sequence may be only a small portion of the DNA or RNA in question, so it may be difficult to detect the presence of a target nucleic acid sequence without amplification.

The overall folding pattern of DNA polymerases resembles the human right hand and contains three distinct subdomains of palm, fingers, and thumb. (*See* Beese *et al.*, *Science* 260:352-355, 1993); Patel *et al.*, *Biochemistry* 34:5351-5363, 1995). While the structure of the fingers and thumb subdomains vary greatly between polymerases that differ in size and in cellular functions, the catalytic palm subdomains are all superimposable. For example, motif A, which interacts with the incoming dNTP and stabilizes the transition state during chemical catalysis, is superimposable with a mean deviation of about one Å amongst mammalian pol α and prokaryotic pol I family DNA polymerases (Wang *et al.*, *Cell* 89:1087-1099, 1997). Motif A begins structurally at an antiparallel β-strand containing predominantly hydrophobic residues and continues to an α-helix. The primary amino acid sequence of DNA polymerase active sites is exceptionally conserved. In the case of motif A, for example, the sequence DYSQIELR (SEQ ID NO:28) is retained in polymerases from organisms separated by many millions years of evolution, including, *e.g.*, *Thermus aquaticus*, *Chlamydia trachomatis*, and *Escherichia coli*.

In addition to being well-conserved, the active site of DNA polymerases has also been shown to be relatively mutable, capable of accommodating certain amino acid substitutions without reducing DNA polymerase activity significantly. (*See, e.g.*, U.S. Patent No. 6,602,695) Such mutant DNA polymerases can offer various selective advantages in, *e.g.*, diagnostic and research applications comprising nucleic acid synthesis reactions. Thus, there is a need in the art for identification of amino acid positions amenable to mutation to yield improved polymerase activities. The present invention, as set forth herein, meets these and other needs.

BRIEF SUMMARY

Provided herein are DNA polymerases having increased 3′-mismatch discrimination relative to a corresponding, unmodified control polymerase, and methods of making and using such DNA polymerases. In some embodiments, the polymerase is a thermostable DNA polymerase. In some embodiments, the DNA polymerase is a thermoactive DNA polymerase. In some embodiments, the DNA polymerase of the present invention is derived from a *Thermus* species. In some embodiments, the DNA polymerase is derived from a *Thermotoga* species.

In a first aspect, there is provided a DNA polymerase having an amino acid sequence that is at least 80% sequence identity with SEQ ID NO:1 or with SEQ ID NO: 2 and having increased 3′-mismatch discrimination activity compared with a control DNA polymerase, wherein the DNA polymerase comprises a motif in the polymerase domain comprising

A-G-H-P-F-N-L-N-S-R-D-Q-L-X_{10}-R-V-L-F-D-E-L, wherein:

X_{10} is A, G, K, or R (SEQ ID NO:10)

wherein the control DNA polymerase has the same amino acid sequence as the DNA polymerase except that amino acid X_{10} of the control DNA polymerase is E.

According to the disclosure it was discovered that the amino acid of the DNA polymerase corresponding to position 493 of SEQ ID NO:1 can be mutated from the native amino acid at that position to produce an enzyme having increased 3′-mismatch discrimination relative to a corresponding, unmodified control polymerase. In DNA polymerases derived from a *Thermus* species, the native amino acid corresponding to position 493 of SEQ ID NO:1 is E. In DNA polymerases derived from a *Thermotoga* species, the native amino acid corresponding to position 493 of SEQ ID NO:1 is S. In DNA polymerases derived from other non-*Thermus*, non-*Thermotoga* species, the amino acid corresponding to position 493 of SEQ ID NO:1 can be an amino acid other than E or S, for example, A or G. *See*, Figure 1.

