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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
3 June 2010 (03.06.2010)

PCT

(10) International Publication Number
WO 2010/062776 A2

(51) International Patent Classification:
C12N 9/12 (2006.01) *C12N 15/54* (2006.01)

(21) International Application Number:
PCT/US2009/063166

(22) International Filing Date:
3 November 2009 (03.11.2009)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/110,862 3 November 2008 (03.11.2008) US

(71) Applicant *(for all designated States except US)*: **KAPABIOSYSTEMS** [US/US]; 600 West Cummings Park, Suite 5350, Woburn, MA 01801 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **FAURHOLM, Bjarne** [DK/ZA]; 9 Mulvihal Road, Rondebosch, Western Cape 7700 (ZA). **MCEWAN, Paul** [GB/ZA]; 3 Sonnekus, Victoria Road, Camps Bay, Western Cape 8005 (ZA). **BOURN, William** [GB/ZA]; 59 Severn Road, Plumstead, Western Cape 7801 (ZA). **RUSH, Gavin** [ZA/ZA]; 48 Liesbeek Road, Rosebank, Western Cape 7700 (ZA).

(74) Agent: **CHEN, Fangli**; Choate, Hall & Stewart LLP, Two International Place, Boston, MA 02110 (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, 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, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, 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, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (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, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: CHIMERIC DNA POLYMERASES

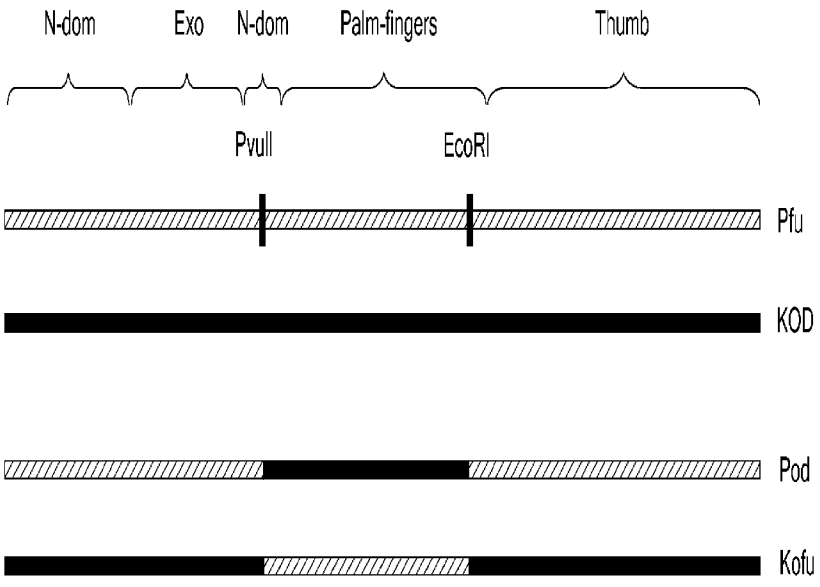


FIG. 2

(57) Abstract: The present invention provides, among other things, chimeric DNA polymerases containing heterologous domains having sequences derived from at least two DNA polymerases that have at least one distinct functional characteristics (e.g., elongation rate, processivity, error rate or fidelity, salt tolerance or resistance) and methods of making and using the same. In some embodiments, the present invention can combine desired functional characteristics (e.g., high processivity; high elongation rate; thermostability; resistance to salt, PCR additives (e.g., PCR enhancers) and other impurities; and high fidelity) of different DNA polymerases in a chimeric polymerase.

WO 2010/062776 A2

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

— *with sequence listing part of description (Rule 5.2(a))*

CHIMERIC DNA POLYMERASES

[0001] The present application claims priority to United States Provisional patent application serial number 61/110,862, filed on November 3, 2008, the entire disclosure of which is incorporated herein by reference.

BACKGROUND OF THE INVENTION

[0002] DNA polymerases are enzymes that use single-stranded DNA as a template to synthesize the complementary DNA strand. In particular, DNA polymerases can add free nucleotides to the 3' end of a newly-forming strand resulting in elongation of the new strand in a 5'-3' direction. Some DNA polymerases can correct mistakes in newly-synthesized DNA. This process is known as error correction. These polymerases can recognize an incorrectly incorporated nucleotide and the 3'->5' exonuclease activity of the enzyme allows the incorrect nucleotide to be excised (this activity is known as proofreading). Following base excision, the polymerase can re-insert the correct base and replication can continue. The proofreading function gives the DNA replication much higher fidelity than it would have if synthesis were the result of only a base-pairing selection step. Brutlag, D. and Kornberg, A., *J. Biol. Chem.*, 247:241-248 (1972). DNA polymerases with 3'-5' proofreading exonuclease activity have a substantially lower error rate when compared with a non-proofreading exonuclease-possessing polymerase. Chang, L. M. S., *J. Biol. Chem.*, 252:1873-1880 (1977). However, sometimes, the advantage of these polymerases is offset by its relatively low processivity that reduces the yield of DNA amplification products.

SUMMARY OF THE INVENTION

[0003] The present invention encompasses the discovery that domain swapping can combine desired functional characteristics (e.g., high processivity, high elongation rate, thermostability, resistance to salt, PCR additives (e.g., PCR enhancers) and other impurities, and high fidelity) of different DNA polymerases in a chimeric enzyme. Thus, the present invention provides, among other things, robust, fast and accurate DNA polymerases for DNA

amplification, synthesis, detection, sequencing and other important recombinant DNA techniques.

[0004] In one aspect, the present invention provides chimeric polymerases containing a first domain having a sequence at least 80% (e.g., at least 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%) identical to an amino acid sequence found in a first DNA polymerase characterized with high processivity, elongation rate, salt resistance, thermostability or TMAC tolerance; and a second domain having a sequence at least 80% (e.g., at least 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%) identical to an amino acid sequence found in a second DNA polymerase characterized with high fidelity, wherein the chimeric polymerases are characterized with both high fidelity and high processivity, elongation rate, or salt resistance. As used herein, the term "high processivity" refers to a processivity higher than 20nts (e.g., higher than 40nts, 60nts, 80nts, 100nts, 120nts, 140nts, 160nts, 180nts, 200nts, 220nts, 240nts, 260nts,280nts, 300nts, 320nts, 340nts, 360nts, 380nts, 400nts, or higher) per association/disassociation with the template. As used herein, the term "high elongation rate" refers to an elongation rate higher than 25 nt/s (e.g., higher than 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140 nt/s). As used herein, the term "high salt resistance" refers to the ability of a DNA polymerase to substantially maintain its enzymatic activity at a salt concentration higher than 30 mM (e.g., higher than 35 mM, 40mM, 45mM, or 50mM). As used herein, the term "high fidelity" refers to an error rate less than 4.45 X 10^{-6} (e.g., less than 4.0 X 10^{-6}, 3.5 X 10^{-6}, 3.0 X 10^{-6}, 2.5 X 10^{-6}, 2.0 X 10^{-6}, 1.5 X 10^{-6}, 1.0 X 10^{-6}, 0.5 X 10^{-6}) mutations/nt/doubling. As used herein, the term "high TMAC tolerance" refers to the ability of a DNA polymerase to substantially maintain its enzymatic activity at a TMAC (tetra-methyl ammonium chloride) concentration higher than 10 mM (e.g., higher than 15 mM, 20 mM, 25 mM, 30 mM). As used herein, the term "high thermostability" refers to the ability of a DNA polymerase to substantially maintain its enzymatic activity after more than 30 minutes incubation at 98 ^{o}C (e.g., 45 min, 60 min, 90min, 180 min, 210 min, 240 min). The terms of "processivity," "elongation rate," "fidelity," "salt resistance," "TMAC tolerance," and "thermostability" are further defined in the Definitions section.

[0005] In some embodiments, exemplary first DNA polymerases suitable for the present invention include, but are not limited to, KOD polymerase, TNA1 polymerase, *Thermococcus* sp. 9 degrees N-7, T4, T7, or phi29. In some embodiments, the first DNA

polymerase is KOD polymerase. In some embodiments, exemplary second DNA polymerases suitable for the invention include, but are not limited to, polymerases isolated from *Pyrococcus furiosus*, *P. abyssi*, *T. gorgonarius*, *T. litoralis, T. zilligii, T.* sp. GT, *or P.* sp. GB-D. In some embodiments, the second DNA polymerase is Pfu polymerase. In particular embodiments, the first DNA polymerase is KOD polymerase and the second DNA polymerase is Pfu polymerase.

[0006] In some embodiments, suitable first domain is an exonuclease domain, N-terminal domain, and/or a thumb domain. In some embodiments, suitable second domain is palm and/or fingers domain.

[0007] In some embodiments, amino acid sequences found in the first DNA polymerase correspond to amino acid residues 26 to 105 of KOD polymerase (SEQ ID NO:11), amino acid residues 156 to 301 of KOD polymerase (SEQ ID NO:11), and/or amino acid residues 612 to 749 of KOD polymerase (SEQ ID NO:11).

[0008] In some embodiments, amino acid sequences found in the second DNA polymerase correspond to amino acid residues 394 to 563 of Pfu polymerase (SEQ ID NO:9).

[0009] In some embodiments, a chimeric polymerase in accordance with the present invention include a first domain having a consensus sequence selected from the group consisting of

XXLXXXXXXXEGXRXXXXXXVXXXXXDXXXTXXXXXXXXXXXVVKXXXXXVLIX XXXXNXXXAXXKXXCXXXXXNFALXXXXXXXXXXXXXIXXMXXRFXXXXXXXXX XXXXPXXRXXXXXXXXXXXXXXXXXVXXQXXXXXXXEXXTTXXXT (SEQ ID NO:30), wherein X is any amino acid or a peptide bond;

XXEXXXXYXXXXEXXFXXXXKXXXAXXXXXXXXXAXXXXTVXTVKRXXXXQXXX XXRXVEXXXXXFTXXXXXXAXXDXIXXXXX (SEQ ID NO:31), wherein X is any amino acid or a peptide bond;

XXXXXXXXXXXXXXXXALXXDXXXXKXXXXXXXXTEXXSKXXVXXXXXVXHX XXXXDXKDXXXTXXXXXXXXRXXXRXXXXRXXTXXSXXXXKXSXRXGDXXXPF DXFXXTXXXXXXXXXXXXXXXXXXEXXXRAXX (SEQ ID NO:32), wherein X is any amino acid or a peptide bond;

NGX$_1$FKIEX$_2$DRTFX$_3$PYX$_4$YALLX$_5$DDSX$_6$IEEVKKITX$_7$ERHGX$_8$X$_9$VX$_{10}$X$_{11}$X$_{12}$X$_{13}$VEK VX$_{14}$KKFLGX$_{15}$PX$_{16}$X$_{17}$VWKLYX$_{18}$X$_{19}$HPQDVPX$_{20}$IRX$_{21}$KX$_{22}$REHPA (SEQ ID NO:33), wherein X$_1$ is not K; X$_2$ is not H; X$_3$ is not R; X$_4$ is not I; X$_5$ is not R; X$_6$ is not K; X$_7$ is not G; X$_8$ is not K; X$_9$ is not I; X$_{10}$ is not R; X$_{11}$ is not I; X$_{12}$ is not V; X$_{13}$ is not D; X$_{14}$ is not E; X$_{15}$ is not K; X$_{16}$ is not I; X$_{17}$ is not T; X$_{18}$ is not L; X$_{19}$ is not E; X$_{20}$ is not T; X$_{21}$ is not E; and X$_{22}$ is not V;

PIX$_1$MISYADEX$_2$X$_3$AX$_4$VITWKNX$_5$DLPYVX$_6$VVSX$_7$EREMIKRFLRX$_8$X$_9$X$_{10}$EKDPDX$_{11}$ X$_{12}$X$_{13}$TYNGDX$_{14}$FDFX$_{15}$YLX$_{16}$KRX$_{17}$EKLGIX$_{18}$X$_{19}$X$_{20}$X$_{21}$GRDGSEPKX$_{22}$QRX$_{23}$GDX$_{24}$ X$_{25}$AVEVKGRIHFDLYX$_{26}$VIX$_{27}$RTINLPTYTLEAVYEAX$_{28}$FGX$_{29}$PKEKVYAX$_{30}$EIX$_{31}$X$_3$ $_2$AWEX$_{33}$ (SEQ ID NO:34), wherein X$_1$ is not I; X$_2$ is not N; X$_3$ is not E; X$_4$ is not K; X$_5$ is not I; X$_6$ is not E; X$_7$ is not S; X$_8$ is not I; X$_9$ is not I; X$_{10}$ is not R; X$_{11}$ is not I; X$_{12}$ is not I; X$_{13}$ is not V; X$_{14}$ is not S; X$_{15}$ is not P; X$_{16}$ is not A; X$_{17}$ is not A; X$_{18}$ is not K; X$_{19}$ is not L; X$_{20}$ is not T; X$_{21}$ is not I; X$_{22}$ is not M; X$_{23}$ is not I; X$_{24}$ is not M; X$_{25}$ is not T; X$_{26}$ is not H; X$_{27}$ is not T; X$_{28}$ is not I; X$_{29}$ is not K; X$_{30}$ is not D; X$_{31}$ is not A; X$_{32}$ is not K; and X$_{33}$ is not S;

RDWSEIAKETQARVLEX$_1$X$_2$LKX$_3$GDVEX$_4$AVRIVKEVX$_5$X$_6$KLX$_7$X$_8$YEX$_9$PPEKLX$_{10}$IX $_{11}$EQITRX$_{12}$LX$_{13}$X$_{14}$YKAX$_{15}$GPHVAVAKX$_{16}$LAAX$_{17}$GVKIX$_{18}$PGX$_{19}$VIX$_{20}$YIVLX$_{21}$GX$_{22}$ GX$_{23}$IX$_{24}$X$_{25}$RAIX$_{26}$X$_{27}$X$_{28}$EX$_{29}$DPX$_{30}$KHKYDAEYYIENQVLPAVX$_{31}$RILX$_{32}$X$_{33}$FG (SEQ ID NO:35), wherein X$_1$ is not T; X$_2$ is not I; X$_3$ is not H; X$_4$ is not E; X$_5$ is not I; X$_6$ is not Q; X$_7$ is not A; X$_8$ is not N; X$_9$ is not I; X$_{10}$ is not A; X$_{11}$ is not Y; X$_{12}$ is not P; X$_{13}$ is not H; X$_{14}$ is not E; X$_{15}$ is not I; X$_{16}$ is not K; X$_{17}$ is not K; X$_{18}$ is not K; X$_{19}$ is not M; X$_{20}$ is not G; X$_{21}$ is not R; X$_{22}$ is not D; X$_{23}$ is not P; X$_{24}$ is not S; X$_{25}$ is not N; X$_{26}$ is not L; X$_{27}$ is not A; X$_{28}$ is not E; X$_{29}$ is not Y; X$_{30}$ is not K; X$_{31}$ is not L; X$_{32}$ is not E; and X$_{33}$ is not G;

and combinations thereof;

and a second domain having a consensus sequence selected from the group consisting of

XKXXXXXXXXXXXXAXXXXXXXXXXXXXXXXXXLXXXXNXXIXXXXXXKXXXXI XXXXXXXXXHXXXXXXXXXTXXXEXQXXXXKIXXXXXXKXXXLXXXXFXXXXX XXKXXXXXXXXXXXXXXXXXKXXELVWXXLXXXFXXXXLXIXXXXLYXXXXXG ESXEIXXXXLX (SEQ ID NO:36), wherein X is any amino acid or a peptide bond;

EX$_1$GLWENIVYLDFRX$_2$LYPSIIITHNVSPDTLNX$_3$EGCKX$_4$YDX$_5$APQVGHX$_6$FCKDX$_7$P GFIPSLLGX$_8$LLEERQKIKX$_9$KMKX$_{10}$TX$_{11}$DPIEX$_{12}$X$_{13}$LLDYRQX$_{14}$AIKX$_{15}$LANSX$_{16}$YG

YYGYAX$_{17}$ARWYCKECAESVTAWGRX$_{18}$YIX$_{19}$X$_{20}$X$_{21}$X$_{22}$KEX$_{23}$EEKX$_{24}$GFKVX$_{25}$YX$_{26}$ DTDGX$_{27}$X$_{28}$ATIPGX$_{29}$X$_{30}$X$_{31}$EX$_{32}$X$_{33}$KKKAX$_{34}$E (SEQ ID NO:37), wherein X$_1$ is not R; X$_2$ is not S; X$_3$ is not R; X$_4$ is not E; X$_5$ is not V; X$_6$ is not R; X$_7$ is not F; X$_8$ is not D; X$_9$ is not K; X$_{10}$ is not A; X$_{11}$ is not I; X$_{12}$ is not R; X$_{13}$ is not K; X$_{14}$ is not R; X$_{15}$ is not I; X$_{16}$ is not Y; X$_{17}$ is not R; X$_{18}$ is not E; X$_{19}$ is not T; X$_{20}$ is not M; X$_{21}$ is not T; X$_{22}$ is not I; X$_{23}$ is not I; X$_{24}$ is not Y; X$_{25}$ is not I; X$_{26}$ is not S; X$_{27}$ is not F; X$_{28}$ is not F; X$_{29}$ is not A; X$_{30}$ is not D; X$_{31}$ is not A; X$_{32}$ is not T; X$_{33}$ is not V; X$_{34}$ is not M,

and combinations thereof,

wherein the chimeric polymerase is characterized with high fidelity and high processivity, elongation rate, salt resistance, TMAC or other PCR enhancer tolerance or thermostability.

[0010] In some embodiments, chimeric polymerases in accordance with the present invention are defined by consensus sequence
XXXXTXXXXXDXXXXXXIXXXXXXEXXXXYXXXXEXXFXXXXKXXXAXXXXXX
XXAXXXXTVXTVKRXXXXQXXXXXRXVEXXXXXFTXXXXXXAXXDXIXXXXXXI
XXYXXXXXXXXXXXXXXXXVXXXXDXXXXMXXXXXXXXXXXXXXXXAEXXXLX
XXXXXXEGXRXXXXXXVXXXXXDXXXTXXXXXXXXXXVVKXXXXXVLIXXXXX
NXXXAXXKXXCXXXXXNFALXXXXXXXXXXIXXMXXRFXXXXXXXXXXXXXPX
XRXXXXXXXXXXXXXXXXVXXQXXXXXXXEXXTTXXXTXXXXXXXXRXXXXX
XXVXXXXXXXXXXXXAXXXXXVXXPXXXXXXXXXXXXXXXXXXXXXXXXXV
XXXXXSXEXYQXXXXEXXTXXFXXXXXKXXXXXXXXXXXXAXXXXXXXXXXX
XXXXXLXXXXNXXIXXXXXXKXXXXIXXXXXXXXXHXXXXXXXXTXXXEXQX
XXXKIXXXXXXKXXXLXXXXFXXXXXXXKXXXXXXXXXXXXXXXXXKXXELVW
XXLXXXFXXXXLXIXXXXLYXXXXXGESXEIXXXXLXXLXXXXAXXXXAXXXXX
XXXXXXXXXXXXXKXXXXXXXXXITXXXXXXXXXXXXXXXXXXXXXXXALX
XDXXXXKXXXXXXXXTEXXSKXXVXXXXXVXHXXXXXDXKDXXXTXXXXXXX
XRXXXRXXXXRXXTXXSXXXXKXSXRXGDXXXPFDXFXXTXXXXXXXXXXXXX
XXXXXEXXXRAXXXXXXXXXXXXXXXXXXSAXXKPXGT (SEQ ID NO:38),
wherein X is any amino acid or a peptide bond, and wherein the chimeric polymerase has a fidelity higher than that of KOD and a processivity, an elongation rate, a salt resistance, a TMAC or other PCR enhancer tolerance or a thermostability higher than that of Pfu.

[0011] In some embodiments, chimeric polymerases in accordance with the present invention are defined by consensus sequence XIXDTDYXTXDGXPXXRIFXKXXGEFXXXYDXXFEPYFYALLKDDSAIXXXXXXXA XRHGTVXTVKRXXXXQXKFLXRXVEVWXLXFTHPQDVPAXXDXIXXHXXVIDIYE YDIPFAKRYLIDXGLVPMEGDEXLXMXXXDIETXYHEGXEFAEGXXLMISYADXEG ARVITWKXVDLPYVDVVSTEXEMIKRXXXVVKEKDPDVLIXYXGDNFDXAYLKXR CEXLGXNFALXRXXXXXEPKIXXMGXRFAVEXKGRXHFDLXPXXRXTXNLPTYXL XXVYEXVXGQXKXKXXXEEITTXWETXXXXXXXARYSMEDAXVTXELGXEFXPM EAXLXXLVGXPXWDVXRSSTGNLVEWXLLXXAYXRNEVAPNKPSXEEYQXRXXE XYTGXFVXEPEKGLWXXXXXLDXXALYPSIIXXHNVSPDTLXLEXCXNYDIAPXVG XKFCKDIPGFIPSXLXHLXXXRQXXKTXMXEXQDPXEKIXLDYRQKAXKLLXNSFY GYXGYXKARWYXXECAESVTXWGRKYIELVWXELEXXFGFKXLYIDTDGLYATIP GGESXEIKXXXLXFLXYINAXLPGALELEYEXFYXRGFFVXKKKYAXIDEEXXITTR GLEXVRRDWSXXAKETXAXVLEALLXDXXVXKAVXXVXXXTEXXSKYXVPXEKL VIHEQITRDXKDYXATGPHVAXAKRLXXRGXXXRPGTXISYXXLKGSGRXGDRXIPF DEFXXTKHXYDXXYYIENQVLPAVERXLRAFGYXXXXLXXQXXXQXGLSAWXKP XGT (SEQ ID NO:39), wherein X is any amino acid or a peptide bond.

[0012] In some embodiments, the present invention further provides chimeric polymerases containing a first domain having a sequence at least 80% (e.g., at least 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%) identical to an amino acid sequence found in an exonuclease domain, an N-terminal domain, and/or a thumb domain of a first DNA polymerase; and a second domain having a sequence at least 80% (e.g., at least 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%) identical to an amino acid sequence found in palm and/or fingers domain of a second DNA polymerase. In some embodiments, the chimeric polymerase has a fidelity higher than that of the second DNA polymerase and a processivity, an elongation rate, a salt resistance, a TMAC or other PCR enhancer tolerance or a thermostability higher than that of the first DNA polymerase.

[0013] In another aspect, the present invention provides methods of engineering chimeric polymerases. Inventive methods in accordance with the present invention include steps of: (a) providing an N-terminal domain, an exonuclease domain, and/or a thumb domain based on a first DNA polymerase; (b) providing a palm and/or fingers domain based on a second DNA polymerase; (c) combining the domains from step (a) and step (b) to form a

chimeric polymerase; wherein the chimeric polymerase has a fidelity higher than that of the first DNA polymerase and a processivity, an elongation rate, a salt resistance, a TMAC or other PCR enhancer tolerance or a thermostability higher than that of the second DNA polymerase. In some embodiments, a chimeric polymerase engineered according to the present invention has a processivity, an elongation rate, a salt resistance, a TMAC or other PCR enhancer tolerance or a thermostability substantially similar to that of the first DNA polymerase and a fidelity substantially similar to that of the second DNA polymerase.

[0014] In some embodiments, exemplary first DNA polymerases suitable for the present invention include, but are not limited to, KOD polymerase, TNA1 polymerase, *Thermococcus* sp. 9 degrees N-7, T4, T7, or phi29. In some embodiments, the first DNA polymerase is KOD polymerase. In some embodiments, exemplary second DNA polymerases suitable for the invention include, but are not limited to, polymerases isolated from *Pyrococcus furiosus*, *P. abyssi*, *T. gorgonarius*, *T. litoralis, T. zilligii, T.* sp. GT, *or P.* sp. GB-D. In some embodiments, the second DNA polymerase is Pfu polymerase.

[0015] In some embodiments, the first DNA polymerase is KOD polymerase and the second DNA polymerase is Pfu polymerase. In some embodiments, the first DNA polymerase is Pfu polymerase and the second DNA polymerase is KOD polymerase.

[0016] In some embodiments, the present invention provides methods of improving the fidelity of a DNA polymerase. In particular embodiments, inventive methods in accordance with the invention include a step of replacing a sequence within the palm and/or fingers domain of the DNA polymerase of interest with a corresponding sequence from a different DNA polymerase that is characterized with higher fidelity relative to the DNA polymerase of interest.

[0017] In some embodiments, the present invention provides methods of improving the processivity, elongation rate, salt resistance, TMAC or other PCR enhancer tolerance or thermostability of a DNA polymerase. In particular embodiments, inventive methods in accordance with the present invention include a step of replacing a sequence within the N-terminal domain, the exonuclease domain and/or the thumb domain of the DNA polymerase of interest with a corresponding sequence from a different DNA polymerase that is characterized with higher processivity, elongation rate, salt resistance, TMAC or other PCR enhancer tolerance or thermostability relative to the DNA polymerase of interest.

[0018] The present invention provides various chimeric polymerases described herein including chimeric polymerases engineered and/or improved using inventive methods as described herein. In some embodiments, chimeric polymerases in accordance with the present invention contain an amino acid sequence at least 80% identical to SEQ ID NO:16 (the Kofu amino acid sequence as shown in the Sequences section). In particular embodiments, a chimeric polymerase in accordance with the present invention contains the amino acid sequence of SEQ ID NO:16. In some embodiments, chimeric polymerases in accordance with the present invention contain an amino acid sequence at least 80% identical to SEQ ID NO:15 (the Pod amino acid sequence as shown in the Sequences section). In particular embodiments, a chimeric polymerase in accordance with the present invention contains the amino acid sequence of SEQ ID NO:15.

[0019] The present invention also provides kits and compositions containing various chimeric polymerases described herein and uses thereof (e.g., methods of amplifying DNA fragments using chimeric DNA polymerases of the invention). In addition, the present invention provides nucleotide sequences encoding various chimeric polymerases described herein and vectors and/or cells containing the nucleotide sequences according to the invention.

BRIEF DESCRIPTION OF THE DRAWINGS

[0020] The drawings are for illustration purposes only, not for limitation.

[0021] Figure 1 depicts an alignment of domains in exemplary naturally-occurring type B DNA polymerases and exemplary chimeric DNA polymerases, Kofu and Pod. The KOD and Pfu polymerase domains that were swapped in the Kofu and Pod chimeras are indicated above the alignment.

[0022] Figure 2 depicts that an exemplary chimeric polymerase Pod contains the N-terminal domain, the 3'-5' exonuclease domain and the thumb domain of Pfu and the palm and fingers domain of KOD and the reciprocal chimeric polymerase Kofu contains the N-terminal domain, the 3'-5' exonuclease domain and the thumb domain of KOD and the palm and fingers domain of Pfu.

[0023] Figure 3 depicts exemplary results showing the thermostability of KOD, Pfu, Kofu and Pod.

[0024] Figure 4 depicts exemplary results showing the salt resistance of KOD, Pfu, Kofu and Pod.

[0025] Figure 5 depicts exemplary results showing the TMAC tolerance of KOD, Pfu, Kofu and Pod.

DEFINITIONS

[0026] *Amino acid*: As used herein, term "amino acid," in its broadest sense, refers to any compound and/or substance that can be incorporated into a polypeptide chain. In some embodiments, an amino acid has the general structure H_2N–C(H)(R)–COOH. In some embodiments, an amino acid is a naturally-occurring amino acid. In some embodiments, an amino acid is a synthetic amino acid; in some embodiments, an amino acid is a D-amino acid; in some embodiments, an amino acid is an L-amino acid. "Standard amino acid" refers to any of the twenty standard L-amino acids commonly found in naturally occurring peptides. "Nonstandard amino acid" refers to any amino acid, other than the standard amino acids, regardless of whether it is prepared synthetically or obtained from a natural source. As used herein, "synthetic amino acid" encompasses chemically modified amino acids, including but not limited to salts, amino acid derivatives (such as amides), and/or substitutions. Amino acids, including carboxy- and/or amino-terminal amino acids in peptides, can be modified by methylation, amidation, acetylation, and/or substitution with other chemical groups. Amino acids may participate in a disulfide bond. The term "amino acid" is used interchangeably with "amino acid residue," and may refer to a free amino acid and/or to an amino acid residue of a peptide. It will be apparent from the context in which the term is used whether it refers to a free amino acid or a residue of a peptide. It should be noted that all amino acid residue sequences are represented herein by formulae whose left and right orientation is in the conventional direction of amino-terminus to carboxy-terminus.

[0027] *Base Pair (bp)*: As used herein, base pair refers to a partnership of adenine (A) with thymine (T), or of cytosine (C) with guanine (G) in a double stranded DNA molecule.

[0028] *Chimeric polymerase*: As used herein, the term "chimeric polymerase" (also referred to as "chimera") refers to any polymerase containing two or more heterologous domains, amino acid sequences, peptides, and/or proteins joined either covalently or non-covalently to produce a polymerase that does not occur in nature. Typically, a chimeric polymerase contains a first domain joined to a second domain, wherein the first and second domains are not found in the same relationship in nature. Typically, the first domain is derived from a first DNA polymerase and a second domain is derived from a second DNA polymerase. Typically, the first and second DNA polymerases are characterized with at least one distinct functional characteristics (e.g., processivity, elongation rate, fidelity, salt tolerance, tolerance to PCR additives or thermostability). As used herein, a sequence derived from a DNA polymerase of interest refers to any sequence found in the DNA polymerase of interest, or any sequence having at least 70% (e.g., at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%) identical to an amino acid sequence found in the DNA polymerase of interest. A "chimeric polymerase" according to the invention may contain two or more amino acid sequences from related or similar polymerases (e.g., proteins sharing similar sequences and/or structures), joined to form a new functional protein. A "chimeric polymerase" according to the invention may contain two or more amino acid sequences from unrelated polymerases, joined to form a new functional protein. For example, a chimeric polymerase of the invention may be an "interspecies" or "intergenic" fusion of protein structures expressed by different kinds of organisms.

[0029] *Complementary*: As used herein, the term "complementary" refers to the broad concept of sequence complementarity between regions of two polynucleotide strands or between two nucleotides through base-pairing. It is known that an adenine nucleotide is capable of forming specific hydrogen bonds ("base pairing") with a nucleotide which is thymine or uracil. Similarly, it is known that a cytosine nucleotide is capable of base pairing with a guanine nucleotide.

[0030] *DNA binding affinity:* As used herein, the term "DNA-binding affinity" typically refers to the activity of a DNA polymerase in binding DNA nucleic acid. In some embodiments, DNA binding activity can be measured in a two band-shift assay. For example, in some embodiments (based on the assay of Guagliardi *et al.* (1997) *J. Mol. Biol.* 267:841-848), double-stranded nucleic acid (the 452-bp HindIII-EcoRV fragment from the *S. solfataricus lacS* gene) is labeled with ^{32}P to a specific activity of at least about 2.5 X 10^7

cpm/μg (or at least about 4000 cpm/fmol) using standard methods. See, e.g., Sambrook *et al.* (2001) Molecular Cloning: A Laboratory Manual ($_3^{rd}$ ed., Cold Spring Harbor Laboratory Press, NY) at 9.63-9.75 (describing end-labeling of nucleic acids). A reaction mixture is prepared containing at least about 0.5 μg of the polypeptide in about 10 μl of binding buffer (50 mM sodium phosphate buffer (pH 8.0), 10% glycerol, 25 mM KCl, 25 mM $MgCl_2$). The reaction mixture is heated to 37 °C. for 10 min. About 1 X 10^4 to 5 X 10^4 cpm (or about 0.5-2 ng) of the labeled double-stranded nucleic acid is added to the reaction mixture and incubated for an additional 10 min. The reaction mixture is loaded onto a native polyacrylamide gel in 0.5 X Tris-borate buffer. The reaction mixture is subjected to electrophoresis at room temperature. The gel is dried and subjected to autoradiography using standard methods. Any detectable decrease in the mobility of the labeled double-stranded nucleic acid indicates formation of a binding complex between the polypeptide and the double-stranded nucleic acid. Such nucleic acid binding activity may be quantified using standard densitometric methods to measure the amount of radioactivity in the binding complex relative to the total amount of radioactivity in the initial reaction mixture. Other methods of measuring DNA binding affinity are known in the art (see, e.g., Kong et al. (1993) *J. Biol. Chem.* 268(3):1965-1975).

[0031] *Domain*: As used herein, the term "Domain" as used herein refers to an amino acid sequence of a polypeptide (*e.g.*, polymerase) comprising one or more defined functions or properties.

[0032] *Elongation rate*: As used herein, the term "elongation rate" refers to the average speed at which a DNA polymerase extends a polymer chain. As used herein, a high elongation rate refers to an elongation rate higher than 25 nt/s (e.g., higher than 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140 nt/s).

[0033] *Enzyme activity*: As used herein, the term "enzyme activity" refers to the specificity and efficiency of a DNA polymerase. Enzyme activity of a DNA polymerase is also referred to as "polymerase activity," which typically refers to the activity of a DNA polymerase in catalyzing the template-directed synthesis of a polynucleotide. Enzyme activity of a polymerase can be measured using various techniques and methods known in the art. For example, serial dilutions of polymerase can be prepared in dilution buffer (e.g., 20 mM Tris.Cl, pH 8.0, 50 mM KCl, 0.5% NP 40, and 0.5% Tween-20). For each dilution, 5 μl can be removed and added to 45 μl of a reaction mixture containing 25 mM TAPS (pH 9.25),

50 mM KCl, 2 mM MgCl2, 0.2 mM dATP, 0.2 mM dGTP, 0.2 mM dTTP, 0.1 mM dCTP, 12.5 μg activated DNA, 100 μM [α-^{32}P]dCTP (0.05 μCi/nmol) and sterile deionized water. The reaction mixtures can be incubated at 37 °C. (or 74 °C. for thermostable DNA polymerases) for 10 minutes and then stopped by immediately cooling the reaction to 4 °C and adding 10 μl of ice-cold 60 mM EDTA. A 25 μl aliquot can be removed from each reaction mixture. Unincorporated radioactively labeled dCTP can be removed from each aliquot by gel filtration (Centri-Sep, Princeton Separations, Adelphia, N.J.). The column eluate can be mixed with scintillation fluid (1 ml). Radioactivity in the column eluate is quantified with a scintillation counter to determine the amount of product synthesized by the polymerase. One unit of polymerase activity can be defined as the amount of polymerase necessary to synthesize 10 nmole of product in 30 minutes (Lawyer et al. (1989) *J. Biol. Chem.* 264:6427-647). Other methods of measuring polymerase activity are known in the art (see, e.g., Sambrook et al. (2001) Molecular Cloning: A Laboratory Manual (3rd ed., Cold Spring Harbor Laboratory Press, NY)).

