



(51) International Patent Classification:

C12N 9/12 (2006.01) C12N 15/63 (2006.01)
C12N 15/54 (2006.01)

(21) International Application Number:

PCT/KR2012/000894

(22) International Filing Date:

7 February 2012 (07.02.2012)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

10-2011-0011639 9 February 2011 (09.02.2011) KR

(71) Applicant (for all designated States except US): **BIONEER CORPORATION** [KR/KR]; 8-11, Munpyeongseo-ro, Daedeok-gu, Daejeon 306-220 (KR).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **PARK, Han Oh** [KR/KR]; 208-601, Expo Apt., Jeonmin-dong, Yuseong-gu, Daejeon 305-761 (KR). **YANG, Sung Jun** [KR/KR];

102-603, Dasarang Apt., Moksang-dong, Daedeok-gu, Daejeon 306-760 (KR). **JOO, Sung Mo** [KR/KR]; 14-1306, Samsung Apt., Oryu-dong, Jung-gu, Daejeon 301-787 (KR). **HWANG, Byoung Oh** [KR/KR]; 105-1501, LG Deokdong Apt, Bijeon 2-dong, Pyeongtaek-si, Gyeonggi-do 450-729 (KR).

(74) Agent: **BAE, KIM & LEE IP GROUP**; 11th Floor, Shinduk Bldg., 1329-7 Secho 2-dong, Secho-gu, Seoul 137-858 (KR).

(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, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

[Continued on next page]

(54) Title: REVERSE TRANSCRIPTASE HAVING IMPROVED THERMOSTABILITY

[Fig. 1]

```

1 mlniedehr1 hetskepdvs lgstwlstdfp qawaetggmg lavrqaplii plkatstpv
61 ikrypmgsqa rigikphiqr lldqgilvpc qspwntpllp vkkpgtndyr pvqdzrevnk
  →L
121 rvedihptvp npynllsglp pshqnytvld lkdaaffclrl hptsqplfaf ewrdpengis
181 gqtlwtrlpq gfknsptlfd ealhrdladf riqtpdlill qyvddillaa tseldeqggt
241 rallqtlgnl gyasakkaq icqkqvkylg yllkeggzwl tearketvmg qptpkptrql
  →L →Q
301 reflgtagfc rlwpgfaem aaplyplkt gtlfmwpdq qkayqeikga lltapalglp
  →L →H
361 dltpkfelfv dekqgyakgv ltqklgpwrr pvaylskkld pvaagwppcl rmvaaiavlt
  →E
421 kdagklmgq plvilaphav ealvkppdx wlsnamthy gallldtdrv qfgpvvalnp
  →Y →F
481 atllplpeeg lqhnclidila eahgtrpdlr dqplpdadht wytdgssllq egqkagaav
541 ttetewiwak alpagsaqr aelialtgal kmaegkklmv ytdsryafat ahiehgyrr
601 rglitsegke iknkdeilal lkalfpkrl siihcyghqk ghsaeargnr madgaarkaa
661 itetpdtstl li

```

(57) Abstract: The present invention relates to a reverse transcriptase having improved thermostability, more precisely a mutant reverse transcriptase with improved thermostability by substitution of one or more amino acids selected from the group consisting of the 63rd glutamine (Q63), the 264th lysine (K264), the 295th lysine (K295), the 306th threonine (T306), the 346th glutamic acid (E346), the 408th proline (P408), the 438th histidine (H438), and the 454th asparagin (N454) of the amino acid sequence of M-MLV originated reverse transcriptase represented by SEQ. ID. NO: 1 with other amino acids. The mutant reverse transcriptase of the present invention demonstrates excellent thermostability, compared with the wild type reverse transcriptase. Therefore, it is advantageous to obtain the target cDNA with stable reverse transcription activity even in the presence of RNA that can form the stable secondary structure at a high temperature.



(84) **Designated States** (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS,

SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

- *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))*

Description

Title of Invention: REVERSE TRANSCRIPTASE HAVING IMPROVED THERMOSTABILITY

Technical Field

- [1] The present invention relates to a reverse transcriptase, more precisely a reverse transcriptase having improved thermostability by mutating a specific amino acid residue composing the conventional reverse transcriptase.

[2]

Background Art

- [3] Many studies have been undergoing about RNA tumor virus, particularly Moloney-Murine Leukemia Virus (M-MLV), human HIV, or Avian Myeloblastosis Virus (AMV) originated reverse transcriptases and accordingly their functions and properties have been disclosed. Reverse transcriptase has been used for many molecular biological methods such as cDNA library construction, reverse transcription and polymerase chain reaction (PCR), etc, owing to its unique characteristics favoring the synthesis of complementary DNA (cDNA) by using RNA as a template.
- [4] Three prototypes of retrovirus reverse transcriptase have been mainly studied. Moloney-Murine Leukemia Virus originated reverse transcriptase contains 78 kDa single subunit having RNA dependent DNA polymerase and RNase H activity. The said enzyme is cloned and expressed in *E. coli* as an authentic active form. HIV originated reverse transcriptase is hetero-dimer of p66 and p51 subunits. P51 subunit is generated by proteolytic cleavage of p66 subunit. P66 subunit contains both RNA dependent DNA polymerase and RNase H domains, but p51 subunit contains only DNA polymerase domain. Active HIV originated p66/p51 reverse transcriptase is cloned and expressed in many expression hosts including *E. coli*. In HIV p66/p51 hetero-dimer, 51 kDa subunit is catalytically inactive and 66 kDa subunit shows both DNA polymerase activity and RNase H activity. In the meantime, Avian Sarcoma-Leukosis Virus (ASLV) reverse transcriptases such as Rous Sarcoma Virus (RSV) reverse transcriptase, Avian Myeloblastosis Virus reverse transcriptase, Avian Erythroblastosis Virus (AEV) helper virus MCAV reverse transcriptase, Avian Myelocytomatosis Virus MC29 helper virus MCAV reverse transcriptase, Avian Reticuloendotheliosis Virus (REV-T) helper virus REV-A reverse transcriptase, Avian Sarcoma Virus UR2 helper virus UR2AV reverse transcriptase, Avian Sarcoma Virus Y73 helper virus YAV reverse transcriptase, Rous Associated Virus (RAV) reverse transcriptase, and Myeloblastosis Associated Virus (MAV) reverse transcriptase are hetero-dimer of two subunits of alpha (approximately 62 kDa) and beta (approximately

94 kDa). Alpha subunit is generated by proteolytic cleavage of beta subunit. ASLV reverse transcriptase can exist in two catalytically active structures of alpha/beta and beta/beta. From the sediment analysis, it was confirmed that the said alpha/beta and beta/beta structures are dimers and the alpha herein exists in the equilibrium state between monomer and dimer. ASLV alpha/beta or beta/beta reverse transcriptase is the only retrovirus reverse transcriptase informed up to date to have three different activities such as DNA polymerase activity, RNase H activity, and DNA endonuclease (integrase) activity. The alpha structure has neither integrase domain nor its activity.

