKALAEHLVRKAAAIRALEKEFIHDLQE      178
SEQ ID NO. 21    EDVASIAKTKQYAGVQSDEAVYGKGAKQKVPDEQVLAEHLVRKAAAIRALEQGFIHDLQE      178
SEQ ID NO. 22    EDVASIAKTKQYVGVQSDEAVYGKGAKQKVPDEQVLAEHLVRKAAAIRALEQGFIHDLQE      178
SEQ ID NO. 23    GDVAAVAKTKQYTDVQSDEEVYGKGAKQAIPPTNVLAEHLVRKAKAIASLKETYIQELKR      178

                 -+---------+---------+---------+---------+---------+--------
                  180       190       200       210       220       230
                 -+---------+---------+---------+---------+---------+--------
SEQ ID NO. 1     NEQDQLLTKLEQPLAAILAEMEFTGVNVDTKRLEQMGSELAEQLRAIEQRIYELAGQEFN      238
SEQ ID NO. 2     NEQDQLLTKLEQPLAAILAEMEFTGVNVDTKRLEQMGSELAEQLRAIEQRIYEHAGQEFN      238
SEQ ID NO. 3     NEQDGLLIKLEQPLATILAEMEFTGIKVDTKRLEQMGSELAEQLRAVEQRIYELAGQEFN      238
SEQ ID NO. 4     NEQDGLLIKLEQPLATILAEMEFTGIKVDTKRLEQMGSELAEQLRAVEQRIYELAGQEFN      238
SEQ ID NO. 5     NEQDELLIKLEQPLATILAEMEFTGVKVDTKRLEQMGSELAEQLGAIEQRIYELAGQEFN      238
SEQ ID NO. 6     NEQDELLIKLEQPLATILAEMEFTGVKVDTKRLEQMGSELAEQLGAVEQRIYELAGQEFN      238
SEQ ID NO. 7     NEQDELLIKLEQPLATILAEMEFTGVKVDTKRLEQMGSELAEQLGAVEQRIYELAGQEFN      238
SEQ ID NO. 8     NEQDELLIKLEQPLATILAEMEFTGVKVDTKRLEQMGSELAEQLGAIEQRIYELAGQEFN      238
SEQ ID NO. 9     NEQDELLIKLEQPLATILAEMEFTGVKVDTKRLEQMGSELAEQLGAVEQRIYELAGQEFN      238
SEQ ID NO. 10    NEQDRLLTELEQPLAGILANMEFTGVKVDTKRLEQMGAELTEQLQAVERRIYELAGQEFN      238
SEQ ID NO. 11    NEQDRLLTELEQPLAGILANMEFTGVKVDTKRLEQMGAELTEQLQAVERRIYELAGQEFN      238
SEQ ID NO. 12    NEQDRLLVELEQPLSSILAEMEFAGVKVDTKRLEQMGEELAEQLRTVEQRIYELAGQEFN      238
SEQ ID NO. 13    NEQDRLLVELEQPLSSILAEMEFAGVKVDTKRLEQMGEELAEQLRTVEQRIYELAGQEFN      238
SEQ ID NO. 14    NEQDRLLVELEQPLSSILAEMEFAGVKVDTKRLEQMGEELAEQLRTVEQRIYELAGQEFN      238
SEQ ID NO. 15    NEQDRLLVELEQPLSSILAEMEFAGVKVDTKRLEQMGEELAEQLGTVEQRIYELAGQEFN      238
SEQ ID NO. 16    NEQDRLLVELEQPLSSILAEMEFAGVKVDTKRLEQMGKELAEQLGTVEQRIYELAGQEFN      238
SEQ ID NO. 17    NEQDRLLVELEQPLSSILAEMEFAGVKVDTKRLEQMGEELAEQLRTVEQRIYELAGQEFN      238
SEQ ID NO. 18    NEQDRLLTELEHALAGILANMEFTGVKVDTKRLEQMGAELTEQLQAVERRIYELAGQEFN      240
SEQ ID NO. 19    NEQYSLFTDLELPLSSILAEMEFTGVKIDVKRLKEMGEELTEQLKEIEQEIYELAGQEFN      238
SEQ ID NO. 20    NEQYALFTDLELPLSSILAEMEFTGVKVDVERLKEMGEELTEQLRAVEQEIYELAGQEFN      238
SEQ ID NO. 21    NEQYSLFTDLELPLSTILAEMEFAGVKVDVERLKEMGEELAEQLKEVEQEIYRLAGQEFN      238
SEQ ID NO. 22    NEQYSLFTDLELPLSTILAEMEFAGVKVDVERLKEMGEELAEQLKEVEQEIYRLAGQEFN      238
SEQ ID NO. 23    NEQFELLVHLELPLTFILAQMEFYGVKVDVDRLEQMGKEFTAQLEQIEQRIYELAGTTFN      238
```

3/18

FIG. 1A-3

```
               -+---------+---------+---------+---------+------ ---+-------
                240       250       260       270       280        290
               -+---------+---------+---------+---------+------ ---+-------
SEQ ID NO. 1   INSPKQLGVILFEKLQLPVLKKTKTGYSTSADVLEKLAPHHEIVENIL-HYRQLGKLQST   297
SEQ ID NO. 2   INSPKQLGVILFEKLQLPVLKKTKTGYSTSADVLEKLAPHHEIVENIL-HYRQLGKLQST   297
SEQ ID NO. 3   INSPKQLGIILFEKLQLPVLKKTKTGYSTSADVLEKLAPHHEIVENIL-HYRQLGKLQST   297
SEQ ID NO. 4   INSPKQLGIILFEKLQLPVLKKTKTGYSTSADVLEKLAPHHEIVENIL-HYRQLGKLQST   297
SEQ ID NO. 5   INSPKQLGIILFEKLQLPVLKKTKTGYSTSADVLEKLAPHHEIVENIL-HYRQLGKLQST   297
SEQ ID NO. 6   INSPKQLGIILFEKLQLPVLKKTKTGYSTSADVLEKLAPHHEIVENIL-HYRQLGKLQST   297
SEQ ID NO. 7   INSPKQLGIILFEKLQLPVLKKTKTGYSTSADVLEKLAPHHEIVENIL-HYRQLGKLQST   297
SEQ ID NO. 8   INSPKQLGIILFEKLQLPVLKKTKTGYSTSADVLEKLAPHHEIVENIL-HYRQLGKLQST   297
SEQ ID NO. 9   INSPKQLGIILFEKLQLPVLKKTKTGYSTSADVLEKLAPHHEIVENIL-HYRQLGKLQST   297
SEQ ID NO. 10  INSPKQLGTVLFDKLQLPVLKKTKTGYSTSADVLEKLAPHHEIVEHIL-HYRQLGKLQST   297
SEQ ID NO. 11  INSPKQLGTVLFDKLQLPVLKKTKTGYSTSADVLEKLAPHHEIVEHIL-HYRQLGKLQST   297
SEQ ID NO. 12  INSPKQLGVILFEKLQLPILKKTKTGYSTSADVLEKLAPYHEIVENIL-HYRQLGKLQST   297
SEQ ID NO. 13  INSPKQLGVILFEKLQLPVLKKTKTGYSTSADVLEKLAPYHEIVENIL-HYRQLGKLQST   297
SEQ ID NO. 14  INSPKQLGVILFEKLQLPVLKKTKTGYSTSADVLEKLAPYHEIVENIL-HYRQLGKLQST   297
SEQ ID NO. 15  INSPKQLGVILFEKLQLPVLKKTKTGYSTSADVLEKLAPYHEIVENIL-HYRQLGKLQST   297
SEQ ID NO. 16  INSPKQLGVILFEKLQLPVLKKTKTGYSTSADVLEKLAPYHEIVENIL-HYRQLGKLQST   297
SEQ ID NO. 17  INSPKQLGVILFEKLQLPVLKKSKTGYSTSADVLEKLAPYHEIVENILQHYRQLGKLQST   298
SEQ ID NO. 18  INSPKQLGTVLFDKLQLPVLKKTKTGYSTSADVLEKLAPHHEIVEHIL-HYRQLGKLQST   299
SEQ ID NO. 19  INSPKQLGVILFEKLQLPVLKKTKTGYSTSAEVLEKLAPQHEIVEKIL-HYRQLGKLQST   297
SEQ ID NO. 20  INSPKQLGVILFEKLQLPVLKKTKTGYSTSADVLEKLAPQHEIVEKIL-HYRQLGKLQST   297