In some embodiments, the amino acid of the DNA polymerase corresponding to position 493 of SEQ ID NO:1 is any amino acid other than E, S, A or G, and the control DNA polymerase has the same amino acid sequence as the DNA polymerase except that the amino acid of the control DNA polymerase corresponding to position 493 of SEQ ID NO:1 is E, S, A or G. For example, in some embodiments, the amino acid at the position corresponding to position 493 of SEQ ID NO:1 is selected from V, L, I, M, F, W, P, T, C, Y, N, Q, D, K, R or H.

However, it was further discovered that insertion of A or G (the native amino acid in some non-*Thermus*, non-*Thermotoga* species) in place of E at position 493 of *Thermus* polymerases also increased 3′-mismatch discrimination. Thus, in some embodiments, the DNA polymerase of the disclosure is derived from a *Thermus* species, and the amino acid of the DNA polymerase corresponding to position 493 of SEQ ID NO:1 is any amino acid other than E, and the control DNA polymerase has the same amino acid sequence as the DNA polymerase except that the amino acid of the control DNA polymerase corresponding to position 493 of SEQ ID NO:1 is E. For example, where the DNA polymerase is derived from a *Thermus* species, the amino acid at the position corresponding to position 493 of SEQ ID NO:1 can be selected from S, A, G, V, L, I, M, F, W, P, T, C, Y, N, Q, D, K, R or H. In some embodiments the amino acid at the position corresponding to position 493 of SEQ ID NO:1 is selected from S, A, Q, G, K, or R. In some embodiments the amino acid at the position corresponding to position 493 of SEQ ID NO:1 is selected from A, G, K, or R. In some embodiments the amino acid at the position corresponding to position 493 of SEQ ID NO:1 is K.

In some embodiments, the DNA polymerase of the disclosure is derived from a *Thermus* species, and the amino acid corresponding to position 493 of SEQ ID NO:1 is an amino acid

having a polar, uncharged side-chain (e.g., N, Q, H, S, T, or Y), a nonpolar, uncharged side-chain (e.g., G, A, L, M, W, P, F, C, V, or I), or a polar, positively charged side-chain (e.g., R or K) at neutral pH. In some embodiments, the amino acid corresponding to position 493 of SEQ ID NO:1 having a nonpolar, uncharged side-chain is A or G. In some embodiments, the amino acid corresponding to position 493 of SEQ ID NO:1 having a polar, positively charged side-chain is R or K. In some embodiments, the amino acid corresponding to position 493 of SEQ ID NO:1 having a polar, positively charged side-chain is K.

Further, according to the disclosure other amino acids, located nearby to the amino acid corresponding to position 493 of SEQ ID NO:1, can also be mutated to produce an enzyme having increased 3′-mismatch discrimination relative to a corresponding, unmodified control polymerase. For example, mutations at amino acids corresponding to positions 488 and/or 497 of SEQ ID NO:1 also produce an enzyme having increased 3′-mismatch discrimination relative to a corresponding, unmodified control polymerase.

In some embodiments, the DNA polymerase having improved 3′-mismatch discrimination comprises a motif in the polymerase domain comprising

A-G-H-P-F-N-L-N-X_6-R-D-Q-L-X_{10}-R-V-L-X_{13}-D-E-L, wherein:

X_6 is S, G, A, D, F, K, C, T, or Y;

X_{10} is any amino acid other than E;

X_{13} is F, A, G, S, T, Y, D or K (SEQ ID NO:8).

In some embodiments, X_{10} is selected from L, M, W, P, T, F, Y, C, N, D, V, I, R, K, Q, H, S, A, or G.

In some embodiments, X_{10} is selected from R, K, A, or G (SEQ ID NO:49).

In some embodiments, the DNA polymerase having improved 3′-mismatch discrimination comprises a motif in the polymerase domain comprising

A-G-H-P-F-N-L-N-S-R-D-Q-L-X_{10}-R-V-L-F-D-E-L, wherein:

X_{10} is any amino acid other than E (SEQ ID NO:9).