- [5] The conversion of mRNA to cDNA by reverse transcriptase mediated reverse transcription is very important in the studies on gene expression. However, it is not preferred in many ways to use untransformed reverse transcriptase as a mediator for reverse transcription. Reverse transcriptase happens to degrade RNA template by RNase H activity before the first strand reaction begins or when it is completed. There is also a chance of error in the first strand of cDNA by mis-priming of mRNA template. In fact, during cDNA synthesis, HIV reverse transcriptase makes errors as many as 1 nucleotide error per 3,000 - 6,000 nucleotides, while AMV reverse transcriptase makes 1 nucleotide error per 10,000 nucleotides.
- [6] Another factor that affects the effect of reverse transcriptase is whether RNA can form the secondary structure or not. Such secondary structure can be formed when RNA molecular has enough complementarity to generate double-stranded RNA. In general, the formation of RNA secondary structure can be reduced by increasing the temperature of RNA molecule containing solution. So, it is preferred to perform reverse transcription of RNA at higher than 37°C. The reverse transcriptase of the present invention loses its activity when the culture is performed at much higher temperature than 37°C (for example, at 50°C).
- [7] A variety of methods to process thermostable reverse transcriptase are informed in the prior art, which include the method using thermostable DNA polymerase having reverse transcriptase activity, the method increasing reverse transcriptase activity by inducing mutation in thermostable DNA polymerase, the method inducing mutation in thermo-unstable reverse transcriptase, the method using Mn^{2+} instead of Mg^{2+} in the presence of *Taq/Tth* DNA polymerase, and the method using such additive as trehalose along with thermo-unstable reverse transcriptase.
- [8] To increase fidelity of polymerization of DNA or RNA template, those people skilled in the art have been using various enzyme compositions and methods. For example, Shevelev et al. provide a review article about 3'-5' exonuclease (Shevelev et al., *Nature Rev. Mol. Cell Biol.* 3:364 (2002)). Perrino et al. use the epsilon subunit of *E. coli* DNA polymerase III to increase fidelity of calf thymus DNA polymerase alpha (Perrino et al., PNAS USA, 86:3085 (1989)). Bakhanashvili explains proofreading

activity of p53 protein (Bakhanashvili, *Eur. J. Biochem.* 268:2047 (2001)), while Huang et al. describe the use of p53 to increase fidelity of DNA replication (Huang et al., *Oncogene*, 17:261 (1998)). US Patent Publication No. 2003/0198944A1 and US Patent No. 6,518,19 describe the enzyme mixture containing one or more reverse transcriptases (each reverse transcriptase has different transcription termination site) and if necessary containing one or more DNA polymerases additionally. US Patent Publication No. 2002/0119465A1 describes the composition containing mutant thermostable DNA polymerase and mutant reverse transcriptase (for example, mutant Taq DNA polymerase and mutant MMLV-RT). US Patent No. 6,485,917B1, US Patent Publication No. 2003/0077762, and European Patent Publication No. EP1132470 describe the method of cDNA synthesis in the presence of α -type DNA polymerase having 3'-5' exonuclease activity and the enzyme having reverse transcriptase activity.

- [9] When RNase H activity of reverse transcriptase is eliminated, the problem of RNA degradation of RNA template can be excluded, and further reverse transcription efficiency can be improved. However, such reverse transcriptase ('RNase H-' type) cannot solve the problems of mis-priming and the generation of mRNA secondary structure. The conventional reverse transcriptase has low thermostability, and thus reverse transcription is only induced at comparatively low temperature. Accordingly, reverse transcription product cannot be efficiently obtained by the interruption of RNA secondary structure formed generally at as high temperature as at least 65°C. Such limitation is a major barrier not only for the synthesis of cDNA from the full-length RNA but also for various biochemical experiments requiring reverse transcription such as RNA detection and profiling. Therefore, it is highly required to develop a novel reverse transcriptase having stable reverse transcription activity even at high temperature.

[10]

Disclosure of Invention

Technical Problem

- [11] The present invention is presented to solve the problem of the conventional reverse transcriptase as explained above. Therefore, it is an object of the present invention to provide a reverse transcriptase having improved thermostability that is effective in producing cDNA stably prepared by mutating a specific amino acid residue of M-MLV originated reverse transcriptase in order to give excellent reverse transcription activity even at higher temperature than the temperature allowing structural changes of RNA.

[12]

Solution to Problem

- [13] Terms and techniques described in this invention indicate the general meaning un-

derstood by those in the art, unless stated otherwise. References described in this invention to explain the present invention are all included in this description.

- [14] In this invention, the term "reverse transcriptase activity" or "reverse transcription" indicates the capability of enzyme to synthesize DNA strand (that is, complementary DNA or cDNA) using RNA as a template. In this invention, "reverse transcriptase" is an enzyme that shows the reverse transcriptase activity when the enzyme activity is measured by the method described in this description or by any conventional method informed to those in the art.
- [15] In this invention, the term "reverse transcription activity" or "reverse transcriptase activity" is reciprocally used to indicate the capability of enzyme to synthesize DNA strand (that is, cDNA) using RNA as a template.
- [16] In this invention, the term "variation" or "mutation" indicates the change introduced in wild type DNA sequence to change amino acid sequence encoded by DNA. Herein, substitution, insertion, deletion, and point mutation, etc, are included, but not always limited thereto.
- [17] In this invention, the term "wild type" indicates the gene or gene product having the same properties as the gene or gene product separated from the natural source. On the contrary, the term "modified" or 'mutant' indicates the gene or gene product showing changed or different properties from the wild type gene or gene product.
- [18] In this invention, the term "expression vector" indicates the expression cassette in which one or more transcription regulatory regions necessarily including coding sequence and promoter are operatively linked or the vector containing the said expression cassette.
- [19] In this invention, the term "coding sequence" indicates the DNA sequence encoding specific amino acids or functional RNA.
- [20] In this invention, the term "promoter" indicates the region inducing transcription of coding sequence to RNA, particularly the region where polymerase and transcription factors are combined.
- [21] In this invention, the amino acid residues composing reverse transcriptase are presented as 3 or 1 alphabet abbreviation as shown in Table 1.
- [22] Table 1

[Table 1]

alanine	A	Ala
cysteine	C	Cys
aspartic acid	D	Asp
glutamic acid	E	Glu
phenylalanine	F	Phe
glycine	G	Gly
histidine	H	His
isoleucine	I	Ile
lysine	K	Lys
leucine	L	Leu
methionine	M	Met
asparagin	N	Asn
proline	P	Pro
glutamine	Q	Gln
arginine	R	Arg
serine	S	Ser
threonine	T	Thr
selenocysteine	U	Sec
valine	V	Val
tryptophan	W	Trp
tyrosine	Y	Tyr

[23] Hereinafter, the present invention is described in detail.