SEQ ID NO. 21  INSPKQLGVILFEKLQLPVLKKTKTGYSTSAEVLEKLAPQHEIVEKIL-HYRQLGKLQST   297
SEQ ID NO. 22  INSPKQLGVILFEKLQLPVLKKTKTGYSTSAEVLEKLAPQHEIVEKIL-HYRQLGKLQST   297
SEQ ID NO. 23  INSPKQLGTILFEKLQLPIVKKTKTGYSTSADVLEKLAPYHEIIEQIL-HYRQLGKLQST   297

               --+---------+---------+---------+---------+---------+-------
                 300       310       320       330       340       350
               --+---------+---------+---------+---------+---------+-------
SEQ ID NO. 1   YIEGLLKVVRPDTGKVHTMFNQALTQTGRLSSAEPNLQNIPIRLEEGRKIRQAFVPSEPD   357
SEQ ID NO. 2   YIEGLLKVVRPDTGKVHTMFNQTLTQTGRLSSAEPNLQNIPIRLEEGRKIRQAFVPSEPD   357
SEQ ID NO. 3   YIEGLLKVVHPDTGKVHTMFNQALTQTGRLSSAEPNLQNIPIRLEEGRKIRQAFVPSEPD   357
SEQ ID NO. 4   YIEGLLKVVHPDTGKVHTMFNQALTQTGRLSSAEPNLQNIPIRLEEGRKIRQAFVPSEPD   357
SEQ ID NO. 5   YIEGLLKVVHHDTGKVHTMFNQALTQTGRLSSAEPNLQNIPIRLEEGRKIRQAFVPSEPG   357
SEQ ID NO. 6   YIEGLLKVVHHDTGKVHTMFNQALTQTGRLSSAEPNLQNIPIRLEEGRKIRQAFVPSEPG   357
SEQ ID NO. 7   YIEGLLKVVHHDTGKVHTMFNQALTQTGRLSSAEPNLQNIPIRLEEGRKIRQAFVPSEPG   357
SEQ ID NO. 8   YIEGLLKVVHHDTGKVHTMFNQALTQTGRLSSAEPNLQNIPIRLEEGRKIRQAFVPSEPG   357
SEQ ID NO. 9   YIEGLLKVVHHDTGKVHTMFNQALTQTGRLSSAEPNLQNIPIRLEEGRKIRQAFVPSEPG   357
SEQ ID NO. 10  YIEGLLKVVHPVTGKVHTMFNQALTQTGRLSSVEPNLQNIPIRLEEGRKIRQAFVPSEPD   357
SEQ ID NO. 11  YIEGLLKVVHPVTGKVHTMFNQALTQTGRLSSVEPNLQNIPIRLEEGRKIRQAFVPSEPD   357
SEQ ID NO. 12  YIEGLLKVVRPDTKKVHTIFNQALTQTGRLSSTEPNLQNIPIRLEEGRKIRQAFVPSESD   357
SEQ ID NO. 13  YIEGLLKVVRPDTKKVHTIFNQALTQTGRLSSTEPNLQNIPIRLEEGRKIRQAFVPSESD   357
SEQ ID NO. 14  YIEGLLKVVRPDTKKVHTIFNQALTQTGRLSSTEPNLQNIPIRLEEGRKIRQAFVPSESD   357
SEQ ID NO. 15  YIEGLLKVVRPDTKKVHTIFNQALTQTGRLSSTEPNLQNIPIRLEEGRKIRQAFVPSESD   357
SEQ ID NO. 16  YIEGLLKVVRPDTKKVHTIFNQALTQTGRLSSTEPNLQNIPIRLEEGRKIRQAFVPSESD   357
SEQ ID NO. 17  YIEGLLKVVRPDTKKVHTIFNQALTQTGRLSSTEPNLQNIPIRLEEGRKIRQAFVPSESD   358
SEQ ID NO. 18  YIEGLLKVVHPVTGKVHTMFNQALTQTGRLSSVEPNLQNIPIRLEEGRKIRQAFVPSEPD   359
SEQ ID NO. 19  YIEGLLKVVHHDTNKVHTIFNQALTQTGRLSSTEPNLQNIPIRLEEGRKIRQAFVPSEPD   357
SEQ ID NO. 20  YIEGLLKVVHCDTHKVHTIFNQALTQTGRLSSTEPNLQNIPIRLEEGRKIRQAFVPSKPD   357
SEQ ID NO. 21  YIEGLLKVVHRDTNKVHTIFNQALTQTGRLSSTEPNLQNIPIRLEEGRKIRQAFIPSEPD   357
SEQ ID NO. 22  YIEGLLKVVHRDTNKVHTIFNQALTQTGRLSSTEPNLQNIPIRLEEGRKIRQAFIPSEPD   357
SEQ ID NO. 23  YVEGLMKVVRKDTGKVHTIFQQALTQTGRLSSTEPNLQNIPIRIEEGRKIRQAFVPSSDD   357
```

FIG. 1A-4

```
              --+---------+---------+---------+---------+---------+-------
                360       370       380       390       400       410
              --+---------+---------+---------+---------+---------+-------
SEQ ID NO. 1  WLIFAADYSQIELRVLAHIADDDNLIEAFQRDLDIHTKTAMDIFHVSEEEVTANMRRQAK     417
SEQ ID NO. 2  WLIFAADYSQIELRVLAHIADDDNLIEAFQRDLDIHTKTAMDIFHVSEEEVTANMRRQAK     417
SEQ ID NO. 3  WLIFAADYSQIELRVLAHIADDDNLIEAFRRDLDIHTKTAMDIFHVSEEEVTATMRRQAK     417
SEQ ID NO. 4  WLIFAADYSQIELRVLAHIADDDNLIEAFRRDLDIHTKTAMDIFHVSEEEVTATMRRQAK     417
SEQ ID NO. 5  WLIFAADYSQIELRVLAHIADDDNLIEAFRRDLDIHAKTAMDIFHVSEEEVTATMRRQAK     417
SEQ ID NO. 6  WLIFAADYSQIELRVLAHIADDDNLIEAFRRDLDIHTKTAMDIFHVSEEEVTATMRRQAK     417
SEQ ID NO. 7  WLIFAADYSQIELRVLAHIADDDNLIEAFRRDLDIHTKTAMDIFHVSEEEVTATMRRQAK     417
SEQ ID NO. 8  WLIFAADYSQIELRVLAHIADDDNLIEAFRRDLDIHTKTAMDIFHVSEEEVTATMRRQAK     417
SEQ ID NO. 9  WLIFAADYSQIELRVLAHIADDDNLIEAFRRDLDIHTKTAMDIFHVSEEEVTATMRRQAK     417
SEQ ID NO. 10 WLIFAADYSQIELRVLAHIAEDDNLIEAFRRGLDIHTKTAMDIFHVSEEDVTANMRRQAK     417
SEQ ID NO. 11 WLIFAADYSQIELRVLAHIAEDDNLIEAFRRGLDIHTKTAMDIFHVSEEDVTANMRRQAK     417
SEQ ID NO. 12 WLIFAADYSQIELRVLAHIAEDDNLMEAFRRDLDIHTKTAMDIFQVSEDEVTPNMRRQAK     417
SEQ ID NO. 13 WLIFAADYSQIELRVLAHIAEDDNLMEAFRRDLDIHTKTAMDIFQVSEDEVTPNMRRQAK     417
SEQ ID NO. 14 WLIFAADYSQIELRVLAHIAEDDNLMEAFRRDLDIHTKTAMDIFQVSEDEVTPNMRRQAK     417
SEQ ID NO. 15 WLIFAADYSQIELRVLAHIAEDDNLMEAFRRDLDIHTKTAMDIFQVSEDEVTPNMRRQAK     417
SEQ ID NO. 16 WLIFAADYSQIELRVLAHIAEDDNLMEAFRRDLDIHTKTAMDIFQVSEDEVTPNMRRQAK     417
SEQ ID NO. 17 WLIFAADYSQIELRVLAHIAEDDNLMEAFRRDLDIHTKTAMDIFQVSEDEVTPNMRRQAK     418
SEQ ID NO. 18 WLIFAADYSQIELRVLAHIAEDDNLIEAFRRWLDIHTKTAMDIFHVSEEDVTANMRRQAK     419
SEQ ID NO. 19 WVIFSADYSQIELRVLAHIANDENLIDAFRHDLDIHTKTAMDIFHVNEDEVTPNMRRQAK     417
SEQ ID NO. 20 WVIFSADYSQIELRVLAHIANDENLIDAFRHDLDIHTKTAMDIFHVTADEVTPNMRRQAK     417
SEQ ID NO. 21 WVIFSADYSQIELRVLAHIANDENLIDAFRHGLDIHTKTAMDIFHVSKEEVTPNMRRQAK     417
SEQ ID NO. 22 WVIFSADYSQIELRVLAHIANDENLIDAFRHGLDIHTKTAMDIFHVSKEEVTPNMRRQAK     417