[0034] *Fidelity*: As used herein, the term "fidelity" refers to the accuracy of DNA polymerization by template-dependent DNA polymerase. The fidelity of a DNA polymerase is typically measured by the error rate (the frequency of incorporating an inaccurate nucleotide, i.e., a nucleotide that is not complementary to the template nucleotide). The accuracy or fidelity of DNA polymerization is maintained by both the polymerase activity and the 3'-5' exonuclease activity of a DNA polymerase. The term "high fidelity" refers to an error rate less than 4.45 X 10^{-6} (e.g., less than 4.0 X 10^{-6}, 3.5 X 10^{-6}, 3.0 X 10^{-6}, 2.5 X 10^{-6}, 2.0 X 10^{-6}, 1.5 X 10^{-6}, 1.0 X 10^{-6}, 0.5 X 10^{-6}) mutations/nt/doubling. The fidelity or error rate of a DNA polymerase may be measured using assays known to the art. For example, the error rates of DNA polymerases can be tested using the *lacI* PCR fidelity assay described in Cline, J. et al. (1996) NAR 24: 3546-3551. Briefly, a 1.9 kb fragment encoding the *lacIOlacZα* target gene is amplified from pPRIAZ plasmid DNA using 2.5U DNA polymerase (i.e., amount of enzyme necessary to incorporate 25 nmoles of total dNTPs in 30 min. at 72°C) in the appropriate PCR buffer. The *lacI*-containing PCR products are then cloned into lambda GT10 arms, and the percentage of *lacI* mutants (MF, mutation frequency) is determined in a color screening assay, as described (Lundberg, K. S. , Shoemaker, D. D. , Adams, M. W. W. , Short, J. M. , Sorge, J. A., and Mathur, E. J. (1991) Gene 180: 1-8). Error rates are expressed as mutation frequency per bp per duplication (MF/bp/d), where bp is the number of detectable sites in the lacI gene sequence (349) and d is the number of

effective target doublings. Similar to the above, any plasmid containing the *lacIOlacZα* target gene can be used as template for the PCR. The PCR product may be cloned into a vector different from lambda GT (e.g., plasmid) that allows for blue/white color screening.

[0035] *Joined*: As used herein, "joined" refers to any method known in the art for functionally connecting polypeptide domains, including without limitation recombinant fusion with or without intervening domains, inter-mediated fusion, non-covalent association, and covalent bonding, including disulfide bonding, hydrogen bonding, electrostatic bonding, and conformational bonding.

[0036] *Nucleotide*: As used herein, a monomeric unit of DNA or RNA consisting of a sugar moiety (pentose), a phosphate, and a nitrogenous heterocyclic base. The base is linked to the sugar moiety via the glycosidic carbon (1' carbon of the pentose) and that combination of base and sugar is a nucleoside. When the nucleoside contains a phosphate group bonded to the 3' or 5' position of the pentose it is referred to as a nucleotide. A sequence of operatively linked nucleotides is typically referred to herein as a "base sequence" or "nucleotide sequence," and is represented herein by a formula whose left to right orientation is in the conventional direction of 5'-terminus to 3'-terminus.

[0037] *Oligonucleotide or Polynucleotide*: As used herein, the term "oligonucleotide" is defined as a molecule including two or more deoxyribonucleotides and/or ribonucleotides, preferably more than three. Its exact size will depend on many factors, which in turn depend on the ultimate function or use of the oligonucleotide. The oligonucleotide may be derived synthetically or by cloning. As used herein, the term "polynucleotide" refers to a polymer molecule composed of nucleotide monomers covalently bonded in a chain. DNA (deoxyribonucleic acid) and RNA (ribonucleic acid) are examples of polynucleotides.

[0038] *Polymerase*: As used herein, a "polymerase" refers to an enzyme that catalyzes the polymerization of nucleotide (i.e., the polymerase activity). Generally, the enzyme will initiate synthesis at the 3'- end of the primer annealed to a polynucleotide template sequence, and will proceed towards the 5' end of the template strand. A "DNA polymerase" catalyzes the polymerization of deoxynucleotides.

[0039] *Processivity*: As used herein, "processivity" refers to the ability of a polymerase to remain attached to the template and perform multiple modification reactions.

"Modification reactions" include but are not limited to polymerization, and exonucleolytic cleavage. In some embodiments, "processivity" refers to the ability of a DNA polymerase to perform a sequence of polymerization steps without intervening dissociation of the enzyme from the growing DNA chains. Typically, "processivity" of a DNA polymerase is measured by the length of nucleotides (for example 20 nts, 300 nts, 0.5-l kb, or more) that are polymerized or modified without intervening dissociation of the DNA polymerase from the growing DNA chain. "Processivity" can depend on the nature of the polymerase, the sequence of a DNA template, and reaction conditions, for example, salt concentration, temperature or the presence of specific proteins. As used herein, the term "high processivity" refers to a processivity higher than 20nts (e.g., higher than 40nts, 60nts, 80nts, 100nts, 120nts, 140nts, 160nts, 180nts, 200nts, 220nts, 240nts, 260nts,280nts, 300nts, 320nts, 340nts, 360nts, 380nts, 400nts, or higher) per association/disassociation with the template. Processivity can be measured according the methods defined herein and in WO 01/92501 Al.

[0040] *Primer*: As used herein, the term "primer" refers to an oligonucleotide, whether occurring naturally or produced synthetically, which is capable of acting as a point of initiation of nucleic acid synthesis when placed under conditions in which synthesis of a primer extension product which is complementary to a nucleic acid strand is induced, *e.g.*, in the presence of four different nucleotide triphosphates and thermostable enzyme in an appropriate buffer ("buffer" includes appropriate pH, ionic strength, cofactors, etc.) and at a suitable temperature. The primer is preferably single-stranded for maximum efficiency in amplification, but may alternatively be double-stranded. If double-stranded, the primer is first treated to separate its strands before being used to prepare extension products. Preferably, the primer is an oligodeoxyribonucleotide. The primer must be sufficiently long to prime the synthesis of extension products in the presence of the thermostable enzyme. The exact lengths of the primers will depend on many factors, including temperature, source of primer and use of the method. For example, depending on the complexity of the target sequence, the oligonucleotide primer typically contains 15-25 nucleotides, although it may contain more or few nucleotides. Short primer molecules generally require lower temperatures to form sufficiently stable hybrid complexes with template.

[0041] *Salt resistance*: As used herein, the term "salt resistance" (also referred to as salt tolerance) refers to the ability of a DNA polymerase to substantially maintain its enzymatic activity in the presence of salt or PCR additives (e.g., TMAC). In some

embodiments, resistance to salt or PCR additives is measured by the maximum salt concentration at which a DNA polymerase is still active. The maximum salt concentration differs for each polymerase and is known in the art, or can be experimentally determined according to methods in the art. For example, Pfu is inhibited at 30 mM salt (in a PCR reaction).

[0042] *Synthesis*: As used herein, the term "synthesis" refers to any *in vitro* method for making new strand of polynucleotide or elongating existing polynucleotide (i.e., DNA or RNA) in a template dependent manner. Synthesis, according to the invention, includes amplification, which increases the number of copies of a polynucleotide template sequence with the use of a polymerase. Polynucleotide synthesis (e.g., amplification) results in the incorporation of nucleotides into a polynucleotide (i.e., a primer), thereby forming a new polynucleotide molecule complementary to the polynucleotide template. The formed polynucleotide molecule and its template can be used as templates to synthesize additional polynucleotide molecules. "DNA synthesis," as used herein, includes, but is not limited to, PCR, the labeling of polynucleotide (i.e., for probes and oligonucleotide primers), polynucleotide sequencing.

[0043] *Template DNA molecule*: As used herein, the term "template DNA molecule" refers to a strand of a nucleic acid from which a complementary nucleic acid strand is synthesized by a DNA polymerase, for example, in a primer extension reaction.

[0044] *Template dependent manner*: As used herein, the term "template dependent manner" refers to a process that involves the template dependent extension of a primer molecule (e.g., DNA synthesis by DNA polymerase). The term "template dependent manner" typically refers to polynucleotide synthesis of RNA or DNA wherein the sequence of the newly synthesized strand of polynucleotide is dictated by the well-known rules of complementary base pairing (see, for example, Watson, J. D. et al., In: Molecular Biology of the Gene, 4th Ed. , W. A. Benjamin, Inc., Menlo Park, Calif. (1987)).

[0045] *Thermostable enzyme*: As used herein, the term "thermostable enzyme" refers to an enzyme which is stable to heat (also referred to as heat-resistant) and catalyzes (facilitates) polymerization of nucleotides to form primer extension products that are complementary to a polynucleotide template sequence. Typically, thermostable stable polymerases are preferred in a thermocycling process wherein double stranded nucleic acids

are denatured by exposure to a high temperature (e.g., about 95 C) during the PCR cycle. A thermostable enzyme described herein effective for a PCR amplification reaction satisfies at least one criteria, i.e., the enzyme do not become irreversibly denatured (inactivated) when subjected to the elevated temperatures for the time necessary to effect denaturation of double-stranded nucleic acids. Irreversible denaturation for purposes herein refers to permanent and complete loss of enzymatic activity. The heating conditions necessary for denaturation will depend, e.g., on the buffer salt concentration and the length and nucleotide composition of the nucleic acids being denatured, but typically range from about 90 °C to about 98 °C for a time depending mainly on the temperature and the nucleic acid length, typically about 0.2 to four minutes. Higher temperatures may be tolerated as the buffer salt concentration and/or GC composition of the nucleic acid is increased. In some embodiments, thermostable enzymes will not become irreversibly denatured at about 90 °C -100 °C. Typically, a thermostable enzyme suitable for the invention has an optimum temperature at which it functions that is higher than about 40 °C, which is the temperature below which hybridization of primer to template is promoted, although, depending on (1) magnesium and salt, concentrations and (2) composition and length of primer, hybridization can occur at higher temperature (e.g., 45 °C - 70 °C). The higher the temperature optimum for the enzyme, the greater the specificity and/or selectivity of the primer-directed extension process. However, enzymes that are active below 40 °C (e.g., at 37 °C) are also with the scope of this invention provided they are heat-stable. In some embodiments, the optimum temperature ranges from about 50 °C to 90 °C (e.g., 60 °C – 80 °C).

[0046] *TMAC or other PCR enhancer tolerance:* As used herein, the term "TMAC or other PCR enhancer tolerance" (also referred to as TMAC or other PCR enhancer resistance) refers to the ability of a DNA polymerase to substantially maintain its enzymatic activity in the presence of TMAC or other PCR enhancers (e.g., glycerol, DMSO, betaine, amides, other tetramethyl ammonium salts).

DETAILED DESCRIPTION OF THE INVENTION

[0047] The present invention provides, among other things, chimeric DNA polymerases containing heterologous domains having sequences derived from at least two DNA polymerases that have at least one distinct functional characteristics (e.g., elongation rate, processivity, error rate or fidelity, salt tolerance or resistance) and methods of making and using the same.

DNA polymerases

[0048] Chimeric DNA polymerases in accordance with the present invention may be engineered from any DNA polymerases, in particular, thermostable polymerases. Typically, DNA polymerases are grouped into six families: A, B, C, D, X and Y. Families A, B, C are grouped based on their amino acid sequence homologies to *E. coli* polymerases I, II, and III, respectively. Family X has no homologous *E. coli* polymerases. In some embodiments, DNA polymerases suitable for the present invention are family B DNA polymerases. Family B polymerases include, but are not limited to, *E. coli* pol II, archaeal polymerases, PRD1, phi29, M2, T4 bacteriophage DNA polymerases, eukaryotic polymerases α, Δ, ε, and many viral polymerases. In some embodiments, DNA polymerases suitable for the invention are archaeal polymerases (e.g., euryarchaeal polymerases).

[0049] Suitable exemplary archaeal polymerases include, but are not limited to, DNA polymerases from archaea (e.g., *Thermococcus litoralis* (Vent™, GenBank: AAA72101), *Pyrococcus furiosus* (Pfu, GenBank: D12983, BAA02362), *Pyrococcus woesii, Pyrococcus* GB-D (Deep Vent™, GenBank: AAA67131), *Thermococcus kodakaraensis KODI* (KOD, GenBank: BD175553, BAA06142; *Thermococcus* sp. strain KOD (Pfx, GenBank: AAE68738)), *Thermococcus gorgonarius* (*Tgo*, Pdb: 4699806), *Sulfolobus solataricus* (GenBank: NC002754, P26811), *Aeropyrum pernix* (GenBank: BAA81109), *Archaeglobus fulgidus* (GenBank: O29753), *Pyrobaculum aerophilum* (GenBank: AAL63952), *Pyrodictium occultum* (GenBank: BAA07579, BAA07580), *Thermococcus* 9 degree Nm (GenBank: AAA88769, Q56366), *Thermococcus fumicolans* (GenBank: CAA93738, P74918), *Thermococcus hydrothermalis* (GenBank: CAC18555), *Thermococcus* sp. GE8 (GenBank: CAC12850), *Thermococcus* sp. JDF-3 (GenBank: AX135456; WO0132887), *Thermococcus* sp. TY (GenBank: CAA73475), *Pyrococcus abyssi* (GenBank: P77916),

Pyrococcus glycovorans (GenBank: CAC12849), *Pyrococcus horikoshii* (GenBank: NP 143776), *Pyrococcus* sp. GE23 (GenBank: CAA90887), *Pyrococcus* sp. ST700 (GenBank: CAC12847), *Thermococcus pacificus* (GenBank: AX411312.1), *Thermococcus zilligii* (GenBank: DQ3366890), *Thermococcus aggregans*, *Thermococcus barossii*, *Thermococcus celer* (GenBank: DD259850.1), *Thermococcus profundus* (GenBank: E14137), *Thermococcus siculi* (GenBank: DD259857.1), *Thermococcus thioreducens*, *Thermococcus onnurineus* NA1, *Sulfolobus acidocaldarium*, *Sulfolobus tokodaii*, *Pyrobaculum calidifontis*, *Pyrobaculum islandicum* (GenBank: AAF27815), *Methanococcus jannaschii* (GenBank: Q58295), *Desulforococcus species TOK*, *Desulfurococcus, Pyrolobus, Pyrodictium, Staphylothermus, Vulcanisaetta, Methanococcus* (GenBank: P52025) and other archaeal B polymerases, such as GenBank AAC62712, P956901, BAAA07579)). Additional representative temperature-stable family A and B polymerases include, e.g., polymerases extracted from the thermophilic bacteria *Thermus* species (e.g., *flavus*, *ruber*, *thermophilus*, *lacteus*, *rubens*, *aquaticus*), *Bacillus stearothermophilus*, *Thermotoga maritima*, *Methanothermus fervidus*.

[0050] DNA polymerases suitable for the present invention include DNA polymerases that have not yet been isolated. Suitable polymerases for the present invention include fusion polymerases. Fusion polymerases generally contain an additional protein domain at the N- or C-terminus that changes the phenotype of the fusion polymerase compared to the polymerase without the extra domain. Exemplary polymerases include, but are not limited to, polymerases with double-stranded DNA-binding domains fused at the C- or N-terminus. Further examples of fusion polymerases include those with dUTPase fused to the N-or C-terminus (U.S. patent application 20070190538).

[0051] In some embodiments, chimeric DNA polymerases according to the invention contain sequences derived from two or more DNA polymerases that have at least one distinct functional characteristic. Exemplary functional characteristics include, but are not limited to, processivity, elongation rate, fidelity, resistance to salt or PCR additive (e.g., PCR enhancers), thermostability, strand displacement activity, exonuclease activity, uracil read-ahead function, nucleotide selectivity, ability to incorporate modified analogs, and reverse transcriptase activity. For example, some DNA polymerases are characterized with high fidelity. As used herein, the term "high fidelity" refers to an error rate less than 4.45×10^{-6} (e.g., less than 4.0×10^{-6}, 3.5×10^{-6}, 3.0×10^{-6}, 2.5×10^{-6}, 2.0×10^{-6}, 1.5×10^{-6}, 1.0×10^{-6},

0.5 X 10^{-6}) mutations/nt/doubling. Some DNA polymerases are characterized with high processivity. As used herein, the term "high processivity" refers to a processivity higher than 20nts (e.g., higher than 40nts, 60nts, 80nts, 100nts, 120nts, 140nts, 160nts, 180nts, 200nts, 220nts, 240nts, 260nts,280nts, 300nts, 320nts, 340nts, 360nts, 380nts, 400nts, or higher) per association/disassociation with the template. Some DNA polymerases are characterized with high elongation rate. As used herein, the term "high elongation rate" refers to an elongation rate higher than 25 nt/s (e.g., higher than 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140 nt/s). Some enzymes are characterized with high resistance to salt (also referred to as salt tolerance). As used herein, the term "high resistance to salt" (also referred to as high salt tolerance) refers to the ability of a DNA polymerase to substantially maintain its activity at a salt concentration higher than 30mM (e.g., higher than 35mM, 40mM, 45mM, 50mM). In addition, some enzymes are characterized with resistance to PCR additives. Certain PCR additives are PCR enhancers. For example, Kovarova et al. showed that TMA salts, DMSO, betaine and formamide act as PCR enhancers (Kovarova and Draber. (2000) *Nucl. Acids. Res*. 28(13), e70). Another example of PCR enhancers is glycerol. Some enzymes are characterized with resistance to PCR enhancers, in particular, TMAC (also referred to as TMAC tolerance). As used herein, the term "high TMAC tolerance" refers to the ability of a DNA polymerase to substantially maintain its enzymatic activity at a TMAC (tetra-methyl ammonium chloride) concentration higher than 10 mM (e.g., higher than 15 mM, 20 mM). Certain characteristics of exemplary DNA polymerases are shown in Table 1.

Table 1. Characteristics of exemplary DNA polymerases

Polymerases	Fidelity/ Error rate	Processivity (nts)	Elongation rate (nts/s)	Salt tolerance
Pfu	2.0 X 10^{-6}	>20	25	30mM
KOD	4.45 X 10^{-6}	~300	106-138	
TNA1		150		
T. zilligii	2.0 X 10^{-6}			
P. abyssi	0.66 X 10^{-6}			
T. gorgonarius	2.2-3.4 X 10^{-6}			

[0052] Typically, enzymes with high salt tolerance are also characterized with high processivity and/or elongation rate. Without wishing to be bound by any theories, it is thought that salt tolerance affects the binding affinity between polymerase and DNA which,

in turn, affects processivity or elongation rate. Typically, binding of polymerases to DNA involves binding interaction between positively charged amino acid residues and negatively charged DNA. At high salt concentrations, competition from the anions of the salt for the positively charged amino acid residues on the polymerases lead to diminished DNA binding affinity. See, Pavlov et al. (2002) *Proc. Natl. Acad. Sci.* 99(21): 13510-13515, which is incorporated by reference herein. On the other hand, increasing the contact points between DNA and polymerase may increase the salt resistance of the polymerase as well as the processivity or elongation rate because the additional contact points between DNA and polymerase may increase binding affinity of the polymerase for DNA and decrease the rate of dissociation so that the polymerase will remain associated with DNA longer, which will in turn lead to an increase in processivity. For example, Pavlov et al. added helix-hairpin-helix (HhH) motifs from topoisomerase V to Taq and Pfu. These motifs are involved in DNA binding in topoisomerase V. Pavlov et al. showed that both Pfu and Taq become more salt resistant when fused to the HhH motifs. Pavlov et al. also showed that HhH fusion to both Taq and Pfu increased the processivity of the polymerases. As another example, dsDNA binding proteins, e.g., Sso7d, can be fused to DNA polymerases to increase the number of contact points between DNA and polymerases (Wang et al. (2004) *Nucl. Acids Res.* 32(3): 1197-1207, which is incorporated by reference herein). Sso7d is a sequence non-specific dsDNA binding protein involved in ensuring DNA stability and/or DNA packing in *Sulfolobus solfataricus*. Fusion of Sso7d to both Taq and Pfu increased the salt resistance and processivity of the polymerases.

[0053] Exemplary DNA polymerases characterized with high processivity, elongation rate, thermostability, salt or PCR enhancer tolerance include, but are not limited to, KOD polymerase, TNA1 polymerase, *Thermococcus* sp. 9 degrees N-7, T4, T7, or phi29. Exemplary DNA polymerases characterized with high fidelity include, but are not limited to, polymerases isolated from *Pyrococcus furiosus*, *P. abyssi*, *T. gorgonarius*, *T. litoralis, T. zilligii, T.* sp. GT, *or P.* sp. GB-D.

[0054] As non-limiting examples, KOD, Pfu, *T. gorgonarius, T. zilligii, T. litoralis* and *Thermococcus* sp. 9N-7 polymerases are used to engineer chimeric DNA polymerases (see the Example sections).

Domains of DNA polymerases

[0055] Typically, archaeal DNA polymerases include at least the following domains: N-terminal domain, exonuclease domain (e.g., 3' -> 5' exonuclease domain), palm, fingers, and thumb domain (see Figure 1). Knowledge of domain structure, function and coordination is primary based on crystal structure studies and site-directed mutagenesis of various DNA polymerases, in particular, archaeal DNA polymerases. For example, among the first crystal structures of family B DNA polymerases obtained was that of bacteriophage RB69 DNA polymerase (Wang et al. (1997) *Cell,* 89:1087-1099, which is incorporated by reference herein). Among the first crystal structures of archaeal DNA polymerases solved was Tgo DNA polymerase (see, Hopfner et al. 1999 *Proc. Natl. Acad. Sci.* 96(7), 3600-3605, which is incorporated by reference herein). Recently, crystal structures of the following archaeal family B DNA polymerases have been reported: DNA polymerase from Thermococcus sp. 9°N-7 (Rodriguez et al. (2000) *J. Mol. Biol.* 299:447-462, which is incorporated by reference herein), KOD1 DNA polymerase (Hashimoto et al. 2001 *J. Mol. Biol.* 306(3), 469-477, which is incorporated by reference herein), Pfu DNA polymerase (see, U. S. Patent Nos. 5,948, 663; 5,866, 395; 5,545, 552; 5,556, 772 and Kim et al. (2008) *Int. J. Biol. Macromol. 42*(4), 356-61, all of which are hereby incorporated by reference).

[0056] Various functions, such as substrate binding, nucleotide transfer, catalytic activity, proofreading, have been assigned to various domains based on the structural-functional analysis of DNA polymerases. It has also been suggested that the domains tightly coordinate with each other to complete the DNA replication process.

[0057] For example, the polymerase activity has been associated with palm, fingers and thumb domains. In particular, the palm subdomain is thought to be the catalytic site of the polymerase. The polymerase catalyzes a phosphoryl transfer reaction in which the alpha phosphate of the incoming dNTP undergoes nucleophilic attack from the OH primer terminus. Typically, three carboxylate side chains are important to this active site. These residues may bind two metal ions (Mg++) which may facilitate deprotonation of the OH terminus and formation of a transition state at the alpha phosphate of the dNTP. The thumb domain is believed to interact with the minor grove of the newly synthesized dsDNA and also with the incoming nucleotide. The thumb domain is less conserved but typically has a largely helical structure. The fingers domain may play a role in template fixation and nucleotide specificity. Like the thumb domain, it is likely to interact with the incoming nucleotide. The

thumb domain may contain α helices, and/or β strands. It is thought that unbound DNA polymerases form open conformations of the fingers and thumb domains, and when the DNA is bound, the two domains move towards the palm domain to hold the DNA template and primer more tightly and to probe for Watson-Crick base pairing between the incoming nucleotide and the template nucleotide. The presence of a nucleotide that forms a Watson-Crick base pair with the template facilitates formation of an appropriate conformation of the active site of the polymerase and subsequent incorporation of this nuleotide. For review see Hamilton et al. (2001) BioTechniques 31:370-383. It was reported that mutagenesis in the palm/fingers domain may affects the nucleotide selectivity and affinity and mutagenesis in the thumb domain may affect the binding affinity to dsDNA. Important amino acids in the palm, fingers and thumb domain are described in U. S. Application Publication No. 20060281109, which is hereby incorporated by reference.

[0058] The uracil read-ahead function has been associated with the N-terminal domain. For example, archaeal family B DNA polymerases are able to recognize unrepaired uracil in a template strand and stall polymerization upstream of the lesion to prevent an A-T mutation. A "pocket" in the N-terminal domains of archaeal DNA polymerases was identified to be positioned to interact with the template strand and provide this uracil read-ahead function (Fogg et al. (2002) Nature Structural Biology 9(12), 922-927).

[0059] The exonuclease domain is associated with either 5'-> 3' exonuclease activity, 3'-> 5" exonuclease activity or both, which is required to remove incorrectly inserted nucleotide. When a mismatched nucleotide is incorporated, the template/primer strand binds to the polymerase more weakly and/or is misaligned with respect to the polymerase active site causing the mismatched nucleotide to be moved to the active site of the exonuclease domain and excised.

[0060] It is thought that the fidelity is affected by the ratio of the polymerase and the exonuclease activity, which may be influenced by the rate of dissociation, conformational change, and the rate of nucleotide incorporation in the presence of mismatched nucleotides. It has also been suggested that the balance between the 3' -> 5' exonuclease activity and the polymerase activity is mediated by a flexible loop containing the Y-GG/A motif located between the N-terminal and exonuclease domains and the C-terminal polymerase domains (i.e., the palm, fingers and thumb domains). See, Bohlke et al. (2000) *Nucl. Acids Res.* 28(20), 3910-3917. A unique loop of the exonuclease domain, and the tip of the thumb are

important for the coordination of proofreading and polymerase activities in DNA polymerases. Site-directed mutagenesis in this loop, especially at H147 in KOD DNA polymerase, suggested that electrostatic and hydrophobic interactions between this loop and the thumb affect the ratio between exonuclease activity and polymerase activity and hence fidelity. See, Kuroita et al. *J. Mol. Biol.* (2005) 351, 291-298.

Domain swapping

[0061] According to the present invention, heterologous domains from different DNA polymerases (e.g., polymerases with at least one distinct functional characteristic) may be combined to form a chimeric polymerase. Suitable domains include naturally-occurring N-terminal domains, exonuclease domains, palm, fingers, and/or thumb domains found in various DNA polymerases. Naturally-occurring N-terminal domains, exonuclease domains, palm, fingers, and/or thumb domains in various DNA polymerases are well defined. For example, an N-terminal domain may include a sequence corresponding to amino acid residues 26 to 105 of KOD polymerase (SEQ ID NO:11); an exonuclease domain may include a region corresponding to amino acid residues 156 to 301 of KOD polymerase (SEQ ID NO:11); a thumb domain may include a region corresponding to amino acid residues 612 to 749 of KOD polymerase (SEQ ID NO:11); and palm and fingers domain may include a region corresponding to amino acid residues 394 to 563 of Pfu polymerase (SEQ ID NO:9).

[0062] Corresponding domains or positions in various DNA polymerases can be determined by alignment of amino acid sequences. Alignment of amino acid sequences can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST, ALIGN or Megalign (DNASTAR) software. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared. Preferably, the WU-BLAST-2 software is used to determine amino acid sequence identity (Altschul et al., *Methods in Enzymology* 266, 460-480 (1996); http://blast.wustl/edu/blast/README.html). WU-BLAST-2 uses several search parameters, most of which are set to the default values. The adjustable parameters are set with the following values: overlap span=1, overlap fraction=0.125, word threshold (T)=11. HSP score (S) and HSP S2 parameters are dynamic values and are established by the program itself,

depending upon the composition of the particular sequence, however, the minimum values may be adjusted and are set as indicated above. An example of an alignment is shown in Figure 1.

[0063] In some embodiments, a suitable domain may be a variant (e.g., mutant or fragment) of a naturally-occurring domain sequence. For example, a suitable domain may have a sequence having at least 70% (e.g., at least 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%) identical to an amino acid sequence of a naturally-occurring domain found in a DNA polymerase of interest.

[0064] It is further contemplated that sequences defining the N-terminal domain, exonuclease domain, palm, fingers, and/or thumb domains may correlate with certain enzymatic characteristics of DNA polymerases, such as, fidelity or error rate, elongation rate, processivity, and salt resistance. For example, as described in the Examples section, the present inventors have demonstrated that sequences defining the N-terminal, exonuclease, and/or thumb domain may correlate with the characteristics associated with elongation rate, processivity, thermostability, TMAC tolerance and/or salt resistance; and that sequences defining the palm and/or fingers domain may correlate with the characteristics associated with fidelity or error rate of DNA polymerases.

[0065] In addition, based on sequence alignments between various DNA polymerases (see, e.g., Figure 1), it is further contemplated that domains correlative with high processivity, elongation rate, thermostability, TMAC tolerance and/or salt resistance may be defined by one or more of the following positive consensus sequences:

Positive consensus sequence 1 (defining an N-terminal domain)

XXLXXXXXXXEGXRXXXXXXVXXXXXDXXXTXXXXXXXXXXXVVKXXXXXVLIX
XXXXNXXXAXXKXXCXXXXXNFALXXXXXXXXXXXXXIXXMXXRFXXXXXXXXX
XXXXPXXRXXXXXXXXXXXXXXXXXVXXQXXXXXXXEXXTTXXXT (SEQ ID NO:30), wherein X is any amino acid or a peptide bond;

Positive consensus sequence 2 (defining an exonuclease domain)

XXEXXXXYXXXXEXXFXXXXKXXXAXXXXXXXXXAXXXXTVXTVKRXXXXQXXX XXRXVEXXXXXFTXXXXXXAXXDXIXXXXX (SEQ ID NO:31), wherein X is any amino acid or a peptide bond; and

Positive consensus sequence 3 (defining a thumb domain)

XXXXXXXXXXXXXXXXALXXDXXXXKXXXXXXXXTEXXSKXXVXXXXXVXHX XXXXDXKDXXXTXXXXXXXXRXXXRXXXXRXXTXXSXXXXKXSXRXGDXXXPF DXFXXTXXXXXXXXXXXXXXXXXXEXXXRAXX (SEQ ID NO:32), wherein X is any amino acid or a peptide bond.

[0066] Additionally or alternatively, a domain or domains correlative with high processivity, elongation rate, thermostability, TMAC tolerance and/or salt resistance may be defined by one or more of the following negative consensus sequences:

Negative consensus sequence 1 (defining an N-terminal domain)

$NGX_1FKIEX_2DRTFX_3PYX_4YALLX_5DDSX_6IEEVKKITX_7ERHGX_8X_9VX_{10}X_{11}X_{12}X_{13}VEK$ $VX_{14}KKFLGX_{15}PX_{16}X_{17}VWKLYX_{18}X_{19}HPQDVPX_{20}IRX_{21}KX_{22}REHPA$ (SEQ ID NO:33), wherein X_1 is not K; X_2 is not H; X_3 is not R; X_4 is not I; X_5 is not R; X_6 is not K; X_7 is not G; X_8 is not K; X_9 is not I; X_{10} is not R; X_{11} is not I; X_{12} is not V; X_{13} is not D; X_{14} is not E; X_{15} is not K; X_{16} is not I; X_{17} is not T; X_{18} is not L; X_{19} is not E; X_{20} is not T; X_{21} is not E; and X_{22} is not V;

Negative consensus sequence 2 (defining an exonuclease domain)

$PIX_1MISYADEX_2X_3AX_4VITWKNX_5DLPYVX_6VVSX_7EREMIKRFLRX_8X_9X_{10}EKDPDX_{11}$ $X_{12}X_{13}TYNGDX_{14}FDFX_{15}YLX_{16}KRX_{17}EKLGIX_{18}X_{19}X_{20}X_{21}GRDGSEPKX_{22}QRX_{23}GDX_{24}$ $X_{25}AVEVKGRIHFDLYX_{26}VIX_{27}RTINLPTYTLEAVYEAX_{28}FGX_{29}PKEKVYAX_{30}EIX_{31}X_{32}AWEX_{33}$ (SEQ ID NO:34), wherein X_1 is not I; X_2 is not N; X_3 is not E; X_4 is not K; X_5 is not I; X_6 is not E; X_7 is not S; X_8 is not I; X_9 is not I; X_{10} is not R; X_{11} is not I; X_{12} is not I; X_{13} is not V; X_{14} is not S; X_{15} is not P; X_{16} is not A; X_{17} is not A; X_{18} is not K; X_{19} is not L; X_{20} is not T; X_{21} is not I; X_{22} is not M; X_{23} is not I; X_{24} is not M; X_{25} is not T; X_{26} is not H; X_{27} is not T; X_{28} is not I; X_{29} is not K; X_{30} is not D; X_{31} is not A; X_{32} is not K; and X_{33} is not S; and

Negative consensus sequence 3 (defining a thumb domain)

RDWSEIAKETQARVLEX$_1$X$_2$LKX$_3$GDVEX$_4$AVRIVKEVX$_5$X$_6$KLX$_7$X$_8$YEX$_9$PPEKLX$_{10}$IX$_{11}$EQITRX$_{12}$LX$_{13}$X$_{14}$YKAX$_{15}$GPHVAVAKX$_{16}$LAAX$_{17}$GVKIX$_{18}$PGX$_{19}$VIX$_{20}$YIVLX$_{21}$GX$_{22}$GX$_{23}$IX$_{24}$X$_{25}$RAIX$_{26}$X$_{27}$X$_{28}$EX$_{29}$DPX$_{30}$KHKYDAEYYIENQVLPAVX$_{31}$RILX$_{32}$X$_{33}$FG (SEQ ID NO:35), wherein X$_1$ is not T; X$_2$ is not I; X$_3$ is not H; X$_4$ is not E; X$_5$ is not I; X$_6$ is not Q; X$_7$ is not A; X$_8$ is not N; X$_9$ is not I; X$_{10}$ is not A; X$_{11}$ is not Y; X$_{12}$ is not P; X$_{13}$ is not H; X$_{14}$ is not E; X$_{15}$ is not I; X$_{16}$ is not K; X$_{17}$ is not K; X$_{18}$ is not K; X$_{19}$ is not M; X$_{20}$ is not G; X$_{21}$ is not R; X$_{22}$ is not D; X$_{23}$ is not P; X$_{24}$ is not S; X$_{25}$ is not N; X$_{26}$ is not L; X$_{27}$ is not A; X$_{28}$ is not E; X$_{29}$ is not Y; X$_{30}$ is not K; X$_{31}$ is not L; X$_{32}$ is not E; and X$_{33}$ is not G.