[24] The present invention provides a reverse transcriptase having improved thermostability by substituting one or more amino acid residues selected from the group consisting of the 63rd glutamine (Q63), the 264th lysine (K264), the 295th lysine (K295), the 306th threonine (T306), the 346th glutamic acid (E346), the 408th proline (P408), the 438th histidine (H438), and the 454th asparagin (N454) of the amino acid sequence of M-MLV originated reverse transcriptase represented by SEQ. ID. NO: 1 with other amino acid residues.

[25] The substitution of amino acid herein includes one or more substitutions selected from the group consisting of substitution of the 63rd glutamine with leucine (Q63L),

substitution of the 264th lysine with leucine (K264L), substitution of the 295th lysine with glutamine (K295Q), substitution of the 306th threonine with leucine (T306L), substitution of the 346th glutamic acid with methionine (E346M), substitution of the 408th proline with glutamic acid (P408E), substitution of the 438th histidine with tyrosine (H438Y), and substitution of the 454th asparagin with phenylalanine (N454F), but not always limited thereto (see Figure 1). In a preferred embodiment of the present invention, the thermostability of 8 mutant reverse transcriptases having the amino acid sequences each represented by SEQ. ID. NOs: 2 - 9, in which the said 8 amino acid sites were substituted, was confirmed. Even if at least 2 amino acid sites were simultaneously mutated in the reverse transcriptase, it can be included in the spirit and scope of the present invention as long as it showed thermostability at high temperature.

[26] In a preferred embodiment of the present invention, the mutant reverse transcriptases of the present invention demonstrated higher thermostability than the wild type reverse transcriptase. Among those mutants, K295Q, T306L, and P408E demonstrated excellent thermostability (see Figure 6). In particular, the mutant reverse transcriptase T306L showing the highest thermostability demonstrated higher thermostability at the temperatures of 60°C, 65°C and 70°C with producing cDNA than the wild type M-MLV originated reverse transcriptase (see Figure 7).

[27] Those skilled in the art are well aware of that even when the amino acid substituted once is replaced again with another amino acid having similar characteristics (that is, conservative amino acid substitution), similar physiological and biochemical properties are still observed. The effect of amino acid substitution on the various amino acid properties and protein structure and functions has been targeted by those in the art.

[28] For example, hydropathic index of amino acid can be considered (Kyte & Doolittle, 1982, *J. Mol. Biol.*, 157:105-132). Relative hydropathic property of amino acid is attributed to the secondary structure of the generated protein, which is used for the determination of interaction of other molecules. Each amino acid has hydropathic index based on its hydrophobicity and charge characteristics (Kyte & Doolittle, 1982, *J. Mol. Biol.*, 157:105-132), which is as follows: isoleucine (+4.5); valine (+4.2); leucine (+3.8); phenylalanine (+2.8); cysteine/cystine (+2.5); methionine (+1.9); alanine (+1.8); glycine (-0.4); threonine (-0.7); serine (-0.8); tryptophan (-0.9); tyrosine (-1.3); proline (-1.6); histidine (-3.2); glutamic acid (-3.5); glutamine (-3.5); aspartic acid (-3.5); asparagin (-3.5); lysine (-3.9); and arginine (-4.5). In conservative substitution, it is preferred to use the amino acid having the hydropathic index of less than ± 2 , and the amino acid having the hydropathic index of less than ± 1 is more preferred and the amino acid having the hydropathic index of less than ± 0.5 is most preferred.

[29] For the amino acid substitution, hydrophilicity of an amino acid residue can be also considered (see US Patent No. 4,554,101). Each amino acid residue has its own hy-

drophilicity value, which is as follows: arginine (+3.0); lysine (+3.0); aspartic acid (+3.0); glutamic acid (+3.0); serine (+0.3); asparagin (+0.2); glutamine (+0.2); glycine (0); threonine (-0.4); proline (-0.5 .+-1); alanine (-0.5); histidine (-0.5); cysteine (-1.0); methionine (-1.3); valine (-1.5); leucine (-1.8); isoleucine (-1.8); tyrosine (-2.3); phenylalanine (-2.5); and tryptophan (-3.4). It is preferred to replace an amino acid residue with the one having similar hydrophilicity, but not always limited thereto.

- [30] The size of amino acid side chain can be also considered. For example, it is not preferred to replace such amino acid having dense side chain such as glycine or serine with the amino acid having big side chain such as tryptophan or tyrosine. The effect of various amino acid residues on the secondary structure of protein can be also considered. Through the experience obtained from the previous studies, it has been disclosed and measured that amino acid residues have effect on the selection of secondary structure of protein domain such as alpha-helix, beta-sheet, or reverse turn (Chou & Fasman, 1974, *Biochemistry*, 13:222-245; 1978, *Ann. Rev. Biochem.*, 47:251-276; 1979, *Biophys. J.*, 26:367-384).
- [31] Based on the broad experimental studies and the said considerations, the conservative amino acid substitution table has been made, which has been well informed to those in the art. Some examples are as follows: arginine and lysine; glutamic acid and aspartic acid; serine and threonine; glutamine and asparagin; and valine, leucine and isoleucine. Other examples are as follows: Ala(A) leu, ile, val; Arg(R) gln, asn, lys; Asn(N) his, asp, lys, arg, gln; Asp(D) asn, glu; Cys(C) ala, ser; Gln(Q) glu, asn; Glu(E) gln, asp; Gly(G) ala; His(H) asn, gln, lys, arg; Ile(I) val, met, ala, phe, leu; Leu(L) val, met, ala, phe, ile; Lys(K) gln, asn, arg; Met(M) phe, ile, leu; Phe(F) leu, val, ile, ala, tyr; Pro(P) ala; Ser(S), thr; Thr(T) ser; Trp(W) phe, tyr; Tyr(Y) trp, phe, thr, ser; Val(V) ile, leu, met, phe, ala.
- [32] For the amino acid substitution, it can be also considered whether the target residue is located in the inside of the protein or exposed to the solvent. If the residue is in the inside, conservative substitution includes as follows: Asp and Asn; Ser and Thr; Ser and Ala; Thr and Ala; Ala and Gly; Ile and Val; Val and Leu; Leu and Ile; Leu and Met; Phe and Tyr; and Tyr and Trp. If the residue is exposed on the solvent, conservative substitution includes as follows: Asp and Asn; Asp and Glu; Glu and Gln; Glu and Ala; Gly and Asn; Ala and Pro; Ala and Gly; Ala and Ser; Ala and Lys; Ser and Thr; Lys and Arg; Val and Leu; Leu and Ile; Ile and Val; and Phe and Tyr.
- [33] Before determining amino acid substitution, intramolecular or intermolecular conjugation such as disulfide bond between neighboring cysteine residues or ionic bond (salt bridge) between positive charged residue (for example, His, Arg, and Lys) and negative charged residue (for example Asp, and Glu) is preferably considered.
- [34] The present invention also provides a gene encoding the said reverse transcriptase

having improved thermostability.