In some embodiments, X_{10} is selected from L, M, W, P, T, F, Y, C, N, D, V, I, R, K, Q, H, S, A, or G.

In some embodiments, X_{10} is selected from R, K, A, or G (SEQ ID NO:50).

In some embodiments, the DNA polymerase having improved 3′-mismatch discrimination comprises a motif in the polymerase domain comprising

A-G-H-P-F-N-L-N-S-R-D-Q-L-X_{10}-R-V-L-F-D-E-L, wherein:

X_{10} is A, G, K or R (SEQ ID NO:10).

In some embodiments, of said DNA polymeraseX_{10} is selected from L, M, W, P, T, F, Y, C, N, D, V, I, R, K, Q, H, S, A, or G.

In some embodiments, of said DNA polymerase is selected from A, G, K or R (SEQ ID NO:50).

In some embodiments, of said DNA polymerase is selected from A or G.

In some embodiments, the DNA polymerase having improved 3′-mismatch discrimination comprises a motif in the polymerase domain comprising

A-G-H-P-F-N-L-N-S-R-D-Q-L-X_{10}-R-V-L-F-D-E-L, wherein:

X_{10} is K (SEQ ID NO:11).

In some embodiments, the amino acid of the DNA polymerase corresponding to position 580 of SEQ ID NO:1 is any amino acid other than D or E. In some embodiments, the amino acid of the DNA polymerase corresponding to position 580 of SEQ ID NO:1 is any amino acid other than D. In some embodiments, the amino acid of the DNA polymerase corresponding to position 580 of SEQ ID NO:1 is selected from the group consisting of L, G, T, Q, A, S, N, R and K. In some embodiments, the amino acid of the DNA polymerase corresponding to position 580 of SEQ ID NO:1 is G.

In some embodiments, the DNA polymerase according to the disclosure further comprise a motif comprising

T-G-R-L-S-S-X_7-X_8-P-N-L-Q-N; wherein

X_7 is Ser (S) or Thr (T); and

X_8 is any amino acid other than D or E (SEQ ID NO:27).

In some embodiments, the amino acid corresponding to position X_8 is selected from the group consisting of L, G, T, Q, A, S, N, R, and K (SEQ ID NO:51).

In some embodiments, the DNA polymerase further comprises a mutation at one or more amino acids corresponding to a position selected from 488 and/or 497 SEQ ID NO:1. In some embodiments, the amino acid of the DNA polymerase corresponding to position 488 of SEQ ID NO:1 is any amino acid other than S. In some embodiments, the amino acid of the DNA polymerase corresponding to position 488 of SEQ ID NO:1 is selected from G, A, V, L, I, M, F, W, P, E, T, C, N, Q, D, Y, K, R or H. In some embodiments, the amino acid of the DNA polymerase corresponding to position 488 of SEQ ID NO:1 is G, A, D, F, K, C, T, or Y. In some embodiments, the amino acid of the DNA polymerase corresponding to position 497 of SEQ ID NO:1 is any amino acid other than F or Y. In some embodiments, the amino acid of the DNA polymerase corresponding to position 497 of SEQ ID NO:1 is selected from R, V, L, I, M, W, P, T, C, N, D, E, S, A, G, K, E or H. In some embodiments, the DNA polymerase is derived from a *Thermus* species, and the amino acid of the DNA polymerase corresponding to position 497 of SEQ ID NO:1 is any amino acid other than F. Thus, in some embodiments, the amino acid of the DNA polymerase corresponding to position 497 of SEQ ID NO:1 is A, G, S, T, Y, D, or K.