[0067] In some embodiments, a domain correlative with high fidelity may be defined by the following positive consensus sequence (defining palm and fingers domain):

XKXXXXXXXXXXXXXAXXXXXXXXXXXXXXXXXXLXXXXNXXIXXXXXXKXXXXI
XXXXXXXXXHXXXXXXXXXTXXXEXQXXXXKIXXXXXXKXXXLXXXXFXXXXX
XXKXXXXXXXXXXXXXXXXXKXXELVWXXLXXXFXXXXLXIXXXXLYXXXXXG
ESXEIXXXXLX (SEQ ID NO:36), wherein X is any amino acid or a peptide bond.

[0068] Additionally or alternatively, a domain correlative with high fidelity may be defined by the following negative consensus sequence (defining palm and fingers domain):

EX$_1$GLWENIVYLDFRX$_2$LYPSIIITHNVSPDTLNX$_3$EGCKX$_4$YDX$_5$APQVGHX$_6$FCKDX$_7$PGFIPSLLGX$_8$LLEERQKIKX$_9$KMKX$_{10}$TX$_{11}$DPIEX$_{12}$X$_{13}$LLDYRQX$_{14}$AIKX$_{15}$LANSX$_{16}$YGYYGYAX$_{17}$ARWYCKECAESVTAWGRX$_{18}$YIX$_{19}$X$_{20}$X$_{21}$X$_{22}$KEX$_{23}$EEKX$_{24}$GFKVX$_{25}$YX$_{26}$DTDGX$_{27}$X$_{28}$ATIPGX$_{29}$X$_{30}$X$_{31}$EX$_{32}$X$_{33}$KKKAX$_{34}$E (SEQ ID NO:37), wherein X$_1$ is not R; X$_2$ is not S; X$_3$ is not R; X$_4$ is not E; X$_5$ is not V; X$_6$ is not R; X$_7$ is not F; X$_8$ is not D; X$_9$ is not K; X$_{10}$ is not A; X$_{11}$ is not I; X$_{12}$ is not R; X$_{13}$ is not K; X$_{14}$ is not R; X$_{15}$ is not I; X$_{16}$ is not Y; X$_{17}$ is not R; X$_{18}$ is not E; X$_{19}$ is not T; X$_{20}$ is not M; X$_{21}$ is not T; X$_{22}$ is not I; X$_{23}$ is not I; X$_{24}$ is not Y; X$_{25}$ is not I; X$_{26}$ is not S; X$_{27}$ is not F; X$_{28}$ is not F; X$_{29}$ is not A; X$_{30}$ is not D; X$_{31}$ is not A; X$_{32}$ is not T; X$_{33}$ is not V; X$_{34}$ is not M.

[0069] Therefore, appropriate domains may be taken or derived from DNA polymerases with distinct functional characteristics to engineer a chimeric DNA polymerase with desirable combinations of functional features. In some embodiments, inventive methods in accordance with the present invention include steps of: (a) providing an N-terminal

domain, an exonuclease domain, and/or a thumb domain based on a first DNA polymerase; (b) providing palm and/or fingers domain based on a second DNA polymerase; (c) combining the domains from step (a) and step (b) to form a chimeric polymerase. In some embodiments, the first and the second DNA polymerases are characterized with at least one distinct characteristic. For example, the first DNA polymerase may be characterized with high processivity, elongation rate, thermostability, TMAC tolerance and/or salt resistance and the second DNA polymerase may be characterized with high fidelity. In some embodiments, the first DNA polymerase may be characterized with high fidelity and the second DNA polymerase may be characterized with high processivity, elongation rate, thermostability, TMAC tolerance and/or salt resistance. In some embodiments, a chimeric polymerase engineered according to the invention has a processivity, elongation rate, thermostability, TMAC tolerance or salt resistance substantially similar to that of the first DNA polymerase and a fidelity substantially similar to that of the second DNA polymerase. In some embodiments, a chimeric polymerases engineered according to the present invention has the fidelity higher than that of the first DNA polymerase and the processivity, elongation rate or salt resistance higher than that of the second DNA polymerase.

[0070] The present invention further contemplates methods of improving the fidelity, processivity, elongation rate, thermostability, TMAC tolerance and/or salt resistance of a DNA polymerase. In some embodiments, inventive methods in accordance with the invention include a step of replacing a sequence within the palm-fingers domain of the DNA polymerase of interest with a corresponding sequence from a different DNA polymerase that is characterized with higher fidelity relative to the DNA polymerase of interest.

[0071] Additionally or alternatively, in some embodiments, inventive methods in accordance with the present invention include a step of replacing a sequence within the N-terminal domain, the exonuclease domain and/or the thumb domain of the DNA polymerase of interest with a corresponding sequence from a different DNA polymerase that is characterized with higher processivity, elongation rate, thermostability, TMAC tolerance or salt resistance relative to the DNA polymerase of interest.

[0072] As a non-limiting example, the present inventors have engineered a chimeric DNA polymerase Kofu and its reciprocal chimera POD based on KOD polymerase and Pfu polymerase (see the Examples section). As discussed in the example section, Kofu contains the N-terminal domain, the exonuclease domain and the thumb domain from KOD

polymerase and the palm-fingers domain from Pfu polymerase. The sequence of Kofu polymerase is provided in SEQ ID NO:16. The reciprocal chimera POD contains the N-terminal domain, the exonuclease domain and the thumb domain from Pfu polymerase and the palm-fingers domain from KOD polymerase. The sequence of POD polymerase is provided in SEQ ID NO:15.

[0073] As discussed in the examples section, the Kofu chimeric polymerase displays the approximate replication fidelity of Pfu but the elongation speed, processivity, thermostability, TMAC tolerance and PCR performance similar to KOD. Alternatively, the Pod chimeric polymerase displays the approximate replication fidelity of KOD but the elongation speed, processivity, thermostability, TMAC tolerance and PCR performance similar to Pfu.

[0074] In some embodiments, the present invention provides variants of Kofu chimeric polymerase that contain an amino acid sequence at least 80% (e.g., at least 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%) identical to SEQ ID NO:16 (Kofu amino acid sequence). In particular embodiments, variants of Kofu chimeric polymerase in accordance with the invention have processivity, elongation rate, thermostability, TMAC tolerance and/or fidelity substantially similar to Kofu.

[0075] In some embodiments, variants of Kofu chimeric polymerases in accordance with the present invention are defined by consensus sequence

XXXXTXXXXXDXXXXXXIXXXXXXEXXXXYXXXXEXXFXXXXKXXXAXXXXXX
XXAXXXXTVXTVKRXXXXQXXXXXRXVEXXXXXFTXXXXXXAXXDXIXXXXXXI
XXYXXXXXXXXXXXXXXXXVXXXXDXXXXMXXXXXXXXXXXXXXXXAEXXXLX
XXXXXXEGXRXXXXXXVXXXXXDXXXTXXXXXXXXXXVVKXXXXXVLIXXXXX
NXXXAXXKXXCXXXXXNFALXXXXXXXXXXIXXMXXRFXXXXXXXXXXXXXPX
XRXXXXXXXXXXXXXXXXVXXQXXXXXXXEXXTTXXXTXXXXXXXXRXXXXX
XXVXXXXXXXXXXXXXAXXXXXVXXPXXXXXXXXXXXXXXXXXXXXXXXXXV
XXXXXSXEXYQXXXXEXXTXXFXXXXXKXXXXXXXXXXXXXAXXXXXXXXXXX
XXXXXLXXXXNXXIXXXXXXKXXXXIXXXXXXXXXHXXXXXXXXTXXXEXQX
XXXKIXXXXXXKXXXLXXXXFXXXXXXXKXXXXXXXXXXXXXXXXXKXXELVW
XXLXXXFXXXXLXIXXXXLYXXXXXGESXEIXXXXLXXLXXXXAXXXXAXXXXX
XXXXXXXXXXXXXKXXXXXXXXXITXXXXXXXXXXXXXXXXXXXXXXALX
XDXXXXKXXXXXXXXTEXXSKXXVXXXXXVXHXXXXXDXKDXXXTXXXXXXX

XRXXXRXXXXRXXTXXSXXXXKXSXRXGDXXXPFDXFXXTXXXXXXXXXXXXXX
XXXXXEXXXRAXXXXXXXXXXXXXXXXXXXSAXXKPXGT (SEQ ID NO:38),
wherein X is any amino acid or a peptide bond.

[0076] In some embodiments, variants of Kofu chimeric polymerases in accordance with the present invention are defined by consensus sequence
XIXDTDYXTXDGXPXXRIFXKXXGEFXXXYDXXFEPYFYALLKDDSAIXXXXXXXA
XRHGTVXTVKRXXXXQXKFLXRXVEVWXLXFTHPQDVPAXXDXIXXHXXVIDIYE
YDIPFAKRYLIDXGLVPMEGDEXLXMXXXDIETXYHEGXEFAEGXXLMISYADXEG
ARVITWKXVDLPYVDVVSTEXEMIKRXXXVVKEKDPDVLIXYXGDNFDXAYLKXR
CEXLGXNFALXRXXXXXEPKIXXMGXRFAVEXKGRXHFDLXPXXRXTXNLPTYXL
XXVYEXVXGQXKXKXXXEEITTXWETXXXXXXXARYSMEDAXVTXELGXEFXPM
EAXLXXLVGXPXWDVXRSSTGNLVEWXLLXXAYXRNEVAPNKPSXEEYQXRXXE
XYTGXFVXEPEKGLWXXXXXLDXXALYPSIIXXHNVSPDTLXLEXCXNYDIAPXVG
XKFCKDIPGFIPSXLXHLXXXRQXXKTXMXEXQDPXEKIXLDYRQKAXKLLXNSFY
GYXGYXKARWYXXECAESVTXWGRKYIELVWXELEXXFGFKXLYIDTDGLYATIP
GGESXEIKXXXLXFLXYINAXLPGALELEYEXFYXRGFFVXKKKYAXIDEEXXITTR
GLEXVRRDWSXXAKETXAXVLEALLXDXXVXKAVXXVXXXTEXXSKYXVPXEKL
VIHEQITRDXKDYXATGPHVAXAKRLXXRGXXXRPGTXISYXXLKGSGRXGDRXIPF
DEFXXTKHXYDXXYYIENQVLPAVERXLRAFGYXXXXLXXQXXXQXGLSAWXKP
XGT (SEQ ID NO:39), wherein X is any amino acid or a peptide bond

[0077] In some embodiments, the present invention provide variants of POD chimeric polymerases that contain an amino acid sequence at least 80% (e.g., at least 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%) identical to SEQ ID NO:15 (Pod amino acid sequence). In particular embodiments, variants of POD chimeric polymerases in accordance with the present invention have processivity, elongation rate, thermostability, TMAC tolerance and/or fidelity substantially similar to POD.

Expression of chimeric DNA polymerases of the invention

[0078] Standard recombinant DNA techniques (e.g., restriction enzyme digestion, ligation, PCR) can be used to engineer chimeric DNA polymerases in accordance with the present invention. Methods well known in the art may be applied to express and isolate

chimeric DNA polymerases. Many bacterial expression vectors contain sequence elements or combinations of sequence elements allowing high level inducible expression of the protein encoded by a foreign sequence. Expression vectors are commercially available from, for example, Novagen (http://www.emdbiosciences.com/html/NVG/AllTables.html#).

[0079] In addition, bacteria expressing an integrated inducible form of the T7 RNA polymerase gene may be transformed with an expression vector bearing a chimeric DNA polymerase gene linked to the T7 promoter. Induction of the T7 RNA polymerase by addition of an appropriate inducer, for example, isopropyl-p-D-thiogalactopyranoside (IPTG) for a lac-inducible promoter, induces the high level expression of the chimeric gene from the T7 promoter.

[0080] Appropriate host strains of bacteria may be selected from those available in the art by one of skill in the art. As a non-limiting example, *E. coli* strain BL-21 is commonly used for expression of exogenous proteins since it is protease deficient relative to other strains of *E. coli*. For situations in which codon usage for the particular polymerase gene differs from that normally seen in E. coli genes, there are strains of BL-21 that are modified to carry tRNA genes encoding tRNAs with rarer anticodons (for example, argU, ileY, leuW, and proL tRNA genes), allowing high efficiency expression of cloned chimeric genes (several BL21-CODON PLUSTM cell strains carrying rare-codon tRNAs are available from Stratagene, for example). Additionally or alternatively, genes encoding DNA polymerases may be codon optimized to facilitate expression in E. coli. Codon optimized sequences can be chemically synthesized.

[0081] There are many methods known to those of skill in the art that are suitable for the purification of a chimeric DNA polymerase of the invention. For example, the method of Lawyer et al. (1993, PCR Meth. & App. 2: 275) is well suited for the isolation of DNA polymerases expressed in *E. coli*, as it was designed originally for the isolation of Taq polymerase. Alternatively, the method of Kong et al. (1993, J. Biol. Chem. 268: 1965, incorporated herein by reference) may be used, which employs a heat denaturation step to destroy host proteins, and two column purification steps (over DEAE-Sepharose and heparin-Sepharose columns) to isolate highly active and approximately 80% pure DNA polymerase.

[0082] Further, DNA polymerase mutants may be isolated by an ammonium sulfate fractionation, followed by Q Sepharose and DNA cellulose columns, or by adsorption of

contaminants on a HiTrap Q column, followed by gradient elution from a HiTrap heparin column.

Uses of chimeric DNA polymerases of the invention

[0083] Chimeric DNA polymerases of the present invention may be used for any methods involving polynucleotide synthesis. Polynucleotide synthesis methods are well known to a person of ordinary skill in the art and can be found, for example, in Molecular Cloning second edition, Sambrook et al., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y. (1989). For example, chimeric DNA polymerases of the present invention have a variety of uses in recombinant DNA technology including, but not limited to, labeling of DNA by nick translation, second-strand cDNA synthesis in cDNA cloning, DNA sequencing, and amplifying, detecting, and/or cloning nucleic acid sequences using polymerase chain reaction (PCR).

[0084] In some embodiments, the invention provides robust, fast, and accurate enzymes for PCR. PCR refers to an *in vitro* method for amplifying a specific polynucleotide template sequence. The technique of PCR is described in numerous publications, including, PCR: A Practical Approach, M. J. McPherson, et al., IRL Press (1991), PCR Protocols: A Guide to Methods and Applications, by Innis, et al. , Academic Press (1990), and PCR Technology : Principals and Applications for DNA Amplification, H. A. Erlich, Stockton Press (1989). PCR is also described in many U. S. Patents, including U. S. Patent Nos. 4,683, 195; 4,683, 202; 4,800, 159; 4,965, 188; 4,889, 818; 5,075, 216 ; 5,079, 352; 5,104, 792; 5,023, 171; 5,091, 310; and 5,066, 584, each of which is herein incorporated by reference.

[0085] Chimeric DNA polymerases with higher processivity, elongation rate and/or fidelity are expected to reduce error rate, improve efficiency and success rate of long-range amplification (higher yield, longer targets amplified), and/or reduce the amount of required DNA template.

[0086] Various specific PCR amplification applications are available in the art (for reviews, see for example, Erlich, 1999, Rev Immunogenet., 1: 127-34; Prediger 2001, *Methods Mol. Biol.* 160: 49-63; Jurecic et al., 2000, *Curr. Opin. Microbiol.* 3: 316-21; Triglia, 2000, *Methods Mol. Biol.* 130: 79-83; MaClelland et al. , 1994, *PCR Methods Appl.*

4: S66-81 ; Abramson and Myers, 1993, *Current Opinion in Biotechnology* 4: 41-47; each of which is incorporated herein by references).

[0087] As non-limiting examples, the present invention can be used in PCR applications including, but are not limited to, i) hot-start PCR which reduces non-specific amplification; ii) touch-down PCR which starts at high annealing temperature, then decreases annealing temperature in steps to reduce non-specific PCR product; iii) nested PCR which synthesizes more reliable product using an outer set of primers and an inner set of primers; iv) inverse PCR for amplification of regions flanking a known sequence. In this method, DNA is digested, the desired fragment is circularized by ligation, then PCR using primer complementary to the known sequence extending outwards; v) AP-PCR (arbitrary primed) /RAPD (random amplified polymorphic DNA). These methods create genomic fingerprints from species with little-known target sequences by amplifying using arbitrary oligonucleotides; vi) RT-PCR which uses RNA-directed DNA polymerase (e.g., reverse transcriptase) to synthesize cDNAs which is then used for PCR. This method is extremely sensitive for detecting the expression of a specific sequence in a tissue or cells. It may also be use to quantify mRNA transcripts ; vii) RACE (rapid amplification of cDNA ends). This is used where information about DNA/protein sequence is limited. The method amplifies 3' or 5' ends of cDNAs generating fragments of cDNA with only one specific primer each (plus one adaptor primer). Overlapping RACE products can then be combined to produce full length cDNA; viii) DD-PCR (differential display PCR) which is used to identify differentially expressed genes in different tissues. First step in DD-PCR involves RT-PCR, then amplification is performed using short, intentionally nonspecific primers; ix) Multiplex-PCR in which two or more unique targets of DNA sequences in the same specimen are amplified simultaneously. One DNA sequence can be use as control to verify the quality of PCR; x) Q/C-PCR (Quantitative comparative) which uses an internal control DNA sequence (but of different size) which compete with the target DNA (competitive PCR) for the same set of primers; xi) Recusive PCR which is used to synthesize genes. Oligonucleotides used in this method are complementary to stretches of a gene (>80 bases), alternately to the sense and to the antisense strands with ends overlapping (-20 bases); xii) Asymmetric PCR; xiii) In Situ PCR; xiv) Site-directed PCR Mutagenesis; xv) DOP-PCR that uses partially degenerate primers for whole-genome amplification; xvi) quantitative PCR using SYBR green or oligonucleotide probes to detect amplification; xvii) whole-genome amplification using

adaptor-ligated DNA fragment libraries as template, and xviii) error-prone PCR in which conditions are optimized to give an increased number of mutations in the PCR product.

[0088] It should be understood that this invention is not limited to any particular amplification system. As other systems are developed, those systems may benefit by practice of this invention.

Kits

[0089] The invention also contemplates kit formats which include a package unit having one or more containers containing chimeric DNA polymerases of the invention and compositions thereof. In some embodiments, the present invention provides kits further including containers of various reagents used for polynucleotide synthesis, including synthesis in PCR.

[0090] Inventive kits in accordance with the present invention may also contain one or more of the following items: polynucleotide precursors, primers, buffers, instructions, and controls. Kits may include containers of reagents mixed together in suitable proportions for performing the methods in accordance with the invention. Reagent containers preferably contain reagents in unit quantities that obviate measuring steps when performing the subject methods.

EXAMPLES

Example 1. Designs of chimeras of KOD and Pfu DNA polymerases

[0091] The two enzymes we chose to include in this experiment were Pyroccocus furiosus DNA polymerase (*Pfu*) and Thermococcus Kodarensis (KOD) DNA polymerases. The two enzymes have similar domain structure and have a 79% identity at the amino acid level using blastP alignments (see Table 2). The domain structures of Pfu and KOD are illustrated in Figure 1.

Table 2. ClustalW alignment of Pfu and KOD

```
PFU  1   MILDVDYITEEGKPVIRLFKKENGKFKIEHDRTFRPYIYALLRDDSKIEEVKKITGERHG 60
KOD  1   ....T.....D......I......E....Y....E..F....K...A........A.... 60

PFU  61  KIVRIVDVEKVEKKFLGKPITVWKLYLEHPQDVPTIREKVREHPAVVDIFEYDIPFAKRY 120
KOD  61  TV.TVKR....Q.....R.VE.....FT......A..D.I......I..Y.......... 120

PFU  121 LIDKGLIPMEGEEELKILAFDIETLYHEGEEFGKGPIIMISYADENEAKVITWKNIDLPY 180
KOD  121 ......V....D....M..............AE...L.......EG.R......V.... 180

PFU  181 VEVVSSEREMIKRFLRIIREKDPDIIVTYNGDSFDFPYLAKRAEKLGIKLTIGRDGSEPK 240
KOD  181 .D...T..........VVK.....VLI.....N...A..K..C.....NFAL........ 240

PFU  241 MQRIGDMTAVEVKGRIHFDLYHVITRTINLPTYTLEAVYEAIFGKPKEKVYADEIAKAWE 300
KOD  241 I..M..RF.............P..R................V..Q.......E..TT... 300

PFU  301 SGENLERVAKYSMEDAKATYELGKEFLPMEIQLSRLVGQPLWDVSRSSTGNLVEWFLLRK 360
KOD  301 T........R.......V............A.....I..S.................... 360

PFU  361 AYERNEVAPNKPSEEEYQRRLRESYTGGFVKEPEKGLWENIVYLDFRALYPSIIITHNVS 420
KOD  361 ......L.....D.K.LA..~.Q..E.. Y.....R............S............ 419

PFU  421 PDTLNLEGCKNYDIAPQVGHKFCKDIPGFIPSLLGHLLEERQKIKTKMKETQDPIEKILL 480
KOD  420 .....R....E..V......R....F.........D.........K...A.I....RK.. 479

PFU  481 DYRQKAIKLLANSFYGYYGYAKARWYCKECAESVTAWGRKYIELVWKELEEKFGFKVLYI 540
KOD  480 ....R...I....Y.......R.................E..TMTI..I...Y....I.S 539

PFU  541 DTDGLYATIPGGESEEIKKKALEFVKYINSKLPGLLELEYEGFYKRGFFVTKKRYAVIDE 600
KOD  540 ....FF.....ADA.TV....M..L....A....A.................K...... 599

PFU  601 EGKVITRGLEIVRRDWSEIAKETQARVLETILKHGDVEEAVRIVKEVIQKLANYEIPPEK 660
KOD  600 ...IT.......................AL..D....K........TE..SK..V.... 659

PFU  661 LAIYEQITRPLHEYKAIGPHVAVAKKLAAKGVKIKPGMVIGYIVLRGDGPISNRAILAEE 720
KOD  660 .V.H.....D.KD...T........R...R....R..T..S....K.S.R.GD...PFD. 719

PFU  721 YDPKKHKYDAEYYIENQVLPAVLRILEGFGYRKEDLRYQKTRQVGLTSWLNIKKS* 776
KOD  720 F..T..................E...RA..................SA..KP.GT* 775

PFU  (SEQ ID NO:9)
KOD  (SEQ ID NO:11)
```

[0092] Pfu and KOD have very distinct phenotypic characteristics, in particular, with respect to elongation rate, processivity and error rate (See Table 3):

Table 3

	Pfu	KOD
Elongation Rate:	25nt/s	106-138nt/s (Takagi et al. 1997)
Processivity:	>20nt	~300nt (Takagi et al. 1997)
Error Rate (mutations/nt/doubling):	1.5×10^{-6}	4.45×10^{-6} (internal data)

[0093] Thus, the goal was to find chimeric combinations of these two enzymes which exhibited the error rate comparable to Pfu ($2.0x10^{-6}$) with the processivity and/or elongation rate comparable to KOD (~300nt/s and 106-138nt/s respectively). An enzyme with the above mentioned characteristics has utility as a robust, fast, and accurate enzyme for PCR.

[0094] Restriction sites were inserted into the codon-optimized nucleotide sequence of KOD and Pfu polymerases at positions that approximately flank the polymerase domain of the enzymes (see Example 2). For example, PvuII and EcoRI sites flanking the polymerase domain (the palm and fingers domain) were used to replace the polymerase domain of Pfu with that of KOD to generate the chimera deemed Pod (Figure 2). This chimera contains the N-terminal domain, the 3'-5' exonuclease domain and the thumb domain of Pfu and the palm and fingers domain of KOD. The reciprocal swap, yielding the chimera Kofu, was generated by replacing the polymerase domain (palm and fingers) of KOD with that of Pfu.

Example 2. Codon optimization and synthesis of *Pyrococcus furiosus* and *Thermococcus kodakarensis* DNA Polymerases

[0095] Native DNA sequences for *Pyrococcus furiosus* polymerase I (SEQ ID NO:1) and *Thermococcus kodakarensis* polymerase I (SEQ ID NO:2) were retrieved from Genbank. These two DNA sequences were *in silico* codon optimized by Codon Devices (Cambridge, Massachusetts) for expression in *E. Coli* resulting in SEQ ID NO:3 for the Pfu polymerase I codon optimized gene DNA sequence and SEQ ID NO:4 for the KOD polymerase I codon optimized gene DNA sequence. The two codon optimized genes were chemically synthesized and cloned into pUC19 by Codon Devices (Cambridge, Massachusetts) resulting in SEQ ID NO:7 for Pfu polymerase I and SEQ ID NO:8 for KOD polymerase I.

Example 3: Cloning of codon optimized KOD and Pfu polymerase I sequences into expression vector pKBexp.

[0096] KOD (SEQ ID NO:8) and Pfu (SEQ ID NO:7) polymerase codon optimized pUC 19 constructs were cloned into the pKBexp vector as follows:

[0097] The pKBexp vector contains two Eco31I sites with non-complementary overhangs enabling directional cloning of inserts using a single restriction enzyme. KOD and

Pfu polymerase genes were designed with two flanking Eco31I sites that enabled directional and in-frame cloning into pKBexp.

[0098] Purified DNA from the pKBexp vector was digested with Eco31I and purified from an agarose gel. KOD and Pfu codon optimized pUC DNA constructs (SEQ ID NO.8 and SEQ ID NO.7) were likewise digested with Eco31I and the roughly 2.3 kilobase insert fragments were cut out from an agarose gel and purified. 30 ng of KOD or Pfu polymerase genes were ligated with 15 ng of digested pKBexp using T4 DNA ligase. The ligation reactions were purified and used to transform competent *E. coli* DH10B. DNA minipreps were made of ampicillin resistant clones. The presence of inserts was confirmed by digestion of the minipreps with XbaI and HindIII, two enzymes that flank the insert. The cloning of the KOD polymerase gene sequence in pKBexp deemed pKB11 and the Pfu polymerase gene in pKBexp deemed pKB14 were confirmed by DNA sequencing.

Example 4: Domain swapping of DNA sequences from KOD and Pfu polymerase I genes

[0099] The codon-optimized sequences of KOD (SEQ ID NO:5) and Pfu (SEQ ID NO:3) polymerase I genes were designed with restriction sites that approximately flank the finger and palm domains of KOD and Pfu polymerases. The KOD codon optimized sequence contains a PvuII restriction site and an EcoRI restriction site. The Pfu codon optimized sequence contains a PvuII restricition site and an EcoRI restriction site.

[0100] Purified DNA from pKB11 and pKB14 were each digested the restriction enzymes EcoRI and PvuII. The large fragment (4.7 kb) and small fragment (0.7 kb) from each digest were separately extracted and purified from an agarose gel. The small fragments from each restriction digest contained the finger and palm domains of KOD and Pfu respectively. The digested and purified large fragments (containing the expression vector and remaining polymerase fragments) were dephosphorylated using Shrimp Alkaline Phospate. The construct deemed POD was created by ligation of 30 ng of the 4.7 kb Pfu large fragment (aa residues 1 to 335 and 567 to 778 of Pfu DNA polymerase with 10 ng of the 0.7 kb KOD small fragment (corresponding to amino acid residues 336 to 565 of KOD DNA polymerase SEQ ID NO:11). POD thus includes N-terminal, exonuclease and thumb domains from Pfu DNA polymerase and palm and finger domains from KOD. The construct deemed Kofu was made by ligation of 30 ng of the 4.7 kb KOD large fragment (corresponding to amino acid

residues 1 to 335 and 566 to 777 of KOD DNA polymerase SEQ ID NO:11) with 10 ng of the 0.7 kb Pfu small fragment (corresponding to amino acid residues 336 to 566 of Pfu DNA polymerase SEQ ID NO:11). Kofu thus includes N-terminal, exonuclease and thumb domains from KOD DNA polymerase and palm and finger domains from Pfu. The ligation reactions were used to transform *E. coli* DH10B. The construction of Pod (SEQ ID NO:13) and Kofu (SEQ ID NO:14) was confirmed by DNA sequencing. The domain structures of POD and Kofu are illustrated in Figure 1. Expression and purification of chimeric polymerases are done using methods known in the art, for example, as reviewed in "Detailed description of the invention."

Example 5. Thermostability of KOD, Pfu, Kofu and Pod

[0101] 10 ng of each enzyme was incubated at 98 °C for 240, 120, 60, 30, 15, 8, 4, 2, 1 or 0 min in a 10 µl volume containing the following: 20 mM Tris-HCl pH 8.0, 2 mM $MgCl_2$, 6 mM $(NH4)_2SO_4$, 25 or 50 mM KCl (25 mM for Pfu and Pod, 50 mM for KOD and Kofu). 10 µl of primer/template mix was added to each tube after the heat incubation. The primer template mix contained the following: 20 mM Tris-HCl pH 8.0, 2 mM $MgCl_2$, 6 mM $(NH4)_2SO_4$, 0.6 mM dNTP, 0.6 µM each of primers HPRT1-F1 (5'-tttggaaacatctggagtcct -3' (SEQ ID NO:40)) and HPRT1-R1 (5'- gcccaaagggaactgatagtc -3' (SEQ ID NO:41)), 2 ng human genomic DNA per µl, and 25 or 50 mM KCl (25 mM for Pfu and Pod, 50 mM for KOD and Kofu). The amplifications were performed with the following cycling protocol: 3 minutes at 95 °C, 35 x (20 seconds at 98 °C, 20 seconds at 60 °C, 20 seconds at 72 °C), 20 seconds at 72 °C. The PCR products were analysed on an agarose gel (see Figure 3). As illustrated in Figure 3, no amplification was observed for Pfu after pre-incubation of the enzyme for 4 hours at 98 °C. In contrast, KOD, Kofu and Pod were able to amplify a PCR product for all time points tested.

Example 6. Fidelity assays

[0102] The fidelity of enzymes was determined by a method similar to that described by Cline et al. and references therein (*Nucl. Acids Res.*, 1996, 24(18): 3546-3551). LacI was PCR amplified from *E. coli* and cloned into pUC19 to degenerate plasmid pKB-LacIQZalpha

(SEQ ID NO:17). pKB-LacIQZalpha served both as template for PCR amplification of LacI in the fidelity assays and as vector for cloning the amplified LacI into for blue/white colony screening.

[0103] 2 x 50 µl PCR reactions (for each enzyme) were set-up, using 70ng of pKB-LacIQZalfa plasmid template (equivalent to 25 ng of lacI target) and 2.5U of each enzyme to amplify the 1.386 Kb lacIOZalpha fragment. The PCR conditions were as follows: amplification with Pfu and Pod were done in Pfu buffer (Fermentas); KOD and Kofu in Novagen KOD buffer 1. Final concentrations of 2 mM MgCl2, 0.4 µM each of primers M13-40 (GTTTTCCCAGTCACGAC (SEQ ID NO:42)) and PKBlac-1R (GGTATCTTTATAGTCCTGTCG (SEQ ID NO:43)) and 0.2 mM each dNTP. Cycling parameters for Pfu and Pod were: 94 °C 4 minutes, 30 x (94 °C 15 seconds, 55 °C 15 seconds, 72 °C 3 minutes), 72 °C 6 minutes. Cycling parameters for KOD and Kofu were: 94 °C 2 minutes, 30 x (98 °C 15 seconds, 55 °C 2 seconds, 72 °C 20 seconds), 72 °C 30 seconds.

[0104] PCR product yields were quantitated by means of gel electrophoresis and the number of template doublings were calculated. PCR products were digested with XbaI, NcoI and DpnI, gel-purified (without exposure to UV light) and ligated into XbaI-NcoI-digested pKB-LacIQZalpha. *E.coli* was transformed with the ligation mixtures and the cells were plated onto LB-Amp-X-gal plates. The number of blue colonies, white colonies and total number of colonies were recorded. The error rate *f* was calculated as $f = -\ln(F)/(d \times (bp))$, where F = fraction of white colonies ((total colonies minus blue colonies)/total colonies), d = number of template doublings and b = 349 (only 349 bp of the *lacI* amplicon are scored). Exemplary results are summarized in Table 4. As shown in Table 4, Pfu and Kofu have similar fidelity and that their fidelity is higher than that of KOD and Pod.

Table 4: Fidelity of KOD, Pfu, Kofu and Pod

	White colonies	**Doublings d**	**Blue colonies**	**Total colonies**	**Fidelity f (x 10^{-6})**
KOD	21130	7.77	246	21376	4.27
Pfu	19270	7.76	77	19347	1.47
Kofu	12817	5.8	39	12856	1.50
Pod	22039	7.19	221	22260	3.98

Example 7. Processivity assays

[0105] Processivity can be determined and calculated using assays described in (Wang et al. *Nucl. Acids Res*, 2004, 32(3): 1197-1207; and Von Hippel et al. *NY Acad Sci* 1994; 726:118-131). Briefly, 0.8 pmoles of a 5'FAM-labelled primer (-40M13LFF, 5'FAM- GTTTTCCCAGTCACGACGTTGTAAAACGACGGCC-3' (SEQ ID NO:44)) is added to 1.6 pmoles of ssM13mp18 DNA in the presence of 20 mM Tris-HCl pH 8.0, 25 mM KCl, 2.5 mM MgCl2, 0.3 mM dNTP in a 16 microL volume. The primer is annealed to the template by heating to 95 °C for 2 minutes followed by slow cooling to 72 °C in a thermocycler at a rate of 0.1 °C/second, incubation for 10 minutes at 72 °C and further cooling at 0.1 °C/second to 4 °C. The polymerases are diluted in 20 mM Tris-HCl pH 8.0, 25 mM KCl. The primed template and the diluted polymerases are heated to 72 °C and the reaction is started by adding 4 µl diluted polymerase to 16 µl of primed template. The polymerases are diluted to give polymerase:template ratios of 1:10 – 1:10000. The reactions are terminated after various timepoints by adding EDTA to a final concentration of 10 mM.