- [35] The said gene is the gene encoding the mutant reverse transcriptase harboring one or more amino acid substations selected from the group consisting of substitution of the 63rd glutamine of the wild type reverse transcriptase represented by SEQ. ID. NO: 1 with leucine (Q63L), substitution of the 264th lysine of the wild type reverse transcriptase represented by SEQ. ID. NO: 1 with leucine (K264L), substitution of the 295th lysine of the wild type reverse transcriptase represented by SEQ. ID. NO: 1 with glutamine (K295Q), substitution of the 306th threonine of the wild type reverse transcriptase represented by SEQ. ID. NO: 1 with leucine (T306L), substitution of the 346th glutamic acid of the wild type reverse transcriptase represented by SEQ. ID. NO: 1 with methionine (E346M), substitution of the 408th proline of the wild type reverse transcriptase represented by SEQ. ID. NO: 1 with glutamic acid (P408E), substitution of the 438th histidine of the wild type reverse transcriptase represented by SEQ. ID. NO: 1 with tyrosine (H438Y), and substitution of the 454th asparagin of the wild type reverse transcriptase represented by SEQ. ID. NO: 1 with phenylalanine (N454F). In a preferred embodiment of the present invention, the gene of the invention is preferably the gene having one of the nucleotide sequences each represented by SEQ. ID. NOs: 11 - 18 encoding 8 mutant reverse transcriptases having the amino acid sequences each represented by SEQ. ID. NOs: 2 - 9 in which the said 8 amino acid sites are substituted. Any random gene encoding the mutant reverse transcriptase having at least two amino acid substitutions selected from the group consisting of the 63rd glutamine (Q63), the 264th lysine (K264), the 295th lysine (K295), the 306th threonine (T306), the 346th glutamic acid (E346), the 408th proline (P408), the 438th histidine (H438), and the 454th asparagin (N454) of the M-MLV originated reverse transcriptase represented by SEQ. ID. NO: 1 can be included in the spirit and scope of the present invention (see Figure 2 - Figure 5).

- [36] In addition to the said genes hereinabove, polynucleotides which are actually identical or have same functions as the said genes can be included in the present invention. The phrase "actually identical or have same functions" herein indicates that two polynucleotides demonstrate at least 70% homology, preferably at least 80%, more preferably at least 90%, and most preferably at least 95% homology when they are properly arranged by well-informed computerized algorithm or investigation.

- [37] The present invention also provides an expression vector containing the said gene.

- [38] The mother vector used for the construction of the expression vector of the present invention is not limited, and any conventional vector for the transformation of prokaryotes or eukaryotes can be used. In a preferred embodiment of the present invention, recombinant expression vectors were constructed by inserting each mutant gene represented by SEQ. ID. NO: 11 - NO: 18 into PET 22b(+), the extranuclear gene

vector.

[39] The present invention also provides a transformant transformed by the said expression vector.

[40] The transformant of the present invention can be easily constructed by inserting the said expression vector into random prokaryotic cells or eukaryotic cells. The method to introduce a specific vector into cells is well-known to those in the art. Lipofectamine method is an example of the methods. In a preferred embodiment of the present invention, the PET 22b(+) vector introduced with the mutant gene was introduced in *E. coli* DH5 α , leading to the construction of a transformant.

[41] The present invention also provides a kit for reverse transcription comprising the said reverse transcriptase.

[42] The kit for reverse transcription of the present invention can additionally include, in addition to the said reverse transcriptase, any conventional constituent necessary for reverse transcription such as a primer pair binding specifically to the target gene or oligonucleotide for amplification, dNTP, and reaction buffer, and DNA polymerase if necessary.

[43]

Advantageous Effects of Invention

[44] The mutant reverse transcriptase of the present invention demonstrates excellent thermostability, compared with the wild type reverse transcriptase. Therefore, even in the presence of specific RNA that can form a very stable secondary structure at a high temperature which is a barrier for the efficient reverse transcription, the stable reverse transcription activity can be induced to give wanted cDNA with the reverse transcriptase of the present invention.

[45]

Brief Description of Drawings

[46] The application of the preferred embodiments of the present invention is best understood with reference to the accompanying drawings, wherein:

[47]

[48] Figure 1 is a diagram showing the sites and kinds of amino acids mutated in the mutant reverse transcriptase of the present invention.

[49] Figure 2 - Figure 5 are diagrams showing the sites and sequences of nucleotides mutated in the gene encoding the mutant reverse transcriptase of the present invention.

[50] Figure 6 is a set of electrophoresis photographs showing the results of reverse transcription induced by using the wild type reverse transcriptase (SEQ. ID. NO: 1) and the mutant reverse transcriptases of the present invention (SEQ. ID. NOs: 2 - 9) at 42°C, at 50°C, at 60°C, at 65°C, and at 70°C.

[51] Figure 7 is a set of electrophoresis photographs showing the results of reverse transcription induced by using the wild type reverse transcriptase (SEQ. ID. NO: 1) and the T306L mutant reverse transcriptase of the present invention (SEQ. ID. NO: 5) at 37°C, at 42°C, at 50°C, at 60°C, at 65°C, and at 70°C.

[52]

Best Mode for Carrying out the Invention

[53] Practical and presently preferred embodiments of the present invention are illustrative as shown in the following Examples.

[54] However, it will be appreciated that those skilled in the art, on consideration of this disclosure, may make modifications and improvements within the spirit and scope of the present invention.