SEQ ID NO. 23 WVIFAADYSQIELRVLAHIANDENLIAAFHHDLDIHTKTAMDIFHVKEDEVTAHMRRQAK     417

              --+---------+---------+---------+---------+---------+-------
                420       430       440       450       460       470
              --+---------+---------+---------+---------+---------+-------
SEQ ID NO. 1  AVNFGIVYGISDYGLAQNLNITRKEAAEFIERYFASFPGVKQYMENIVQEAKQKGYVTTL     477
SEQ ID NO. 2  AVNFGIVYGISDYGLAQNLNITRKEAAEFIERYFASFPGVRRYMENIVQEAKQKGYVTTL     477
SEQ ID NO. 3  AVNFGIVYGISDYGLAQNLNITRKEAAEFIERYFASFPGVKRYMETIVQEAKQKGYVTTL     477
SEQ ID NO. 4  AVNFGVVYGISDYGLAQNLNITRKEAAEFIERYFASFPGVKRYMETIVQEAKQKGYVTTL     477
SEQ ID NO. 5  AVNFGIVYGISDYGLAQNLNITRKEAAEFIERYFASFPGVKRYMETIVQEAKQKGYVTTL     477
SEQ ID NO. 6  AVNFGIVYGISDYGLAQNLNITRKEAAEFIERYFASFPGVKRYMETIVQEAKQKGYVTTL     477
SEQ ID NO. 7  AVNFGIVYGISDYGLAQNLNITRKEAAEFIERYFASFPGVKRYMENIVQEAKQKGYVTTL     477
SEQ ID NO. 8  AVNFGIVYGISDYGLAQNLNITRKEAAEFIERYFASFPGVKRYMETIVQEAKQKGYVTTL     477
SEQ ID NO. 9  AVNFGIVYGISDYGLAQNLNITRKEAAEFIERYFASFPGVKRYMETIVQEAKQKGYVTTL     477
SEQ ID NO. 10 AVNFGIVYGISDYGLAQNLNITRKEAAEFIERYFASFPGVKQYMDNIVQEAKQKGYVTTL     477
SEQ ID NO. 11 AVNFGIVYGISDYGLAQNLNITRKEAAEFIERYFASFPGVKQYMDNIVQEAKQKGYVTTL     477
SEQ ID NO. 12 AVNFGIVYGISDYGLAQNLNISRKEAAEFIERYFESFPGVKRYMENIVQEAKQKGYVTTL     477
SEQ ID NO. 13 AVNFGIVYGISDYGLAQNLNISRKEAAEFIERYFESFPGVKRYMENIVQEAKQKGYVTTL     477
SEQ ID NO. 14 AVNFGIVYGISDYGLAQNLNISRKEAAEFIERYFESFPGVKRYMENIVQEAKQKGYVTTL     477
SEQ ID NO. 15 AVNFGIVYGISDYGLAQNLNISRKEAAEFIERYFESFPGVKRYMENIVQEAKQKGYVTTL     477
SEQ ID NO. 16 AVNFGIVYGISDYGLAQNLNISRKEAAEFIERYFESFPGVKRYMENIVQEAKQKGYVTTL     477
SEQ ID NO. 17 AVNFGIVYGISDYGLAQNLNISRKEAAEFIERYFESFPGVKRYMENIVQEAKQKGYVTTL     478
SEQ ID NO. 18 AVNFGIVYGISDYGLAQNLNITRKEAAEFIERYFASFPGVKQYMDNIVQEAKQKGYVTTL     479
SEQ ID NO. 19 AVNFGIVYGISDYGLSQNLNITRKEASEFIKRYFEIFPGVKQYMKDIVQEAKQKGYVTTL     477
SEQ ID NO. 20 AVNFGIVYGISDYGLSQNLNIPRKEAAEFIRRYFEIFPGVKQYMENIVQEAKQKGYVTTL     477
SEQ ID NO. 21 AVNFGIVYGISDYGLSQNLNITRKEAAEFIKRYFEIFPGVKQYMKDIVQEAKQKGYVTTL     477
SEQ ID NO. 22 AVNFGIVYGISDYGLSQNLNITRKEAAEFIKRYFEIFPGVKQYMKDIVQEAKQKGYVTTL     477
SEQ ID NO. 23 AVNFGIVYGISDYGLSQNLGITRKEAAEFIERYFRSYPGVKRYMEEVVQDAKQKGYVTTL     477
```

5/18

FIG. 1A-5

```
             --+---------+---------+---------+---------+---------+--------
               480       490       500       510       520       530
             --+---------+---------+---------+---------+---------+-------
SEQ ID NO. 1  LHRRRYLPDITSRNFNVRSFAERTAMNTPIQGSAADIIKKAMIDLAARLKEEQLQARLLL   537
SEQ ID NO. 2  LHRRRYLPDITSRNFNVRSFAERTAMNTPIQGSAADIIKKAMIDLAARLKEEQLQARLLL   537
SEQ ID NO. 3  LHRRRYLPDITSRNFNVRSFAERTAMNTPIQGSAADIIKKAMIDLAARLKEERLQARLLL   537
SEQ ID NO. 4  LHRRRYLPDITSRNFNVRSFAERTAMNTPIQGSAADIIKKAMIDLAARLKEERLQARLLL   537
SEQ ID NO. 5  LHRRRYLPDITSRNFNVRSFAERTAMNTPIQGSAADIIKKAMIDLAARLKEERLQARLLL   537
SEQ ID NO. 6  LHRRRYFPDITSRNFNVRSFAERTAMNTPIQGSAADIIKKAMIDLAARLKEERLQARLLL   537
SEQ ID NO. 7  LHRRRYLPDITSRNFNVRSFAERMAMNTPIQGSAADIIKKAMIDLAARLKEERLQARLLL   537
SEQ ID NO. 8  LHRRRYFPDITSRNFNVRSFAERTAMNTPIQGSAADIIKKAMIDLAARLKEERLQARLLL   537
SEQ ID NO. 9  LHRRRYFPDITSRNFNVRSFAERTAMNTPIQGSAADIIKKAMIDLAARLKEERLQARLLL   537
SEQ ID NO. 10 LHRRRYLPDITSRNFNVRSFAERTAMNTPIQGSAADIIKKAMIDLSVRLREERLQARLLL   537
SEQ ID NO. 11 LHRRRYLPDITSRNFNVRSFAERTAMNTPIQGSAADIIKKAMIDLSVRLREERLQARLLL   537
SEQ ID NO. 12 LHRRRYLPDITSRNFNVRSFAERMAMNTPIQGSAADIIKKAMIDLNARLKEERLQARLLL   537
SEQ ID NO. 13 LHRRRYLPDITSRNFNVRSFAERMAMNTPIQGSAADIIKKAMIDLNARLKEERLQARLLL   537
SEQ ID NO. 14 LHRRRYLPDITSRNFNVRSFAERMAMNTPIQGSAADIIKKAMIDLNARLKEERLQARLLL   537
SEQ ID NO. 15 LHRRRYLPDITSRNFNVRSFAERMAMNTPIQGSAADIIKKAMIDLNARLKEERLQARLLL   537
SEQ ID NO. 16 LHRRRYLPDITSRNFNVRSFAERMAMNTPIQGSAADIIKKAMIDLNARLKEERLQARLLL   537
SEQ ID NO. 17 LHRRRYLPDITSRNFNVRSFAERMAMNTPIQGSAADIIKKAMIDLNARLKEERLQARLLL   538
SEQ ID NO. 18 LHRRRYLPDITSRNFNVRTFAERTAMNTPIQGSAADIIKKAMIDLSVSVREERLQARLLL   539
SEQ ID NO. 19 LHRRRYLPDITSRNFNLRSFAERTAMNTPIQGSAADIIKKAMIDLSNRLKKENMKARMLL   537
SEQ ID NO. 20 LHRRRYLPDITSRNFNLRSFAERTAMNTPIQGSAADIIKKAMIDLADRLKQEKLQARMLL   537
SEQ ID NO. 21 LHRRRYLPDITSRNFNLRSFAERTAMNTPIQGSAADIIKKAMIDLSKRLQKENMKARMLL   537
SEQ ID NO. 22 LHRRRYLPDITSRNFNLRSFAERTAMNTPIQGSAADIIKKAMIDLSKRLQKENMKARMLL   537
SEQ ID NO. 23 LHRRRYLPDITSGNFNVRSFAERTAMNTPIQGSAADIIKKAMIDLANRLHEERLQTRLLL   537

             --+---------+---------+---------+---------+-------
               540       550       560       570       580