[0106] The extension reactions are analyzed on an ABI 3130XL Genetic Analyzer. The median product length is determined for each reaction. The median product length is defined as the length of the product at which the total fluorescence intensity of all products up to that length equals 50% of the sum of fluorescence intensities of all detectable products. The traces for those samples where the median product length does not change with a change in polymerase concentration or incubation time are used to calculate the processivity according to Von Hippel et al. (Von Hippel et al. *NY Acad Sci* 1994; 726:118-131). Each peak (I) with a fluorescence level significantly above background level is integrated to give the fluorescence intensity of that peak (ni). The total fluorescence intensity (nT) is the sum of the fluorescence of all peaks. The integration data are plotted as log(ni/nT) vs n-1, where n is the number of nucleotides incorporated. The data is fitted to the following equation: log(ni/nT) = (n-1)logPi + log(1-Pi). Pi, the microscopic processivity factor, is defined as the probability of not terminating extension at position i. The average primer extension length is determined from 1/(1-Pi).

Example 8. Salt resistance of KOD, Pfu, Kofu and Pod

[0107] Previous studies (Pavlov et al. (2002) *Proc Natl Acad Sci.* 99(21), 13510-13515; Wang et al. (2004) *Nucl Acids Res.* 32(3), 1197-1207) have shown that there is a direct correlation between increased tolerance of polymerases to salt and the processivity of polymerases. For all polymerases tested (from family A or family B), it was found that polymerases with increased salt tolerance also have increased processivity. We therefore compared the salt tolerance of our chimeras with that of the parental polymerases as a proxy for processivity.

[0108] The protein concentration of the purified KOD, Pfu, Kofu and Pod where determined using a Bioanalyzer 2100 (Agilent, Santa Clara, CA, USA) with the Protein 230 Kit from the same supplier. The polymerases were tested in real-time PCR with increasing amounts of KCl added. The reactions were performed in a 20 µl volume containing 20 mM Tris-HCl pH 8.0, 6 mM $(NH_4)_2SO_4$, 2 mM $MgCl_2$, 3% DMSO, 10 ng polymerase, 20 ng human genomic DNA, 0.3 mM each dNTP, 0.25X SYBR Green (Invitrogen, Carlsbad, CA, USA). A diluted stock 20X SYBR Green in DMSO was made), 0.3 µM forward primer HPRT1-F1 (5'-tttggaaacatctggagtcct -3' (SEQ ID NO:40)) and 0.3 µM reverse primer HPRT1-R1(5'- gcccaaagggaactgatagtc -3' (SEQ ID NO:41)). KCl was added to final concentrations of 10, 25, 50, 75, 100 or 125 mM. PCR amplification was performed in a Corbett 6000 HRM real-time thermocycler (Corbett Life Science, Sidney, Australia) with the following cycling protocol: 3 minutes at 95 °C, 40 cycles of (10 seconds at 95 °C, 20 seconds at 60 °C, 20 seconds at 72 °C, data acquisition), followed by a melting curve analysis step of: ramp from 72 °C to 95 °C in 1 °C steps, wait for 5 seconds before data acquisition at the end of each step. 8 µl of each sample was analysed on a 1.5% agarose gel. 5 µl of Fermentas GeneRuler™ Mix, cat no. SM0333 (Fermentas, Vilnius, Lithuania) was loaded onto the gel as a DNA marker. Exemplary results are shown in Figure 4.

Example 9. TMAC tolerance of KOD, Pfu, Kofu and Pod

[0109] Tetra-methyl ammonium-containing salts enhance PCR reactions as shown by Kovarova et al. (Kovarova, M. and Draber, P.; *Nucl. Acids Res.* (2000) 28(13) e70-). One

such salt is tetra-methyl ammonium chloride (TMAC). We therefore compared the TMAC tolerance of our chimeras with that of the parental polymerases.

[0110] The polymerases were tested in real-time PCR with increasing amounts of TMAC added. The reactions were performed in a 20 µl volume containing 20 mM Tris-HCl pH 8.0, 6 mM $(NH_4)_2SO_4$, 2 mM $MgCl_2$, 25 mM KCl, 10 ng polymerase, 20 ng human genomic DNA, 0.3 mM each dNTP, 0.25X SYBR Green (Invitrogen, Carlsbad, CA, USA. A diluted stock 20X SYBR Green in DMSO was made), 0.3 µM forward primer HPRT1-F1 (5'-tttggaaacatctggagtcct -3' (SEQ ID NO:40)) and 0.3 µM reverse primer HPRT1-R1(5'- gcccaaagggaactgatagtc -3' (SEQ ID NO:41)). TMAC was added to final concentrations of 0, 10, 20, 40, 60, 80, 100 or 120 mM. PCR amplification was performed in a Corbett 6000 HRM real-time thermocycler (Corbett Life Science, Sidney, Australia) with the following cycling protocol: 3 minutes at 95 °C, 40 cycles of (10 seconds at 95 °C, 20 seconds at 50 °C, 20 seconds at 72 °C, data acquisition), followed by a melting curve analysis step of: ramp from 72 °C to 95 °C in 1 °C steps, wait for 5 seconds before data acquisition at the end of each step. 8 µl of each sample was analysed on a 1.5% agarose gel. 5 µl of Fermentas GeneRuler™ Mix, cat no. SM0333 (Fermentas, Vilnius, Lithuania) was loaded onto the gel as a DNA marker. Exemplary results are shown in Figure 5.

Example 10. Additional chimeras of KOD and Pfu polymerases

[0111] This example is designed to show that the positions where the swapping between domains take place may vary.

[0112] Additional chimeras are made by swapping the palm and finger domains of KOD and Pfu polymerases where the exact position of the swap varies slightly compared to positions for Kofu and Pod. Kofu-II (SEQ ID NO:26) is made by replacing amino acid residues 305 to 615 of KOD (SEQ ID NO: 12) with amino acids 305 to 616 of Pfu (SEQ ID NO:10). Pod-II (SEQ ID NO:27) is made by replacing amino acids 305 to 616 of Pfu (SEQ ID NO:10) with amino acids 305 to 615 of KOD (SEQ ID NO:12).

[0113] Kofu-III (SEQ ID NO:28) is made by replacing amino acid residues 396 to 564 of KOD (SEQ ID NO: 12) with amino acids 397 to 565 of Pfu (SEQ ID NO:10). Pod-III (SEQ ID NO:29) is made by replacing amino acids 397 to 565 of Pfu (SEQ ID NO:10) whith amino acids 396 to 564 of KOD (SEQ ID NO:12).

[0114] The amino acid sequence of chimeras Kofu-II, Pod-II, Kofu-III and Pod-III are reverse translated and codon-optimized for expression in E.coli. Additional nucleotide sequences containing Eco31I restriction sites are added to the 5' and 3' ends of the construct to facilitate cloning into an expression vector. More specifically, the 5' and 3' sequences can be designed so that the overhangs, after digestion of the DNA with Eco31I, are complementary to the overhangs in a particular expression vector (e.g., pKB). Codon optimization and gene synthesis is performed by GeneArt Gmbh. Expression and purification of chimeric polymerases are done using methods known in the art, for example, as reviewed in "Detailed description of the invention". The thermostability, fidelity, processivity, salt resistance and TMAC resistance of the chimeric polymerses are determined as described in Examples 5 through 9.

Example 11. Chimeras of *T. litoralis* and 9 degrees N-7 polymerases

[0115] Chimeras 9Nli and Li9N are designed based on the alignment in Figure 1. They are made by swapping the palm and finger domains between the DNA polymerases of *T. litoralis* and *Thermococcus* sp. 9 degrees N-7. The overall sequence identity between these two polymerases are 77% on the amino acid level.

[0116] Chimera 9Nli can be made by replacing the palm and finger region of the 9N polymerase with the palm and finger region of the *T. litoralis* polymerase. In this particular example, 9Nli is made by replacing amino acids 347 to 580 of 9N polymerase (SEQ ID NO:18) with amino acids 349 to 583 of *T. litoralis* polymerase (SEQ ID NO:19). The sequence of the coding region of 9Nli is provided as SEQ ID NO:20.

[0117] Chimera LiN9 can be made by replacing the palm and finger domain of the DNA polymerase of *T. litoralis* with the finger domain of the DNA polymerase of 9 degrees North. In this particular example, LiN9 is made by replacing amino acids 349 to 583 of *T. litoralis* polymerase (SEQ ID NO:19) with amino acids 347 to 580 of 9 degrees N-7 polymerase (SEQ ID NO:18). The sequence of the coding region of LiN9 is provided as SEQ ID NO:21.

Example 12. Chimeras of *T. gorgonarius* and *T. zilligii* type B DNA polymerases

[0118] Chimerase GoZi and ZiGo are designed based on the alignment in Figure 1. They are made by swapping the palm and finger domains between the DNA polymerases of *T. gorgonarius* and *T. zilligii*. The overall sequence identity between these two polymerases are 94% on the amino acid level.

[0119] Chimera GoZi can be made by replacing the palm and finger region of the *T. gorgonarius* polymerase with the palm and finger region of the *T.zilligii* polymerase. In this particular example, GoZi is made by replacing amino acids 391 to 559 of *T. gorgonarius* polymerase (SEQ ID NO:22) with amino acids 391 to 559 of *T. zilligii* polymerase (SEQ ID NO:23). The sequence of the resulting chimera GoZi is provided as SEQ ID NO:24.

[0120] Chimera ZiGo can be made by replacing the palm and finger domain of the DNA polymerase of *T. zilligii* with the finger domain of the DNA polymerase of *T. gorgonarius*. In this particular example, ZiGo is made by replacing amino acids 391 to 559 of *T. zilligii* polymerase (SEQ ID NO:23) with amino acids 391 to 559 of *T. gorgonarius* polymerase (SEQ ID NO:22). The sequence of the coding region of ZiGo is provided as SEQ ID NO:25.

Table 5. Sequences

Native DNA sequences of Pfu and KOD

Sequence 1 (SEQ ID NO:1)

>Native Pfu nucleotide sequence from genomic sequence (Acc. No. AE010147)

```
1       ATGATTTTAG ATGTGGATTA CATAACTGAA GAAGGAAAAC CTGTTATTAG GCTATTCAAA
61      AAAGAGAACG GAAAATTTAA GATAGAGCAT GATAGAACTT TTAGACCATA CATTTACGCT
121     CTTCTCAGGG ATGATTCAAA GATTGAAGAA GTTAAGAAAA TAACGGGGGA AAGGCATGGA
181     AAGATTGTGA GAATTGTTGA TGTAGAGAAG GTTGAGAAAA AGTTTCTCGG CAAGCCTATT
241     ACCGTGTGGA AACTTTATTT GGAACATCCC CAAGATGTTC CCACTATTAG AGAAAAAGTT
301     AGAGAACATC CAGCAGTTGT GGACATCTTC GAATACGATA TTCCATTTGC AAAGAGATAC
361     CTCATCGACA AAGGCCTAAT ACCAATGGAG GGGGAAGAAG AGCTAAAGAT TCTTGCCTTC
421     GATATAGAAA CCCTCTATCA CGAAGGAGAA GAGTTTGGAA AAGGCCCAAT TATAATGATT
481     AGTTATGCAG ATGAAAATGA AGCAAAGGTG ATTACTTGGA AAAACATAGA TCTTCCATAC
541     GTTGAGGTTG TATCAAGCGA GAGAGAGATG ATAAAGAGAT TTCTCAGGAT TATCAGGGAG
601     AAGGATCCTG ACATTATAGT TACTTATAAT GGAGACTCAT TCGACTTCCC ATATTTAGCG
661     AAAAGGGCAG AAAAACTTGG GATTAAATTA ACCATTGGAA GAGATGGAAG CGAGCCCAAG
721     ATGCAGAGAA TAGGCGATAT GACGGCTGTA GAAGTCAAGG GAAGAATACA TTTCGACTTG
781     TATCATGTAA TAACAAGGAC AATAAATCTC CCAACATACA CACTAGAGGC TGTATATGAA
841     GCAATTTTTG GAAAGCCAAA GGAGAAGGTA TACGCCGACG AGATAGCAAA AGCCTGGGAA
901     AGTGGAGAGA ACCTTGAGAG AGTTGCCAAA TACTCGATGG AAGATGCAAA GGCAACTTAT
961     GAACTCGGGA AAGAATTCCT TCCAATGGAA ATTCAGCTTT CAAGATTAGT TGGACAACCT
1021    TTATGGGATG TTTCAAGGTC AAGCACAGGG AACCTTGTAG AGTGGTTCTT ACTTAGGAAA
1081    GCCTACGAAA GAAACGAAGT AGCTCCAAAC AAGCCAAGTG AAGAGGAGTA TCAAAGAAGG
1141    CTCAGGGAGA GCTACACAGG TGGATTCGTT AAAGAGCCAG AAAAGGGGTT GTGGGAAAAC
1201    ATAGTATACC TAGATTTTAG AGCCCTATAT CCCTCGATTA TAATTACCCA CAATGTTTCT
1261    CCCGATACTC TAAATCTTGA GGGATGCAAG AACTATGATA TCGCTCCTCA AGTAGGCCAC
1321    AAGTTCTGCA AGGACATCCC TGGTTTTATA CCAAGTCTCT TGGGACATTT GTTAGAGGAA
1381    AGACAAAAGA TTAAGACAAA AATGAAGGAA ACTCAAGATC CTATAGAAAA AATACTCCTT
1441    GACTATAGAC AAAAAGCGAT AAAACTCTTA GCAAATTCTT TCTACGGATA TTATGGCTAT
1501    GCAAAAGCAA GATGGTACTG TAAGGAGTGT GCTGAGAGCG TTACTGCCTG GGGAAGAAAG
1561    TACATCGAGT TAGTATGGAA GGAGCTCGAA GAAAAGTTTG GATTTAAAGT CCTCTACATT
1621    GACACTGATG GTCTCTATGC AACTATCCCA GGAGGAGAAA GTGAGGAAAT AAAGAAAAAG
1681    GCTCTAGAAT TTGTAAAATA CATAAATTCA AAGCTCCCTG GACTGCTAGA GCTTGAATAT
1741    GAAGGGTTTT ATAAGAGGGG ATTCTTCGTT ACGAAGAAGA GGTATGCAGT AATAGATGAA
1801    GAAGGAAAAG TCATTACTCG TGGTTTAGAG ATAGTTAGGA GAGATTGGAG TGAAATTGCA
1861    AAAGAAACTC AAGCTAGAGT TTTGGAGACA ATACTAAAAC ACGGAGATGT TGAAGAAGCT
1921    GTGAGAATAG TAAAAGAAGT AATACAAAAG CTTGCCAATT ATGAAATTCC ACCAGAGAAG
1981    CTCGCAATAT ATGAGCAGAT AACAAGACCA TTACATGAGT ATAAGGCGAT AGGTCCTCAC
2041    GTAGCTGTTG CAAAGAAACT AGCTGCTAAA GGAGTTAAAA TAAAGCCAGG AATGGTAATT
2101    GGATACATAG TACTTAGAGG CGATGGTCCA ATTAGCAATA GGGCAATTCT AGCTGAGGAA
2161    TACGATCCCA AAAAGCACAA GTATGACGCA GAATATTACA TTGAGAACCA GGTTCTTCCA
2221    GCGGTACTTA GGATATTGGA GGGATTTGGA TACAGAAAGG AAGACCTCAG ATACCAAAAG
2281    ACAAGACAAG TCGGCCTAAC TTCCTGGCTT AACATTAAAA AATCCTAG
```

Sequence 2 (SEQ ID NO:2)

>Native KOD nucleotide sequence (from genomic sequence, Acc. no. AP006878)

```
1       ATGATCCTCG ACACTGACTA CATAACCGAG GATGGAAAGC CTGTCATAAG AATTTTCAAG
61      AAGGAAAACG GCGAGTTTAA GATTGAGTAC GACCGGACTT TTGAACCCTA CTTCTACGCC
121     CTCCTGAAGG ACGATTCTGC CATTGAGGAA GTCAAGAAGA TAACCGCCGA GAGGCACGGG
181     ACGGTTGTAA CGGTTAAGCG GGTTGAAAAG GTTCAGAAGA AGTTCCTCGG GAGACCAGTT
241     GAGGTCTGGA AACTCTACTT TACTCATCCG CAGGACGTCC CAGCGATAAG GGACAAGATA
301     CGAGAGCATC CAGCAGTTAT TGACATCTAC GAGTACGACA TACCCTTCGC CAAGCGCTAC
361     CTCATAGACA AGGGATTAGT GCCAATGGAA GGCGACGAGG AGCTGAAAAT GCTCGCCTTC
421     GACATTGAAA CTCTCTACCA TGAGGGCGAG GAGTTCGCCG AGGGGCCAAT CCTTATGATA
481     AGCTACGCCG ACGAGGAAGG GGCCAGGGTG ATAACTTGGA AGAACGTGGA TCTCCCCTAC
541     GTTGACGTCG TCTCGACGGA GAGGGAGATG ATAAAGCGCT TCCTCCGTGT TGTGAAGGAG
601     AAAGACCCGG ACGTTCTCAT AACCTACAAC GGCGACAACT TCGACTTCGC CTATCTGAAA
661     AAGCGCTGTG AAAAGCTCGG AATAAACTTC GCCCTCGGAA GGGATGGAAG CGAGCCGAAG
721     ATTCAGAGGA TGGGCGACAG GTTTGCCGTC GAAGTGAAGG GACGGATACA CTTCGATCTC
781     TATCCTGTGA TAAGACGGAC GATAAACCTG CCCACATACA CGCTTGAGGC CGTTTATGAA
841     GCCGTCTTCG GTCAGCCGAA GGAGAAGGTT TACGCTGAGG AAATAACCAC AGCCTGGGAA
901     ACCGGCGAGA ACCTTGAGAG AGTCGCCCGC TACTCGATGG AAGATGCGAA GGTCACATAC
961     GAGCTTGGGA AGGAGTTCCT TCCGATGGAG GCCCAGCTTT CTCGCTTAAT CGGCCAGTCC
```

```
1021   CTCTGGGACG TCTCCCGCTC CAGCACTGGC AACCTCGTTG AGTGGTTCCT CCTCAGGAAG
1081   GCCTATGAGA GGAATGAGCT GGCCCCGAAC AAGCCCGATG AAAAGGAGCT GGCCAGAAGA
1141   CGGCAGAGCT ATGAAGGAGG CTATGTAAAA GAGCCCGAGA GAGGGTTGTG GGAGAACATA
1201   GTGTACCTAG ATTTTAGATC CCTGTACCCC TCAATCATCA TCACCCACAA CGTCTCGCCG
1261   GATACGCTCA ACAGAGAAGG ATGCAAGGAA TATGACGTTG CCCCACAGGT CGGCCACCGC
1321   TTCTGCAAGG ACTTCCCAGG ATTTATCCCG AGCCTGCTTG GAGACCTCCT AGAGGAGAGG
1381   CAGAAGATAA AGAAGAAGAT GAAGGCCACG ATTGACCCGA TCGAGAGGAA GCTCCTCGAT
1441   TACAGGCAGA GGGCCATCAA GATCCTGGCA AACAGCTACT ACGGTTACTA CGGCTATGCA
1501   AGGGCGCGCT GGTACTGCAA GGAGTGTGCA GAGAGCGTAA CGGCCTGGGG AAGGGAGTAC
1561   ATAACGATGA CCATCAAGGA GATAGAGGAA AAGTACGGCT TTAAGGTAAT CTACAGCGAC
1621   ACCGACGGAT TTTTTGCCAC AATACCTGGA GCCGATGCTG AAACCGTCAA AAAGAAGGCT
1681   ATGGAGTTCC TCAAGTATAT CAACGCCAAA CTTCCGGGCG CGCTTGAGCT CGAGTACGAG
1741   GGCTTCTACA AACGCGGCTT CTTCGTCACG AAGAAGAAGT ATGCGGTGAT AGACGAGGAA
1801   GGCAAGATAA CAACGCGCGG ACTTGAGATT GTGAGGCGTG ACTGGAGCGA GATAGCGAAA
1861   GAGACGCAGG CGAGGGTTCT TGAAGCTTTG CTAAAGGACG GTGACGTCGA GAAGGCCGTG
1921   AGGATAGTCA AAGAAGTTAC CGAAAAGCTG AGCAAGTACG AGGTTCCGCC GGAGAAGCTG
1981   GTGATCCACG AGCAGATAAC GAGGGATTTA AAGGACTACA AGGCAACCGG TCCCCACGTT
2041   GCCGTTGCCA AGAGGTTGGC CGCGAGAGGA GTCAAAATAC GCCCTGGAAC GGTGATAAGC
2101   TACATCGTGC TCAAGGGCTC TGGGAGGATA GGCGACAGGG CGATACCGTT CGACGAGTTC
2161   GACCCGACGA AGCACAAGTA CGACGCCGAG TACTACATTG AGAACCAGGT TCTCCCAGCC
2221   GTTGAGAGAA TTCTGAGAGC CTTCGGTTAC CGCAAGGAAG ACCTGCGCTA CCAGAAGACG
2281   AGACAGGTTG GTTTGAGTGC TTGGCTGAAG CCGAAGGGAA CTTGA
```

Codon optimized sequences of Pfu and KOD

Sequence 3 (SEQ ID NO:3)

```
>Pfu codon optimized nucleotide sequence
1      ATGATTCTGG ATGTGGACTA TATCACCGAA GAGGGCAAAC CGGTTATACG TTTATTTAAG
61     AAAGAGAATG GTAAATTCAA GATCGAGCAT GACCGCACGT TCCGTCCATA CATTTACGCG
121    TTGCTTCGGG ATGATAGCAA AATTGAGGAA GTCAAAAAGA TCACCGGGGA ACGTCATGGA
181    AAAATAGTAA GAATTGTGGA CGTTGAAAAA GTCGAAAAGA AATTTCTGGG CAAACCGATC
241    ACTGTATGGA AGCTCTATCT GGAACATCCT CAGGATGTGC CCACAATTCG AGAAAAAGTT
301    CGTGAGCACC CAGCCGTCGT GGATATATTT GAATATGACA TCCCTTTTGC AAAACGCTAC
361    TTAATTGATA AAGGCCTGAT CCCGATGGAG GGGGAAGAAG AACTTAAAAT TCTGGCTTTT
421    GACATAGAAA CGCTCTATCA TGAGGGAGAA GAATTTGGCA AAGGTCCCAT CATTATGATT
481    TCTTACGCGG ATGAGAACGA AGCCAAGGTA ATCACTTGGA AAAATATTGA CCTGCCGTAC
541    GTTGAAGTGG TCAGTTCAGA GCGGGAAATG ATTAAACGTT TTTTACGCAT CATTAGAGAG
601    AAAGATCCAG ATATAATCGT TACATATAAC GGCGACTCCT TCGATTTTCC TTACCTGGCA
661    AAACGAGCTG AAAAATTGGG TATTAAACTT ACCATCGGGC GTGACGGATC GGAACCGAAA
721    ATGCAACGCA TTGGCGATAT GACGGCGGTA GAGGTGAAAG GTCGGATACA CTTTGATCTG
781    TATCATGTCA TCACCCGTAC TATTAATCTC CCCACATACA CGTTAGAAGC CGTTTATGAG
841    GCAATATTCG GCAAGCCGAA AGAAAAAGTG TACGCTGACG AAATCGCGAA GGCATGGGAG
901    AGCGGCGAAA ACCTGGAGCG CGTAGCAAAA TATTCTATGG AAGATGCTAA AGCGACCTAC
961    GAATTGGGGA AAGAATTTCT TCCAATGGAA ATTCAGCTGA GTCGTTTAGT CGGACAACCT
1021   CTGTGGGACG TTTCACGCTC CTCGACTGGC AATCTCGTGG AGTGGTTCCT GTTGAGAAAA
1081   GCCTATGAAC GAAACGAAGT AGCACCGAAT AAACCAAGCG AGGAAGAATA TCAGCGTCGC
1141   CTTCGCGAGT CTTACACAGG TGGGTTTGTT AAGGAACCGG AGAAAGGTCT TTGGGAAAAC
1201   ATCGTGTATT TAGATTTCCG TGCGCTGTAC CCCAGTATTA TAATCACCCA CAATGTCTCA
1261   CCTGACACGC TCAACTTGGA AGGTTGCAAA AATTATGATA TTGCTCCGCA AGTTGGACAT
1321   AAGTTTTGTA AAGATATTCC GGGCTTCATC CCGTCCCTGC TTGGTCACTT ACTGGAAGAG
1381   CGCCAAAAAA TTAAGACCAA AATGAAAGAG ACTCAGGATC CCATTGAAAA GATCCTGCTC
1441   GATTACCGGC AAAAAGCCAT TAAATTGCTT GCAAACTCGT TTTATGGGTA CTATGGCTAT
1501   GCGAAGGCTC GTTGGTACTG CAAAGAATGT GCCGAGAGCG TGACAGCATG GGGTCGCAAA
1561   TATATAGAAT TAGTATGGAA GGAGCTGGAA GAAAAATTCG GATTCAAAGT CCTGTACATC
1621   GATACGGATG GCCTCTATGC GACCATTCCT GGTGGGGAGT CTGAAGAAAT CAAGAAAAAA
1681   GCCTTGGAAT TCGTTAAGTA CATTAATAGT AAATTACCGG GACTGCTTGA ACTGGAGTAT
1741   GAAGGCTTCT ACAAAAGAGG TTTTTTCGTT ACTAAGAAAC GATATGCCGT AATAGATGAA
1801   GAGGGGAAAG TCATCACACG TGGCCTCGAG ATTGTTCGCC GGGACTGGTC AGAGATAGCA
1861   AAGGAAACGC AGGCGCGCGT GCTCGAAACC ATCTTGAAAC ATGGTGATGT AGAGGAAGCC
1921   GTCCGCATTG TTAAAGAGGT GATCCAGAAG TTAGCAAACT ATGAAATTCC ACCGGAAAAA
1981   CTGGCGATAT ACGAGCAAAT CACTCGTCCC CTTCACGAAT ATAAAGCTAT TGGACCTCAT
2041   GTAGCCGTCG CGAAGAAACT GGCTGCAAAA GGCGTTAAGA TAAAACCAGG TATGGTGATC
2101   GGGTACATTG TACTCCGCGG CGACGGTCCG ATTTCCAATA GAGCCATCTT GGCGGAGGAA
2161   TATGATCCTA AAAAGCATAA ATACGACGCT GAATATTACA TTGAGAACCA GGTCTTGCCG
2221   GCAGTTCTGC GGATACTTGA AGGATTTGGC TATCGTAAAG AAGATCTGCG CTATCAAAAG
2281   ACGCGACAGG TGGGTCTGAC TAGCTGGTTG AATATCAAAA AATCGTAA
```

Sequence 4 (SEQ ID NO:4)

```
>Pfu codon optimized nucleotide sequence, extra 9 nt in 5' area.
1       ATGGCTAGCG CCATTCTGGA TGTGGACTAT ATCACCGAAG AGGGCAAACC GGTTATACGT
61      TTATTTAAGA AAGAGAATGG TAAATTCAAG ATCGAGCATG ACCGCACGTT CCGTCCATAC
121     ATTTACGCGT TGCTTCGGGA TGATAGCAAA ATTGAGGAAG TCAAAAAGAT CACCGGGGAA
181     CGTCATGGAA AAATAGTAAG AATTGTGGAC GTTGAAAAAG TCGAAAAGAA ATTTCTGGGC
241     AAACCGATCA CTGTATGGAA GCTCTATCTG GAACATCCTC AGGATGTGCC CACAATTCGA
301     GAAAAAGTTC GTGAGCACCC AGCCGTCGTG GATATATTTG AATATGACAT CCCTTTTGCA
361     AAACGCTACT TAATTGATAA AGGCCTGATC CCGATGGAGG GGGAAGAAGA ACTTAAAATT
421     CTGGCTTTTG ACATAGAAAC GCTCTATCAT GAGGGAGAAG AATTTGGCAA AGGTCCCATC
481     ATTATGATTT CTTACGCGGA TGAGAACGAA GCCAAGGTAA TCACTTGGAA AAATATTGAC
541     CTGCCGTACG TTGAAGTGGT CAGTTCAGAG CGGGAAATGA TTAAACGTTT TTTACGCATC
601     ATTAGAGAGA AAGATCCAGA TATAATCGTT ACATATAACG GCGACTCCTT CGATTTTCCT
661     TACCTGGCAA AACGAGCTGA AAAATTGGGT ATTAAACTTA CCATCGGGCG TGACGGATCG
721     GAACCGAAAA TGCAACGCAT TGGCGATATG ACGGCGGTAG AGGTGAAAGG TCGGATACAC
781     TTTGATCTGT ATCATGTCAT CACCCGTACT ATTAATCTCC CCACATACAC GTTAGAAGCC
841     GTTTATGAGG CAATATTCGG CAAGCCGAAA GAAAAAGTGT ACGCTGACGA AATCGCGAAG
901     GCATGGGAGA GCGGCGAAAA CCTGGAGCGC GTAGCAAAAT ATTCTATGGA AGATGCTAAA
961     GCGACCTACG AATTGGGGAA AGAATTTCTT CCAATGGAAA TTCAGCTGAG TCGTTTAGTC
1021    GGACAACCTC TGTGGGACGT TTCACGCTCC TCGACTGGCA ATCTCGTGGA GTGGTTCCTG
1081    TTGAGAAAAG CCTATGAACG AAACGAAGTA GCACCGAATA AACCAAGCGA GGAAGAATAT
1141    CAGCGTCGCC TTCGCGAGTC TTACACAGGT GGGTTTGTTA AGGAACCGGA GAAAGGTCTT
1201    TGGGAAAACA TCGTGTATTT AGATTTCCGT GCGCTGTACC CCAGTATTAT AATCACCCAC
1261    AATGTCTCAC CTGACACGCT CAACTTGGAA GGTTGCAAAA ATTATGATAT TGCTCCGCAA
1321    GTTGGACATA AGTTTTGTAA AGATATTCCG GGCTTCATCC CGTCCCTGCT TGGTCACTTA
1381    CTGGAAGAGC GCCAAAAAAT TAAGACCAAA ATGAAAGAGA CTCAGGATCC CATTGAAAAG
1441    ATCCTGCTCG ATTACCGGCA AAAAGCCATT AAATTGCTTG CAAACTCGTT TTATGGGTAC
1501    TATGGCTATG CGAAGGCTCG TTGGTACTGC AAAGAATGTG CCGAGAGCGT GACAGCATGG
1561    GGTCGCAAAT ATATAGAATT AGTATGGAAG GAGCTGGAAG AAAAATTCGG ATTCAAAGTC
1621    CTGTACATCG ATACGGATGG CCTCTATGCG ACCATTCCTG GTGGGGAGTC TGAAGAAATC
1681    AAGAAAAAAG CCTTGGAATT CGTTAAGTAC ATTAATAGTA AATTACCGGG ACTGCTTGAA
1741    CTGGAGTATG AAGGCTTCTA CAAAAGAGGT TTTTTCGTTA CTAAGAAACG ATATGCCGTA
1801    ATAGATGAAG AGGGGAAAGT CATCACACGT GGCCTCGAGA TTGTTCGCCG GGACTGGTCA
1861    GAGATAGCAA AGGAAACGCA GGCGCGCGTG CTCGAAACCA TCTTGAAACA TGGTGATGTA
1921    GAGGAAGCCG TCCGCATTGT TAAAGAGGTG ATCCAGAAGT TAGCAAACTA TGAAATTCCA
1981    CCGGAAAAAC TGGCGATATA CGAGCAAATC ACTCGTCCCC TTCACGAATA TAAAGCTATT
2041    GGACCTCATG TAGCCGTCGC GAAGAAACTG GCTGCAAAAG GCGTTAAGAT AAAACCAGGT
2101    ATGGTGATCG GGTACATTGT ACTCCGCGGC GACGGTCCGA TTTCCAATAG AGCCATCTTG
2161    GCGGAGGAAT ATGATCCTAA AAAGCATAAA TACGACGCTG AATATTACAT TGAGAACCAG
2221    GTCTTGCCGG CAGTTCTGCG GATACTTGAA GGATTTGGCT ATCGTAAAGA AGATCTGCGC
2281    TATCAAAAGA CGCGACAGGT GGGTCTGACT AGCTGGTTGA ATATCAAAAA ATCGTAA
```