[55]

Example 1: Construction of gene encoding the mutant reverse transcriptase

[57] To induce mutation in the M-MLV originated reverse transcriptase having the amino acid sequence represented by SEQ. ID. NO: 1, nucleotide sequence of the gene encoding the M-MLV originated reverse transcriptase was substituted. Particularly, 8 mutant reverse transcriptases each harboring substitution of the 63rd glutamine of the M-MLV originated reverse transcriptase represented by SEQ. ID. NO: 1 with leucine (Q63L), substitution of the 264th lysine of the M-MLV originated reverse transcriptase represented by SEQ. ID. NO: 1 with leucine (K264L), substitution of the 295th lysine of the M-MLV originated reverse transcriptase represented by SEQ. ID. NO: 1 with glutamine (K295Q), substitution of the 306th threonine of the M-MLV originated reverse transcriptase represented by SEQ. ID. NO: 1 with leucine (T306L), substitution of the 346th glutamic acid of the M-MLV originated reverse transcriptase represented by SEQ. ID. NO: 1 with methionine (E346M), substitution of the 408th proline of the M-MLV originated reverse transcriptase represented by SEQ. ID. NO: 1 with glutamic acid (P408E), substitution of the 438th histidine of the M-MLV originated reverse transcriptase represented by SEQ. ID. NO: 1 with tyrosine (H438Y), and substitution of the 454th asparagin of the M-MLV originated reverse transcriptase represented by SEQ. ID. NO: 1 with phenylalanine (N454F) were constructed (see Figure 1). To do so, mutation was induced in the gene represented by SEQ. ID. NO: 10 encoding the reverse transcriptase having the amino acid sequence represented by SEQ. ID. NO: 1 (see Figure 2 to Figure 5). The said mutant gene has one of the nucleotide sequences each represented by SEQ. ID. NOs: 11 - 18. The procedure of the construction of the said mutant gene is as follows:

[58] To increase protein expression in *Escherichia coli*, gene sequence of the M-MLV reverse transcriptase represented by SEQ. ID. NO: 10 was optimized to fit codon usage

of *E. coli*, and as a result, the nucleotide sequence represented by SEQ. ID. NO: 19 was obtained. Based on the optimized nucleotide sequence, wild type gene and K295Q mutant gene were synthesized. First, to obtain the nucleotide sequence of the wild type gene represented by SEQ. ID. NO: 19, sense and antisense strands were designed to have the annealing temperature of about 60°C (20 - 40 bases). Then, the designed oligonucleotides were synthesized (Bioneer, Korea). The synthesized oligonucleotides had the nucleotide sequences each represented by SEQ. ID. NOs: 20 - 149. Ligation of the synthesized oligonucleotides was induced by Kination-LCR (Korean Patent Application No. 2008-0025050) by repeated reaction at 40°C and at 60°C, leading to the synthesis of double-stranded total gene having the nucleotide sequence represented by SEQ. ID. NO: 19. K295Q mutant gene was also constructed by the same manner as described above except that the nucleotide sequences each represented by SEQ. ID. NO: 150 and NO: 151 were used instead of those nucleotide sequences each represented by SEQ. ID. NO: 77 and NO: 78. The synthesized genes were heat-treated at 95°C for 5 minutes, followed by reaction with the primers represented by SEQ. ID. NO: 152 and NO: 153 at 95°C for 1 minute, at 65°C for 1 minute, and at 72°C for 2 minutes and 30 seconds. This reaction cycle was repeated 30 times and additional reaction was induced at 72°C for 10 minutes to amplify the gene. The amplified gene was ligated in pGEM-T easy vector (Promega, USA), which was used for the transformation of *E. coli* DH5a (RBC, USA). As a result, wild type and K295Q mutant gene clones were obtained at a large scale.

- [59] Mutation was induced by site-directed mutagenesis by using the synthesized wild type gene as a template to produce mutations of Q63L, K264L, T306L, E346M, P408E, H438Y, and N454F. The site-directed mutagenesis was performed as follows: First, oligonucleotides having the nucleotide sequences each represented by SEQ. ID. NOs: 154 - 167 were synthesized. The mutant gene was heat-treated at 95°C for 5 minutes, followed by reaction with the synthesized oligonucleotides as primers and pfu DNA polymerase at 95°C for 1 minute, at 65°C for 1 minute, and at 72°C for 5 minutes and 30 seconds. This reaction cycle was repeated 30 times and additional reaction was induced at 72°C for 10 minutes to amplify the gene. The amplified mutant gene product was self-ligated by Kination-self ligation, which was used for the transformation of *E. coli* DH5a (RBC, USA). As a result, clones of Q63L, K264L, T306L, E346M, P408E, H438Y, and N454F mutant genes were obtained at a large scale.

[60]

[61] Example 2: Mutant gene transformation

- [62] The 8 mutant reverse transcriptase genes constructed in Example 1 were cloned by using the extranuclear gens PET 22b(+) (NOVAGEN CO.) as a recombination vector. To do so, PCR was performed with the mutant genes (Q63L, K264L, T306L, and

E346M) synthesized by being introduced in pGEM-T easy vector (Promega) using the following primer set.

[63] Sense strand: 5'-GCG CGC CAT ATG CTG AAC ATC GAA GAC GAA CAC CGT CTG CAC GAA AC-3'(Nde I) (SEQ. ID. NO: 168)

[64] Antisense strand: 5'-GCG CGC GCG GCC GCT TAG ATC AGC AGG GTA GAG GTG TCC GGG GTT TC-3'(Not I) (SEQ. ID. NO: 169)

[65] PCR was also performed with other mutant genes (K295Q, P408E, H438Y, and N454Y) using the following primer set.

[66] Sense strand: 5'-GCG CGC CAT ATG CTG AAC ATC GAA GAC GAA CAC CGT CTG CAC GAA AC-3'(Nde I) (SEQ. ID. NO: 170)

[67] Antisense strand: 5'-GCG CGC GCG GCC GCG ATC AGC AGG GTA GAG GTG TCC GGG GTT TC-3'(Not I) (SEQ. ID. NO: 171)

[68] The extranuclear gene vector PET 22b(+) and each M-MLV reverse transcriptase gene fragment obtained by PCR were digested with NdeI (5' end) and NotI (3' end). The digested fragments were purified by using gel extraction kit (BIONEER CO.). The fragments were reacted at 16°C for 2 hours by using T4 DNA ligase (TAKARA CO.). Then, each mutant gene was inserted in the extranuclear gene vector PET 22b(+).

[69] *Escherichia coli* DH5 α was transformed with the extranuclear gene vector PET 22b(+) introduced with the mutant gene, and as a result recombinant host cells were prepared. The transformation herein was performed by the following steps: DH5 α host cells frozen at -70°C were thawed. The thawed host cells were mixed with 1 $\mu\text{g}/\mu\ell$ of the DNA recombinant vector, which stood in ice for 30 minutes, followed by heat-treatment at 42°C for 90 seconds. The heat-treated host cells were smeared on LB medium plate containing ampicillin (50 mg/ml), followed by culture at 37°C for overnight. As a result, clones were obtained. Each gene sequence of the obtained clones was confirmed by using the primer set comprising the following nucleotide sequences. Each plasmid of the confirmed clones was introduced in BL21(DE3) host cells for transformation and then expression strains were obtained.