             --+---------+---------+---------+---------+-------
SEQ ID NO. 1  QVHDELILEAPKEEIERLCELVPEVMEQAVTLRVPLKVDYHYGPTWYDAK   587
SEQ ID NO. 2  QVHDELILEAPKEEIERLCELVPEVMEQAVSS-VPLKVDYHYGPTWYDAK   586
SEQ ID NO. 3  QVHDELILEAPKEEMERLCQLVPEVMEQAVALRVPLKVDYHYGPTWYDAK   587
SEQ ID NO. 4  QVHDELILEAPKEEMERLCQLVPEVMEQAVALRVPLKVDYHYGPTWYDAK   587
SEQ ID NO. 5  QVHDELILEAPKEEMERLCQLVPEVMEQAVELRVPLKVDYHYGPTWYDAK   587
SEQ ID NO. 6  QVHDELILEAPKEEMERLCQLVPEVMEQAVELRVPLKVDYHYGPTWYDPK   587
SEQ ID NO. 7  QVHDELILEAPKEEMERLCRLVPEVMEQAVELRVPLKVDYHYGPTWYDPK   587
SEQ ID NO. 8  QVHDELILEAPKEEMERLCQLVPEVMEQAVELRVPLKVDYHYGPTWYDPK   587
SEQ ID NO. 9  QVHDELILEAPKEEMERLCQLVPEVMEQAVELRVPLKVDYHYGPTWYDPK   587
SEQ ID NO. 10 QVHDELILEAPKEEIERLCRLVPEVMEQAVTLRVPLKVDYHYGPTWYDAK   587
SEQ ID NO. 11 QVHDELILEAPKEEIERLCRLVPEVMEQAVALRVPLKVDYHYGPTWYDAK   587
SEQ ID NO. 12 QVHDELILEAPKEEMERLCRLVPEVMEQAVTLRVPLKVDYHYGPTWYDAK   587
SEQ ID NO. 13 QVHDELILEAPKEEMERLCRLVPEVMEQAVTLRVPLKVDYHYGSTWYDAK   587
SEQ ID NO. 14 QVHDELILEAPKEEMERLCRLVPEVMEQAVTLRVPLKVDYHYGSTWYDAK   587
SEQ ID NO. 15 QVHDELILEAPKEEMERLCRLVPEVMEQAVTLRVPLKVDYHYGSTWYDAK   587
SEQ ID NO. 16 QVHDELILEAPKEEMERLCRLVPEVMEQAVTLRVPLKVDYHYGSTWYDAK   587
SEQ ID NO. 17 QVHDELILEAPKEEMERLCRLVPEVMEQAVTLRVPLKVDYHYGSTWYDAK   588
SEQ ID NO. 18 QGHDELILEAPKEEIGRLCRLVPEVMEQAVTLRVPLKVDYHYGPTWYDAK   589
SEQ ID NO. 19 QVHDELILEAPKEEVERLQQIVPEVMENAVQLRVPLKVDYHFGPTWYDAK   587
SEQ ID NO. 20 QVHDELILEAPKEEVKRLQQIVPEVMENAIQLKVPLLIDFRFGPTWDDAK   587
SEQ ID NO. 21 QVHDELILEAPKDEIERLQQIVPEVMENAVQLRVPLKVDYHFGPTWYDAK   587
SEQ ID NO. 22 QVHDELILEAPKDEIERLQQIVPEVMENAVQLRVPLKVDYHFGPTWYDAK   587
SEQ ID NO. 23 QVHDELILEAPKEEIELLKKIVPDVMENAVSLRVPLKVDYHFGPTWYDAK   587
```

FIG. 1B

PERCENT IDENTITY

DIVERGENCE

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	
1		98.3	94.1	93.7	93.2	92.9	93.1	92.9	92.7	89.7	89.3	89.0	89.0	88.8	88.6	88.5	88.5	87.3	79.0	78.8	78.1	78.3	76.6	1
2	1.6		92.9	92.5	92.0	91.7	91.9	91.7	91.5	88.0	87.8	87.6	87.6	87.5	87.3	87.1	87.1	85.6	77.6	77.5	76.9	77.1	75.8	2
3	6.2	7.4		99.7	98.0	98.0	97.6	97.6	97.8	89.5	89.5	89.7	89.7	89.8	89.7	89.5	89.5	87.1	78.1	78.1	77.3	77.5	76.6	3
4	6.6	7.7	0.3		97.6	97.6	97.3	97.3	97.5	89.3	89.3	89.3	89.3	89.5	89.3	89.2	89.2	86.9	77.8	77.8	76.9	77.1	76.4	4
5	7.2	8.3	2.1	2.4		99.3	99.0	99.3	99.2	88.6	88.5	89.2	89.2	89.3	89.5	89.3	89.0	86.3	78.0	77.6	76.6	76.8	76.6	5
6	7.5	8.7	2.1	2.4	0.7		99.3	99.7	99.8	88.6	88.5	89.2	89.2	89.3	89.5	89.3	89.0	86.3	77.6	77.6	76.6	76.8	76.3	6
7	7.3	8.5	2.4	2.8	1.0	0.7		99.0	99.2	89.0	88.8	89.8	89.8	90.0	90.2	90.0	89.7	86.6	77.5	77.6	76.4	76.6	76.3	7
8	7.5	8.7	2.4	2.8	0.7	0.3	1.0		99.8	88.3	88.1	88.8	88.8	89.0	89.2	89.0	88.6	85.9	77.6	77.3	76.3	76.4	76.3	8
9	7.7	8.9	2.2	2.6	0.9	0.2	0.9	0.2		88.5	88.3	89.0	89.0	89.2	89.3	89.2	88.8	86.1	77.5	77.5	76.4	76.6	76.1	9
10	11.2	13.0	11.4	11.6	12.4	12.4	12.0	12.8	12.6		99.7	88.8	88.8	88.6	88.6	88.5	88.3	97.5	77.3	77.5	76.6	76.6	75.1	10
11	11.6	13.2	11.4	11.6	12.6	12.6	12.2	13.0	12.8	0.3		88.5	88.5	88.3	88.3	88.1	88.0	97.1	77.1	77.3	76.4	76.4	74.9	11
12	12.0	13.5	11.2	11.6	11.8	11.8	11.0	12.2	12.0	12.2	12.6		99.7	99.5	99.3	99.0	99.2	86.6	78.8	79.0	77.8	77.6	76.6	12
13	12.0	13.5	11.2	11.6	11.8	11.8	11.0	12.2	12.0	12.2	12.6	0.3		99.8	99.7	99.3	99.5	86.6	78.8	79.0	77.8	77.6	76.3	13
14	12.2	13.7	11.0	11.4	11.6	11.6	10.8	12.0	11.8	12.4	12.8	0.5	0.2		99.8	99.5	99.7	86.4	78.6	78.8	77.6	77.5	76.1	14
15	12.4	13.9	11.2	11.6	11.4	11.4	10.6	11.8	11.6	12.4	12.8	0.7	0.3	0.2		99.7	99.5	86.4	78.6	78.6	77.6	77.5	76.1	15
16	12.6	14.1	11.4	11.8	11.6	11.6	10.8	12.0	11.8	12.6	13.0	1.0	0.7	0.5	0.3		99.2	86.3	78.3	78.3	77.3	77.1	76.3	16
17	12.4	13.9	11.2	11.6	11.8	11.8	11.0	12.2	12.0	12.6	13.0	0.7	0.3	0.2	0.3	0.7		86.1	78.3	78.5	77.3	77.1	75.8	17
18	13.6	15.5	13.8	14.0	14.9	14.9	14.5	15.3	15.1	2.2	2.6	14.5	14.5	14.7	14.7	14.9	14.9		75.1	75.4	74.2	74.2	73.1	18
19	24.9	26.6	26.1	26.5	26.3	26.8	27.0	26.8	27.0	27.3	27.5	25.1	25.1	25.3	25.3	25.8	25.6	30.0		88.8	93.9	93.9	76.6	19
20	25.1	26.8	26.1	26.5	26.8	26.8	26.8	27.3	27.0	27.0	27.3	24.9	24.9	25.1	25.3	25.8	25.3	29.5	12.2		88.0	88.0	75.6	20
21	26.1	27.6	27.3	27.8	28.3	28.3	28.5	28.8	28.5	28.3	28.5	26.5	26.5	26.8	26.8	27.3	27.0	31.3	6.4	13.2		99.3	74.7	21
22	25.8	27.3	27.0	27.5	28.0	28.0	28.3	28.5	28.3	28.3	28.5	26.8	26.8	27.0	27.0	27.5	27.3	31.3	6.4	13.2	0.7		74.7	22
23	28.3	29.3	28.3	28.5	28.3	28.8	28.8	28.8	29.0	30.5	30.8	28.3	28.8	29.0	29.0	28.8	29.3	33.1	28.3	29.8	31.0	31.0		23
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	