Sequence 5 (SEQ ID NO:5)

```
>KOD codon optimized nucleotide sequence
1       ATGATTCTGG ATACCGACTA TATCACGGAA GATGGCAAAC CGGTGATACG TATTTTTAAG
61      AAAGAGAATG GTGAGTTCAA AATCGAGTAC GACCGCACTT TTGAGCCATA TTTCTACGCG
121     TTACTGAAGG ACGATAGCGC CATTGAAGAA GTTAAAAAAA TCACCGCAGA GCGGCATGGG
181     ACAGTGGTAA CCGTGAAGAG AGTTGAAAAA GTCCAGAAAA AATTTTTGGG ACGACCTGTA
241     GAAGTGTGGA AACTTTATTT CACTCACCCC CAAGATGTTC CGGCTATACG TGATAAAATT
301     CGCGAACATC CAGCGGTCAT TGATATTTAC GAATATGATA TACCTTTTGC CAAGCGTTAC
361     CTCATCGACA AAGGCCTGGT GCCGATGGAA GGTGATGAAG AATTAAAAAT GTTGGCATTC
421     GACATTGAAA CACTTTATCA CGAGGGGGAA GAGTTTGCTG AGGGTCCCAT CCTGATGATT
481     TCTTATGCGG ATGAAGAGGG TGCCCGCGTA ATAACCTGGA AGAACGTTGA TCTCCCGTAC
541     GTGGACGTCG TTAGTACGGA ACGGGAAATG ATCAAACGTT TCCTGCGCGT AGTGAAAGAG
601     AAAGATCCAG ACGTCTTAAT TACCTATAAT GGTGATAACT TTGATTTTGC ATACCTGAAA
661     AAAAGATGCG AAAAGTTGGG CATAAATTTC GCTCTTGGTC GAGACGGGTC AGAGCCTAAA
721     ATCCAGCGTA TGGGAGATCG CTTTGCGGTT GAAGTGAAAG GCCGGATTCA TTTCGACCTG
781     TATCCGGTAA TTCGTCGCAC TATCAACCTC CCCACATACA CGTTAGAAGC CGTCTATGAG
841     GCAGTTTTTG GTCAACCGAA GGAAAAAGTT TACGCTGAGG AAATTACCAC TGCGTGGGAA
901     ACAGGCGAGA ATCTGGAACG TGTAGCCCGC TATTCTATGG AGGATGCAAA AGTTACCTAT
961     GAATTGGGTA AGGAATTTCT TCCAATGGAG GCGCAGCTGT CGAGATTAAT AGGGCAGAGC
1021    CTGTGGGACG TGTCTCGAAG TTCAACGGGA AACCTCGTCG AATGGTTTCT GTTGCGGAAA
1081    GCATACGAGC GTAATGAACT TGCCCCTAAC AAACCGGATG AAAAGGAGCT GGCACGCCGT
1141    CGCCAATCCT ATGAAGGCGG TTACGTTAAA GAACCAGAGC GGGGGTTATG GGAAAATATC
1201    GTGTATCTGG ATTTCCGTTC GCTCTACCCG AGCATTATCA TTACCCACAA CGTATCTCCC
```

```
1261   GACACTTTGA ATCGCGAGGG CTGTAAAGAA TATGATGTCG CGCCGCAGGT TGGTCATAGA
1321   TTTTGCAAGG ACTTCCCGGG ATTTATACCA AGTCTGCTTG GCGATTTACT GGAAGAGCGA
1381   CAAAAAATCA AAAAGAAAAT GAAAGCTACA ATCGATCCGA TAGAACGTAA GCTGCTCGAC
1441   TACCGCCAGC GGGCCATCAA AATTTTGGCA AACTCATATT ATGGTTACTA TGGGTACGCG
1501   CGTGCTCGCT GGTATTGTAA AGAGTGCGCC GAATCCGTGA CGGCATGGGG CCGTGAATAC
1561   ATCACCATGA CTATTAAGGA GATAGAAGAG AAATATGGTT TCAAAGTAAT CTACTCGGAT
1621   ACAGACGGAT TCTTTGCGAC GATTCCCGGT GCCGATGCAG AAACCGTCAA GAAAAAAGCG
1681   ATGGAATTCC TTAAGTATAT AAATGCTAAA TTACCTGGTG CCCTGGAGCT GGAATACGAA
1741   GGGTTTTACA AACGCGGATT CTTTGTTACT AAGAAAAAAT ATGCGGTGAT CGACGAGGAA
1801   GGCAAGATTA CGACCAGAGG CCTCGAGATT GTACGGCGTG ATTGGAGCGA AATCGCTAAA
1861   GAAACACAGG CACGTGTCTT GGAGGCATTA CTGAAAGATG GGGACGTTGA AAAGGCGGTG
1921   CGAATTGTAA AAGAAGTCAC CGAAAAACTT TCTAAGTACG AAGTTCCGCC AGAGAAACTG
1981   GTGATACACG AACAAATCAC TCGTGATCTG AAAGACTATA AGGCTACAGG CCCGCATGTA
2041   GCAGTCGCCA AACGCCTCGC GGCTCGGGGT GTTAAAATTC GTCCCGGAAC GGTGATCAGT
2101   TACATTGTAT TGAAGGGCTC AGGTCGCATA GGGGATAGAG CAATCCCTTT CGACGAGTTT
2161   GATCCAACCA AACACAAATA TGATGCCGAA TACTATATTG AAAACCAGGT CTTGCCGGCG
2221   GTTGAGCGTA TACTGCGCGC TTTCGGCTAT CGAAAGGAAG ATCTTCGTTA CCAAAAAACT
2281   AGACAGGTGG GTCTGTCCGC ATGGCTCAAA CCTAAGGGAA CGTAA
```

Sequence 6 (SEQ ID NO:6)

```
>KOD codon optimized nucleotide sequence, extra 9 nt in 5' area.
1      ATGGCTAGCG CCATTCTGGA TACCGACTAT ATCACGGAAG ATGGCAAACC GGTGATACGT
61     ATTTTTAAGA AAGAGAATGG TGAGTTCAAA ATCGAGTACG ACCGCACTTT TGAGCCATAT
121    TTCTACGCGT TACTGAAGGA CGATAGCGCC ATTGAAGAAG TTAAAAAAAT CACCGCAGAG
181    CGGCATGGGA CAGTGGTAAC CGTGAAGAGA GTTGAAAAAG TCCAGAAAAA ATTTTTGGGA
241    CGACCTGTAG AAGTGTGGAA ACTTTATTTC ACTCACCCCC AAGATGTTCC GGCTATACGT
301    GATAAAATTC GCGAACATCC AGCGGTCATT GATATTTACG AATATGATAT ACCTTTTGCC
361    AAGCGTTACC TCATCGACAA AGGCCTGGTG CCGATGGAAG GTGATGAAGA ATTAAAAATG
421    TTGGCATTCG ACATTGAAAC ACTTTATCAC GAGGGGGAAG AGTTTGCTGA GGGTCCCATC
481    CTGATGATTT CTTATGCGGA TGAAGAGGGT GCCCGCGTAA TAACCTGGAA GAACGTTGAT
541    CTCCCGTACG TGGACGTCGT TAGTACGGAA CGGGAAATGA TCAAACGTTT CCTGCGCGTA
601    GTGAAAGAGA AAGATCCAGA CGTCTTAATT ACCTATAATG GTGATAACTT TGATTTTGCA
661    TACCTGAAAA AAAGATGCGA AAAGTTGGGC ATAAATTTCG CTCTTGGTCG AGACGGGTCA
721    GAGCCTAAAA TCCAGCGTAT GGGAGATCGC TTTGCGGTTG AAGTGAAAGG CCGGATTCAT
781    TTCGACCTGT ATCCGGTAAT TCGTCGCACT ATCAACCTCC CCACATACAC GTTAGAAGCC
841    GTCTATGAGG CAGTTTTTGG TCAACCGAAG GAAAAAGTTT ACGCTGAGGA AATTACCACT
901    GCGTGGGAAA CAGGCGAGAA TCTGGAACGT GTAGCCCGCT ATTCTATGGA GGATGCAAAA
961    GTTACCTATG AATTGGGTAA GGAATTTCTT CCAATGGAGG CGCAGCTGTC GAGATTAATA
1021   GGGCAGAGCC TGTGGGACGT GTCTCGAAGT TCAACGGGAA ACCTCGTCGA ATGGTTTCTG
1081   TTGCGGAAAG CATACGAGCG TAATGAACTT GCCCCTAACA AACCGGATGA AAAGGAGCTG
1141   GCACGCCGTC GCCAATCCTA TGAAGGCGGT TACGTTAAAG AACCAGAGCG GGGGTTATGG
1201   GAAAATATCG TGTATCTGGA TTTCCGTTCG CTCTACCCGA GCATTATCAT TACCCACAAC
1261   GTATCTCCCG ACACTTTGAA TCGCGAGGGC TGTAAAGAAT ATGATGTCGC GCCGCAGGTT
1321   GGTCATAGAT TTTGCAAGGA CTTCCCGGGA TTTATACCAA GTCTGCTTGG CGATTTACTG
1381   GAAGAGCGAC AAAAAATCAA AAAGAAAATG AAAGCTACAA TCGATCCGAT AGAACGTAAG
1441   CTGCTCGACT ACCGCCAGCG GGCCATCAAA ATTTTGGCAA ACTCATATTA TGGTTACTAT
1501   GGGTACGCGC GTGCTCGCTG GTATTGTAAA GAGTGCGCCG AATCCGTGAC GGCATGGGGC
1561   CGTGAATACA TCACCATGAC TATTAAGGAG ATAGAAGAGA AATATGGTTT CAAAGTAATC
1621   TACTCGGATA CAGACGGATT CTTTGCGACG ATTCCCGGTG CCGATGCAGA AACCGTCAAG
1681   AAAAAAGCGA TGGAATTCCT TAAGTATATA AATGCTAAAT TACCTGGTGC CCTGGAGCTG
1741   GAATACGAAG GGTTTTACAA ACGCGGATTC TTTGTTACTA AGAAAAAATA TGCGGTGATC
1801   GACGAGGAAG GCAAGATTAC GACCAGAGGC CTCGAGATTG TACGGCGTGA TTGGAGCGAA
1861   ATCGCTAAAG AAACACAGGC ACGTGTCTTG GAGGCATTAC TGAAAGATGG GGACGTTGAA
1921   AAGGCGGTGC GAATTGTAAA AGAAGTCACC GAAAAACTTT CTAAGTACGA AGTTCCGCCA
1981   GAGAAACTGG TGATACACGA ACAAATCACT CGTGATCTGA AAGACTATAA GGCTACAGGC
2041   CCGCATGTAG CAGTCGCCAA ACGCCTCGCG GCTCGGGGTG TTAAAATTCG TCCCGGAACG
2101   GTGATCAGTT ACATTGTATT GAAGGGCTCA GGTCGCATAG GGGATAGAGC AATCCCTTTC
2161   GACGAGTTTG ATCCAACCAA ACACAAATAT GATGCCGAAT ACTATATTGA AAACCAGGTC
2221   TTGCCGGCGG TTGAGCGTAT ACTGCGCGCT TTCGGCTATC GAAAGGAAGA TCTTCGTTAC
2281   CAAAAAACTA GACAGGTGGG TCTGTCCGCA TGGCTCAAAC CTAAGGGAAC GTAA
```

Sequence 7 (SEQ ID NO:7)

```
>pKB13 - Pfu codon optimized nucleotide sequence in pUC19 vector
1      TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG GAGACGGTCA
61     CAGCTTGTCT GTAAGCGGAT GCCGGGAGCA GACAAGCCCG TCAGGGCGCG TCAGCGGGTG
121    TTGGCGGGTG TCGGGGCTGG CTTAACTATG CGGCATCAGA GCAGATTGTA CTGAGAGTGC
181    ACCATATGCG GTGTGAAATA CCGCACAGAT GCGTAAGGAG AAAATACCGC ATCAGGCGCC
```

241 ATTCGCCATT CAGGCTGCGC AACTGTTGGG AAGGGCGATC GGTGCGGGCC TCTTCGCTAT
301 TACGCCAGCT GGCGAAAGGG GGATGTGCTG CAAGGCGATT AAGTTGGGTA ACGCCAGGGT
361 TTTCCCAGTC ACGACGTTGT AAAACGACGG CCAGTGAATT CGGTCTCAGC GCCATTCTGG
421 ATACCGACTA TATCACGGAA GATGGCAAAC CGGTGATACG TATTTTTAAG AAAGAGAATG
481 GTGAGTTCAA AATCGAGTAC GACCGCACTT TTGAGCCATA TTTCTACGCG TTACTGAAGG
541 ACGATAGCGC CATTGAAGAA GTTAAAAAAA TCACCGCAGA GCGGCATGGG ACAGTGGTAA
601 CCGTGAAGAG AGTTGAAAAA GTCCAGAAAA AATTTTTGGG ACGACCTGTA GAAGTGTGGA
661 AACTTTATTT CACTCACCCC CAAGATGTTC CGGCTATACG TGATAAAATT CGCGAACATC
721 CAGCGGTCAT TGATATTTAC GAATATGATA TACCTTTTGC CAAGCGTTAC CTCATCGACA
781 AAGGCCTGGT GCCGATGGAA GGTGATGAAG AATTAAAAAT GTTGGCATTC GACATTGAAA
841 CACTTTATCA CGAGGGGGAA GAGTTTGCTG AGGGTCCCAT CCTGATGATT TCTTATGCGG
901 ATGAAGAGGG TGCCCGCGTA ATAACCTGGA AGAACGTTGA TCTCCCGTAC GTGGACGTCG
961 TTAGTACGGA ACGGGAAATG ATCAAACGTT TCCTGCGCGT AGTGAAAGAG AAAGATCCAG
1021 ACGTCTTAAT TACCTATAAT GGTGATAACT TTGATTTTGC ATACCTGAAA AAAAGATGCG
1081 AAAAGTTGGG CATAAATTTC GCTCTTGGTC GAGACGGGTC AGAGCCTAAA ATCCAGCGTA
1141 TGGGAGATCG CTTTGCGGTT GAAGTGAAAG GCCGGATTCA TTTCGACCTG TATCCGGTAA
1201 TTCGTCGCAC TATCAACCTC CCCACATACA CGTTAGAAGC CGTCTATGAG GCAGTTTTTG
1261 GTCAACCGAA GGAAAAAGTT TACGCTGAGG AAATTACCAC TGCGTGGGAA ACAGGCGAGA
1321 ATCTGGAACG TGTAGCCCGC TATTCTATGG AGGATGCAAA AGTTACCTAT GAATTGGGTA
1381 AGGAATTTCT TCCAATGGAG GCGCAGCTGT CGAGATTAAT AGGGCAGAGC CTGTGGGACG
1441 TGTCTCGAAG TTCAACGGGA AACCTCGTCG AATGGTTTCT GTTGCGGAAA GCATACGAGC
1501 GTAATGAACT TGCCCCTAAC AAACCGGATG AAAAGGAGCT GGCACGCCGT CGCCAATCCT
1561 ATGAAGGCGG TTACGTTAAA GAACCAGAGC GGGGGTTATG GGAAAATATC GTGTATCTGG
1621 ATTTCCGTTC GCTCTACCCG AGCATTATCA TTACCCACAA CGTATCTCCC GACACTTTGA
1681 ATCGCGAGGG CTGTAAAGAA TATGATGTCG CGCCGCAGGT TGGTCATAGA TTTTGCAAGG
1741 ACTTCCCGGG ATTTATACCA AGTCTGCTTG GCGATTTACT GGAAGAGCGA CAAAAAATCA
1801 AAAAGAAAAT GAAAGCTACA ATCGATCCGA TAGAACGTAA GCTGCTCGAC TACCGCCAGC
1861 GGGCCATCAA AATTTTGGCA AACTCATATT ATGGTTACTA TGGGTACGCG CGTGCTCGCT
1921 GGTATTGTAA AGAGTGCGCC GAATCCGTGA CGGCATGGGG CCGTGAATAC ATCACCATGA
1981 CTATTAAGGA GATAGAAGAG AAATATGGTT TCAAAGTAAT CTACTCGGAT ACAGACGGAT
2041 TCTTTGCGAC GATTCCCGGT GCCGATGCAG AAACCGTCAA GAAAAAAGCG ATGGAATTCC
2101 TTAAGTATAT AAATGCTAAA TTACCTGGTG CCCTGGAGCT GGAATACGAA GGGTTTTACA
2161 AACGCGGATT CTTTGTTACT AAGAAAAAAT ATGCGGTGAT CGACGAGGAA GGCAAGATTA
2221 CGACCAGAGG CCTCGAGATT GTACGGCGTG ATTGGAGCGA AATCGCTAAA GAAACACAGG
2281 CACGTGTCTT GGAGGCATTA CTGAAAGATG GGGACGTTGA AAAGGCGGTG CGAATTGTAA
2341 AAGAAGTCAC CGAAAAACTT TCTAAGTACG AAGTTCCGCC AGAGAAACTG GTGATACACG
2401 AACAAATCAC TCGTGATCTG AAAGACTATA AGGCTACAGG CCCGCATGTA GCAGTCGCCA
2461 AACGCCTCGC GGCTCGGGGT GTTAAAATTC GTCCCGGAAC GGTGATCAGT TACATTGTAT
2521 TGAAGGGCTC AGGTCGCATA GGGGATAGAG CAATCCCTTT CGACGAGTTT GATCCAACCA
2581 AACACAAATA TGATGCCGAA TACTATATTG AAAACCAGGT CTTGCCGGCG GTTGAGCGTA
2641 TACTGCGCGC TTTCGGCTAT CGAAAGGAAG ATCTTCGTTA CCAAAAAACT AGACAGGTGG
2701 GTCTGTCCGC ATGGCTCAAA CCTAAGGGAA CGTAATGATA TGAGACCGGA TCCTCTAGAG
2761 TCGACCTGCA GGCATGCAAG CTTGGCGTAA TCATGGTCAT AGCTGTTTCC TGTGTGAAAT
2821 TGTTATCCGC TCACAATTCC ACACAACATA CGAGCCGGAA GCATAAAGTG TAAAGCCTGG
2881 GGTGCCTAAT GAGTGAGCTA ACTCACATTA ATTGCGTTGC GCTCACTGCC CGCTTTCCAG
2941 TCGGGAAACC TGTCGTGCCA GCTGCATTAA TGAATCGGCC AACGCGCGGG GAGAGGCGGT
3001 TTGCGTATTG GGCGCTCTTC CGCTTCCTCG CTCACTGACT CGCTGCGCTC GGTCGTTCGG
3061 CTGCGGCGAG CGGTATCAGC TCACTCAAAG GCGGTAATAC GGTTATCCAC AGAATCAGGG
3121 GATAACGCAG GAAAGAACAT GTGAGCAAAA GGCCAGCAAA AGGCCAGGAA CCGTAAAAAG
3181 GCCGCGTTGC TGGCGTTTTT CCATAGGCTC CGCCCCCCTG ACGAGCATCA CAAAAATCGA
3241 CGCTCAAGTC AGAGGTGGCG AAACCCGACA GGACTATAAA GATACCAGGC GTTTCCCCCT
3301 GGAAGCTCCC TCGTGCGCTC TCCTGTTCCG ACCCTGCCGC TTACCGGATA CCTGTCCGCC
3361 TTTCTCCCTT CGGGAAGCGT GGCGCTTTCT CATAGCTCAC GCTGTAGGTA TCTCAGTTCG
3421 GTGTAGGTCG TTCGCTCCAA GCTGGGCTGT GTGCACGAAC CCCCCGTTCA GCCCGACCGC
3481 TGCGCCTTAT CCGGTAACTA TCGTCTTGAG TCCAACCCGG TAAGACACGA CTTATCGCCA
3541 CTGGCAGCAG CCACTGGTAA CAGGATTAGC AGAGCGAGGT ATGTAGGCGG TGCTACAGAG
3601 TTCTTGAAGT GGTGGCCTAA CTACGGCTAC ACTAGAAGAA CAGTATTTGG TATCTGCGCT
3661 CTGCTGAAGC CAGTTACCTT CGGAAAAAGA GTTGGTAGCT CTTGATCCGG CAAACAAACC
3721 ACCGCTGGTA GCGGTGGTTT TTTTGTTTGC AAGCAGCAGA TTACGCGCAG AAAAAAAGGA
3781 TCTCAAGAAG ATCCTTTGAT CTTTTCTACG GGGTCTGACG CTCAGTGGAA CGAAAACTCA
3841 CGTTAAGGGA TTTTGGTCAT GAGATTATCA AAAAGGATCT TCACCTAGAT CCTTTTAAAT
3901 TAAAAATGAA GTTTTAAATC AATCTAAAGT ATATATGAGT AAACTTGGTC TGACAGTTAC
3961 CAATGCTTAA TCAGTGAGGC ACCTATCTCA GCGATCTGTC TATTTCGTTC ATCCATAGTT
4021 GCCTGACTCC CCGTCGTGTA GATAACTACG ATACGGGAGG GCTTACCATC TGGCCCCAGT
4081 GCTGCAATGA TACCGCGAGA CCCACGCTCA CCGGCTCCAG ATTTATCAGC AATAAACCAG
4141 CCAGCCGGAA GGGCCGAGCG CAGAAGTGGT CCTGCAACTT TATCCGCCTC CATCCAGTCT
4201 ATTAATTGTT GCCGGGAAGC TAGAGTAAGT AGTTCGCCAG TTAATAGTTT GCGCAACGTT
4261 GTTGCCATTG CTACAGGCAT CGTGGTGTCA CGCTCGTCGT TTGGTATGGC TTCATTCAGC

```
4321   TCCGGTTCCC AACGATCAAG GCGAGTTACA TGATCCCCCA TGTTGTGCAA AAAAGCGGTT
4381   AGCTCCTTCG GTCCTCCGAT CGTTGTCAGA AGTAAGTTGG CCGCAGTGTT ATCACTCATG
4441   GTTATGGCAG CACTGCATAA TTCTCTTACT GTCATGCCAT CCGTAAGATG CTTTTCTGTG
4501   ACTGGTGAGT ACTCAACCAA GTCATTCTGA GAATAGTGTA TGCGGCGACC GAGTTGCTCT
4561   TGCCCGGCGT CAATACGGGA TAATACCGCG CCACATAGCA GAACTTTAAA AGTGCTCATC
4621   ATTGGAAAAC GTTCTTCGGG GCGAAAACTC TCAAGGATCT TACCGCTGTT GAGATCCAGT
4681   TCGATGTAAC CCACTCGTGC ACCCAACTGA TCTTCAGCAT CTTTTACTTT CACCAGCGTT
4741   TCTGGGTGAG CAAAAACAGG AAGGCAAAAT GCCGCAAAAA AGGGAATAAG GGCGACACGG
4801   AAATGTTGAA TACTCATACT CTTCCTTTTT CAATATTATT GAAGCATTTA TCAGGGTTAT
4861   TGTCTCATGA GCGGATACAT ATTTGAATGT ATTTAGAAAA ATAAACAAAT AGGGGTTCCG
4921   CGCACATTTC CCCGAAAAGT GCCACCTGAC GTCTAAGAAA CCATTATTAT CATGACATTA
4981   ACCTATAAAA ATAGGCGTAT CACGAGGCCC TTTCGTC
```

Sequence 8 (SEQ ID NO:8)

```
>pKB8 - KOD codon optimized nucleotide sequence in pUC19 vector
1      TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG GAGACGGTCA
61     CAGCTTGTCT GTAAGCGGAT GCCGGGAGCA GACAAGCCCG TCAGGGCGCG TCAGCGGGTG
121    TTGGCGGGTG TCGGGGCTGG CTTAACTATG CGGCATCAGA GCAGATTGTA CTGAGAGTGC
181    ACCATATGCG GTGTGAAATA CCGCACAGAT GCGTAAGGAG AAAATACCGC ATCAGGCGCC
241    ATTCGCCATT CAGGCTGCGC AACTGTTGGG AAGGGCGATC GGTGCGGGCC TCTTCGCTAT
301    TACGCCAGCT GGCGAAAGGG GGATGTGCTG CAAGGCGATT AAGTTGGGTA ACGCCAGGGT
361    TTTCCCAGTC ACGACGTTGT AAAACGACGG CCAGTGAATT CGGTCTCAGC GCCATTCTGG
421    ATACCGACTA TATCACGGAA GATGGCAAAC CGGTGATACG TATTTTTAAG AAAGAGAATG
481    GTGAGTTCAA AATCGAGTAC GACCGCACTT TTGAGCCATA TTTCTACGCG TTACTGAAGG
541    ACGATAGCGC CATTGAAGAA GTTAAAAAAA TCACCGCAGA GCGGCATGGG ACAGTGGTAA
601    CCGTGAAGAG AGTTGAAAAA GTCCAGAAAA AATTTTTGGG ACGACCTGTA GAAGTGTGGA
661    AACTTTATTT CACTCACCCC CAAGATGTTC CGGCTATACG TGATAAAATT CGCGAACATC
721    CAGCGGTCAT TGATATTTAC GAATATGATA TACCTTTTGC CAAGCGTTAC CTCATCGACA
781    AAGGCCTGGT GCCGATGGAA GGTGATGAAG AATTAAAAAT GTTGGCATTC GACATTGAAA
841    CACTTTATCA CGAGGGGGAA GAGTTTGCTG AGGGTCCCAT CCTGATGATT TCTTATGCGG
901    ATGAAGAGGG TGCCCGCGTA ATAACCTGGA AGAACGTTGA TCTCCCGTAC GTGGACGTCG
961    TTAGTACGGA ACGGGAAATG ATCAAACGTT TCCTGCGCGT AGTGAAAGAG AAAGATCCAG
1021   ACGTCTTAAT TACCTATAAT GGTGATAACT TTGATTTTGC ATACCTGAAA AAAAGATGCG
1081   AAAAGTTGGG CATAAATTTC GCTCTTGGTC GAGACGGGTC AGAGCCTAAA ATCCAGCGTA
1141   TGGGAGATCG CTTTGCGGTT GAAGTGAAAG GCCGGATTCA TTTCGACCTG TATCCGGTAA
1201   TTCGTCGCAC TATCAACCTC CCCACATACA CGTTAGAAGC CGTCTATGAG GCAGTTTTTG
1261   GTCAACCGAA GGAAAAAGTT TACGCTGAGG AAATTACCAC TGCGTGGGAA ACAGGCGAGA
1321   ATCTGGAACG TGTAGCCCGC TATTCTATGG AGGATGCAAA AGTTACCTAT GAATTGGGTA
1381   AGGAATTTCT TCCAATGGAG GCGCAGCTGT CGAGATTAAT AGGGCAGAGC CTGTGGGACG
1441   TGTCTCGAAG TTCAACGGGA AACCTCGTCG AATGGTTTCT GTTGCGGAAA GCATACGAGC
1501   GTAATGAACT TGCCCCTAAC AAACCGGATG AAAAGGAGCT GGCACGCCGT CGCCAATCCT
1561   ATGAAGGCGG TTACGTTAAA GAACCAGAGC GGGGGTTATG GGAAAATATC GTGTATCTGG
1621   ATTTCCGTTC GCTCTACCCG AGCATTATCA TTACCCACAA CGTATCTCCC GACACTTTGA
1681   ATCGCGAGGG CTGTAAAGAA TATGATGTCG CGCCGCAGGT TGGTCATAGA TTTTGCAAGG
1741   ACTTCCCGGG ATTTATACCA AGTCTGCTTG GCGATTTACT GGAAGAGCGA CAAAAAATCA
1801   AAAAGAAAAT GAAAGCTACA ATCGATCCGA TAGAACGTAA GCTGCTCGAC TACCGCCAGC
1861   GGGCCATCAA AATTTTGGCA AACTCATATT ATGGTTACTA TGGGTACGCG CGTGCTCGCT
1921   GGTATTGTAA AGAGTGCGCC GAATCCGTGA CGGCATGGGG CCGTGAATAC ATCACCATGA
1981   CTATTAAGGA GATAGAAGAG AAATATGGTT TCAAAGTAAT CTACTCGGAT ACAGACGGAT
2041   TCTTTGCGAC GATTCCCGGT GCCGATGCAG AAACCGTCAA GAAAAAAGCG ATGGAATTCC
2101   TTAAGTATAT AAATGCTAAA TTACCTGGTG CCCTGGAGCT GGAATACGAA GGGTTTTACA
2161   AACGCGGATT CTTTGTTACT AAGAAAAAAT ATGCGGTGAT CGACGAGGAA GGCAAGATTA
2221   CGACCAGAGG CCTCGAGATT GTACGGCGTG ATTGGAGCGA AATCGCTAAA GAAACACAGG
2281   CACGTGTCTT GGAGGCATTA CTGAAAGATG GGGACGTTGA AAAGGCGGTG CGAATTGTAA
2341   AAGAAGTCAC CGAAAAACTT TCTAAGTACG AAGTTCCGCC AGAGAAACTG GTGATACACG
2401   AACAAATCAC TCGTGATCTG AAAGACTATA AGGCTACAGG CCCGCATGTA GCAGTCGCCA
2461   AACGCCTCGC GGCTCGGGGT GTTAAAATTC GTCCCGGAAC GGTGATCAGT TACATTGTAT
2521   TGAAGGGCTC AGGTCGCATA GGGGATAGAG CAATCCCTTT CGACGAGTTT GATCCAACCA
2581   AACACAAATA TGATGCCGAA TACTATATTG AAAACCAGGT CTTGCCGGCG GTTGAGCGTA
2641   TACTGCGCGC TTTCGGCTAT CGAAAGGAAG ATCTTCGTTA CCAAAAAACT AGACAGGTGG
2701   GTCTGTCCGC ATGGCTCAAA CCTAAGGGAA CGTAATGATA TGAGACCGGA TCCTCTAGAG
2761   TCGACCTGCA GGCATGCAAG CTTGGCGTAA TCATGGTCAT AGCTGTTTCC TGTGTGAAAT
2821   TGTTATCCGC TCACAATTCC ACACAACATA CGAGCCGGAA GCATAAAGTG TAAAGCCTGG
2881   GGTGCCTAAT GAGTGAGCTA ACTCACATTA ATTGCGTTGC GCTCACTGCC CGCTTTCCAG
2941   TCGGGAAACC TGTCGTGCCA GCTGCATTAA TGAATCGGCC AACGCGCGGG GAGAGGCGGT
3001   TTGCGTATTG GGCGCTCTTC CGCTTCCTCG CTCACTGACT CGCTGCGCTC GGTCGTTCGG
3061   CTGCGGCGAG CGGTATCAGC TCACTCAAAG GCGGTAATAC GGTTATCCAC AGAATCAGGG
3121   GATAACGCAG GAAAGAACAT GTGAGCAAAA GGCCAGCAAA AGGCCAGGAA CCGTAAAAAG
```

```
3181   GCCGCGTTGC TGGCGTTTTT CCATAGGCTC CGCCCCCCTG ACGAGCATCA CAAAAATCGA
3241   CGCTCAAGTC AGAGGTGGCG AAACCCGACA GGACTATAAA GATACCAGGC GTTTCCCCCT
3301   GGAAGCTCCC TCGTGCGCTC TCCTGTTCCG ACCCTGCCGC TTACCGGATA CCTGTCCGCC
3361   TTTCTCCCTT CGGGAAGCGT GGCGCTTTCT CATAGCTCAC GCTGTAGGTA TCTCAGTTCG
3421   GTGTAGGTCG TTCGCTCCAA GCTGGGCTGT GTGCACGAAC CCCCCGTTCA GCCCGACCGC
3481   TGCGCCTTAT CCGGTAACTA TCGTCTTGAG TCCAACCCGG TAAGACACGA CTTATCGCCA
3541   CTGGCAGCAG CCACTGGTAA CAGGATTAGC AGAGCGAGGT ATGTAGGCGG TGCTACAGAG
3601   TTCTTGAAGT GGTGGCCTAA CTACGGCTAC ACTAGAAGAA CAGTATTTGG TATCTGCGCT
3661   CTGCTGAAGC CAGTTACCTT CGGAAAAAGA GTTGGTAGCT CTTGATCCGG CAAACAAACC
3721   ACCGCTGGTA GCGGTGGTTT TTTTGTTTGC AAGCAGCAGA TTACGCGCAG AAAAAAAGGA
3781   TCTCAAGAAG ATCCTTTGAT CTTTTCTACG GGGTCTGACG CTCAGTGGAA CGAAAACTCA
3841   CGTTAAGGGA TTTTGGTCAT GAGATTATCA AAAAGGATCT TCACCTAGAT CCTTTTAAAT
3901   TAAAAATGAA GTTTTAAATC AATCTAAAGT ATATATGAGT AAACTTGGTC TGACAGTTAC
3961   CAATGCTTAA TCAGTGAGGC ACCTATCTCA GCGATCTGTC TATTTCGTTC ATCCATAGTT
4021   GCCTGACTCC CCGTCGTGTA GATAACTACG ATACGGGAGG GCTTACCATC TGGCCCCAGT
4081   GCTGCAATGA TACCGCGAGA CCCACGCTCA CCGGCTCCAG ATTTATCAGC AATAAACCAG
4141   CCAGCCGGAA GGGCCGAGCG CAGAAGTGGT CCTGCAACTT TATCCGCCTC CATCCAGTCT
4201   ATTAATTGTT GCCGGGAAGC TAGAGTAAGT AGTTCGCCAG TTAATAGTTT GCGCAACGTT
4261   GTTGCCATTG CTACAGGCAT CGTGGTGTCA CGCTCGTCGT TTGGTATGGC TTCATTCAGC
4321   TCCGGTTCCC AACGATCAAG GCGAGTTACA TGATCCCCCA TGTTGTGCAA AAAAGCGGTT
4381   AGCTCCTTCG GTCCTCCGAT CGTTGTCAGA AGTAAGTTGG CCGCAGTGTT ATCACTCATG
4441   GTTATGGCAG CACTGCATAA TTCTCTTACT GTCATGCCAT CCGTAAGATG CTTTTCTGTG
4501   ACTGGTGAGT ACTCAACCAA GTCATTCTGA GAATAGTGTA TGCGGCGACC GAGTTGCTCT
4561   TGCCCGGCGT CAATACGGGA TAATACCGCG CCACATAGCA GAACTTTAAA AGTGCTCATC
4621   ATTGGAAAAC GTTCTTCGGG GCGAAAACTC TCAAGGATCT TACCGCTGTT GAGATCCAGT
4681   TCGATGTAAC CCACTCGTGC ACCCAACTGA TCTTCAGCAT CTTTTACTTT CACCAGCGTT
4741   TCTGGGTGAG CAAAAACAGG AAGGCAAAAT GCCGCAAAAA AGGGAATAAG GGCGACACGG
4801   AAATGTTGAA TACTCATACT CTTCCTTTTT CAATATTATT GAAGCATTTA TCAGGGTTAT
4861   TGTCTCATGA GCGGATACAT ATTTGAATGT ATTTAGAAAA ATAAACAAAT AGGGGTTCCG
4921   CGCACATTTC CCCGAAAAGT GCCACCTGAC GTCTAAGAAA CCATTATTAT CATGACATTA
4981   ACCTATAAAA ATAGGCGTAT CACGAGGCCC TTTCGTC
```