Various DNA polymerases are amenable to mutation according to the present disclosure. Particularly suitable are thermostable polymerases, including wild-type or naturally occurring thermostable polymerases from various species of thermophilic bacteria, as well as synthetic thermostable polymerases derived from such wild-type or naturally occurring enzymes by amino acid substitution, insertion, or deletion, or other modification. Exemplary unmodified forms of polymerase include, *e.g.*, CS5 (SEQ ID NO:29), CS6 (SEQ ID NO:30) or Z05 DNA polymerase (SEQ ID NO:1), or a functional DNA polymerase having at least 80%, preferably at least 90%, more preferably at least 95% sequence identity thereto. Other unmodified polymerases include, *e.g.*, DNA polymerases from any of the following species of thermophilic bacteria (or a functional DNA polymerase having at least 80%, preferably at least 90%, more preferably at least 95% sequence identity to such a polymerase): *Thermotoga*

maritima (SEQ ID NO:38); *Thermus aquaticus* (SEQ ID NO:2); *Thermus thermophilus* (SEQ ID NO:6); *Thermus flavus* (SEQ ID NO:4); *Thermus filiformis* (SEQ ID NO:3); *Thermus* sp. Sps17 (SEQ ID NO:5); *Thermus* sp. Z05 (SEQ ID NO:1); *Thermotoga neopolitana* (SEQ ID NO:39); *Thermosipho africanus* (SEQ ID NO:37); *Thermus caldophilus* (SEQ ID NO:7), *Deinococcus radiodurans* (SEQ ID NO:36), *Bacillus stearothermophilus* (SEQ ID NO:40) or *Bacillus caldotenax* (SEQ ID NO:41). Suitable polymerases also include those having reverse transcriptase (RT) activity and/or the ability to incorporate unconventional nucleotides, such as ribonucleotides or other 2'-modified nucleotides.

While thermostable DNA polymerases possessing efficient 3'-mismatch discrimination activity are particularly suited for performing PCR, thermoactive, but not thermostable DNA polymerases possessing efficient 3'-mismatch discrimination activity also are amenable to mutation according to the present disclosure.

In some embodiments, the DNA polymerase is a *Thermus* DNA polymerase. For example, in some embodiments, the DNA polymerase has at least 80%, preferably at least 90%, more preferably at least 95% sequence identity to a polymerase selected from the group consisting of:

(a) a *Thermus sp.* Z05 DNA polymerase (Z05) (SEQ ID NO:1);

(b) a *Thermus aquaticus* DNA polymerase (Taq) (SEQ ID NO:2);

(c) a *Thermus filiformis* DNA polymerase (Tfi) (SEQ ID NO:3);

(d) a *Thermus flavus* DNA polymerase (Tfl) (SEQ ID NO:4);

(e) a *Thermus sp.* Sps17 DNA polymerase (Sps17) (SEQ ID NO:5);

(f) a *Thermus thermophilus* DNA polymerase (Tth) (SEQ ID NO:6); and

(g) a *Thermus caldophilus* DNA polymerase (Tca) (SEQ ID NO:7).

In some embodiments, the DNA polymerase is a *Thermotoga* DNA polymerase. For example, in some embodiments, the DNA polymerase has at least 80%, preferably at least 90%, more preferably at least 95% sequence identity to a polymerase selected from the group consisting of:

polynucleotide equal in number or fewer than the number of copies of wild-type target (*e.g.*, 10,000 or fewer copies). Two or more reaction mixtures can have titrated numbers of copies of mutant BRAF V600R target polynucleotide (*e.g.*, 1:5 titrations, 1:10 titrations, *e.g.*, 10,000 copies, 1000 copies, 100 copies, 10 copies, 1 copy, 0 copies in several reaction mixtures). The 3′-mismatch discrimination ability of a polymerase of the disclosure can be compared to the 3′-mismatch discrimination ability of a reference polymerase (*e.g.*, a naturally occurring or unmodified polymerase), over a preselected unit of time, as described herein. Polymerases with increased 3′-mismatch discrimination ability will not amplify the wild-type sequence when contacted with a primer that is perfectly matched to a mutant allele, or will require a greater number of PCR cycles to amplify the wild-type sequence using the mutant allele-specific primer (*i.e.*, exhibit a higher Cp value), in comparison to a naturally occurring or unmodified polymerase.