FIG. 2-1

Segment 1: 1..31
AEGEKPLEEMEFAIVDVITEEMLADKAALVV (SEQ ID NO:24)
AEEEKPLEDIEFEIADEVTEEMLADEAALVV (SEQ ID NO:25)
AEEEVPLEEMEFVIADEITEEMLADKAALVV (SEQ ID NO:26)
AEEEVPLEEMEFTIADEITEEMLADKAALVV (SEQ ID NO:27)
AEEEKPLEEMEFAIADEVTEEMLADKAALVV (SEQ ID NO:28)
AEDEKPLEEIEFAIADEITEEMLADKAALVV (SEQ ID NO:29)
AEEEKPLAEMEFTIADEVTEEMLADKAALVV (SEQ ID NO:30)
AEEEVPLAEMEFVIADEITEEMLADKAALVV (SEQ ID NO:31)
AEGEKPLAEMEFAIVDEITEEMLADKAALVV (SEQ ID NO:32)
AEEEAPLEDIEFDIADEVTEEMLADKAALVV (SEQ ID NO:33)
AEEEKPLAEMEFAIADEVTEEMLADKAALVV (SEQ ID NO:34)
EEEEVPLEEIEFAIADEVTEEMLADKAALVV (SEQ ID NO:35)
AEDEKPLAEMEFVIADGITDEMLADKAALVV (SEQ ID NO:36)
AEKELPLMEMEFADADTITMEMLADKAALVV (SEQ ID NO:37)
AEEEKPLAEMEFVIADGITEEMLADKAALVV (SEQ ID NO:38)
AEEEKPLAEMEFAIADEITEEMLADKAALVV (SEQ ID NO:39)

Segment 2: 32..111
EVMEENYHDAPIVGIALVNEHGRFFMRPETALADSQFLAWLADETKKKSMFDAKRAVVALKWKGIELRGVAFDLLLAAYL
(SEQ ID NO:40)
EVQEENYHDAPIVGIAIVNEHGRFFLRAETALADFQFVAWLEDETKKKSMFDAKRAVVALKWKGIELVGVDFDLLLAAYL
(SEQ ID NO:41)
EVLEENYHDAPIVGIALVNEHGRFFLRPETALADPQFVAWLEDETKKKSMFDAKRAAVALKWKGIELRGVAFDLLLAAYL
(SEQ ID NO:42)
EVMEENYHDAPIVGIALVNEHGRFFLRAETALADPQFVAWLADETKKKSMFDAKRAAVALKWKGIELRGVAFDLLLAAYL
(SEQ ID NO:43)
EVMEENYHDAPIVGIALVNEHGRFFLRPETALADPQFLAWLADETKKKSMFDAKRAAVALKWKGIELRGVAFDLLLAAYL
(SEQ ID NO:44)
EVLEENYHDAPIVGFAIVNEHGRFFLRTELALADFQFVAWLEDETKKKSMFDRKRAVVALKWKGIELVGVDFDLLLAAYL
(SEQ ID NO:45)
EVQEDNYHDAPIVGFAIVNEHGRFFIRTETALASEAFKAWLADETKKKSVFDAKRAIVALKWKGIELRGVDFDLLLAAYL
(SEQ ID NO:46)
EVMEENYHDAPIVGIALVNEHGRFFLRPETALADPQFLAWLADETKKKSMFDAKRAIVALKWKGIELRGVAFDLLLAAYL
(SEQ ID NO:47)
EVMEENYHDAPIVGIANVNEHGRFFLRTELALADFQFVAWLEDETKKKSMFDRKRAAVALKWKGIELVGVDFDLLLAAYL
(SEQ ID NO:48)
EVLEENYHDAPIVGFALVNEHGRFFIRTETALASSQFKAWLEDETKKKSVFDAKRAIVALKWKGIELRGVDFDLLLAAYL
(SEQ ID NO:49)
EVLEENYHDAPIVGFALVNEHGRFFLRPETALADEQFVAWLADETKKKSVFDAKRAVVALKWKGIELRGVDFDLLLAAYL
(SEQ ID NO:50)
EVMEENYHDAPIVGIALVNEHGRFFLRPETALADPQFVAWLADETKKKSMFDAKRAAVALKWKGIELRGVAFDLLLAAYL
(SEQ ID NO:51)
EVMEENYHDAPIVGIALVNEHGRFFLRPETALASPQFKAWLADETKKKSMFDAKRAIVALKWKGIELRGVAFDLLLAAYL
(SEQ ID NO:52)
EVLEENYHDAPIVGFAIVNEHGRFFIRPETALASSQFKAWLEDETKKKSMFDAKRAAVALKWKGIELRGVAFDLLLAAYL
(SEQ ID NO:53)
EVMEENYHDAPIVGIALVNEHGRFFLRTETALADPQFKAWLADETKKKSMFDAKRAIVALKWKGIELRGVDFDLLLAAYL
(SEQ ID NO:54)
EVQEDNYHDAPIVGFANVNEHGRFFLRTETALASEAFAAWLEDETKKKSVFDAKRAVVALKWKGIELVGVAFDLLLAAYL
(SEQ ID NO:55)
EVLEENYHDAPIVGIALVNEHGRFFLRPETALADSQFLAWLEDETKKKSMFDAKRAAVALKWKGIELRGVAFDLLLAAYL
(SEQ ID NO:56)

FIG. 2-2

Segment 3: 112..191
LNPAQDAGDIAAVAKMKQYEAVRSDEAVYGKGVKRSLPDEQTLAEHLVRKAAAIWALEQPFMDDLRNNEQDQLLTKLEQP
(SEQ ID NO:57)
LNPAQDAGDVAAVAKMKQYEAVRSDEAVYGKGAKRSLPDEPVLAEHLVRKAAAIWALERPFLDELRENEQDELLTDLEQP
(SEQ ID NO:58)
LNPAQAAGDVAAVAKMKQYEAVRSDEAVYGKGAKRAVPDEPVLAEHLVRKAAAIWALEEPFIDELRRNEQDRLLTDLEQP
(SEQ ID NO:59)
LNPAQDAGDVAAVAKMKQYEAVRSDEAVYGKGAKRAVPDEPTLAEHLVRKAAAIWALERPFLDELRRNEQDELLIKLEQP
(SEQ ID NO:60)
LAPAQDDGDAAAKAKMKQYEAVREDEAVYGKGAKRPDPDELALAEHLVRKAAAIWALERPFLDELRENEQDLLLLELEQP
(SEQ ID NO:61)
LNPAQDAGDVAAVAKMKQYEAVRSDEAVYGKGAKRSLPDEPTLAEHLVRKAAAIWALERPFLDELRENEQDELLIKLEQP
(SEQ ID NO:62)
LNPAQTADDVAAVAKMKQYHAVRSDEAVYGKGAKRAVPDEPVLAEHLVRKAAAIWALEEPFLDELRKNEQDELFTELELP
(SEQ ID NO:63)
LNPAQDAGDVAAVAKMKQYEAVRSDEAVYGKGAKRSLPDEPTLAEHLVRKAAAIWALEQPFIDELRRNEQDELLTKLEQP
(SEQ ID NO:64)
LNPAQSAEDVAAVAKMKQYEAVRSDEAVYGKGAKRAVPDEPVLAEHLVRKAAAIWALEEPFIDELRENEQDELFTDLEQP
(SEQ ID NO:65)
LNPAQDAGDVAAVAKMKQYEAVRSDEAVYGKGAKRALPDEPTLAEHLVRKAAAIWALEEPFLDELRENEQDELLTELEQP
(SEQ ID NO:66)
LNPAQSAGDVAAVAKMKQYEAVRSDEAVYGKGAKRAVPDEPTLAEHLVRKAAAIWALEEPFIDELRENEQDELFTELEMP
(SEQ ID NO:67)
LNPAQDAGDVAAVAKMKQYEAVRPDEAVYGKGAKRSLPDEPTLAEHLVRKAAAIWALERPFLDELRRNEQDRLLIKLEQP
(SEQ ID NO:68)
LNPAQDAGDVAAVAKMKQYEAVRSDEAVYGKGAKRSLPDEPTLAEHLVRKAAAIWALERPFLDELRRNEQDELLTKLEQP
(SEQ ID NO:69)
LNPAQDAGDVAAVAKMKQYEAVRSDEAVYGKGAKRAVPDEPVLAEHLVRKAAAIWALERPFLDELRNNEQDELLTELEQP
(SEQ ID NO:70)
LNPAQDAGDVAAVAKMKQYEAVRSDEAVYGKGAKRSVPDEPTLAEHLVRKAAAIWALEQPFLDELRRNEQDRLLTKLEQP
(SEQ ID NO:71)
LNPAQSAGDVAAVAKMKQYEAVRSDEAVYGKGAKRAVPDEPTLAEHLVRKAAAIWALEQPFLDELRENEQDELLTKLEQP
(SEQ ID NO:72)

FIG. 2-3

Segment 4: 192..271
LAAILAEMEFTGVNVDTKRLEQMGSELAEQLRAIEQRIYELAGQEFNINSPKQLGVILFEKLQLPVLKKTKTGYSTSADV
(SEQ ID NO:73)
LATILAEMEFTGVKVDTKRLEQMGSELAEQLRAVEQRIYELAGQEFNINSPKQLGVILFEKLQLPVLKKTKTGYSTSADV
(SEQ ID NO:74)
LATILAEMEFTGVKVDTKRLEQMGEELAEQLGAVEQRIYELAGQEFNINSPKQLGVILFEKLQLPVLKKTKTGYSTSADV
(SEQ ID NO:75)
LILILAEMEFTGVDVDTKRLEQMGLELAEQLVEQEQRIYELAGQEFNINSPKQLGLILFEKLQLPVLKKTKTGYSTSADV
(SEQ ID NO:76)