Amino acid sequences of Pfu and KOD

Sequence 9 (SEQ ID NO:9)

```
>Pfu amino acid sequence
1         MILDVDYITE EGKPVIRLFK KENGKFKIEH DRTFRPYIYA LLRDDSKIEE VKKITGERHG
61        KIVRIVDVEK VEKKFLGKPI TVWKLYLEHP QDVPTIREKV REHPAVVDIF EYDIPFAKRY
121       LIDKGLIPME GEEELKILAF DIETLYHEGE EFGKGPIIMI SYADENEAKV ITWKNIDLPY
181       VEVVSSEREM IKRFLRIIRE KDPDIIVTYN GDSFDFPYLA KRAEKLGIKL TIGRDGSEPK
241       MQRIGDMTAV EVKGRIHFDL YHVITRTINL PTYTLEAVYE AIFGKPKEKV YADEIAKAWE
301       SGENLERVAK YSMEDAKATY ELGKEFLPME IQLSRLVGQP LWDVSRSSTG NLVEWFLLRK
361       AYERNEVAPN KPSEEEYQRR LRESYTGGFV KEPEKGLWEN IVYLDFRALY PSIIITHNVS
421       PDTLNLEGCK NYDIAPQVGH KFCKDIPGFI PSLLGHLLEE RQKIKTKMKE TQDPIEKILL
481       DYRQKAIKLL ANSFYGYYGY AKARWYCKEC AESVTAWGRK YIELVWKELE EKFGFKVLYI
541       DTDGLYATIP GGESEEIKKK ALEFVKYINS KLPGLLELEY EGFYKRGFFV TKKRYAVIDE
601       EGKVITRGLE IVRRDWSEIA KETQARVLET ILKHGDVEEA VRIVKEVIQK LANYEIPPEK
661       LAIYEQITRP LHEYKAIGPH VAVAKKLAAK GVKIKPGMVI GYIVLRGDGP ISNRAILAEE
721       YDPKKHKYDA EYYIENQVLP AVLRILEGFG YRKEDLRYQK TRQVGLTSWL NIKKS*
```

Sequence 10 (SEQ ID NO:10)

```
>Pfu amino acid sequence, extra 3 aa in 5' area.
1         MASAILDVDY ITEEGKPVIR LFKKENGKFK IEHDRTFRPY IYALLRDDSK IEEVKKITGE
61        RHGKIVRIVD VEKVEKKFLG KPITVWKLYL EHPQDVPTIR EKVREHPAVV DIFEYDIPFA
121       KRYLIDKGLI PMEGEEELKI LAFDIETLYH EGEEFGKGPI IMISYADENE AKVITWKNID
181       LPYVEVVSSE REMIKRFLRI IREKDPDIIV TYNGDSFDFP YLAKRAEKLG IKLTIGRDGS
241       EPKMQRIGDM TAVEVKGRIH FDLYHVITRT INLPTYTLEA VYEAIFGKPK EKVYADEIAK
301       AWESGENLER VAKYSMEDAK ATYELGKEFL PMEIQLSRLV GQPLWDVSRS STGNLVEWFL
361       LRKAYERNEV APNKPSEEEY QRRLRESYTG GFVKEPEKGL WENIVYLDFR ALYPSIIITH
421       NVSPDTLNLE GCKNYDIAPQ VGHKFCKDIP GFIPSLLGHL LEERQKIKTK MKETQDPIEK
481       ILLDYRQKAI KLLANSFYGY YGYAKARWYC KECAESVTAW GRKYIELVWK ELEEKFGFKV
541       LYIDTDGLYA TIPGGESEEI KKKALEFVKY INSKLPGLLE LEYEGFYKRG FFVTKKRYAV
601       IDEEGKVITR GLEIVRRDWS EIAKETQARV LETILKHGDV EEAVRIVKEV IQKLANYEIP
661       PEKLAIYEQI TRPLHEYKAI GPHVAVAKKL AAKGVKIKPG MVIGYIVLRG DGPISNRAIL
721       AEEYDPKKHK YDAEYYIENQ VLPAVLRILE GFGYRKEDLR YQKTRQVGLT SWLNIKKS*
```

Sequence 11 (SEQ ID NO:11)

```
>KOD amino acid sequence
1      MILDTDYITE DGKPVIRIFK KENGEFKIEY DRTFEPYFYA LLKDDSAIEE VKKITAERHG
61     TVVTVKRVEK VQKKFLGRPV EVWKLYFTHP QDVPAIRDKI REHPAVIDIY EYDIPFAKRY
121    LIDKGLVPME GDEELKMLAF DIETLYHEGE EFAEGPILMI SYADEEGARV ITWKNVDLPY
181    VDVVSTEREM IKRFLRVVKE KDPDVLITYN GDNFDFAYLK KRCEKLGINF ALGRDGSEPK
241    IQRMGDRFAV EVKGRIHFDL YPVIRRTINL PTYTLEAVYE AVFGQPKEKV YAEEITTAWE
301    TGENLERVAR YSMEDAKVTY ELGKEFLPME AQLSRLIGQS LWDVSRSSTG NLVEWFLLRK
361    AYERNELAPN KPDEKELARR RQSYEGGYVK EPERGLWENI VYLDFRSLYP SIIITHNVSP
421    DTLNREGCKE YDVAPQVGHR FCKDFPGFIP SLLGDLLEER QKIKKKMKAT IDPIERKLLD
481    YRQRAIKILA NSYYGYYGYA RARWYCKECA ESVTAWGREY ITMTIKEIEE KYGFKVIYSD
541    TDGFFATIPG ADAETVKKKA MEFLKYINAK LPGALELEYE GFYKRGFFVT KKKYAVIDEE
601    GKITTRGLEI VRRDWSEIAK ETQARVLEAL LKDGDVEKAV RIVKEVTEKL SKYEVPPEKL
661    VIHEQITRDL KDYKATGPHV AVAKRLAARG VKIRPGTVIS YIVLKGSGRI GDRAIPFDEF
721    DPTKHKYDAE YYIENQVLPA VERILRAFGY RKEDLRYQKT RQVGLSAWLK PKGT
```

Sequence 12 (SEQ ID NO:12)

```
>KOD amino acid sequence, extra 3 aa in 5' area.
1          MASAILDTDY ITEDGKPVIR IFKKENGEFK IEYDRTFEPY FYALLKDDSA IEEVKKITAE
61         RHGTVVTVKR VEKVQKKFLG RPVEVWKLYF THPQDVPAIR DKIREHPAVI DIYEYDIPFA
121        KRYLIDKGLV PMEGDEELKM LAFDIETLYH EGEEFAEGPI LMISYADEEG ARVITWKNVD
181        LPYVDVVSTE REMIKRFLRV VKEKDPDVLI TYNGDNFDFA YLKKRCEKLG INFALGRDGS
241        EPKIQRMGDR FAVEVKGRIH FDLYPVIRRT INLPTYTLEA VYEAVFGQPK EKVYAEEITT
301        AWETGENLER VARYSMEDAK VTYELGKEFL PMEAQLSRLI GQSLWDVSRS STGNLVEWFL
361        LRKAYERNEL APNKPDEKEL ARRRQSYEGG YVKEPERGLW ENIVYLDFRS LYPSIIITHN
421        VSPDTLNREG CKEYDVAPQV GHRFCKDFPG FIPSLLGDLL EERQKIKKKM KATIDPIERK
481        LLDYRQRAIK ILANSYYGYY GYARARWYCK ECAESVTAWG REYITMTIKE IEEKYGFKVI
541        YSDTDGFFAT IPGADAETVK KKAMEFLKYI NAKLPGALEL EYEGFYKRGF FVTKKKYAVI
601        DEEGKITTRG LEIVRRDWSE IAKETQARVL EALLKDGDVE KAVRIVKEVT EKLSKYEVPP
661        EKLVIHEQIT RDLKDYKATG PHVAVAKRLA ARGVKIRPGT VISYIVLKGS GRIGDRAIPF
721        DEFDPTKHKY DAEYYIENQV LPAVERILRA FGYRKEDLRY QKTRQVGLSA WLKPKGT*
```

DNA sequences of chimeras Pod and Kofu

Sequence 13 (SEQ ID NO:13)

```
>Pod codon optimized nucleotide sequence
1      ATGGCTAGCG CCATTCTGGA TGTGGACTAT ATCACCGAAG AGGGCAAACC GGTTATACGT
61     TTATTTAAGA AAGAGAATGG TAAATTCAAG ATCGAGCATG ACCGCACGTT CCGTCCATAC
121    ATTTACGCGT TGCTTCGGGA TGATAGCAAA ATTGAGGAAG TCAAAAAGAT CACCGGGGAA
181    CGTCATGGAA AAATAGTAAG AATTGTGGAC GTTGAAAAAG TCGAAAAGAA ATTTCTGGGC
241    AAACCGATCA CTGTATGGAA GCTCTATCTG GAACATCCTC AGGATGTGCC CACAATTCGA
301    GAAAAAGTTC GTGAGCACCC AGCCGTCGTG GATATATTTG AATATGACAT CCCTTTTGCA
361    AAACGCTACT TAATTGATAA AGGCCTGATC CCGATGGAGG GGGAAGAAGA ACTTAAAATT
421    CTGGCTTTTG ACATAGAAAC GCTCTATCAT GAGGGAGAAG AATTTGGCAA AGGTCCCATC
481    ATTATGATTT CTTACGCGGA TGAGAACGAA GCCAAGGTAA TCACTTGGAA AAATATTGAC
541    CTGCCGTACG TTGAAGTGGT CAGTTCAGAG CGGGAAATGA TTAAACGTTT TTTACGCATC
601    ATTAGAGAGA AAGATCCAGA TATAATCGTT ACATATAACG GCGACTCCTT CGATTTTCCT
661    TACCTGGCAA AACGAGCTGA AAAATTGGGT ATTAAACTTA CCATCGGGCG TGACGGATCG
721    GAACCGAAAA TGCAACGCAT TGGCGATATG ACGGCGGTAG AGGTGAAAGG TCGGATACAC
781    TTTGATCTGT ATCATGTCAT CACCCGTACT ATTAATCTCC CCACATACAC GTTAGAAGCC
841    GTTTATGAGG CAATATTCGG CAAGCCGAAA GAAAAAGTGT ACGCTGACGA AATCGCGAAG
901    GCATGGGAGA GCGGCGAAAA CCTGGAGCGC GTAGCAAAAT ATTCTATGGA AGATGCTAAA
961    GCGACCTACG AATTGGGGAA AGAATTTCTT CCAATGGAAA TTCAGCTGTC GAGATTAATA
1021   GGGCAGAGCC TGTGGGACGT GTCTCGAAGT TCAACGGGAA ACCTCGTCGA ATGGTTTCTG
1081   TTGCGGAAAG CATACGAGCG TAATGAACTT GCCCCTAACA AACCGGATGA AAAGGAGCTG
1141   GCACGCCGTC GCCAATCCTA TGAAGGCGGT TACGTTAAAG AACCAGAGCG GGGGTTATGG
1201   GAAAATATCG TGTATCTGGA TTTCCGTTCG CTCTACCCGA GCATTATCAT TACCCACAAC
1261   GTATCTCCCG ACACTTTGAA TCGCGAGGGC TGTAAAGAAT ATGATGTCGC GCCGCAGGTT
1321   GGTCATAGAT TTTGCAAGGA CTTCCCGGGA TTTATACCAA GTCTGCTTGG CGATTTACTG
1381   GAAGAGCGAC AAAAAATCAA AAAGAAAATG AAAGCTACAA TCGATCCGAT AGAACGTAAG
1441   CTGCTCGACT ACCGCCAGCG GGCCATCAAA ATTTTGGCAA ACTCATATTA TGGTTACTAT
1501   GGGTACGCGC GTGCTCGCTG GTATTGTAAA GAGTGCGCCG AATCCGTGAC GGCATGGGGC
1561   CGTGAATACA TCACCATGAC TATTAAGGAG ATAGAAGAGA AATATGGTTT CAAAGTAATC
1621   TACTCGGATA CAGACGGATT CTTTGCGACG ATTCCCGGTG CCGATGCAGA AACCGTCAAG
1681   AAAAAAGCGA TGGAATTCGT TAAGTACATT AATAGTAAAT TACCGGGACT GCTTGAACTG
```

```
1741   GAGTATGAAG GCTTCTACAA AAGAGGTTTT TTCGTTACTA AGAAACGATA TGCCGTAATA
1801   GATGAAGAGG GGAAAGTCAT CACACGTGGC CTCGAGATTG TTCGCCGGGA CTGGTCAGAG
1861   ATAGCAAAGG AAACGCAGGC GCGCGTGCTC GAAACCATCT TGAAACATGG TGATGTAGAG
1921   GAAGCCGTCC GCATTGTTAA AGAGGTGATC CAGAAGTTAG CAAACTATGA AATTCCACCG
1981   GAAAAACTGG CGATATACGA GCAAATCACT CGTCCCCTTC ACGAATATAA AGCTATTGGA
2041   CCTCATGTAG CCGTCGCGAA GAAACTGGCT GCAAAAGGCG TTAAGATAAA ACCAGGTATG
2101   GTGATCGGGT ACATTGTACT CCGCGGCGAC GGTCCGATTT CCAATAGAGC CATCTTGGCG
2161   GAGGAATATG ATCCTAAAAA GCATAAATAC GACGCTGAAT ATTACATTGA GAACCAGGTC
2221   TTGCCGGCAG TTCTGCGGAT ACTTGAAGGA TTTGGCTATC GTAAAGAAGA TCTGCGCTAT
2281   CAAAAGACGC GACAGGTGGG TCTGACTAGC TGGTTGAATA TCAAAAAATC GTAA
```

Sequence 14 (SEQ ID NO:14)

```
>Kofu codon optimized nucleotide sequence
1      ATGGCTAGCG CCATTCTGGA TACCGACTAT ATCACGGAAG ATGGCAAACC GGTGATACGT
61     ATTTTTAAGA AAGAGAATGG TGAGTTCAAA ATCGAGTACG ACCGCACTTT TGAGCCATAT
121    TTCTACGCGT TACTGAAGGA CGATAGCGCC ATTGAAGAAG TTAAAAAAAT CACCGCAGAG
181    CGGCATGGGA CAGTGGTAAC CGTGAAGAGA GTTGAAAAAG TCCAGAAAAA ATTTTTGGGA
241    CGACCTGTAG AAGTGTGGAA ACTTTATTTC ACTCACCCCC AAGATGTTCC GGCTATACGT
301    GATAAAATTC GCGAACATCC AGCGGTCATT GATATTTACG AATATGATAT ACCTTTTGCC
361    AAGCGTTACC TCATCGACAA AGGCCTGGTG CCGATGGAAG GTGATGAAGA ATTAAAAATG
421    TTGGCATTCG ACATTGAAAC ACTTTATCAC GAGGGGGAAG AGTTTGCTGA GGGTCCCATC
481    CTGATGATTT CTTATGCGGA TGAAGAGGGT GCCCGCGTAA TAACCTGGAA GAACGTTGAT
541    CTCCCGTACG TGGACGTCGT TAGTACGGAA CGGGAAATGA TCAAACGTTT CCTGCGCGTA
601    GTGAAAGAGA AAGATCCAGA CGTCTTAATT ACCTATAATG GTGATAACTT TGATTTTGCA
661    TACCTGAAAA AAAGATGCGA AAAGTTGGGC ATAAATTTCG CTCTTGGTCG AGACGGGTCA
721    GAGCCTAAAA TCCAGCGTAT GGGAGATCGC TTTGCGGTTG AAGTGAAAGG CCGGATTCAT
781    TTCGACCTGT ATCCGGTAAT TCGTCGCACT ATCAACCTCC CCACATACAC GTTAGAAGCC
841    GTCTATGAGG CAGTTTTTGG TCAACCGAAG GAAAAAGTTT ACGCTGAGGA AATTACCACT
901    GCGTGGGAAA CAGGCGAGAA TCTGGAACGT GTAGCCCGCT ATTCTATGGA GGATGCAAAA
961    GTTACCTATG AATTGGGTAA GGAATTTCTT CCAATGGAGG CGCAGCTGAG TCGTTTAGTC
1021   GGACAACCTC TGTGGGACGT TTCACGCTCC TCGACTGGCA ATCTCGTGGA GTGGTTCCTG
1081   TTGAGAAAAG CCTATGAACG AAACGAAGTA GCACCGAATA AACCAAGCGA GGAAGAATAT
1141   CAGCGTCGCC TTCGCGAGTC TTACACAGGT GGGTTTGTTA AGGAACCGGA GAAAGGTCTT
1201   TGGGAAAACA TCGTGTATTT AGATTTCCGT GCGCTGTACC CCAGTATTAT AATCACCCAC
1261   AATGTCTCAC CTGACACGCT CAACTTGGAA GGTTGCAAAA ATTATGATAT TGCTCCGCAA
1321   GTTGGACATA AGTTTTGTAA AGATATTCCG GGCTTCATCC CGTCCCTGCT TGGTCACTTA
1381   CTGGAAGAGC GCCAAAAAAT TAAGACCAAA ATGAAAGAGA CTCAGGATCC CATTGAAAAG
1441   ATCCTGCTCG ATTACCGGCA AAAAGCCATT AAATTGCTTG CAAACTCGTT TTATGGGTAC
1501   TATGGCTATG CGAAGGCTCG TTGGTACTGC AAAGAATGTG CCGAGAGCGT GACAGCATGG
1561   GGTCGCAAAT ATATAGAATT AGTATGGAAG GAGCTGGAAG AAAAATTCGG ATTCAAAGTC
1621   CTGTACATCG ATACGGATGG CCTCTATGCG ACCATTCCTG GTGGGGAGTC TGAAGAAATC
1681   AAGAAAAAAG CCTTGGAATT CCTTAAGTAT ATAAATGCTA AATTACCTGG TGCCCTGGAG
1741   CTGGAATACG AAGGGTTTTA CAAACGCGGA TTCTTTGTTA CTAAGAAAAA ATATGCGGTG
1801   ATCGACGAGG AAGGCAAGAT TACGACCAGA GGCCTCGAGA TTGTACGGCG TGATTGGAGC
1861   GAAATCGCTA AAGAAACACA GGCACGTGTC TTGGAGGCAT TACTGAAAGA TGGGGACGTT
1921   GAAAAGGCGG TGCGAATTGT AAAAGAAGTC ACCGAAAAAC TTTCTAAGTA CGAAGTTCCG
1981   CCAGAGAAAC TGGTGATACA CGAACAAATC ACTCGTGATC TGAAAGACTA TAAGGCTACA
2041   GGCCCGCATG TAGCAGTCGC CAAACGCCTC GCGGCTCGGG GTGTTAAAAT TCGTCCCGGA
2101   ACGGTGATCA GTTACATTGT ATTGAAGGGC TCAGGTCGCA TAGGGGATAG AGCAATCCCT
2161   TTCGACGAGT TTGATCCAAC CAAACACAAA TATGATGCCG AATACTATAT TGAAAACCAG
2221   GTCTTGCCGG CGGTTGAGCG TATACTGCGC GCTTTCGGCT ATCGAAAGGA AGATCTTCGT
2281   TACCAAAAAA CTAGACAGGT GGGTCTGTCC GCATGGCTCA AACCTAAGGG AACGTAA
```

Amino acid sequences of chimeras Pod and Kofu

Sequence 15 (SEQ ID NO:15)

```
>Pod amino acid sequence
1          MASAILDVDY ITEEGKPVIR LFKKENGKFK IEHDRTFRPY IYALLRDDSK IEEVKKITGE
61         RHGKIVRIVD VEKVEKKFLG KPITVWKLYL EHPQDVPTIR EKVREHPAVV DIFEYDIPFA
121        KRYLIDKGLI PMEGEEELKI LAFDIETLYH EGEEFGKGPI IMISYADENE AKVITWKNID
181        LPYVEVVSSE REMIKRFLRI IREKDPDIIV TYNGDSFDFP YLAKRAEKLG IKLTIGRDGS
241        EPKMQRIGDM TAVEVKGRIH FDLYHVITRT INLPTYTLEA VYEAIFGKPK EKVYADEIAK
301        AWESGENLER VAKYSMEDAK ATYELGKEFL PMEIQLSRLI GQSLWDVSRS STGNLVEWFL
361        LRKAYERNEL APNKPDEKEL ARRRQSYEGG YVKEPERGLW ENIVYLDFRS LYPSIIITHN
```

```
421       VSPDTLNREG CKEYDVAPQV GHRFCKDFPG FIPSLLGDLL EERQKIKKKM KATIDPIERK
481       LLDYRQRAIK ILANSYYGYY GYARARWYCK ECAESVTAWG REYITMTIKE IEEKYGFKVI
541       YSDTDGFFAT IPGADAETVK KKAMEFVKYI NSKLPGLLEL EYEGFYKRGF FVTKKRYAVI
601       DEEGKVITRG LEIVRRDWSE IAKETQARVL ETILKHGDVE EAVRIVKEVI QKLANYEIPP
661       EKLAIYEQIT RPLHEYKAIG PHVAVAKKLA AKGVKIKPGM VIGYIVLRGD GPISNRAILA
721       EEYDPKKHKY DAEYYIENQV LPAVLRILEG FGYRKEDLRY QKTRQVGLTS WLNIKKS*
```

Sequence 16 (SEQ ID NO:16)

```
>Kofu amino acid sequence
1         MASAILDTDY ITEDGKPVIR IFKKENGEFK IEYDRTFEPY FYALLKDDSA IEEVKKITAE
61        RHGTVVTVKR VEKVQKKFLG RPVEVWKLYF THPQDVPAIR DKIREHPAVI DIYEYDIPFA
121       KRYLIDKGLV PMEGDEELKM LAFDIETLYH EGEEFAEGPI LMISYADEEG ARVITWKNVD
181       LPYVDVVSTE REMIKRFLRV VKEKDPDVLI TYNGDNFDFA YLKKRCEKLG INFALGRDGS
241       EPKIQRMGDR FAVEVKGRIH FDLYPVIRRT INLPTYTLEA VYEAVFGQPK EKVYAEEITT
301       AWETGENLER VARYSMEDAK VTYELGKEFL PMEAQLSRLV GQPLWDVSRS STGNLVEWFL
361       LRKAYERNEV APNKPSEEEY QRRLRESYTG GFVKEPEKGL WENIVYLDFR ALYPSIIITH
421       NVSPDTLNLE GCKNYDIAPQ VGHKFCKDIP GFIPSLLGHL LEERQKIKTK MKETQDPIEK
481       ILLDYRQKAI KLLANSFYGY YGYAKARWYC KECAESVTAW GRKYIELVWK ELEEKFGFKV
541       LYIDTDGLYA TIPGGESEEI KKKALEFLKY INAKLPGALE LEYEGFYKRG FFVTKKKYAV
601       IDEEGKITTR GLEIVRRDWS EIAKETQARV LEALLKDGDV EKAVRIVKEV TEKLSKYEVP
661       PEKLVIHEQI TRDLKDYKAT GPHVAVAKRL AARGVKIRPG TVISYIVLKG SGRIGDRAIP
721       FDEFDPTKHK YDAEYYIENQ VLPAVERILR AFGYRKEDLR YQKTRQVGLS AWLKPKGT*
```

Sequence 17 (SEQ ID NO:17)

>pLACIQZa

```
1         TCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCA
61        CAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTG
121       TTGGCGGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGC
181       ACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCC
241       ATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTAT

                                                                    GT
301       TACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGT

          TTTCCCAGTCACGAC >>> Primer M13-40 (SEQ ID NO:42)
361       TTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAATTCGAGCTCGGTACCCGGGGAT

            XbaI
421       CCTCTAGAGCCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACA
481       ATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTG
541       AGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCG
601       TGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGC
661       CAGGGTGGTTTTTCTTTTCACCAGTGAGACGGGCAACAGCTGATTGCCCTTCACCGCCTG
721       GCCCTGAGAGAGTTGCAGCAAGCGGTCCACGCTGGTTTGCCCCAGCAGGCGAAAATCCTG
781       TTTGATGGTGGTTGACGGCGGGATATAACATGAGCTGTCTTCGGTATCGTCGTATCCCAC
841       TACCGAGATATCCGCACCAACGCGCAGCCCGGACTCGGTAATGGCGCGCATTGCGCCCAG
901       CGCCATCTGATCGTTGGCAACCAGCATCGCAGTGGGAACGATGCCCTCATTCAGCATTTG
961       CATGGTTTGTTGAAAACCGGACATGGCACTCCAGTCGCCTTCCCGTTCCGCTATCGGCTG
1021      AATTTGATTGCGAGTGAGATATTTATGCCAGCCAGCCAGACGCAGACGCGCCGAGACAGA
1081      ACTTAATGGGCCCGCTAACAGCGCGATTTGCTGGTGACCCAATGCGACCAGATGCTCCAC
1141      GCCCAGTCGCGTACCGTCTTCATGGGAGAAAATAATACTGTTGATGGGTGTCTGGTCAGA
1201      GACATCAAGAAATAACGCCGGAACATTAGTGCAGGCAGCTTCCACAGCAATGGCATCCTG
1261      GTCATCCAGCGGATAGTTAATGATCAGCCCACTGACGCGTTGCGCGAGAAGATTGTGCAC
1321      CGCCGCTTTACAGGCTTCGACGCCGCTTCGTTCTACCATCGACACCACCACGCTGGCACC
1381      CAGTTGATCGGCGCGAGATTTAATCGCCGCGACAATTTGCGACGGCGCGTGCAGGGCCAG
1441      ACTGGAGGTGGCAACGCCAATCAGCAACGACTGTTTGCCCGCCAGTTGTTGTGCCACGCG
1501      GTTGGGAATGTAATTCAGCTCCGCCATCGCCGCTTCCACTTTTTCCCGCGTTTTCGCAGA
1561      AACGTGGCTGGCCTGGTTCACCACGCGGGAAACGGTCTGATAAGAGACACCGGCATACTC
1621      TGCGACATCGTATAACGTTACTGGTTTCACATTCACCACCCTGAATTGACTCTCTTCCGG
1681      GCGCTATCATGCCATACCGCGAAAGGTTTTGCGCCATTCGATGGTGTCAACGTAAATGCA

                                          NcoI
```

```
1741   TGCCGCTTCGCCTTCCGGCCACCAGAATAGCCTGCGCCATGGGCTTCCTCGCTCACTGAC
1801   TCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATA
1861   CGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAA
1921   AAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCT
1981   GACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAA
                                  PRIMER PKBLACIR <<< GCTGTCCTGATATT
       TCTATGG (SEQ ID NO:43)
2041   AGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCG
2101   CTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCA
2161   CGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAA
2221   CCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCG
2281   GTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGG
2341   TATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGA
2401   ACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGC
2461   TCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAG
2521   ATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGAC
2581   GCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATC
2641   TTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAG
2701   TAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGT
2761   CTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAG
2821   GGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCA
2881   GATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACT
2941   TTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCA
3001   GTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCG
3061   TTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCC
3121   ATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTG
3181   GCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCA
3241   TCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGT
3301   ATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGC
3361   AGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATC
3421   TTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCA
3481   TCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAA
3541   AAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTAT
3601   TGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAA
3661   AATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAA
3721   ACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTCGTC
```

Amino acid sequences of DNA polymerases from *T. litoralis*, *Thermococcus* sp. 9 degrees N-7 and chimeras thereof.

Sequence 18 (SEQ ID NO:18)

Thermococcus sp. 9 degrees N-7 DNA polymerase amino acid sequence (acc no. U47108)

```
1      MILDTDYITE NGKPVIRVFK KENGEFKIEY DRTFEPYFYA LLKDDSAIED VKKVTAKRHG
61     TVVKVKRAEK VQKKFLGRPI EVWKLYFNHP QDVPAIRDRI RAHPAVVDIY EYDIPFAKRY
121    LIDKGLIPME GDEELTMLAF DIETLYHEGE EFGTGPILMI SYADGSEARV ITWKKIDLPY
181    VDVVSTEKEM IKRFLRVVRE KDPDVLITYN GDNFDFAYLK KRCEELGIKF TLGRDGSEPK
241    IQRMGDRFAV EVKGRIHFDL YPVIRRTINL PTYTLEAVYE AVFGKPKEKV YAEEIAQAWE
301    SGEGLERVAR YSMEDAKVTY ELGREFFPME AQLSRLIGQS LWDVSRSSTG NLVEWFLLRK
361    AYKRNELAPN KPDERELARR RGGYAGGYVK EPERGLWDNI VYLDFRSLYP SIIITHNVSP
421    DTLNREGCKE YDVAPEVGHK FCKDFPGFIP SLLGDLLEER QKIKRKMKAT VDPLEKKLLD
481    YRQRAIKILA NSFYGYYGYA KARWYCKECA ESVTAWGREY IEMVIRELEE KFGFKVLYAD
541    TDGLHATIPG ADAETVKKKA KEFLKYINPK LPGLLELEYE GFYVRGFFVT KKKYAVIDEE
601    GKITTRGLEI VRRDWSEIAK ETQARVLEAI LKHGDVEEAV RIVKEVTEKL SKYEVPPEKL
661    VIHEQITRDL RDYKATGPHV AVAKRLAARG VKIRPGTVIS YIVLKGSGRI GDRAIPADEF
721    DPTKHRYDAE YYIENQVLPA VERILKAFGY RKEDLRYQKT KQVGLGAWLK VKGKK
```

Sequence 19 (SEQ ID NO:19)
T. litoralis DNA polymerase amino acid sequence (acc no. M74198.1)

```
1       MILDTDYITK DGKPIIRIFK KENGEFKIEL DPHFQPYIYA LLKDDSAIEE IKAIKGERHG
61      KTVRVLDAVK VRKKFLGREV EVWKLIFEHP QDVPAMRGKI REHPAVVDIY EYDIPFAKRY
121     LIDKGLIPME GDEELKLLAF DIETFYHEGD EFGKGEIIMI SYADEEEARV ITWKNIDLPY
181     VDVVSNEREM IKRFVQVVKE KDPDVIITYN GDNFDLPYLI KRAEKLGVRL VLGRDKEHPE
241     PKIQRMGDSF AVEIKGRIHF DLFPVVRRTI NLPTYTLEAV YEAVLGKTKS KLGAEEIAAI
301     WETEESMKKL AQYSMEDARA TYELGKEFFP MEAELAKLIG QSVWDVSRSS TGNLVEWYLL
361     RVAYARNELA PNKPDEEEYK RRLRTTYLGG YVKEPEKGLW ENIIYLDFRS LYPSIIVTHN
421     VSPDTLEKEG CKNYDVAPIV GYRFCKDFPG FIPSILGDLI AMRQDIKKKM KSTIDPIEKK
481     MLDYRQRAIK LLANSYYGYM GYPKARWYSK ECAESVTAWG RHYIEMTIRE IEEKFGFKVL
541     YADTDGFYAT IPGEKPELIK KKAKEFLNYI NSKLPGLLEL EYEGFYLRGF FVTKKRYAVI
601     DEEGRITTRG LEVVRRDWSE IAKETQAKVL EAILKEGSVE KAVEVVRDVV EKIAKYRVPL
661     EKLVIHEQIT RDLKDYKAIG PHVAIAKRLA ARGIKVKPGT IISYIVLKGS GKISDRVILL
721     TEYDPRKHKY DPDYYIENQV LPAVLRILEA FGYRKEDLRY QSSKQTGLDA WLKR
```

Sequence 20 (SEQ ID NO:20)
Amino acid sequence of chimeric DNA polymerase 9Nli

```
1       MILDTDYITE NGKPVIRVFK KENGEFKIEY DRTFEPYFYA LLKDDSAIED VKKVTAKRHG
61      TVVKVKRAEK VQKKFLGRPI EVWKLYFNHP QDVPAIRDRI RAHPAVVDIY EYDIPFAKRY
121     LIDKGLIPME GDEELTMLAF DIETLYHEGE EFGTGPILMI SYADGSEARV ITWKKIDLPY
181     VDVVSTEKEM IKRFLRVVRE KDPDVLITYN GDNFDFAYLK KRCEELGIKF TLGRDGSEPK
241     IQRMGDRFAV EVKGRIHFDL YPVIRRTINL PTYTLEAVYE AVFGKPKEKV YAEEIAQAWE
301     SGEGLERVAR YSMEDAKVTY ELGREFFPME AQLSRLIGQS LWDVSRSSTG NLVEWYLLRV
361     AYARNELAPN KPDEEEYKRR LRTTYLGGYV KEPEKGLWEN IIYLDFRSLY PSIIVTHNVS
421     PDTLEKEGCK NYDVAPIVGY RFCKDFPGFI PSILGDLIAM RQDIKKKMKS TIDPIEKKML
481     DYRQRAIKLL ANSYYGYMGY PKARWYSKEC AESVTAWGRH YIEMTIREIE EKFGFKVLYA
541     DTDGFYATIP GEKPELIKKK AKEFLNYINS KLPGLLELEY EGFYVRGFFV TKKKYAVIDE
601     EGKITTRGLE IVRRDWSEIA KETQARVLEA ILKHGDVEEA VRIVKEVTEK LSKYEVPPEK
661     LVIHEQITRD LRDYKATGPH VAVAKRLAAR GVKIRPGTVI SYIVLKGSGR IGDRAIPADE
721     FDPTKHRYDA EYYIENQVLP AVERILKAFG YRKEDLRYQK TKQVGLGAWL KVKGKK
```

Sequence 21 (SEQ ID NO:21)
Amino acid sequence of chimeric DNA polymerase Li9N

```
1       MILDTDYITK DGKPIIRIFK KENGEFKIEL DPHFQPYIYA LLKDDSAIEE IKAIKGERHG
61      KTVRVLDAVK VRKKFLGREV EVWKLIFEHP QDVPAMRGKI REHPAVVDIY EYDIPFAKRY
121     LIDKGLIPME GDEELKLLAF DIETFYHEGD EFGKGEIIMI SYADEEEARV ITWKNIDLPY
181     VDVVSNEREM IKRFVQVVKE KDPDVIITYN GDNFDLPYLI KRAEKLGVRL VLGRDKEHPE
241     PKIQRMGDSF AVEIKGRIHF DLFPVVRRTI NLPTYTLEAV YEAVLGKTKS KLGAEEIAAI
301     WETEESMKKL AQYSMEDARA TYELGKEFFP MEAELAKLIG QSVWDVSRSS TGNLVEWFLL
361     RKAYKRNELA PNKPDERELA RRRGGYAGGY VKEPERGLWD NIVYLDFRSL YPSIIITHNV
421     SPDTLNREGC KEYDVAPEVG HKFCKDFPGF IPSLLGDLLE ERQKIKRKMK ATVDPLEKKL
481     LDYRQRAIKI LANSFYGYYG YAKARWYCKE CAESVTAWGR EYIEMVIREL EEKFGFKVLY
541     ADTDGLHATI PGADAETVKK KAKEFLKYIN PKLPGLLELE YEGFYLRGFF VTKKRYAVID
601     EEGRITTRGL EVVRRDWSEI AKETQAKVLE AILKEGSVEK AVEVVRDVVE KIAKYRVPLE
661     KLVIHEQITR DLKDYKAIGP HVAIAKRLAA RGIKVKPGTI ISYIVLKGSG KISDRVILLT
721     EYDPRKHKYD PDYYIENQVL PAVLRILEAF GYRKEDLRYQ SSKQTGLDAW LKR
```

Amino acid sequences of DNA polymerases from *T. gorgonarius*, *T. zilligii* and chimeras thereof.