[70] T7 promoter sense strand: 5'-TAA TAC GAC TCA CTA TAG GG-3'(SEQ. ID. NO: 172)

[71] T7 terminator antisense strand: 5'-GCT AGT TAT TGC TCA GCG G-3'(SEQ. ID. NO: 173)

[72]

[73] Example 3: Expression of mutant reverse transcriptase

[74] Single colony of the expression strain prepared by transforming BL21(DE3) host cells was taken and transferred in 15 ml of LB (Luria-Bertani) medium containing ampicillin (50 mg/ml), followed by culture at 37°C for overnight. As a result, seed cells were obtained. The obtained seed cells were placed in 1 l of LB medium containing

ampicillin, followed by culture at 37°C for 2 hours, to which 0.5 mM IPTG (isopropyl β -D-1-thiogalactopyranoside) was added. Expression was induced for 2 more hours. The cultured cells were centrifuged and as a result pellet was obtained. The pellet was stored at -50°C until the pellet was purified.

[75]

[76] Example 4: Purification of mutant reverse transcriptase

[77] 5 g of the frozen cell pellet obtained in Example 3 was pulverized, followed by suspension in lysis buffer (25 mM Tris-HCl (pH 7.6), 1 mM EDTA, 10%(v/v) β -mercaptoethanol, and phenylmethylsulfonyl fluoride (PMSF) 20 μ l) for 30 minutes with stirring at 4°C. Then, the cells were lysed by using Fisher sonicator. The cell lysate was centrifuged at 11,000 rpm for 1 hour at 4°C, and pellet was discarded. The remaining clear lysate was dialyzed in a buffer (40 mM Tris-HCl (pH 7.6), 2 mM EDTA, 28.58 mM β -mercaptoethanol, and 100 mM KCl).

[78]

[79] Example 5: Chromatography of mutant reverse transcriptase

[80] FPLC was performed to purify the mutant reverse transcriptase by using Amersham FPLC device. At that time, Amersham XK16 (1.6 cm $\times$ 30.0 cm) was used as the first column. The column was filled with ion exchange resin column chromatography DEAE hyper-D (Pall Co.), which was washed with 250 μ l of column buffer containing 1 M KCl. Then, 60 μ l of bed equilibrated with buffer was obtained. 20 μ l of the clear cell lysate was loaded to the column at the flow rate of 5 μ l/min. Among the loaded lysate, those not absorbed on the column were taken and then applied to the next column. The obtained lysate was applied to the second chromatography. The column Amersham XK16 (1.6 cm $\times$ 30.0 cm) was washed and equilibrated. The column was filled with affinity chromatography heparin hyper-D (Pall Co.). As a result, 50 μ l of bed was obtained. The bed was equilibrated with buffer. The lysate was loaded to the column, followed by elution with 100 - 150 μ l of 0.2 M - 0.6 M KCl linear salt gradient of column buffer. The eluted lysate was applied to 26X60 S-200 gel filtration column (GE healthcare). As a result, high purity reverse transcriptase was obtained. The column fraction proceeded to SDS-polyacrylamide gel electrophoresis (SDS-PAGE), followed by Coomassie brilliant blue staining. 10 μ l was taken from each fraction, which was analyzed on each gel lane. Protein having similar molecular weight of wild type (74 KDa) was collected. To secure the stability of the collected protein, the protein was dialyzed in conservative solution (20 mM Tris-HCl (pH 7.6), 0.1 mM EDTA, 150 mM NaCl, 0.1% IGEPAL CA-630 (Polyethoxyethanol) 1 mM DTT (Dithiothreitol), and 50% glycerol), and as a result the mutant enzyme was obtained and stored at -20°C.

[81]

[82] Example 6: RT-PCR with mutant reverse transcriptase

[83] RT-PCR was performed to confirm the activity of the separated mutant reverse transcriptase. Ingredients used for RT-PCR were as follows, and the total RNA was used at the concentrations of 5 ng/ μ l, 500 pg/ μ l, 50 pg/ μ l, and 5 pg/ μ l.

[84] 5X reaction buffer 4 μ l (Bioneer Co.)

[85] 10 mM dNTP 2 μ l

[86] dT(18) 10 pmol 1 μ l

[87] 100 mM DTT 2 μ l

[88] RNase inhibitor (50 ng/ μ l) 1 μ l

[89] DW 7 μ l

[90] Total RNA 2 μ l

[91] For the comparative example, wild type reverse transcriptase was used for RT-PCR performed at 42°C, 50°C, 60°C, 65°C, and 70°C (in the case of T306L, reaction at 37°C was added) for 1 hour. The produced cDNA was used for PCR. The target gene for the amplification was GAPDH (500 bp) and the primer set having the nucleotide sequences presented below was used.

[92] Sense strand: 5'-GAAGGTGAAGGTCGGAGTCAACG-3' (SEQ. ID. NO: 174)

[93] Antisense strand: 5'-AGTCCTTCCACGATACCAAAGTTG-3' (SEQ. ID. NO: 175)

[94] Preparation having the below composition was used:

[95] PCR premix type (Bioneer)

[96] Sense primer: 5 pmol, 1 μ l

[97] Antisense primer: 5 pmol, 1 μ l

[98] DW 13 μ l

[99] cDNA 5 μ l

[100] PCR was performed as follows: 95°C for 30 seconds, 57°C for 30 seconds, and 72°C for 30 seconds (30 cycles). The PCR product proceeded to agarose gel electrophoresis to confirm the amplified product.

[101] As a result, the wild type M-MLV originated mutant reverse transcriptase of the present invention demonstrated higher thermostability than the wild type reverse transcriptase. In particular, K295Q, T306L, and P408E mutant reverse transcriptases showed excellent thermostability. The wild type M-MLV reverse transcriptase did not show activity at 60°C, while K295Q, T306L, and P408E mutant reverse transcriptases maintained reverse transcription activity at the reaction temperature reaching 60°C (Figure 6). T306L mutant reverse transcriptase showing the highest thermostability demonstrated equal thermostability to the wild type M-MLV reverse transcriptase at 37°C, 42°C, and at 50°C, but showed much higher thermostability than the wild type M-MLV reverse transcriptase at 60°C, 65°C, and at 70°C.

[102]

[103] Those skilled in the art will appreciate that the conceptions and specific embodiments disclosed in the foregoing description may be readily utilized as a basis for modifying or designing other embodiments for carrying out the same purposes of the present invention. Those skilled in the art will also appreciate that such equivalent embodiments do not depart from the spirit and scope of the invention as set forth in the appended claims.