The present disclosure provides a recombinant nucleic acid encoding a mutant or improved DNA polymerase as described herein, a vector comprising the recombinant nucleic acid, and/or a host cell transformed with the vector. In a second aspect, there is provided a recombinant nucleic acid encoding the DNA polymerase as defined herein. In certain embodiments, the vector is an expression vector. In a third aspect, there is provided an expression vector comprising the recombinant nucleic acid as defined herein. Host cells comprising such expression vectors are useful in methods of the disclosure for producing the mutant or improved polymerase by culturing the host cells under conditions suitable for expression of the recombinant nucleic acid. The polymerases of the disclosure may be contained in reaction mixtures and/or kits. The embodiments of the recombinant nucleic acids, host cells, vectors, expression vectors, reaction mixtures and kits are as described above and herein.

In the present disclosure, a method for conducting polynucleotide extension is provided. The method generally includes contacting a DNA polymerase having increased 3′-mismatch discrimination as described herein with a primer, a polynucleotide template, and nucleoside triphosphates under conditions suitable for extension of the primer, thereby producing an extended primer. In a fourth aspect, there is provided a method for conducting primer extension, comprising:

contacting a DNA polymerase as defined herein with a primer, a polynucleotide template, and nucleoside triphosphates under conditions suitable for extension of the primer, thereby producing an extended primer. The polynucleotide template can be, for example, an RNA or DNA template. The nucleoside triphosphates can include unconventional nucleotides such as, *e.g.*, ribonucleotides and/or labeled nucleotides. Further, the primer and/or template can include one or more nucleotide analogs. In some variations, the polynucleotide extension method is a method for polynucleotide amplification that includes contacting the mutant or improved DNA polymerase with a primer pair, the polynucleotide template, and the nucleoside

triphosphates under conditions suitable for amplification of the polynucleotide. The polynucleotide extension reaction can be, *e.g.*, PCR, isothermal extension, or sequencing (*e.g.*, 454 sequencing reaction). In some embodiments the primer extension method is a method for conducting polymerase chain reaction (PCR).

The present disclosure also provides a kit useful in such a polynucleotide extension method. In a fifth aspect, there is provided a kit for producing an extended primer, comprising:
at least one container providing a DNA polymerase as defined herein.
Generally, the kit includes at least one container providing a mutant or improved DNA polymerase as described herein. In certain embodiments, the kit further includes one or more additional containers providing one or more additional reagents. For example, in specific variations, the one or more additional containers provide nucleoside triphosphates; a buffer suitable for polynucleotide extension; and/or a primer hybridizable, under polynucleotide extension conditions, to a predetermined polynucleotide template.

Further provided are reaction mixtures comprising the polymerases of the disclosure. In a sixth aspect, there is provided a reaction mixture comprising a DNA polymerase as defined herein, at least one primer, a polynucleotide template, and nucleoside triphosphates. The reactions mixtures can also contain a template nucleic acid (DNA and/or RNA), one or more primer or probe polynucleotides, nucleoside triphosphates (including, *e.g.*, deoxyribonucleotides, ribonucleotides, labeled nucleotides, unconventional nucleotides), buffers, salts, labels (*e.g.*, fluorophores).

Further embodiments of the disclosure are described herein.

DEFINITIONS

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention pertains. Although essentially any methods and materials similar to those described herein can be used in the practice or testing of the present invention, only exemplary methods and materials are described. For purposes of the present invention, the following terms are defined below.

The terms “a,” “an,” and “the” include plural referents, unless the context clearly indicates otherwise.

An "amino acid" refers to any monomer unit that can be incorporated into a peptide, polypeptide, or protein. As used herein, the term "amino acid" includes the following twenty natural or genetically encoded alpha-amino acids: alanine (Ala or A), arginine (Arg or R), asparagine (Asn or N), aspartic acid (Asp or D), cysteine (Cys or C), glutamine (Gln or Q),