Sequence 22 (SEQ ID NO:22)
T. gorgonarius DNA polymerase amino acid sequence (acc no. 4699806)

```
1       MILDTDYITE DGKPVIRIFK KENGEFKIDY DRNFEPYIYA LLKDDSAIED VKKITAERHG
61      TTVRVVRAEK VKKKFLGRPI EVWKLYFTHP QDVPAIRDKI KEHPAVVDIY EYDIPFAKRY
121     LIDKGLIPME GDEELKMLAF DIETLYHEGE EFAEGPILMI SYADEEGARV ITWKNIDLPY
181     VDVVSTEKEM IKRFLKVVKE KDPDVLITYN GDNFDFAYLK KRSEKLGVKF ILGREGSEPK
241     IQRMGDRFAV EVKGRIHFDL YPVIRRTINL PTYTLEAVYE AIFGQPKEKV YAEEIAQAWE
301     TGEGLERVAR YSMEDAKVTY ELGKEFFPME AQLSRLVGQS LWDVSRSSTG NLVEWFLLRK
361     AYERNELAPN KPDERELARR RESYAGGYVK EPERGLWENI VYLDFRSLYP SIIITHNVSP
421     DTLNREGCEE YDVAPQVGHK FCKDFPGFIP SLLGDLLEER QKVKKKMKAT IDPIEKKLLD
481     YRQRAIKILA NSFYGYYGYA KARWYCKECA ESVTAWGRQY IETTIREIEE KFGFKVLYAD
541     TDGFFATIPG ADAETVKKKA KEFLDYINAK LPGLLELEYE GFYKRGFFVT KKKYAVIDEE
601     DKITTRGLEI VRRDWSEIAK ETQARVLEAI LKHGDVEEAV RIVKEVTEKL SKYEVPPEKL
661     VIYEQITRDL KDYKATGPHV AVAKRLAARG IKIRPGTVIS YIVLKGSGRI GDRAIPFDEF
721     DPAKHKYDAE YYIENQVLPA VERILRAFGY RKEDLRYQKT RQVGLGAWLK PKT
```

Sequence 23 (SEQ ID NO:23)
T. zilligii DNA polymerase amino acid sequence

```
1       MILDADYITE DGKPVIRVFK KEKGEFKIDY DRDFEPYIYA LLKDDSAIED IKKITAERHG
61      TTVRVTRAER VKKKFLGRPV EVWKLYFTHP QDVPAIRDKI REHPAVVDIY EYDIPFAKRY
121     LIDRGLIPME GDEELRMLAF DIETLYHEGE EFGEGPILMI SYADEEGARV ITWKNIDLPY
181     VESVSTEKEM IKRFLKVIQE KDPDVLITYN GDNFDFAYLK KRSETLGVKF ILGRDGSEPK
241     IQRMGDRFAV EVKGRIHFDL YPVIRRTINL PTYTLETVYE AIFGQPKEKV YAEEIARAWE
301     SGEGLERVAR YSMEDAKATY ELGKEFFPME AQLSRLVGQS LWDVSRSSTG NLVEWFLLRK
361     AYERNELAPN KPDERELARR AESYAGGYVK EPEKGLWENI VYLDYKSLYP SIIITHNVSP
421     DTLNREGCRE YDVAPQVGHR FCKDFPGFIP SLLGDLLEER QKVKKKMKAT VDPIERKLLD
481     YRQRAIKILA NSYYGYYGYA NARWYCRECA ESVTAWGRQY IETTMREIEE KFGFKVLYAD
541     TDGFFATIPG ADAETVKKKA KEFLNYINPR LPGLLELEYE GFYRRGFFVT KKKYAVIDEE
601     DKITTRGLEI VRRDWSEIAK ETQARVLEAI LKHGDVEEAV RIVKEVTEKL SRYEVPPEKL
661     VIYEQITRDL RDYRATGPHV AVAKRLAARG IKIRPGTVIS YIVLKGPGRV GDRAIPFDEF
721     DPAKHRYDAE YYIENQVLPA VERILRAFGY RKEDLRYQKT KQAGLGAWLK PKT
```

Sequence 24 (SEQ ID NO:24)
Amino acid sequence of chimeric DNA polymerase GoZi

```
1       MILDTDYITE DGKPVIRIFK KENGEFKIDY DRNFEPYIYA LLKDDSAIED VKKITAERHG
61      TTVRVVRAEK VKKKFLGRPI EVWKLYFTHP QDVPAIRDKI KEHPAVVDIY EYDIPFAKRY
121     LIDKGLIPME GDEELKMLAF DIETLYHEGE EFAEGPILMI SYADEEGARV ITWKNIDLPY
181     VDVVSTEKEM IKRFLKVVKE KDPDVLITYN GDNFDFAYLK KRSEKLGVKF ILGREGSEPK
241     IQRMGDRFAV EVKGRIHFDL YPVIRRTINL PTYTLEAVYE AIFGQPKEKV YAEEIAQAWE
301     TGEGLERVAR YSMEDAKVTY ELGKEFFPME AQLSRLVGQS LWDVSRSSTG NLVEWFLLRK
361     AYERNELAPN KPDERELARR RESYAGGYVK EPEKGLWENI VYLDYKSLYP SIIITHNVSP
421     DTLNREGCRE YDVAPQVGHR FCKDFPGFIP SLLGDLLEER QKVKKKMKAT VDPIERKLLD
481     YRQRAIKILA NSYYGYYGYA NARWYCRECA ESVTAWGRQY IETTMREIEE KFGFKVLYAD
541     TDGFFATIPG ADAETVKKKA KEFLDYINAK LPGLLELEYE GFYKRGFFVT KKKYAVIDEE
601     DKITTRGLEI VRRDWSEIAK ETQARVLEAI LKHGDVEEAV RIVKEVTEKL SKYEVPPEKL
661     VIYEQITRDL KDYKATGPHV AVAKRLAARG IKIRPGTVIS YIVLKGSGRI GDRAIPFDEF
721     DPAKHKYDAE YYIENQVLPA VERILRAFGY RKEDLRYQKT RQVGLGAWLK PKT
```

Sequence 25 (SEQ ID NO:25)
Amino acid sequence of chimeric DNA polymerase ZiGo

```
1       MILDADYITE DGKPVIRVFK KEKGEFKIDY DRDFEPYIYA LLKDDSAIED IKKITAERHG
61      TTVRVTRAER VKKKFLGRPV EVWKLYFTHP QDVPAIRDKI REHPAVVDIY EYDIPFAKRY
121     LIDRGLIPME GDEELRMLAF DIETLYHEGE EFGEGPILMI SYADEEGARV ITWKNIDLPY
181     VESVSTEKEM IKRFLKVIQE KDPDVLITYN GDNFDFAYLK KRSETLGVKF ILGRDGSEPK
241     IQRMGDRFAV EVKGRIHFDL YPVIRRTINL PTYTLETVYE AIFGQPKEKV YAEEIARAWE
301     SGEGLERVAR YSMEDAKATY ELGKEFFPME AQLSRLVGQS LWDVSRSSTG NLVEWFLLRK
361     AYERNELAPN KPDERELARR AESYAGGYVK EPERGLWENI VYLDFRSLYP SIIITHNVSP
421     DTLNREGCEE YDVAPQVGHK FCKDFPGFIP SLLGDLLEER QKVKKKMKAT IDPIEKKLLD
481     YRQRAIKILA NSFYGYYGYA KARWYCKECA ESVTAWGRQY IETTIREIEE KFGFKVLYAD
541     TDGFFATIPG ADAETVKKKA KEFLNYINPR LPGLLELEYE GFYRRGFFVT KKKYAVIDEE
601     DKITTRGLEI VRRDWSEIAK ETQARVLEAI LKHGDVEEAV RIVKEVTEKL SRYEVPPEKL
661     VIYEQITRDL RDYRATGPHV AVAKRLAARG IKIRPGTVIS YIVLKGPGRV GDRAIPFDEF
721     DPAKHRYDAE YYIENQVLPA VERILRAFGY RKEDLRYQKT KQAGLGAWLK PKT
```

Amino acid sequences of additional chimeras of KOD and Pfu DNA polymerases.

Sequence 26 (SEQ ID NO:26)
Amino acid sequence of chimeric DNA polymerase Kofu-II.

```
1       MASAILDTDY ITEDGKPVIR IFKKENGEFK IEYDRTFEPY FYALLKDDSA IEEVKKITAE
61      RHGTVVTVKR VEKVQKKFLG RPVEVWKLYF THPQDVPAIR DKIREHPAVI DIYEYDIPFA
121     KRYLIDKGLV PMEGDEELKM LAFDIETLYH EGEEFAEGPI LMISYADEEG ARVITWKNVD
181     LPYVDVVSTE REMIKRFLRV VKEKDPDVLI TYNGDNFDFA YLKKRCEKLG INFALGRDGS
241     EPKIQRMGDR FAVEVKGRIH FDLYPVIRRT INLPTYTLEA VYEAVFGQPK EKVYAEEITT
301     AWETGENLER VAKYSMEDAK ATYELGKEFL PMEIQLSRLV GQPLWDVSRS STGNLVEWFL
361     LRKAYERNEV APNKPSEEEY QRRLRESYTG GFVKEPEKGL WENIVYLDFR ALYPSIIITH
421     NVSPDTLNLE GCKNYDIAPQ VGHKFCKDIP GFIPSLLGHL LEERQKIKTK MKETQDPIEK
481     ILLDYRQKAI KLLANSFYGY YGYAKARWYC KECAESVTAW GRKYIELVWK ELEEKFGFKV
541     LYIDTDGLYA TIPGGESEEI KKKALEFVKY INSKLPGLLE LEYEGFYKRG FFVTKKRYAV
601     IDEEGKVITR GLEIVRRDWS EIAKETQARV LEALLKDGDV EKAVRIVKEV TEKLSKYEVP
661     PEKLVIHEQI TRDLKDYKAT GPHVAVAKRL AARGVKIRPG TVISYIVLKG SGRIGDRAIP
721     FDEFDPTKHK YDAEYYIENQ VLPAVERILR AFGYRKEDLR YQKTRQVGLS AWLKPKGT
```

Sequence 27 (SEQ ID NO:27)
Amino acid sequence of chimeric DNA polymerase Pod-II.

```
1       MASAILDVDY ITEEGKPVIR LFKKENGKFK IEHDRTFRPY IYALLRDDSK IEEVKKITGE
61      RHGKIVRIVD VEKVEKKFLG KPITVWKLYL EHPQDVPTIR EKVREHPAVV DIFEYDIPFA
121     KRYLIDKGLI PMEGEEELKI LAFDIETLYH EGEEFGKGPI IMISYADENE AKVITWKNID
181     LPYVEVVSSE REMIKRFLRI IREKDPDIIV TYNGDSFDFP YLAKRAEKLG IKLTIGRDGS
241     EPKMQRIGDM TAVEVKGRIH FDLYHVITRT INLPTYTLEA VYEAIFGKPK EKVYADEIAK
301     AWESGENLER VARYSMEDAK VTYELGKEFL PMEAQLSRLI GQSLWDVSRS STGNLVEWFL
361     LRKAYERNEL APNKPDEKEL ARRRQSYEGG YVKEPERGLW ENIVYLDFRS LYPSIIITHN
421     VSPDTLNREG CKEYDVAPQV GHRFCKDFPG FIPSLLGDLL EERQKIKKKM KATIDPIERK
481     LLDYRQRAIK ILANSYYGYY GYARARWYCK ECAESVTAWG REYITMTIKE IEEKYGFKVI
541     YSDTDGFFAT IPGADAETVK KKAMEFLKYI NAKLPGALEL EYEGFYKRGF FVTKKKYAVI
601     DEEGKITTRG LEIVRRDWSE IAKETQARVL ETILKHGDVE EAVRIVKEVI QKLANYEIPP
661     EKLAIYEQIT RPLHEYKAIG PHVAVAKKLA AKGVKIKPGM VIGYIVLRGD GPISNRAILA
721     EEYDPKKHKY DAEYYIENQV LPAVLRILEG FGYRKEDLRY QKTRQVGLTS WLNIKKS
```

Sequence 28 (SEQ ID NO:28)
Amino acid sequence of chimeric DNA polymerase Kofu-III.

```
1       MASAILDTDY ITEDGKPVIR IFKKENGEFK IEYDRTFEPY FYALLKDDSA IEEVKKITAE
61      RHGTVVTVKR VEKVQKKFLG RPVEVWKLYF THPQDVPAIR DKIREHPAVI DIYEYDIPFA
121     KRYLIDKGLV PMEGDEELKM LAFDIETLYH EGEEFAEGPI LMISYADEEG ARVITWKNVD
181     LPYVDVVSTE REMIKRFLRV VKEKDPDVLI TYNGDNFDFA YLKKRCEKLG INFALGRDGS
241     EPKIQRMGDR FAVEVKGRIH FDLYPVIRRT INLPTYTLEA VYEAVFGQPK EKVYAEEITT
301     AWETGENLER VARYSMEDAK VTYELGKEFL PMEAQLSRLI GQSLWDVSRS STGNLVEWFL
361     LRKAYERNEL APNKPDEKEL ARRRQSYEGG YVKEPEKGLW ENIVYLDFRA LYPSIIITHN
421     VSPDTLNLEG CKNYDIAPQV GHKFCKDIPG FIPSLLGHLL EERQKIKTKM KETQDPIEKI
481     LLDYRQKAIK LLANSFYGYY GYAKARWYCK ECAESVTAWG RKYIELVWKE LEEKFGFKVL
541     YIDTDGLYAT IPGGESEEIK KKALEFLKYI NAKLPGALEL EYEGFYKRGF FVTKKKYAVI
601     DEEGKITTRG LEIVRRDWSE IAKETQARVL EALLKDGDVE KAVRIVKEVT EKLSKYEVPP
661     EKLVIHEQIT RDLKDYKATG PHVAVAKRLA ARGVKIRPGT VISYIVLKGS GRIGDRAIPF
721     DEFDPTKHKY DAEYYIENQV LPAVERILRA FGYRKEDLRY QKTRQVGLSA WLKPKGT
```

Sequence 29 (SEQ ID NO:29)
Amino acid sequence of chimeric DNA polymerase Pod-III.

```
1       MASAILDVDY ITEEGKPVIR LFKKENGKFK IEHDRTFRPY IYALLRDDSK IEEVKKITGE
61      RHGKIVRIVD VEKVEKKFLG KPITVWKLYL EHPQDVPTIR EKVREHPAVV DIFEYDIPFA
121     KRYLIDKGLI PMEGEEELKI LAFDIETLYH EGEEFGKGPI IMISYADENE AKVITWKNID
181     LPYVEVVSSE REMIKRFLRI IREKDPDIIV TYNGDSFDFP YLAKRAEKLG IKLTIGRDGS
241     EPKMQRIGDM TAVEVKGRIH FDLYHVITRT INLPTYTLEA VYEAIFGKPK EKVYADEIAK
301     AWESGENLER VAKYSMEDAK ATYELGKEFL PMEIQLSRLV GQPLWDVSRS STGNLVEWFL
361     LRKAYERNEV APNKPSEEEY QRRLRESYTG GFVKEPERGL WENIVYLDFR SLYPSIIITH
421     NVSPDTLNRE GCKEYDVAPQ VGHRFCKDFP GFIPSLLGDL LEERQKIKKK MKATIDPIER
481     KLLDYRQRAI KILANSYYGY YGYARARWYC KECAESVTAW GREYITMTIK EIEEKYGFKV
541     IYSDTDGFFA TIPGADAETV KKKAMEFVKY INSKLPGLLE LEYEGFYKRG FFVTKKRYAV
601     IDEEGKVITR GLEIVRRDWS EIAKETQARV LETILKHGDV EEAVRIVKEV IQKLANYEIP
661     PEKLAIYEQI TRPLHEYKAI GPHVAVAKKL AAKGVKIKPG MVIGYIVLRG DGPISNRAIL
721     AEEYDPKKHK YDAEYYIENQ VLPAVLRILE GFGYRKEDLR YQKTRQVGLT SWLNIKKS
```

EQUIVALENTS

[0121] Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. The scope of the present invention is not intended to be limited to the above Description, but rather is as set forth in the appended claims. The articles "a", "an", and "the" as used herein in the specification and in the claims, unless clearly indicated to the contrary, should be understood to include the plural referents. Claims or descriptions that include "or" between one or more members of a group are considered satisfied if one, more than one, or all of the group members are present in, employed in, or otherwise relevant to a given product or process unless indicated to the contrary or otherwise evident from the context. The invention includes embodiments in which exactly one member of the group is present in, employed in, or otherwise relevant to a given product or process. The invention also includes embodiments in which more than one, or all of the group members are present in, employed in, or otherwise relevant to a given product or process. Furthermore, it is to be understood that the invention encompasses variations, combinations, and permutations in which one or more limitations, elements, clauses, descriptive terms, etc., from one or more of the claims is introduced into another claim dependent on the same base claim (or, as relevant, any other claim) unless otherwise indicated or unless it would be evident to one of ordinary skill in the art that a contradiction or inconsistency would arise. Where elements are presented as lists, e.g., in Markush group or similar format, it is to be understood that each subgroup of the elements is also disclosed, and any element(s) can be removed from the group. It should it be understood that, in general, where the invention, or aspects of the invention, is/are referred to as comprising particular elements, features, etc., certain embodiments of the invention or aspects of the invention consist, or consist essentially of, such elements, features, etc. For purposes of simplicity those embodiments have not in every case been specifically set forth herein. It should also be understood that any embodiment of the invention, e.g., any embodiment found within the prior art, can be explicitly excluded from the claims, regardless of whether the specific exclusion is recited in the specification.

[0122] It should also be understood that, unless clearly indicated to the contrary, in any methods claimed herein that include more than one act, the order of the acts of the method is not necessarily limited to the order in which the acts of the method are recited, but the invention includes embodiments in which the order is so limited. Furthermore, where the

claims recite a composition, the invention encompasses methods of using the composition and methods of making the composition. Where the claims recite a composition, it should be understood that the invention encompasses methods of using the composition and methods of making the composition.

INCORPORATION OF REFERENCES

[0123] All publications and patent documents cited in this application are incorporated by reference in their entirety to the same extent as if the contents of each individual publication or patent document were incorporated herein.

[0124] What is claimed is:

1. A chimeric polymerase comprising:

a first domain having a sequence at least 80% identical to an amino acid sequence found in a first DNA polymerase characterized with high processivity, elongation rate, salt resistance, thermostability or TMAC tolerance; and

a second domain having a sequence at least 80% identical to an amino acid sequence found in a second DNA polymerase characterized with high fidelity,

wherein the chimeric polymerase is characterized with high fidelity and high processivity, elongation rate, salt resistance, thermostability or TMAC tolerance.

2. The chimeric polymerase of claim 1, wherein the first DNA polymerase is selected from KOD polymerase, or TNA1 polymerase, *Thermococcus* sp. 9 degrees N-7, T4, T7, or phi29.

3. The chimeric polymerase of claim 2, wherein the first DNA polymerase is KOD polymerase.

4. The chimeric polymerase of claim 1, wherein the second DNA polymerase is selected from polymerases isolated from *Pyrococcus furiosus*, *P. abyssi*, *T. gorgonarius*, *T. litoralis, T. zilligii, T.* sp. GT, *or P.* sp. GB-D.

5. The chimeric polymerase of claim 4, wherein the second DNA polymerase is Pfu polymerase.

6. The chimeric polymerase of claim 1, wherein the first DNA polymerase is KOD polymerase and the second DNA polymerase is Pfu polymerase.

7. The chimeric polymerase of any one of claims 1-6, wherein the first domain is selected from the group consisting of an exonuclease domain, an N-terminal domain, a thumb domain and combinations thereof.

8. The chimeric polymerase of any one of claims 1-7, wherein the second domain is a palm and/or fingers domain.

9. The chimeric polymerase of any one of claims 1-8, wherein the amino acid sequence found in the first DNA polymerase corresponds to amino acid residues 156 to 301 of KOD polymerase (SEQ ID NO:11).

10. The chimeric polymerase of claim 9 further comprises a third domain having a sequence at least 80% identical to an amino acid sequence corresponding to amino acid residues 26 to 105 of KOD polymerase (SEQ ID NO:11).

11. The chimeric polymerase of claim 10 further comprises a fourth domain having a sequence at least 80% identical to an amino acid sequence corresponding to amino acid residues 612 to 749 of KOD polymerase (SEQ ID NO:11).

12. The chimeric polymerase of any one of claims 1-8, wherein the amino acid sequence found in the first DNA polymerase corresponds to amino acid residues 26 to 105 of KOD polymerase (SEQ ID NO:11).

13. The chimeric polymerase of claim 12 further comprises a third domain having a sequence at least 80% identical to an amino acid sequence corresponding to amino acid residues 612 to 749 of KOD polymerase (SEQ ID NO:11).

14. The chimeric polymerase of any one of claims 1-8, wherein the amino acid sequence found in the first DNA polymerase corresponds to amino acid residues 612 to 749 of KOD polymerase (SEQ ID NO:11).

15. The chimeric polymerase of claim 14 further comprises a third domain having a sequence at least 80% identical to an amino acid sequence corresponding to amino acid residues 156 to 301 of KOD polymerase (SEQ ID NO:11).

16. The chimeric polymerase of any one of claims 1-15, wherein the amino acid sequence found in the second DNA polymerase corresponds to amino acid residues 394 to 563 of Pfu polymerase (SEQ ID NO:9).

17. A chimeric polymerase comprising:

a first domain having a consensus sequence selected from the group consisting of

(1)
XXLXXXXXXXEGXRXXXXXXVXXXXXDXXXTXXXXXXXXXXVVKXXXX
XVLIXXXXXNXXXAXXKXXCXXXXXNFALXXXXXXXXXXXXXIXXMXXRF
XXXXXXXXXXXXXPXXRXXXXXXXXXXXXXXXXVXXQXXXXXXXEXXT
TXXXT (SEQ ID NO:30), wherein X is any amino acid or a peptide bond;

(2)

XXEXXXXYXXXXEXXFXXXXKXXXAXXXXXXXXAXXXXTVXTVKRXXX XQXXXXXRXVEXXXXXFTXXXXXXAXXDXIXXXXX (SEQ ID NO:31), wherein X is any amino acid or a peptide bond;

(3)

XXXXXXXXXXXXXXXXALXXDXXXXKXXXXXXXXTEXXSKXXVXXXXX VXHXXXXXDXKDXXXTXXXXXXXXRXXXRXXXXRXXTXXSXXXXKXSXR XGDXXXPFDXFXXTXXXXXXXXXXXXXXXXXXEXXXRAXX (SEQ ID NO:32), wherein X is any amino acid or a peptide bond;

(4)

$NGX_1FKIEX_2DRTFX_3PYX_4YALLX_5DDSX_6IEEVKKITX_7ERHGX_8X_9VX_{10}X_{11}X_{12}$ $X_{13}VEKVX_{14}KKFLGX_{15}PX_{16}X_{17}VWKLYX_{18}X_{19}HPQDVPX_{20}IRX_{21}KX_{22}REHPA$ (SEQ ID NO:33), wherein X_1 is not K; X_2 is not H; X_3 is not R; X_4 is not I; X_5 is not R; X_6 is not K; X_7 is not G; X_8 is not K; X_9 is not I; X_{10} is not R; X_{11} is not I; X_{12} is not V; X_{13} is not D; X_{14} is not E; X_{15} is not K; X_{16} is not I; X_{17} is not T; X_{18} is not L; X_{19} is not E; X_{20} is not T; X_{21} is not E; and X_{22} is not V;

(5)

$PIX_1MISYADEX_2X_3AX_4VITWKNX_5DLPYVX_6VVSX_7EREMIKRFLRX_8X_9X_{10}EK$ $DPDX_{11}X_{12}X_{13}TYNGDX_{14}FDFX_{15}YLX_{16}KRX_{17}EKLGIX_{18}X_{19}X_{20}X_{21}GRDGSEPKX$ $_{22}QRX_{23}GDX_{24}X_{25}AVEVKGRIHFDLYX_{26}VIX_{27}RTINLPTYTLEAVYEAX_{28}FGX_{29}$ $PKEKVYAX_{30}EIX_{31}X_{32}AWEX_{33}$ (SEQ ID NO:34), wherein X_1 is not I; X_2 is not N; X_3 is not E; X_4 is not K; X_5 is not I; X_6 is not E; X_7 is not S; X_8 is not I; X_9 is not I; X_{10} is not R; X_{11} is not I; X_{12} is not I; X_{13} is not V; X_{14} is not S; X_{15} is not P; X_{16} is not A; X_{17} is not A; X_{18} is not K; X_{19} is not L; X_{20} is not T; X_{21} is not I; X_{22} is not M; X_{23} is not I; X_{24} is not M; X_{25} is not T; X_{26} is not H; X_{27} is not T; X_{28} is not I; X_{29} is not K; X_{30} is not D; X_{31} is not A; X_{32} is not K; and X_{33} is not S;

(6)

$RDWSEIAKETQARVLEX_1X_2LKX_3GDVEX_4AVRIVKEVX_5X_6KLX_7X_8YEX_9PPEK$ $LX_{10}IX_{11}EQITRX_{12}LX_{13}X_{14}YKAX_{15}GPHVAVAKX_{16}LAAX_{17}GVKIX_{18}PGX_{19}VIX_{20}$ $YIVLX_{21}GX_{22}GX_{23}IX_{24}X_{25}RAIX_{26}X_{27}X_{28}EX_{29}DPX_{30}KHKYDAEYYIENQVLPAVX$

$_{31}$RILX$_{32}$X$_{33}$FG (SEQ ID NO:35), wherein X_1 is not T; X_2 is not I; X_3 is not H; X_4 is not E; X_5 is not I; X_6 is not Q; X_7 is not A; X_8 is not N; X_9 is not I; X_{10} is not A; X_{11} is not Y; X_{12} is not P; X_{13} is not H; X_{14} is not E; X_{15} is not I; X_{16} is not K; X_{17} is not K; X_{18} is not K; X_{19} is not M; X_{20} is not G; X_{21} is not R; X_{22} is not D; X_{23} is not P; X_{24} is not S; X_{25} is not N; X_{26} is not L; X_{27} is not A; X_{28} is not E; X_{29} is not Y; X_{30} is not K; X_{31} is not L; X_{32} is not E; and X_{33} is not G;

and combinations thereof, and

a second domain having a consensus sequence selected from the group consisting of

(1)

XKXXXXXXXXXXXXAXXXXXXXXXXXXXXXXXXLXXXXNXXIXXXXXXK
XXXXIXXXXXXXXXHXXXXXXXXXTXXXEXQXXXXKIXXXXXXKXXXLX
XXXFXXXXXXXKXXXXXXXXXXXXXXXXXKXXELVWXXLXXXFXXXXL
XIXXXXLYXXXXXGESXEIXXXXLX (SEQ ID NO:36), wherein X is any amino acid or a peptide bond;

(2)

EX$_1$GLWENIVYLDFRX$_2$LYPSIIITHNVSPDTLNX$_3$EGCKX$_4$YDX$_5$APQVGHX$_6$FC
KDX$_7$PGFIPSLLGX$_8$LLEERQKIKX$_9$KMKX$_{10}$TX$_{11}$DPIEX$_{12}$X$_{13}$LLDYRQX$_{14}$AIKX$_1$
$_5$LANSX$_{16}$YGYYGYAX$_{17}$ARWYCKECAESVTAWGRX$_{18}$YIX$_{19}$X$_{20}$X$_{21}$X$_{22}$KEX$_{23}$E
EKX$_{24}$GFKVX$_{25}$YX$_{26}$DTDGX$_{27}$X$_{28}$ATIPGX$_{29}$X$_{30}$X$_{31}$EX$_{32}$X$_{33}$KKKAX$_{34}$E (SEQ ID NO:37), wherein X_1 is not R; X_2 is not S; X_3 is not R; X_4 is not E; X_5 is not V; X_6 is not R; X_7 is not F; X_8 is not D; X_9 is not K; X_{10} is not A; X_{11} is not I; X_{12} is not R; X_{13} is not K; X_{14} is not R; X_{15} is not I; X_{16} is not Y; X_{17} is not R; X_{18} is not E; X_{19} is not T; X_{20} is not M; X_{21} is not T; X_{22} is not I; X_{23} is not I; X_{24} is not Y; X_{25} is not I; X_{26} is not S; X_{27} is not F; X_{28} is not F; X_{29} is not A; X_{30} is not D; X_{31} is not A; X_{32} is not T; X_{33} is not V; X_{34} is not M,

and combinations thereof,

wherein the chimeric polymerase is characterized with high fidelity and high processivity, elongation rate salt resistance, thermostability or TMAC tolerance.

18. A chimeric polymerase comprising an amino acid sequence at least 85% identical to SEQ ID NO:16 (Kofu amino acid sequence).

19. The chimeric DNA polymerase of claim 18, wherein the amino acid sequence comprises a consensus sequence of

XXXXTXXXXXDXXXXXXIXXXXXXEXXXXYXXXXEXXFXXXXKXXXAXXXXXX
XXAXXXXTVXTVKRXXXXQXXXXXRXVEXXXXXFTXXXXXXAXXDXIXXXXXXI
XXYXXXXXXXXXXXXXXXXVXXXXDXXXXMXXXXXXXXXXXXXXXXAEXXXLX
XXXXXXEGXRXXXXXXVXXXXXDXXXTXXXXXXXXXXVVKXXXXXVLIXXXXX
NXXXAXXKXXCXXXXXNFALXXXXXXXXXXIXXMXXRFXXXXXXXXXXXXXPX
XRXXXXXXXXXXXXXXXXVXXQXXXXXXXEXXTTXXXTXXXXXXXXRXXXXX
XXVXXXXXXXXXXXXAXXXXXVXXPXXXXXXXXXXXXXXXXXXXXXXXXXV
XXXXXSXEXYQXXXXEXXTXXFXXXXXKXXXXXXXXXXXXAXXXXXXXXXXX
XXXXXLXXXXNXXIXXXXXXKXXXXIXXXXXXXXXHXXXXXXXXTXXXEXQX
XXXKIXXXXXXKXXXLXXXXFXXXXXXXKXXXXXXXXXXXXXXXXXKXXELVW
XXLXXXFXXXXLXIXXXXLYXXXXXGESXEIXXXXLXXLXXXXAXXXXAXXXXX
XXXXXXXXXXXXXKXXXXXXXXXITXXXXXXXXXXXXXXXXXXXXXXXALX
XDXXXXKXXXXXXXXTEXXSKXXVXXXXXVXHXXXXXDXKDXXXTXXXXXXX
XRXXXRXXXXRXXTXXSXXXXKXSXRXGDXXXPFDXFXXTXXXXXXXXXXXXX
XXXXXEXXXRAXXXXXXXXXXXXXXXXXXSAXXKPXGT (SEQ ID NO:38)

wherein X is any amino acid or a peptide bond.

20. The chimeric polymerase of claim 18, wherein the amino acid sequence is identical to SEQ ID NO:16 (Kofu amino acid sequence).

21. A chimeric polymerase comprising an amino acid consensus sequence of

XXXXTXXXXXDXXXXXXIXXXXXXEXXXXYXXXXEXXFXXXXKXXXAXXXXXX
XXAXXXXTVXTVKRXXXXQXXXXXRXVEXXXXXFTXXXXXXAXXDXIXXXXXXI
XXYXXXXXXXXXXXXXXXXVXXXXDXXXXMXXXXXXXXXXXXXXXXAEXXXLX
XXXXXXEGXRXXXXXXVXXXXXDXXXTXXXXXXXXXXVVKXXXXXVLIXXXXX
NXXXAXXKXXCXXXXXNFALXXXXXXXXXXIXXMXXRFXXXXXXXXXXXXXPX
XRXXXXXXXXXXXXXXXXVXXQXXXXXXXEXXTTXXXTXXXXXXXXRXXXXX
XXVXXXXXXXXXXXXAXXXXXVXXPXXXXXXXXXXXXXXXXXXXXXXXXXV
XXXXXSXEXYQXXXXEXXTXXFXXXXXKXXXXXXXXXXXXAXXXXXXXXXXX

XXXXXLXXXXNXXIXXXXXXKXXXXIXXXXXXXXXHXXXXXXXXXTXXXEXQX
XXXKIXXXXXXKXXXLXXXXFXXXXXXXKXXXXXXXXXXXXXXXXXXKXXELVW
XXLXXXFXXXXLXIXXXXLYXXXXXGESXEIXXXXLXXLXXXXAXXXXAXXXXX
XXXXXXXXXXXXXKXXXXXXXXXITXXXXXXXXXXXXXXXXXXXXXXXALX
XDXXXXKXXXXXXXXTEXXSKXXVXXXXXVXHXXXXXDXKDXXXTXXXXXXX
XRXXXRXXXXRXXTXXSXXXXKXSXRXGDXXXPFDXFXXTXXXXXXXXXXXXX
XXXXXEXXXRAXXXXXXXXXXXXXXXXXXSAXXKPXGT (SEQ ID NO:38)

wherein X is any amino acid or a peptide bond; and

wherein the chimeric polymerase has a fidelity higher than that of KOD and a processivity, an elongation rate, a salt resistance, a thermostability or a TMAC tolerance higher than that of Pfu.