Claims

- [Claim 1] A reverse transcriptase having improved thermostability by substitution of one or more amino acids selected from the group consisting of the 63rd glutamine (Q63), the 264th lysine (K264), the 295th lysine (K295), the 306th threonine (T306), the 346th glutamic acid (E346), the 408th proline (P408), the 438th histidine (H438), and the 454th asparagin (N454) of the amino acid sequence of M-MLV originated reverse transcriptase represented by SEQ. ID. NO: 1 with other amino acids.
- [Claim 2] The reverse transcriptase according to claim 1, wherein the substitution of amino acid includes one or more substitutions selected from the group consisting of substitution of the 63rd glutamine of the wild type reverse transcriptase represented by SEQ. ID. NO: 1 with leucine (Q63L), substitution of the 264th lysine with leucine (K264L), substitution of the 295th lysine with glutamine (K295Q), substitution of the 306th threonine with leucine (T306L), substitution of the 346th glutamic acid with methionine (E346M), substitution of the 408th proline with glutamic acid (P408E), substitution of the 438th histidine with tyrosine (H438Y), and substitution of the 454th asparagin with phenylalanine (N454F).
- [Claim 3] The reverse transcriptase according to claim 2, wherein the reverse transcriptase has the amino acid sequence selected from the group consisting of amino acid sequences each represented by SEQ. ID. NOs: 2 - 9.
- [Claim 4] The reverse transcriptase according to any one of claims 1 to 3, wherein the reverse transcriptase demonstrates higher reverse transcription activity at the reaction temperature of 60°C - 70°C than the wild type M-MLV reverse transcriptase having the amino acid sequence represented by SEQ. ID. NO: 1.
- [Claim 5] A gene encoding the reverse transcriptase having improved thermostability of claim 1.
- [Claim 6] The gene encoding the reverse transcriptase according to claim 5, wherein the gene is the one encoding the mutant reverse transcriptase containing one or more substitutions of amino acid selected from the group consisting of the 63rd glutamine (Q63), the 264th lysine (K264), the 295th lysine (K295), the 306th threonine (T306), the 346th glutamic acid (E346), the 408th proline (P408), the 438th histidine (H438), and the 454th asparagin (N454) of the amino acid sequence of M-MLV

- originated reverse transcriptase represented by SEQ. ID. NO: 1.
- [Claim 7] The gene encoding the reverse transcriptase according to claim 6, wherein the gene is the one encoding the mutant reverse transcriptase containing one or more amino acid substitutions selected from the group consisting of substitution of the 63rd glutamine of the wild type reverse transcriptase represented by SEQ. ID. NO: 1 with leucine (Q63L), substitution of the 264th lysine with leucine (K264L), substitution of the 295th lysine with glutamine (K295Q), substitution of the 306th threonine with leucine (T306L), substitution of the 346th glutamic acid with methionine (E346M), substitution of the 408th proline with glutamic acid (P408E), substitution of the 438th histidine with tyrosine (H438Y), and substitution of the 454th asparagin with phenylalanine (N454F).
- [Claim 8] The gene encoding the reverse transcriptase according to claim 7, wherein the gene has the nucleotide sequence selected from the group consisting of the nucleotide sequences each represented by SEQ. ID. NOs: 11 - 18.
- [Claim 9] An expression vector comprising the gene of claim 5.
- [Claim 10] A transformant transformed with the expression vector of claim 9.
- [Claim 11] A kit for reverse transcription comprising the reverse transcriptase having improved thermostability of claim 1.

[Fig. 4]

427 tccaccagtggtacactgtgcntgantttaaaggatgcctttttctgacctgagactccac 483
||||| ||||||| ||||||| ||||||| ||||||| ||||| ||| |

421 tccaccagtggtacactgtgcctlga.ttaaaggatgcctttttctgacctgagactccac 480

484 cccaccagtcagcctctcttttgccttcgagtcggagagatccagagatcgggaatctcagga 540
||||| ||||||| ||||||| ||||||| ||||||| ||||| ||| |

481 cccaccagtcagcctctcttttgccttcgagtcggagagatccagagatcgggaatctcagga 540

544 caattgacctggaccagactccacacaggg.ttccaaaaacagtcaccacctg.tttgatgag 603
||||| ||||||| ||||||| ||||||| ||||||| ||||| ||| |

541 caattgacctggaccagactccacacaggg.ttccaaaaacagtcaccacctg.tttgatgag 600

604 ggaactgcacagagacctaccagacttcctggtatccacacccagactcgatcctgctacag 663
||||| ||||||| ||||||| ||||||| ||||||| ||||| ||| |

601 ggaactgcacagagacctaccagacttcctggtatccacacccagactcgatcctgctacag 660

661 taagtggatgaattactgctggcgccacttcctgagctagaactgccaaacaaggtactcgg 723
||||| ||||||| ||||||| ||||||| ||||||| ||||| ||| |

661 taagtggatgaattactgctggcgccacttcctgagctagaactgccaaacaaggtactcgg 720

[Fig. 5]

724 gccctgttataaaacccctaggggaacctogggatcgggcctcgggccaagaaagcccaaatt 783

||||| |||||||||||||||||| ||||||| ||||||| | ||||||| |||||| || ||

721 gccctgttataaaacccctaggggaacctogggatcgggcctcgggccaagaaagcccaaatt 780

781 tgcagaaaacaggtoaagtatcgggggtatcttctaataagagggtcagagatggctgaact 843

||||| |||||||||||||||||| ||||||| ||||||| | ||||||| |||||| || ||

781 tgcagagcttcaggtoaagtatcgggggtatcttctaataagagggtcagagatggctgaact 840**K264L: AAA→CTT**

844 gagggcagaaaagagactctgatggggcagccctactccgaagacccctcgacaactaagg 903

||||| |||||||||||||||||| ||||||| ||||||| | ||||||| |||||| || ||

841 gagggcagaaaagagactctgatggggcagccctactccgcaaacccctcgacaactaagg 900**K295Q: AAG→CAA**

904 gagtccctagggacggcagcctctctgctgcctctggatccctgggctctgcagaaatggca 963

||||| |||||||||||||||||| ||||||| ||||||| | ||||||| |||||| || ||

901 gagtccctagggCTTgcagcctctctgctgcctctggatccctgggctctgcagaaatggca 960**T306L: ACG→CTT**

[Fig. 6]

964 gccscattgtaccctotcaccaaaaacggggactctgtttaattggggcccagaccaacaa 1023

||||| ||||||| ||||||| ||||||| ||||| | |

961 gccscattgtaccctotcaccaaaaacggggactctgtttaattggggcccagaccaacaa 1020

1024 aaggcctatcaagaatatcaagcaagctctcttaactgccccagccctgggggttgcacat 1083

||||| ||||||| ||||||| ||||||| ||||||| ||||| | |

1021 aaggcctatcaaaatgatcaagcaagctctcttaactgccccagccctgggggttgcacat 1080

E346M: GAA→ATG

1084 ttgaactaagccctttgaactctttgtcgacgagaagcagggctacgccaaagggtgccta 1143

||||| ||||||| ||||||| ||||||| ||||||| ||||| | |

1081 ttgaactaagccctttgaactctttgtcgacgagaagcagggctacgccaaagggtgccta 1140

1144 acgcacaaaactgggaccttggcgctggccgggtggcctacctgtccaaaaagctagacca 1203

||||| ||||||| ||||||| ||||||| ||||||| ||||| | |

1141 acgcacaaaactgggaccttggcgctggccgggtggcctacctgtccaaaaagctagacca 1200

1204 gtagcagctgggtggcccccttgccacgggatggtagcagccattgccgtactgacaaag 1263

||||| ||||||| ||||||| ||||||| ||||||| ||||| | |

1201 gtagcagctgggtggccccgaatcctacgggatggtagcagccattgccgtactgacaaag 1260

P408E: CCT→GAA

[Fig. 7]