LSLILAEMEFTGVDVDTKRLEQMGLELAEQLGAQEQRIYELAGQEFNINSPKQLGVILFEKLQLPVLKKTKTGYSTSADV
(SEQ ID NO:77)
LALILAEMEFTGVKVDTKRLEQMGEELAEQLKEVEQRIYELAGQEFNINSPKQLGVILFEKLQLPVLKKTKTGYSTSAQP
(SEQ ID NO:78)
LILILAEMEFTGVDVDTKRLEQMGSELAEQLGAIEQRIYELAGQEFNINSPKQLGLILFEKLQLPVLKKTKTGYSTSADV
(SEQ ID NO:79)
LAAILAEMEFTGVKVDTKRLEQMGEELAEQLKEVEQRIYELAGQEFNINSPKQLGVILFEKLQLPVLKKTKTGYSTSADV
(SEQ ID NO:80)
LILILAEMEFTGVDVDTKRLEQMGLELAEQLVAQEQRIYELAGQEFNINSPKQLGVILFEKLQLPVLKKTKTGYSTSADV
(SEQ ID NO:81)
LSLILAEMEFTGVDVDTKRLEQMGLELAEQLVEIEQRIYELAGQEFNINSPKQLGLILFEKLQLPVLKKTKTGYSTSADV
(SEQ ID NO:82)
LALILAEMEFTGVKVDTKRLEQMGEELAEQLKAIEQRIYELAGQEFNINSPKQLGVILFEKLQLPVLKKTKTGYSTSADV
(SEQ ID NO:83)
LALILAEMEFTGVKVDTKRLEQMGEELAEQLKEIEQRIYELAGQEFNINSPKQLGVILFEKLQLPVLKKTKTGYSTSADV
(SEQ ID NO:84)
LALILAEMEFTGVKVDTKRLEQMGEELAEQLKAVEQRIYELAGQEFNINSPKQLGVILFEKLQLPVLKKTKTGYSTSADV
(SEQ ID NO:85)
LALILAEMEFTGVKVDTKRLEQMGEELAEQLKEIEQRIYELAGQEFNINSPKQLGVILFEKLQLPVLKKTKTGYSTSADV
(SEQ ID NO:86)
LSSILAEMEFTGVKVDTKRLEQMGEELAEQLRAVEQRIYELAGQEFNINSPKQLGVILFEKLQLPVLKKTKTGYSTSADV
(SEQ ID NO:87)

FIG. 2-4

Segment 5: 272..351
LEKLAPHHEIVENILHYRQLGKLQSTYIEGLLKVVRPDTGKVHTMFNQALTQTGRLSSAEPNLQNIPIRLEEGRKIRQAF
(SEQ ID NO:88)
LEKLAPHHEIVENILHYRQLGKLQSTYIEGLLKVVHPDTGKVHTRFNQALTQTGRLSSTDPNLQNIPIRLEEGRKIRQAF
(SEQ ID NO:89)
LEKLAPEHEIVENILHYRQLGKLQSTYIEGLLKVVRTDTGKVHTIFNQALTQTGRLSSAEPNLQNIPIRLEEGRKIRQAF
(SEQ ID NO:90)
LEKLAPEHEIVENILHYRQLGKLQSTYIEGLLKVVRTDTGKVHTMFNQALTQTGRLSSADPNLQNIPIRLEEGRKIRQAF
(SEQ ID NO:91)
LEKLAPHHEIVENILHYRQLGKLQSTYIEGLLKVVHPDTGKVHTMFNQALTQTGRLSSTEPNLQNIPIRLEEGRKIRQAF
(SEQ ID NO:92)
LEKLAPHHEIVENILHYRQLGKLQSTYIEGLLKVVRPDTGKVHTMFNQALTQTGRLSSTEPNLQNIPIRLEEGRKIRQAF
(SEQ ID NO:93)
LEKLAPEHEIVENILHYRQLGKLQSTYIEGLLKVVDTDTGKVHTMFNQALTQTGRLSSAEPNLQNIPIRLEEGRKIRQAF
(SEQ ID NO:94)
LEKLAPHHEIVENILHYRQLGKLQSTYIEGLLKVVHPDTGKVHTIFNQALTQTGRLSSTEPNLQNIPIRLEEGRKIRQAF
(SEQ ID NO:95)
LEKLAPHHEIVENILHYRQLGKLQSTYIEGLLKVVHPDTGKVHTMFNQALTQTGRLSSAEPNLQNIPIRLEEGRKIRQAF
(SEQ ID NO:96)
LEKLAPHHEIVENILHYRQLGKLQSTYIEGLLKVVDTDTGKVHTMFNQALTQTGRLSSAEPNLQNIPIRLEEGRKIRQAF
(SEQ ID NO:97)
LEKLAPHHEIVENILHYRQLGKLQSTYIEGLLKVVHPDTGKVHTRFNQALTQTGRLSSTEPNLQNIPIRLEEGRKIRQAF
(SEQ ID NO:98)
LEKLAPEHEIVENILHYRQLGKLQSTYIEGLLKVVRTDTGKVHTMFNQALTQTGRLSSAEPNLQNIPIRLEEGRKIRQAF
(SEQ ID NO:99)

FIG. 2-5

Segment 6: 352..431
VPSEPDWLIFAADYSQIELRVLAHIADDDNLIEAFQRDLDIHTKTAMDIFHVSEEEVTANMRRQAKAVNFGIVYGISDYG
(SEQ ID NO:100)
VPSEPLWLIFAADYSQIELRVLAHIADDDNLAEAFQRDLDIHTKTAMDIFHVSEEEVTSRMRRQAKAVNFGIVYGISDYG
(SEQ ID NO:101)
VPSEPLWLIFAADYSQIELRVLAHIADDDNLAEAFRRDLDIHTKTAMDIFHVSEEEVTARMRRQAKAVNFGIVYGISDYG
(SEQ ID NO:102)
VPSEPDWLIFAADYSQIELRVLAHIANDENLIEAFRRDLDIHTKTAMDIFHVSEDEVTANMRRQAKAVNFGIVYGISDYG
(SEQ ID NO:103)
VPSEPDWLIFAADYSQIELRVLAHIANDDNLIEAFRRDMDIHTKTAMDVFHVSEDEVTSNMRRQAKAVNFGIVYGISDYG
(SEQ ID NO:104)
VPSEPLWLIFAADYSQIELRVLAHIANDDNLAEAFQRDMDIHTKTAMDVFHVSEEEVTARMRRQAKAVNFGIVYGISDYG
(SEQ ID NO:105)
VPSEPDWLIFAADYSQIELRVLAHIADDDNLIEAFRRDLDIHTKTAMDIFHVSEEEVTANMRRQAKAVNFGIVYGISDYG
(SEQ ID NO:106)
VPSEPDWLIFAADYSQIELRVLAHIADDDNLIEAFRRDLDIHTKTAMDIFHVSEDEVTANMRRQAKAVNFGIVYGISDYG
(SEQ ID NO:107)
VPSEPDWLIFAADYSQIELRVLAHIANDDNLIEAFRRDLDIHTKTAMDIFHVSEDEVTANMRRQAKAVNFGIVYGISDYG
(SEQ ID NO:108)
VPSEPLWLIFAADYSQIELRVLAHIANDENLAEAFQRDLDIHTKTAMDVFHVSEEEVTSRMRRQAKAVNFGIVYGISDYG
(SEQ ID NO:109)
VPSEPDWLIFAADYSQIELRVLAHIADDENLIEAFRRDLDIHTKTAMDIFHVSEEEVTANMRRQAKAVNFGIVYGISDYG
(SEQ ID NO:110)
VPSEPDWLIFAADYSQIELRVLAHIADDDNLIEAFRRDLDIHTKTAMDIFHVSEEEVTANMRRQAKAVNFGIVYGISDYG
(SEQ ID NO:111)

FIG. 2-6

Segment 7: 432..511
LAQNLNITRKEAAEFIERYFASFPGVKQYMENIVQEAKQKGYVTTLLHRRRYLPDITSRNFNVRSFAERTAMNTPIQGSA
(SEQ ID NO:112)
LSQNLNITRKEAAEFIERYFESFPGVKRYMENIVQEAKQKGYVTTLLHRRRYLPDITSRNFNVRSFAERTAMNTPIQGSA
(SEQ ID NO:113)
LAQNLNITRKEAAEFIERYFASFPGVKRYMEDIVQEAKQKGYVTTLLHRRRYLPDITSRNFNVRSFAERTAMNTPIQGSA
(SEQ ID NO:114)
LAQNLNITRKEAAEFIERYFESFPGVKQYMENIVQEAKQKGYVTTLLHRRRYLPDITSRNFNVRSFAERTAMNTPIQGSA
(SEQ ID NO:115)
LAQNLNIKRKEAAEFIERYFASFPGVKRYMEVIVQEAKQKGYVTTLLHRRRYDPDITSRNFNVRSFAERMAMNTPIQGSA
(SEQ ID NO:116)
LSQNLNITRKEAAEFIERYFESFPGVKQYMEDIVQEAKQKGYVTTLLHRRRYLPDITSRNFNVRSFAERTAMNTPIQGSA
(SEQ ID NO:117)
LAQNLGITRKEAAEFIERYFASFPGVKQYMEDIVQEAKQKGYVTTLLHRRRYLPEITSRNFNLRSFAERTAMNTPIQGSA
(SEQ ID NO:118)
LAQNLGITRKEAAEFIERYFASFPGVKEYMEEIVQEAKQKGYVTTLLHRRRYLPDITSRNFNVRSFAERMAMNTPIQGSA
(SEQ ID NO:119)
LAQNLNITRKEAAEFIERYFASFPGVKRYMENIVQEAKQKGYVTTLLHRRRYLPDITSRNFNVRSFAERTAMNTPIQGSA
(SEQ ID NO:120)
LAQNLNITRKEAAEFIERYFESFPGVKRYMENIVQEAKQKGYVTTLLHRRRYLPDITSRNFNVRSFAERTAMNTPIQGSA