22. The method of claim 21, the amino acid consensus sequence is

XIXDTDYXTXDGXPXXRIFXKXXGEFXXXYDXXFEPYFYALLKDDSAIXXXXXXXA
XRHGTVXTVKRXXXXQXKFLXRXVEVWXLXFTHPQDVPAXXDXIXXHXXVIDIYE
YDIPFAKRYLIDXGLVPMEGDEXLXMXXXDIETXYHEGXEFAEGXXLMISYADXEG
ARVITWKXVDLPYVDVVSTEXEMIKRXXXVVKEKDPDVLIXYXGDNFDXAYLKXR
CEXLGXNFALXRXXXXXEPKIXXMGXRFAVEXKGRXHFDLXPXXRXTXNLPTYXL
XXVYEXVXGQXKXKXXXEEITTXWETXXXXXXXARYSMEDAXVTXELGXEFXPM
EAXLXXLVGXPXWDVXRSSTGNLVEWXLLXXAYXRNEVAPNKPSXEEYQXRXXE
XYTGXFVXEPEKGLWXXXXXLDXXALYPSIIXXHNVSPDTLXLEXCXNYDIAPXVG
XKFCKDIPGFIPSXLXHLXXXRQXXKTXMXEXQDPXEKIXLDYRQKAXKLLXNSFY
GYXGYXKARWYXXECAESVTXWGRKYIELVWXELEXXFGFKXLYIDTDGLYATIP
GGESXEIKXXXLXFLXYINAXLPGALELEYEXFYXRGFFVXKKKYAXIDEEXXITTR
GLEXVRRDWSXXAKETXAXVLEALLXDXXVXKAVXXVXXXTEXXSKYXVPXEKL
VIHEQITRDXKDYXATGPHVAXAKRLXXRGXXXRPGTXISYXXLKGSGRXGDRXIPF
DEFXXTKHXYDXXYYIENQVLPAVERXLRAFGYXXXXLXXQXXXQXGLSAWXKP
XGT (SEQ ID NO:39)

wherein X is any amino acid or a peptide bond.

23. A chimeric polymerase comprising:

a first domain having a sequence at least 80% identical to an amino acid sequence found in an exonuclease domain, an N-terminal domain, and/or a thumb domain of a first DNA polymerase; and

a second domain having a sequence at least 80% identical to an amino acid sequence found in palm and/or fingers domain of a second DNA polymerase,

wherein the chimeric polymerase has a processivity, elongation rate, salt resistance, thermostability or TMAC tolerance higher than that of the second DNA polymerase and a fidelity higher than that of the first DNA polymerase.

24. The chimeric polymerase of claim 23, wherein the first DNA polymerase is selected from KOD polymerase, TNA1 polymerase, *Thermococcus* sp. 9 degrees N-7, T4, T7, or phi29.

25. The chimeric polymerase of claim 23, wherein the first DNA polymerase is KOD polymerase.

26. The chimeric polymerase of claim 23, wherein the second DNA polymerase is selected from polymerases isolated from *Pyrococcus furiosus*, *P. abyssi*, *T. gorgonarius*, *T. litoralis, T. zilligii, or P.* sp. GB-D.

27. The chimeric polymerase of claim 26, wherein the second DNA polymerase is Pfu polymerase.

28. The chimeric polymerase of claim 23, wherein the first DNA polymerase is KOD polymerase and the second DNA polymerase is Pfu polymerase.

29. The chimeric polymerase of claim 23, wherein the first DNA polymerase is Pfu polymerase and the second DNA polymerase is KOD polymerase.

30. A chimeric polymerase comprising an amino acid sequence at least 90% identical to SEQ ID NO:15 (Pod amino acid sequence).

31. A nucleotide sequence encoding a chimeric polymerase of any one of claims 1-30.

32. A vector comprising the nucleotide sequence of claim 31.

33. A cell comprising the nucleotide sequence of claim 31.

34. A kit comprising a chimeric polymerase of any one of claims 1-30.

35. A method of DNA synthesis using a chimeric polymerase of any one of claims 1-30.

36. A method of amplifying a DNA fragment using a chimeric polymerase of any one of claims 1-30.

37. A method of engineering a chimeric polymerase, the method comprising steps of:

(a) providing an N-terminal domain, an exonuclease domain, and/or a thumb domain based on a first DNA polymerase

(b) providing a palm and/or fingers domain based on a second DNA polymerase;

(c) combining the domains from step (a) and step (b) to form a chimeric polymerase;

wherein the chimeric polymerase has a processivity, elongation rate, salt resistance, thermostability or TMAC tolerance higher than that of the second DNA polymerase and a fidelity higher than that of the first DNA polymerase.

38. The method of claim 37, wherein the first DNA polymerase is selected from KOD polymerase, or TNA1 polymerase, *Thermococcus* sp. 9 degrees N-7, T4, T7, or phi29.

39. The method of claim 38, wherein the first DNA polymerase is KOD polymerase.

40. The method of claim 38, wherein the second DNA polymerase is selected from polymerases isolated from *Pyrococcus furiosus*, *P. abyssi*, *T. gorgonarius*, *T. litoralis, T. zilligii, T.* sp. GT, *or P.* sp. GB-D.

41. The method of claim 40, wherein the second DNA polymerase is Pfu polymerase.

42. The method of claim 37, wherein the first DNA polymerase is KOD polymerase and the second DNA polymerase is Pfu polymerase.

43. A chimeric polymerase engineered using the method according to any one of claims 35-42.

44. A method of improving the fidelity of a DNA polymerase, the method comprising a step of replacing a sequence within the palm and/or fingers domain of the DNA polymerase of interest with a corresponding sequence from a different DNA polymerase that is characterized with higher fidelity relative to the DNA polymerase of interest.

45. A method of improving the processivity, elongation rate, salt resistance, thermostability or TMAC tolerance of a DNA polymerase, the method comprising a step of replacing a sequence within the N-terminal domain, the exonuclease domain and/or the thumb domain of the DNA polymerase of interest with a corresponding sequence from a different DNA polymerase that is characterized with higher processivity, elongation rate, salt or PCR enhancer resistance relative to the DNA polymerase of interest.

46. An improved DNA polymerase according to the method of claim 44 or 45.

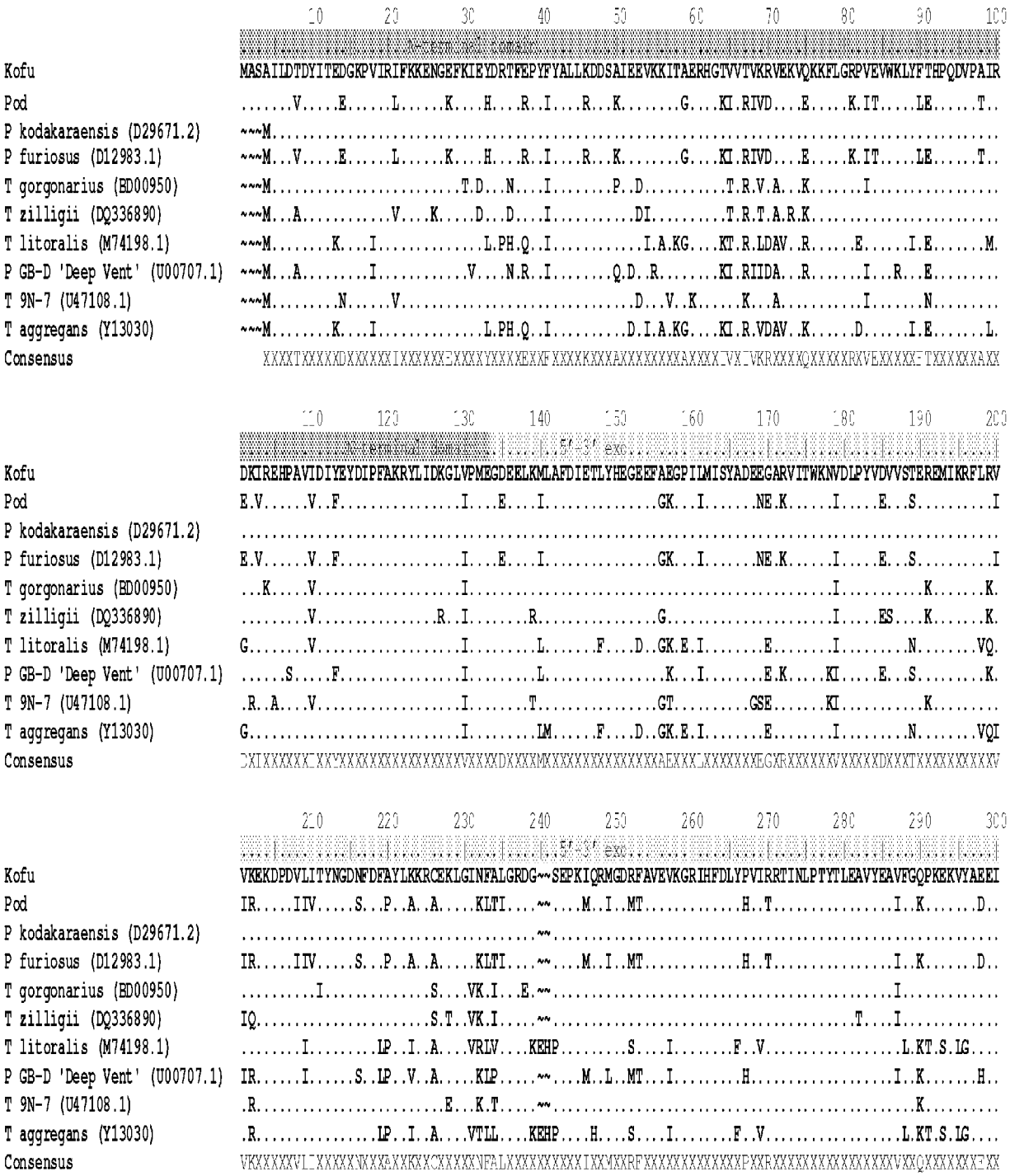
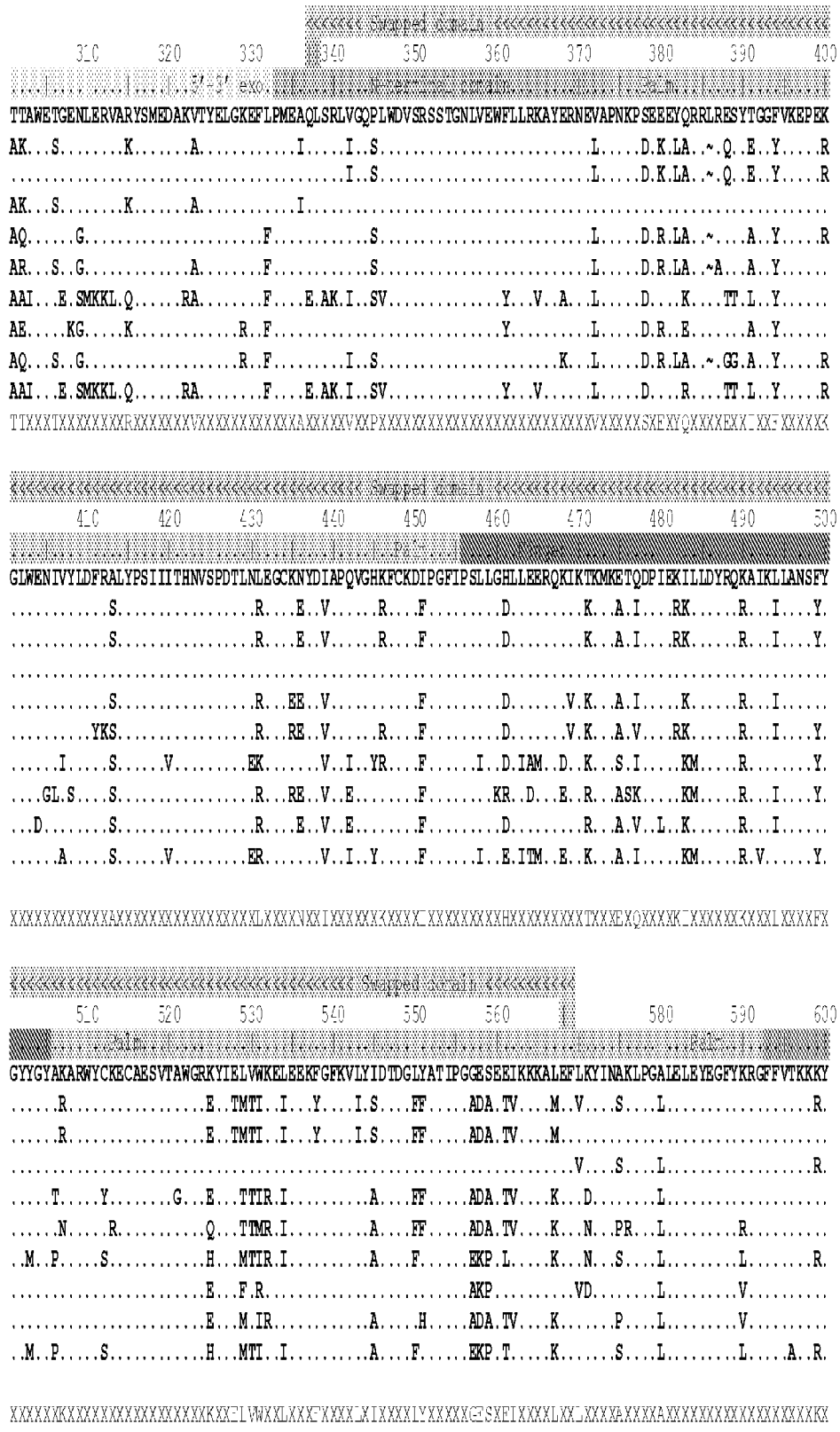
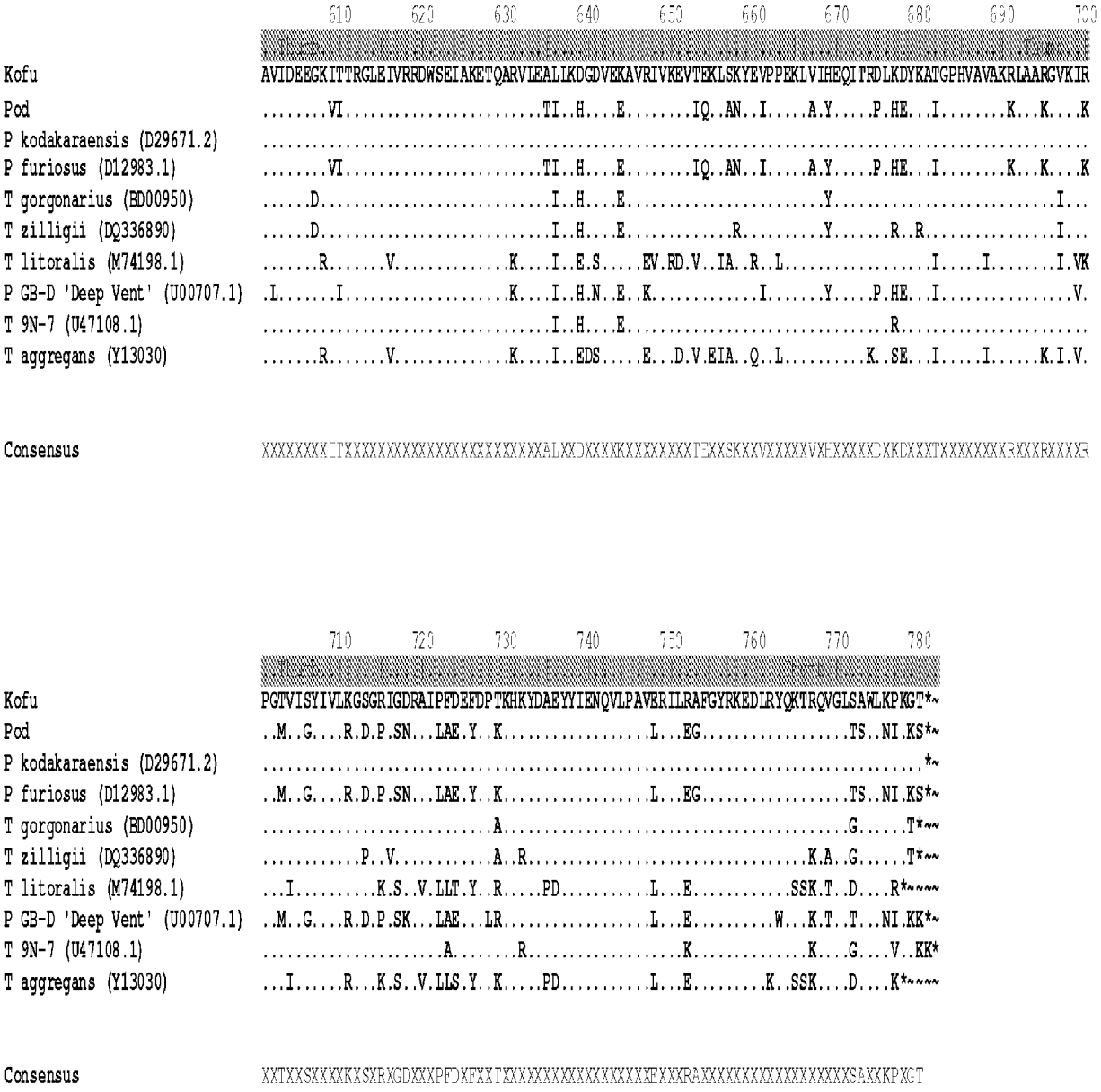

Kofu
Pod
P kodakaraensis (D29671.2)
P furiosus (D12983.1)
T gorgonarius (ED00950)
T zilligii (DQ336890)
T litoralis (M74198.1)
P GB-D 'Deep Vent' (U00707.1)
T 9N-7 (U47108.1)
T aggregans (Y13030)
Consensus

FIG. 1



CONTINUATION OF FIG. 1



CONTINUATION OF FIG. 1

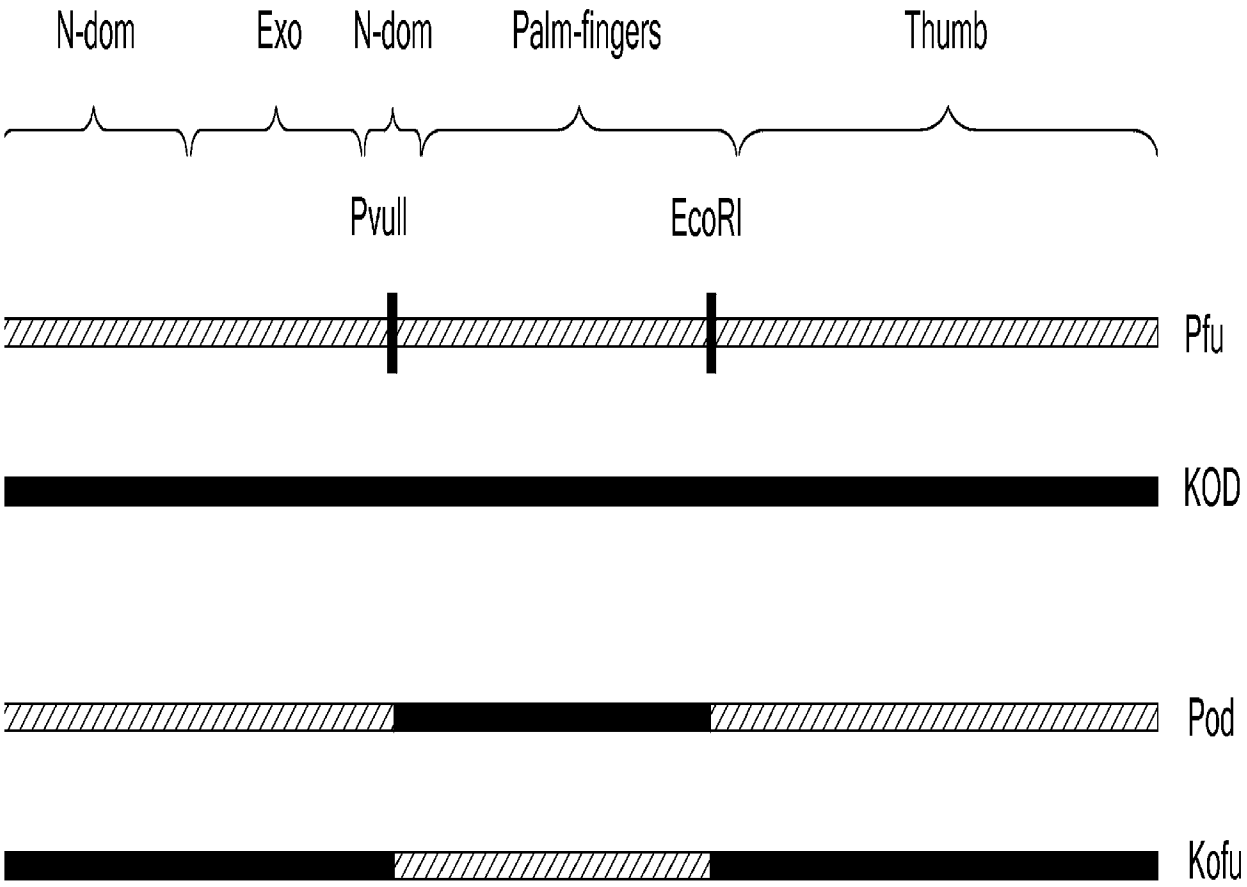

N-dom
Exo
N-dom
Palm-fingers
Thumb
PvuII
EcoRI
Pfu
KOD
Pod
Kofu

FIG. 2

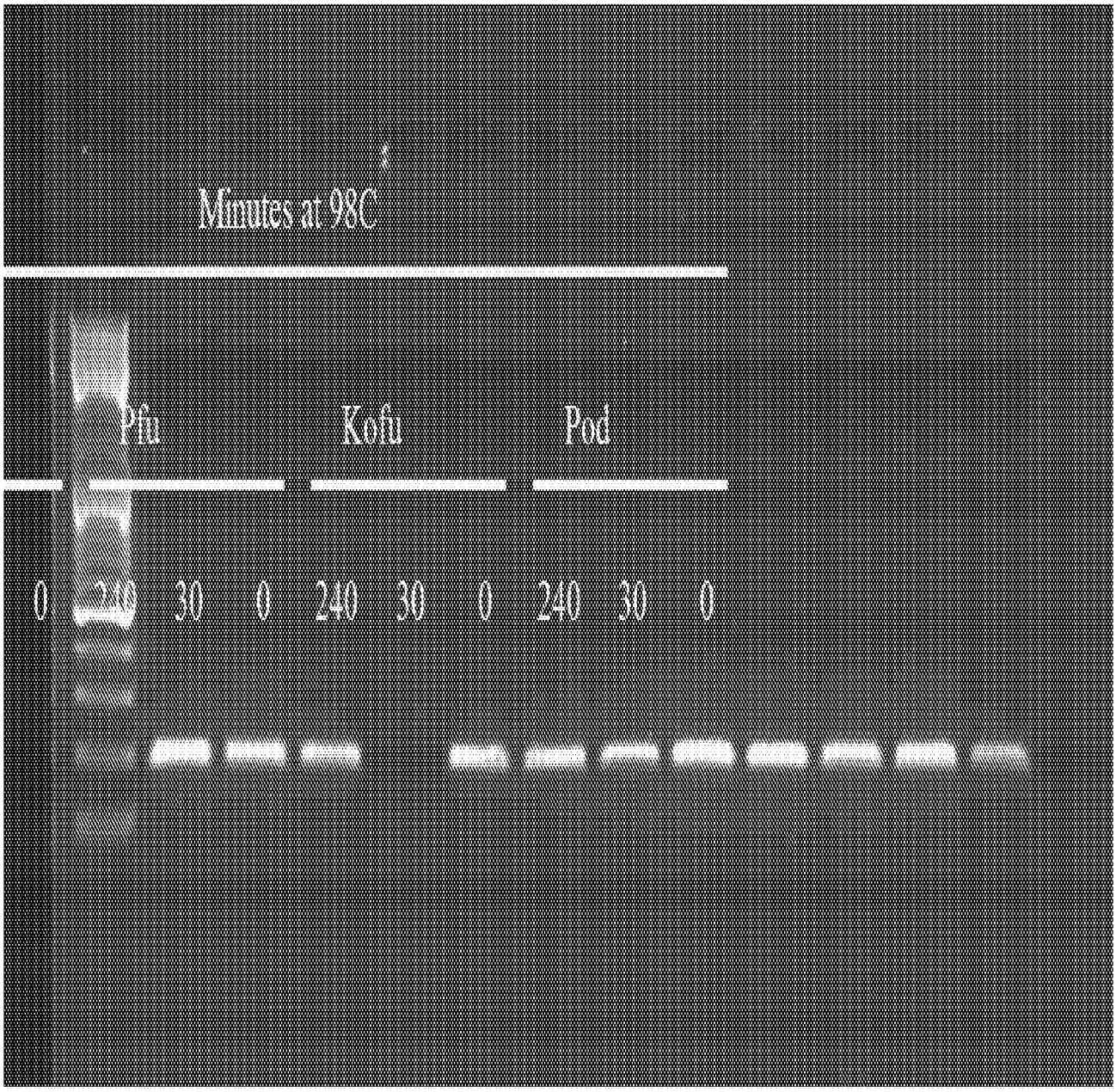


FIG. 3

FIG. 4

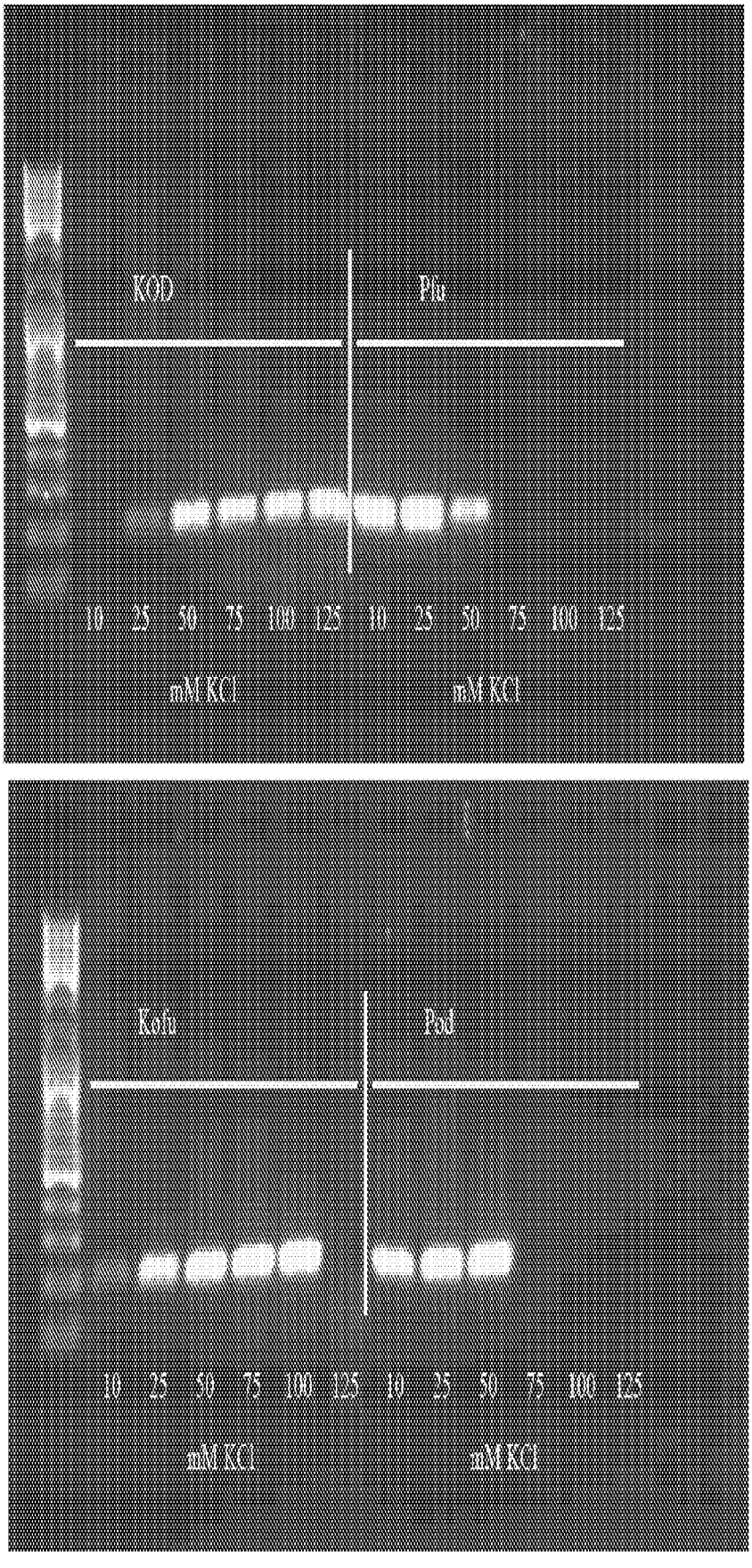
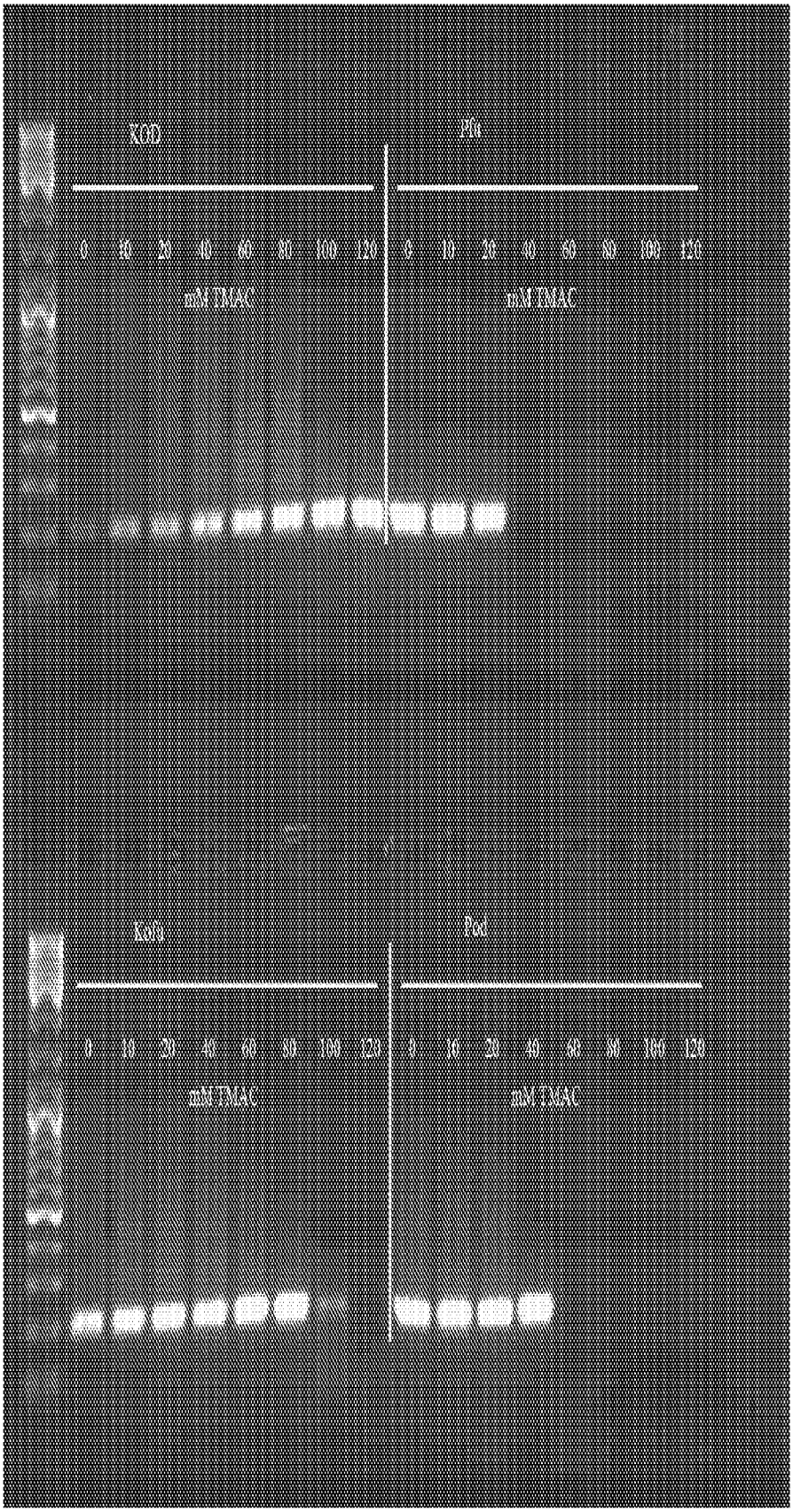

KOD
Pfu
10 25 50 75 100 125 10 25 50 75 100 125
mM KCl
mM KCl
Kofu
Pol
10 25 50 75 100 125 10 25 50 75 100 125
mM KCl
mM KCl

FIG. 5


XOD
Pol
0 10 20 40 60 80 100 120
mM TMAC
Xo1
Pol
0 10 20 40 60 80 100 120
mM TMAC