1264 gatgcaggcaagctaaccatgggacagccactagtcattctggccccccatgcagtagag 1323

||||| ||||||| ||||||| ||||||| ||||||| ||||| | |||||

1261 gatgcaggcaagctaaccatgggacagccactagtcattctggcccccctacgcagtagag 1320**K438Y: CAT→TAC**

1324 gcactagtcaaaacaaccccccgaccgctggctttccaacgcocggatgaactcactatcag 1383

||||||| ||||||| ||||||| ||||||| ||||||| ||||| | |

1321 gcactagtcaaaacaaccccccgaccgctggctttccttcgcocggatgaactcactatcag 1380**N454F: AAC→TTC**

1384 gccttgcttttggacacggacgggtccagttcggacgggtggtagccctgaaccgggt 1443

||||||| ||||||| ||||||| ||||||| ||||||| ||||| | |

1381 gccttgcttttggacacggacgggtccagttcggacgggtggtagccctgaaccgggt 1440

1444 acgctgctccactgctgaggaagggtgcaacacaactgccttgatatcctggccgaa 1503

||||||| ||||||| ||||||| ||||||| ||||||| ||||| | |

1441 acgctgctccactgctgaggaagggtgcaacacaactgccttgatatcctggccgaa 1500

[Fig. 8]

1504 gccacagggaaccogacccgacclaacggaccagccgcgccagacccgaccacaccctgg 1563
|| ||||| ||||||||||||||||||| ||||||||||||||||||| |||

1501 gccacagggaaccogacccgacctaacggaccagccgcgccagacccgaccacaccctgg 1560

1564 tacacggatgggaagcagctctcttacaagagggacagcgtaaggccgggagctgcgggtgacc 1623
|| ||||| ||||||||||||||||||| ||||||||||||||||||| |||

1561 tacacggatgggaagcagctctcttacaagagggacagcgtaaggccgggagctgcgggtgacc 1620

1624 accgagacccgaggtaatctgggctaaagccctggccagccgggacatccgctcagccgggct 1683
|| ||||| ||||||||||||||||||| ||||||||||||||||||| |||

1621 accgagacccgaggtaatctgggctaaagccctggccagccgggacatccgctcagccgggct 1680

1687 gaactgatagcactcaccacaggccctaaagatggccagaaggtaagaagctaatgttttat 1743
|| ||||| ||||||||||||||||||| ||||||||||||||||||| |||

1681 gaactgatagcactcaccacaggccctaaagatggccagaaggtaagaagctaatgttttat 1740

1744 actgatagccglttatgctllltgctacggccatataccatggagaaatatacagaaggcgt 1803
|| ||||| ||||||||||||||||||| ||||||||||||||||||| |||

1741 actgatagccglttatgctllltgctacggccatataccatggagaaatatacagaaggcgt 1800

[Fig. 9]

1804 gggTtggtcacatcagaaggcaagagatcaaaaaaaagacgagatcttggccctacta 1863
 || ||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||

1801 gggTtggtcacatcagaaggcaagagatcaaaaaaaagacgagatcttggccctacta 1860

1864 aaagccctctttctgccccaaaagacttagcataatccattgtccaggacatcaaaaggga 1923
 || ||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||

1861 aaagccctctttctgccccaaaagacttagcataatccattgtccaggacatcaaaaggga 1920

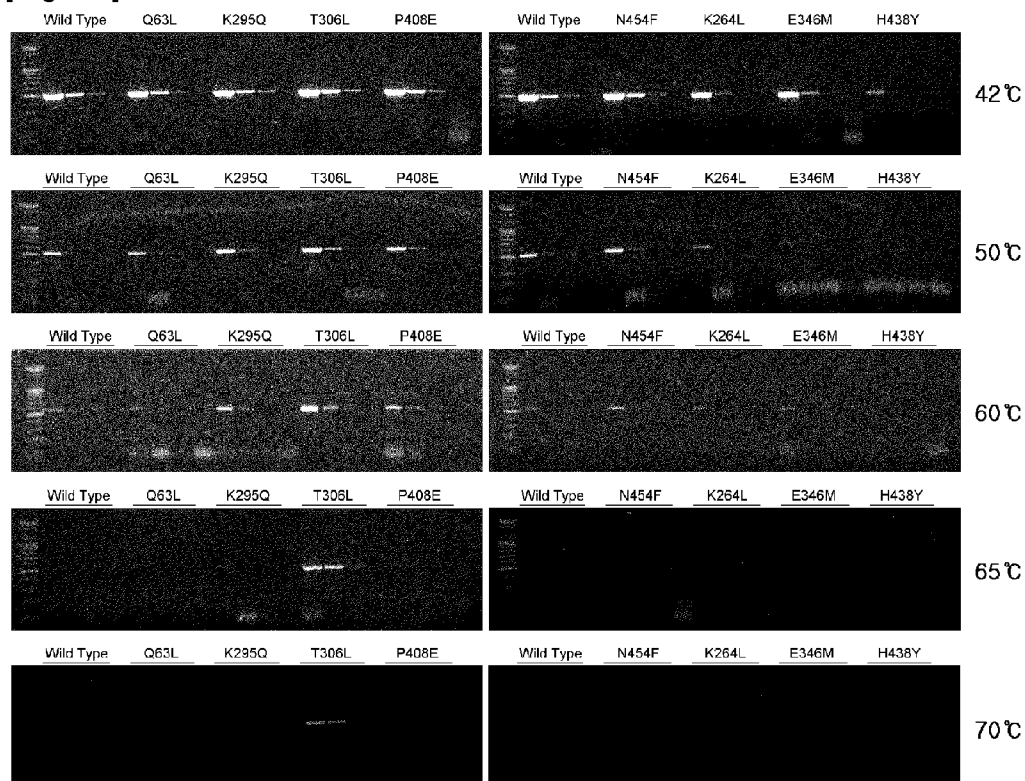
1924 cacagcgccgaggctagaggcaaccggatggctgaccaagcggcccgaaaggcagccato 1983
 || ||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||

1921 cacagcgccgaggctagaggcaaccggatggctgaccaagcggcccgaaaggcagccato 1980

1984 acagagactccagacacotctaccotctctcata 2016
 || ||||| ||||||| ||||||| |||||

1981 acagagactccagacacotctaccotctctcata 2013

[Fig. 10]



[Fig. 11]