(SEQ ID NO:121)
LSQNLNITRKEAAEFIERYFASFPGVKQYMENIVQEAKQKGYVTTLLHRRRYLPDITSRNFNVRSFAERTAMNTPIQGSA
(SEQ ID NO:122)
LAQNLNIKRKEAAEFIERYFASFPGVKQYMEVIVQEAKQKGYVTTLLHRRRYDPDITSRNFNVRSFAERMAMNTPIQGSA
(SEQ ID NO:123)
LSQNLGITRKEAAEFIERYFESFPGVKEYMEDIVQEAKQKGYVTTLLHRRRYLPEITSRNFNLRSFAERTAMNTPIQGSA
(SEQ ID NO:124)
LSQNLNITRKEAAEFIERYFASFPGVKRYMEEIVQEAKQKGYVTTLLHRRRYLPDITSRNFNVRSFAERTAMNTPIQGSA
(SEQ ID NO:125)

FIG. 2-7

Segment B: 512..587
ADIIKKAMIDLAARLKEEQLQARLLLQVHDELILEAPKEEIERLCELVPEVMEQAVTLRVPLKVDYHYGPTWYDAK DYG
(SEQ ID NO:126)
ADIIKKAMIDLAARLKEERLQARLLLQVHDELILEAPKEEIERLCRLVPEVMEQAVTLRVPLKVDYHYGPTWYDAKSDYG
(SEQ ID NO:127)
ADIIKKAMIDLAARLKEERLQARLLLQVHDELILEAPKEEIERLCKLVPEVMENAVTLRVPLKVDYHYGPTWYDAKSDYG
(SEQ ID NO:128)
ADIIKKAMIDLAARLKEEQLQARLLLQVHDELILEAPKEEMERLCVLVPEVMEQAVRLRVPLKVDYHYGWTWYDAKSDYG
(SEQ ID NO:129)
ADIIKKAMIDLAARLKEERLQARLLLQVHDELILEAPKEEMERLCRLVPEVMEQAVTLRVPLKVDYHYGPTWYDAKSDYG
(SEQ ID NO:130)
ADIIKKAMIDMAARLKEEQLQARLLLQVHDELILEAPKEEMERLEELVPEVMEHAVTLRVPLKVDYHYGWTWYDAKSDYG
(SEQ ID NO:131)
ADIIKKAMIDMAARLKEEQLQARLLLQVHDELIFEAPKEEIERLEELVPEVMENAVTLRVPLKVDYHYGWTWYDAKSDYG
(SEQ ID NO:132)
ADIIKKAMIDLAARLKEERLQARLLLQVHDELILEAPKEEIERLCQLVPEVMEQAVELRVPLKVDYHYGPTWYDAKSDYG
(SEQ ID NO:133)
ADIIKKAMIDMAARLKEERLQARLLLQVHDELIFEAPKEEIERLEKLVPEVMEHAVELRVPLKVDYHYGPTWYDAKSDYG
(SEQ ID NO:134)
ADIIKKAMIDLAARLKEEQLQARLLLQVHDELIFEAPKEEMERLEELVPEVMEHAVTLRVPLKVDYHYGWTWYDAKSDYG
(SEQ ID NO:135)
ADIIKKAMIDMAARLKEEQLQARLLLQVHDELIFEAPKEEIERLCELVPEVMENAVTLRVPLKVDYHYGWTWYDAKSDYG
(SEQ ID NO:136)
ADIIKKAMIDLAARLKEERLQARLLLQVHDELILEAPKEEIERLCQLVPEVMEQAVTLRVPLKVDYHYGPTWYDAKSDYG
(SEQ ID NO:137)
ADIIKKAMIDLAARLKEERLQARLLLQVHDELILEAPKEEIERLCKLVPEVMEQAVELRVPLKVDYHYGPTWYDAK
(SEQ ID NO:138)

FIG. 3A

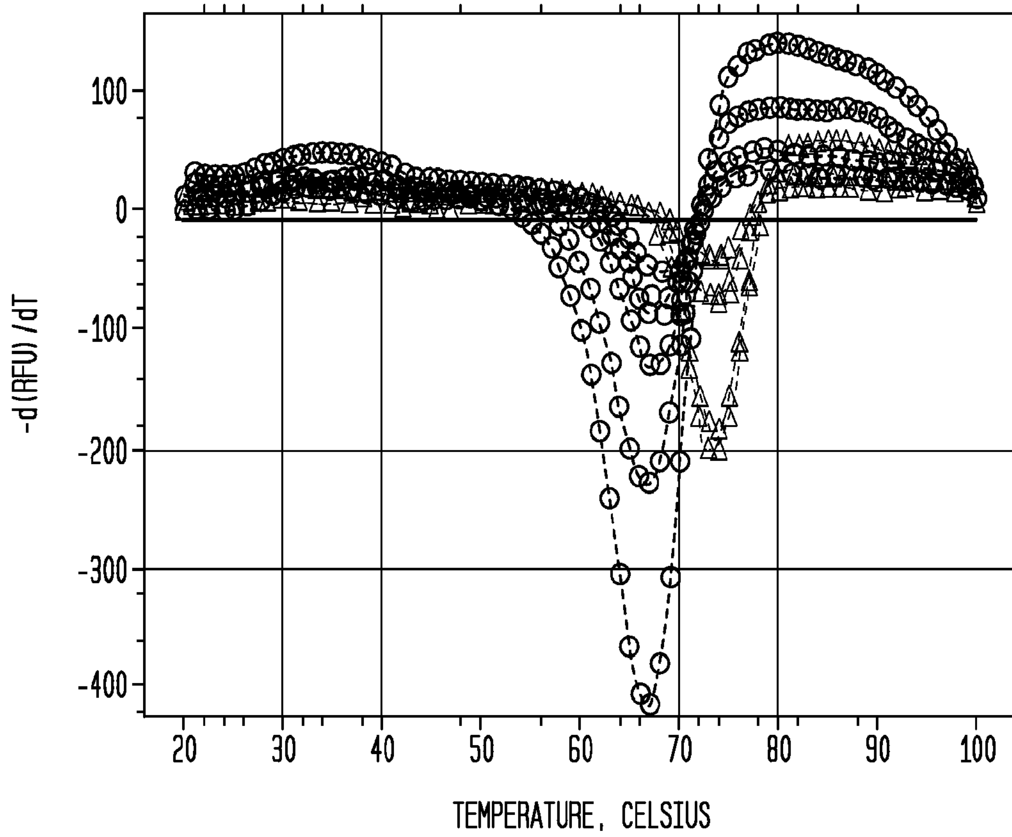


FIG. 3B

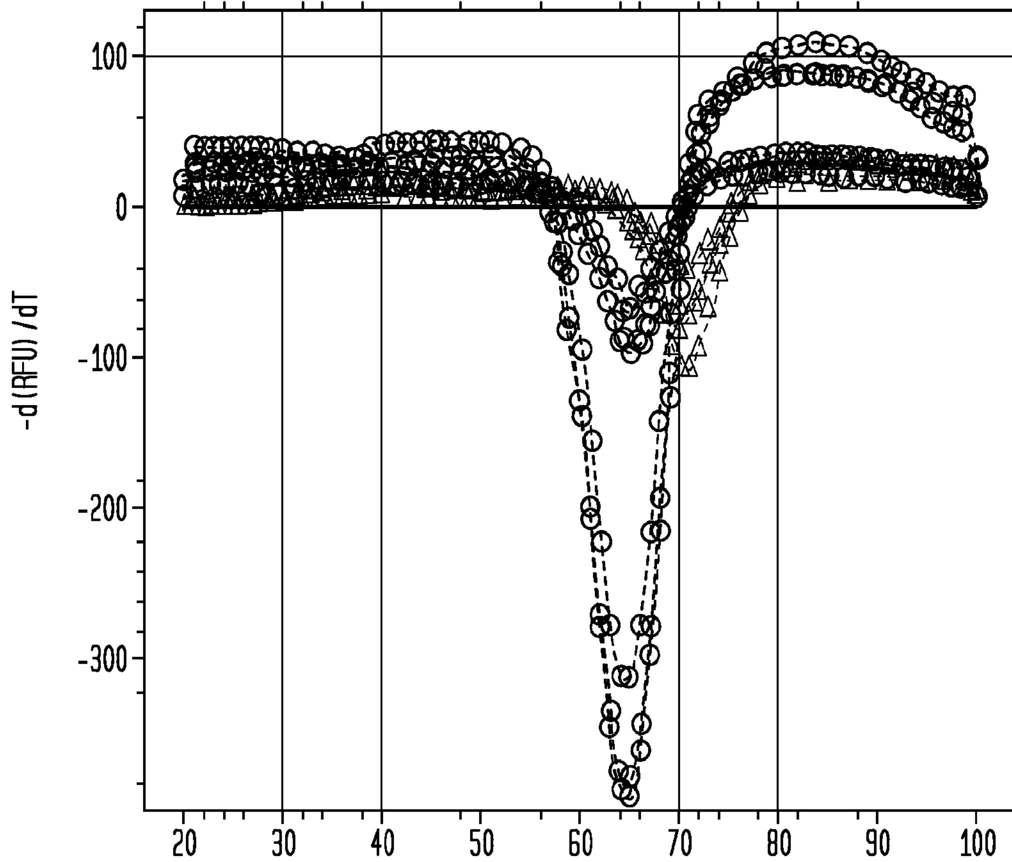


FIG. 4A

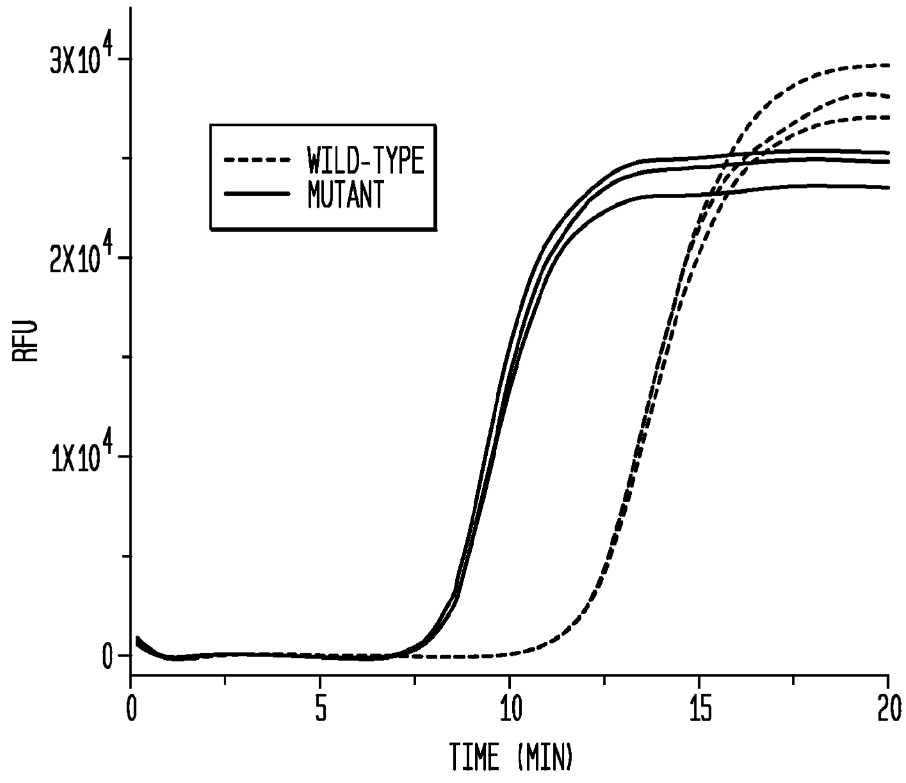
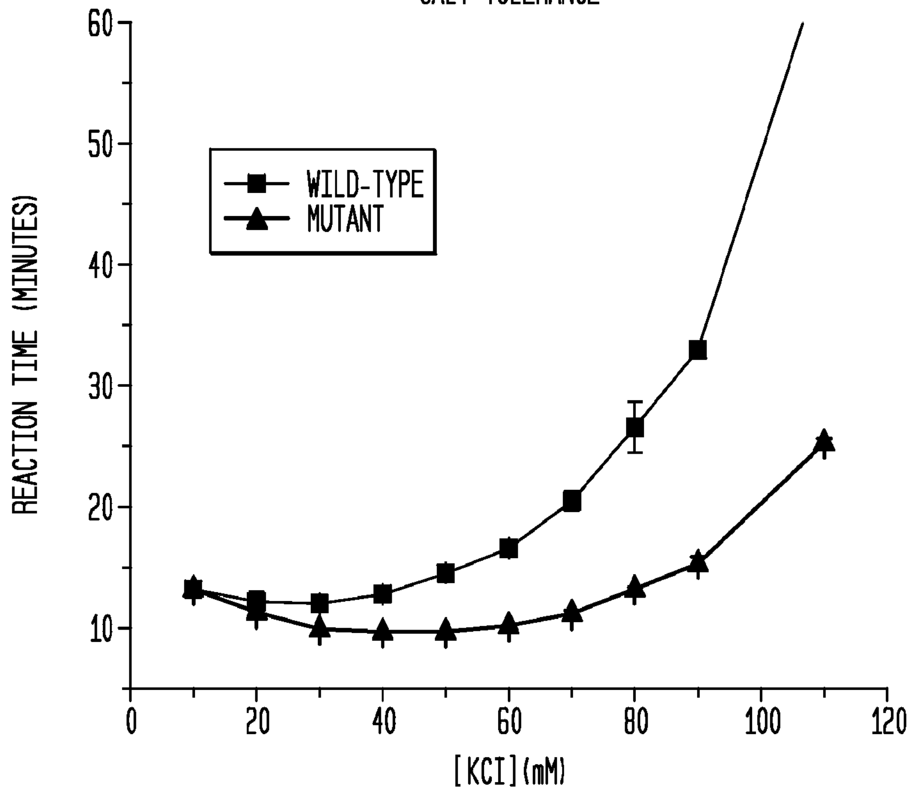


FIG. 4B



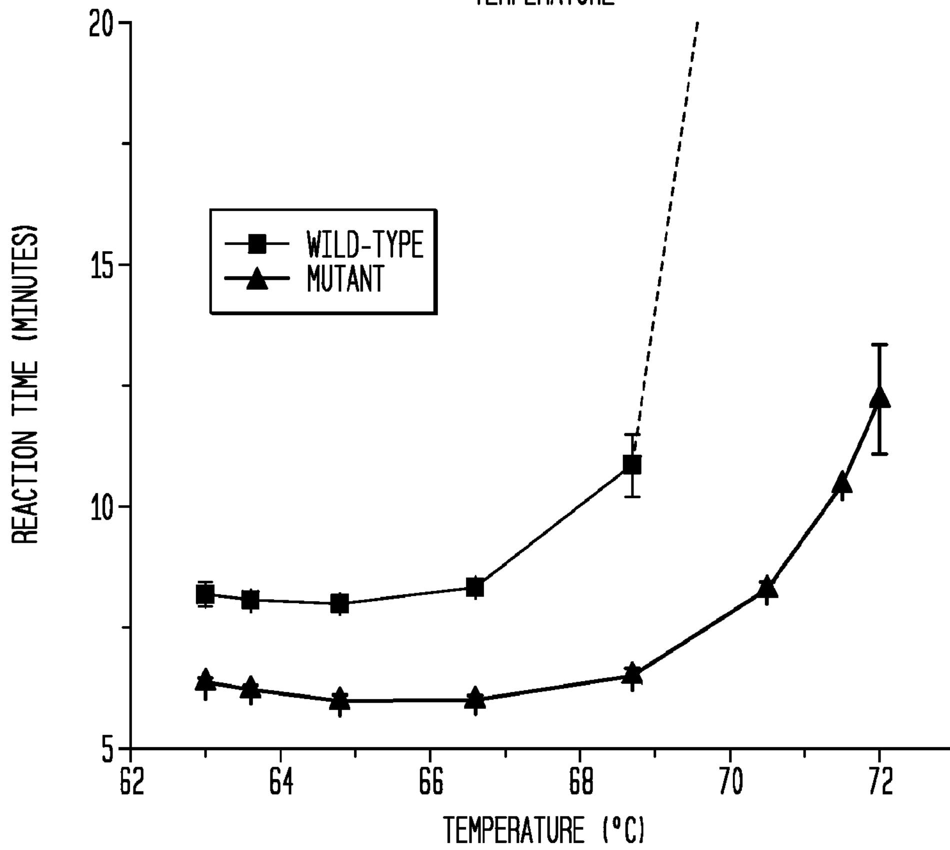


FIG. 4C
TEMPERATURE
REACTION TIME (MINUTES)
20
15
10
5
WILD-TYPE
MUTANT
62
64
66
68
70
72
TEMPERATURE (°C)

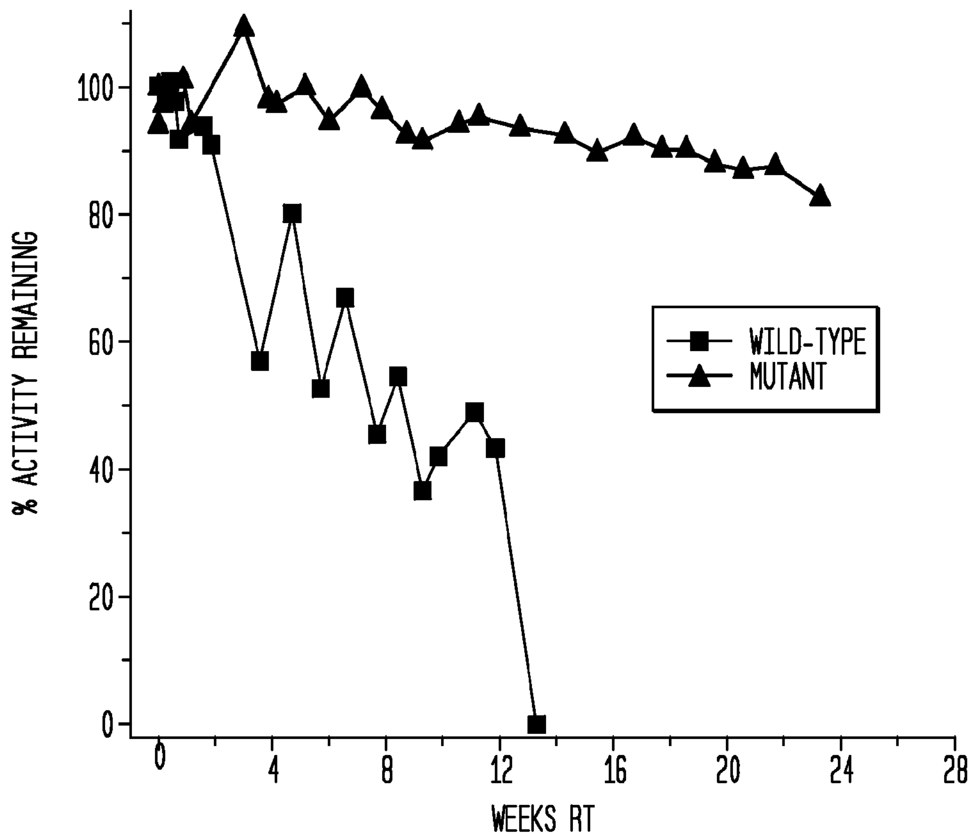


FIG. 4D
STORAGE STABILITY
% ACTIVITY REMAINING
100
80
60
40
20
0
WILD-TYPE
MUTANT
0
4
8
12
16
20
24
28
WEEKS RT

FIG. 4E

dUTP TOLERANCE

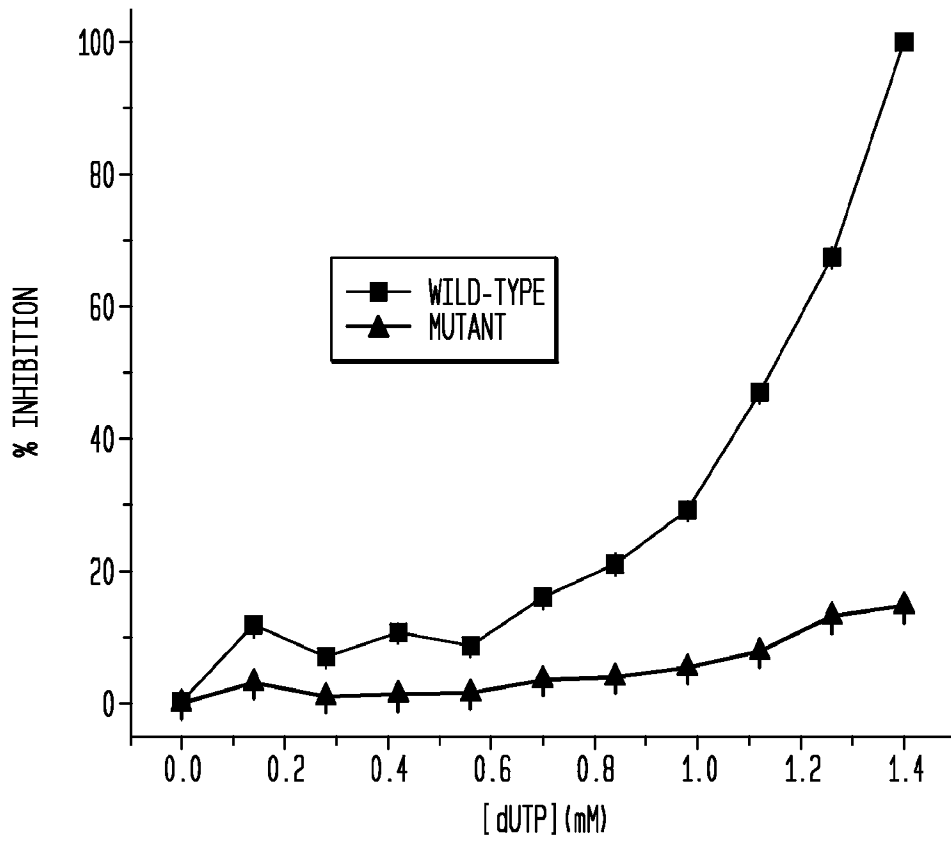


FIG. 5A

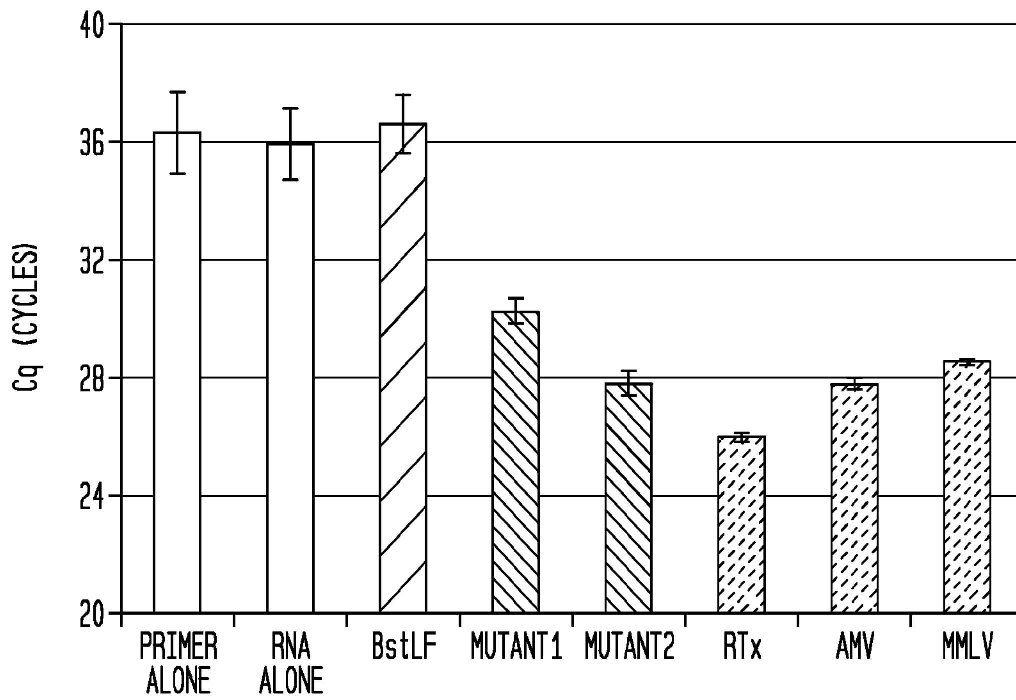
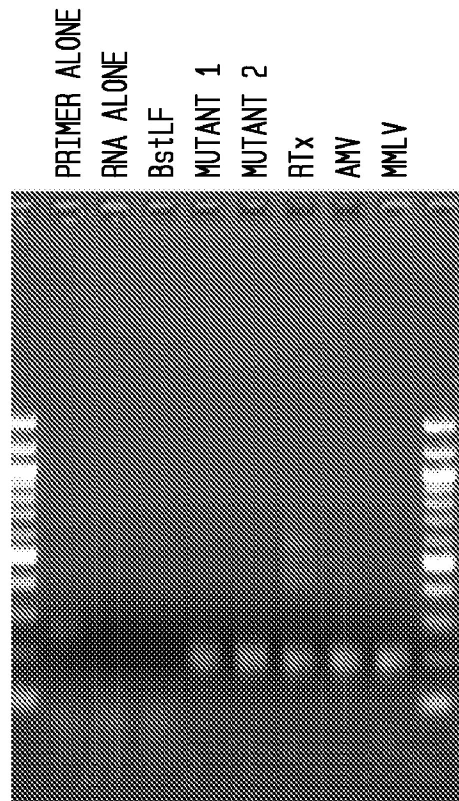


FIG. 5B



INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 12/53330

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - C12N 9/12, C12P 19/34 (2012.01)
USPC - 435/194
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)
IPC(8) - C12N 9/12, C12P 19/34 (2012.01)
USPC - 435/194

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
USPC - 435/91.2, 435/91.5, 435/193

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
PatBase, PubMed
Search terms: DNA polymerase, BST, stearothermophilus, mutant, mutation, variant

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2003/0180737 A1 (GU et al.) 25 September 2003 (25.09.2003) para [0024], [0139]; SEQ ID NOS: 20, 22	1-2
Y	FARASHAHI et al., Uniprot Direct Submission Accession No. A2TF40, 31 May 2011 [online]. [Retrieved on 23 January 2011]. Retrieved from the Internet: <URL: http://www.uniprot.org/uniprot/A2TF40.txt?version=33>	1-2
Y	CAGLAYAN et al., Uniprot Direct Submission Accession No. B6E9X1, 28 June 2011 [online]. [Retrieved on 23 January 2011]. Retrieved from the Internet: <URL: http://www.uniprot.org/uniprot/B6E9X1.txt?version=22>	1-2

☐ Further documents are listed in the continuation of Box C. ☐

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"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	Date of mailing of the international search report
23 January 2013 (23.01.2013)	07 FEB 2013

Name and mailing address of the ISA/US	Authorized officer:
Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	Lee W. Young PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 12/53330

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☒ Claims Nos.: 4-20, 24-27, 32-40, 43-45, 50
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

Group I: Claims 1-2, drawn to a variant Family A DNA polymerase comprising two amino acid sequence motifs, wherein the motifs are 3 .. EEK .. 5 and 15 .. ADE .. 17, wherein the two motifs confer improved reaction speed in an amplification reaction and/or improved stability compared to the reaction speed and/or stability of SEQ ID NO: 1.

---please see continuation on extra sheet---

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-2, limited to motifs 3 .. EEK .. 5 and 15 .. ADE .. 17

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 12/53330

Continuation of: Box No III Observations where unity of invention is lacking

Group II+: Claims 1-3, drawn to a variant Family A DNA polymerase comprising two or more amino acid sequence motifs, wherein the motifs are selected from 65 .. SPQ .. 67, 86 .. RAI..88, 185 .. LTE .. 187, 186 .. TEL .. 188, 222 .. LKE .. 224, 306 .. VHP .. 308, 314 ... HTR .. 316, 555 .. LCK .. 557, 556 .. CKL .. 558 and 567 .. VEL .. 569, wherein the two motifs confer improved reaction speed in an amplification reaction and/or improved stability compared to the reaction speed and/or stability of any of SEQ ID NO: 1-23. If Applicant elects to have this group searched, Applicant must specify the specific sequence motif and sequence to be searched. Each unique sequence motif within a sequence constitutes an inventive concept.

Group III+: Claims 21-23, drawn to variant protein, comprising: an amino acid sequence with at least 75% or 80% or 85% or 90% or 95% but less than 100% sequence identity to any of SEQ ID NOs: 1-23, wherein the variant protein further comprises at least one mutated amino acid. If Applicant elects to have this group searched, Applicant must specify the specific sequence to be searched. Each unique sequence constitutes an inventive concept.

Group IV: Claims 28-30 and 41-42, drawn to a non-naturally occurring synthetic protein, comprising: a fragment 1, a fragment 2, a fragment 3, a fragment 4, a fragment 5, a fragment 6, a fragment 7 and a fragment 8; and a method for synthesizing a protein which is capable of being generated from said protein fragments and assaying the synthetic protein for polymerase activity.

Group V+: Claim 31, drawn to a protein comprising at least 80% sequence identity with SEQ ID NO: 1 and further comprising one or more mutations. If Applicant elects to have this group searched, Applicant must specify the specific mutation to be searched. Each unique mutation constitutes an inventive concept.

Group VI: Claims 46-49, drawn to a DNA polymerase having one or more improved properties for isothermal amplification compared with SEQ ID NO: 1

The inventions listed as Groups I, II+, III+, IV, V+ and VI do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

The shared technical feature of the inventions listed as Groups I, II+, III+, IV, V+ and VI is a sequence having identity with SEQ ID NO:1, such as a sequence comprising the sequence of SEQ ID NO:1 with one or more mutations. This shared technical feature fails to provide a contribution over the prior art, as evidenced by US 2003/0180737 A1 to Gu et al. (hereinafter 'Gu'). Gu discloses a sequence comprising SEQ ID NO:1 (para [0024], Gu SEQ ID NO:20 exhibits 100% identity with SEQ ID NO:1), as well as a sequence comprising SEQ ID NO:1 comprising at least one mutation (Gu SEQ ID NO:22 exhibits 99.2% identity with SEQ ID NO:1). In the absence of a contribution over the prior art, the shared technical feature is not a shared special technical feature.

The shared technical feature of the inventions listed as Groups I and II+ is a variant Family A DNA polymerase comprising two or more amino acid sequence motifs as recited in claim 1, where the number preceding the amino acid in the motif corresponds to the location of that amino acid in the amino acid sequence of Figure 1, wherein the two or more motifs confer improved reaction speed in an amplification reaction and/or improved stability compared to the reaction speed and/or stability of any of SEQ ID NOs: 1-23. This shared technical feature fails to provide a contribution over the prior art, as evidenced by Gu in view of UniProtKB entry A2TF40 (A2TF40 UniProtKB, 06 March 2007 [online]. [Retrieved on 5 November 2012]. Retrieved from the internet: <URL: http://www.uniprot.org/uniprot/A2TF40.txt?version=34>; hereinafter 'A2TF40') and in view of UniProtKB entry B6E9X1 (B6E9X1_9BACI UniProt KB, 25 November 2008 [online]. [Retrieved on 5 November 2012]. Retrieved from the internet: <URL: http://www.uniprot.org/uniprot/B6E9X1.txt?version=22>; hereinafter 'B6E9X1').
Gu discloses a variant Family A DNA polymerase (para [0024] - "Bacillus stearothermophilus DNA polymerase", a member of Family A DNA polymerases) comprising mutations wherein the mutations confer improved reaction speed in an amplification reaction and/or improved stability (para [0139] - "mutant or variant forms of Tvu DNA polymerase. It is possible to modify the structure of a peptide having an activity (e.g. DNA synthesis activity) of Tvu DNA polymerase for such purposes as enhancing stability (e.g., in vitro shelf life, and/or resistance to proteolytic degradation in vivo) or reducing 5' to 3' exonuclease activity") compared to the reaction speed and/or stability of any of SEQ ID NO: 1 (para [0024], SEQ ID NO:20 exhibits 100% identity with SEQ ID NO:1). Gu does not disclose that the mutations comprise the sequence motifs recited in claim 1, however, sequence motifs such ase 3 ..EEK..5 and 567 ..VEL.. 569 in Family A DNA polymerase are known in the art, as evidenced by A2TF40 and B6E9X1. A2TF40 discloses a variant Family A DNA polymerase (A2TF40 sequence "Belongs to the DNA polymerase type-A family" and exhibits 89.1% identity with SEQ ID NO:1) comprising the amino acid sequence motif 3 .. EEK .. 5 where the number preceding the amino acid in the motif corresponds to the location of that amino acid in the amino acid sequence of Figure 1 (A2TF40 sequence amino acids 292 to 294 represent the motif EEK). B6E9X1 discloses a variant Family A DNA polymerase (B6E9X1 sequence "Belongs to the DNA polymerase type-A family" and exhibits 93.2% identity with SEQ ID NO:1) comprising the amino acid sequence motif 567 .. VEL .. 569 where the number preceding the amino acid in the motif corresponds to the location of that amino acid in the amino acid sequence of Figure 1 (B6E9X1 sequence amino acids 858 to 860 represent the motif VEL). Accordingly, it would have been obvious to one of ordinary skill in the art to modify the sequence of Gu to include the motifs disclosed by A2TF40 and B6E9X1 since Gu teaches that mutations incorporated into the sequence can improve polymerase stability. In the absence of a contribution over the prior art, the shared technical feature is not a shared special technical feature.

A further shared technical feature of each of the inventions listed as Group I and II+ is the specific amino acid sequence of SEQ ID NOs:1-23. Significant structural similarities cannot readily be ascertained among the unique amino acid sequences. Without significant structural similarities, the amino acid sequences do not have a shared special technical feature. In the absence of a shared special technical feature, the inventions lack unity with one another.

---please see continuation on next extra sheet---

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 12/53330

Continuation of: Previous extra sheet (Box No III Observations where unity of invention is lacking)

The shared technical feature of the inventions listed as Groups I, II+ and VI is a DNA polymerase having one or more improved properties for isothermal amplification compared with SEQ ID NO: 1. This shared technical feature fails to provide a contribution over Gu, as above, which disclsoes a DNA polymerase having one or more improved properties for isothermal amplification (para [0139] - "enhancing stability (e.g., in vitro shelf life, and/or resistance to proteolytic degradation in vivo) or reducing 5' to 3' exonuclease activity") compared with SEQ ID NO: 1 (para [0024], SEQ ID NO:20 exhibits 100% identity with SEQ ID NO:1). In the absence of a contribution over the prior art, the shared technical feature is not a shared special technical feature.

The shared technical feature of the inventions listed as Groups III+ is a variant protein, comprising: an amino acid sequence with at least 75% or 80% or 85% or 90% or 95% but less than 100% sequence identity to any of SEQ ID NOs: 1-23, wherein the variant protein further comprises at least one mutated amino acid having a position corresponding to SEQ ID NO: 1. This shared technical feature fails to provide a contribution over Gu which discloses a variant protein, comprising: an amino acid sequence with at least 75% or 80% or 85% or 90% or 95% but less than 100% sequence identity to SEQ ID NO: 1 (Gu SEQ ID NO:22 exhibits 99.2% identity with SEQ ID NO:1), wherein the variant protein further comprises at least one mutated amino acid having a position corresponding to SEQ ID NO: 1 (Gu SEQ ID NO:22 has three mutationed amino acids relative to SEQ ID NO:1). In the absence of a contribution over the prior art, the shared technical feature is not a shared special technical feature.

A further shared technical feature of each of the inventions listed as Group III+ is the specific amino acid sequence of SEQ ID NOs:1-23. Significant structural similarities cannot readily be ascertained among the unique amino acid sequences. Without significant structural similarities, the amino acid sequences do not have a shared special technical feature. In the absence of a shared special technical feature, the inventions lack unity with one another.

The shared technical feature of the inventions listed as Groups V+ is a protein comprising at least 80% sequence identity with SEQ ID NO: 1 and further comprising one or more mutations selected from the group consisting of: AIE, G3(K, E or D), K5(L, A or V), E8(M or A), E9D, M1DI, A13(D, E, T or V), I14D, V15A, V17(T, E or G), I18V, E20(M or D) M34(Q or L), E36D, 146F, L48(N or I), M57(L or I), P59(T or A), T61L, D655, 566(F, E or P), Q67A, L69(V or K), A73E, M81V, A84R, V88(A or I), R99V, A102D, NI13A, DI17(T, 5 or A), AI18D, GI19(D or E), I121(A or V), V124K, E131H, 5135(E or P), V144A, 5147(P or A), L148(D or V), Q152(L or P), T153(A or V), Q170(R or E), M173(L or I), D175E, N178(E, K or R), Q183(L, E or R), L185F, T186(L or I), K187(E or D), Q190(L or M), A193(I or S), A194(L, S or T), N205(D or K), S216(L or E), R223(V, K or G), A224E, 1225(Q or V), V247L, R307H, M316R, A330T, D357L, D378N, D380E, I383A, Q387R, L390M, 1400V, E406D, A4105, N411R, A4335, N437G, T439K, A452E, Q459(R or E), N463(V, E or D), L484D, D486E, V494L, T501M, Q530R, 1552M, E557(K, Q or R) and T568(E or R). This shared technical feature fails to provide a contribution over the prior art, as evidenced by A2TF40, which discloses a protein comprising at least 80% sequence identity with SEQ ID NO: 1 (A2TF40 sequence exhibits 89.1% identity with SEQ ID NO:1) and further comprising one or more mutations selected from the group consisting of: G3(K, E or D) (A2TF40 sequence amino acid position 292 represents the mutation from G to E at position 3 of SEQ ID NO:1).

Further, the special technical feature of the inventions listed as Group IV is a synthetic protein comprising eight covalently linked fragments. This special technical feature is not shared by the inventions of Groups I, II+, III+, V+, and VI. The special technical feature of the inventions listed as Group VI is a DNA polymerase having one or more improved properties for isothermal amplification compared with SEQ ID NO: 1. This special technical feature is not shared by the inventions of Groups III+, IV, V+ and VI.

Therefore, the inventions of Groups I, II+, III+, IV, V+ and VI lack unity of invention under PCT Rule 13 because they do not share a same or corresponding special technical feature.