Variants d'une ADN polymérase de la famille polX

Introduction

La présente invention relève du domaine de l'amélioration des enzymes. La présente invention est relative à un variant amélioré d'une ADN polymérase de la famille polX, un acide nucléique codant pour ce variant, la production de ce variant dans une cellule hôte, son utilisation pour la synthèse d'une molécule d'acide nucléique sans brin matrice et un kit pour la synthèse d'une molécule d'acide nucléique sans brin matrice.

La synthèse chimique de fragments d'acide nucléique est une technique largement utilisée dans les laboratoires (Adams et al., 1983, J. Amer. Chem. Soc. 105 :661 ; Froehler et al., 1983, Tetrahedron Lett. 24 :3171). Elle permet d'obtenir de manière rapide des molécules d'acide nucléique comprenant la séquence de nucléotides voulue. A l'inverse des enzymes qui effectuent la synthèse dans le sens 5' vers 3', la synthèse chimique s'effectue dans le sens 3' vers 5'. La synthèse chimique connait cependant certaines limites. En effet, elle nécessite l'utilisation de multiples solvants et réactifs. En outre, elle ne permet d'obtenir que des fragments courts d'acide nucléique, qu'il est ensuite nécessaire d'assembler entre eux pour obtenir les brins d'acides nucléiques finaux souhaités.

Une solution alternative mettant en œuvre des enzymes pouvant réaliser la réaction de couplage entre nucléotides à partir d'un fragment d'acide nucléique initial (amorce) et en l'absence de brin matrice a été développée. Plusieurs enzymes de type polymérase semblent adaptées à ce genre de méthodes de synthèse.

Il existe un très grand nombre d'ADN polymérases capables de catalyser la synthèse d'un brin d'acide nucléique, en présence ou non d'un brin matrice. Ainsi, les ADN polymérases de la famille polX sont impliquées dans un large éventail de processus biologiques, en particulier dans les mécanismes de réparation de l'ADN ou de correction des erreurs apparaissant dans les séquences d'ADN. Ces enzymes sont capables d'introduire des nucléotides dans les brins d'acide nucléique ayant subi des excisions suite à l'identification d'erreurs dans la séquence. Les ADN polymérases de la famille polX regroupent les ADN polymérases β (Pol β), λ (Pol λ), μ (Pol μ), IV de levure (Pol IV) et la desoxyribonucléotidyl-transferase terminale (TdT). La

TdT notamment est très utilisée dans les procédés de synthèse enzymatique de molécules d'acide nucléique.

Cependant, ces ADN polymérases ne permettent le plus souvent que l'incorporation de nucléotides naturels. Dans tous les cas, les ADN polymérases naturelles perdent de leur activité catalytique en présence de nucléotides non naturels, et notamment des nucléotides modifiés en 3'OH, présentant un encombrement stérique plus important que les nucléotides naturels.

Or, l'utilisation de nucléotides modifiés peut s'avérer utile pour certaines applications spécifiques. Il a donc été nécessaire de développer des enzymes capables de catalyser la synthèse d'un brin d'acide nucléique en incorporant de tels nucléotides. Des variants d'ADN polymérase ont ainsi été développés, dans le but de fonctionner avec des nucléotides comportant des modifications structurales importantes.

Les variants actuellement disponibles ne donnent cependant pas entièrement satisfaction, notamment du fait de leur faible activité, compatible seulement avec la synthèse enzymatique à l'échelle du laboratoire. Il existe donc un besoin en ADN polymérases capables de synthétiser, si possible à l'échelle industrielle, un acide nucléique en l'absence de brin matrice et en utilisant des nucléotides modifiés.

Résumé de l'invention

La présente invention lève certains verrous technologiques qui empêchent l'utilisation à l'échelle industrielle d'ADN polymérases pour la synthèse enzymatique d'acides nucléiques.

La présente invention propose ainsi des variants d'ADN polymérases de la famille polX capables de synthétiser un acide nucléique en l'absence de brin matrice et aptes à utiliser des nucléotides modifiés. Les variants développés présentent des capacités d'incorporation de nucléotides modifiés très supérieures à celles des ADN polymérases naturelles dont ils sont dérivés. En particulier, les variants d'ADN polymérases objet de la présente invention sont particulièrement efficaces pour l'incorporation de nucléotides présentant des modifications au niveau du sucre. En effet, les inventeurs ont développé des variants présentant un volume de poche catalytique accru par rapport à celui des ADN polymérases dont ils sont issus, favorisant l'incorporation de nucléotides modifiés plus encombrants que les nucléotides naturels. Plus

particulièrement, les variants d'ADN polymérases de la famille polX objets de la présente invention comprennent au moins une mutation sur un acide aminé intervenant directement au niveau de la cavité catalytique de l'enzyme, ou permettant la déformation des contours de cette cavité afin d'accommoder l'encombrement stérique dû aux modifications présentes au niveau des nucléotides. Par exemple, les mutations introduites permettent l'élargissement de la cavité catalytique de l'enzyme dans laquelle vient se loger l'extrémité 3'-OH des nucléotides modifiés. De manière alternative ou additionnelle, les mutations réalisées permettent l'inflation ou augmentation du volume de la cavité catalytique, l'accroissement de l'accès à la poche catalytique par les nucléotides modifiés en 3'-OH et/ou confèrent la flexibilité nécessaire à la structure de l'enzyme pour lui permettre d'accommoder des modifications stériquement importantes des nucléotides modifiés en 3'-OH. Grâce à de telles mutations, une fois la polymérase fixée au fragment d'acide nucléique à allonger, le nucléotide modifié pénètre au cœur de la poche catalytique dont l'accès est élargi et y adopte une conformation spatiale optimale, une liaison phosphodiester entre l'extrémité 3'-OH du dernier nucléotide du brin d'acide nucléique et l'extrémité 5'-triphosphate du nucléotide modifié se créant.

L'invention a donc pour objet un variant d'une ADN polymérase de la famille polX capable de synthétiser une molécule d'acide nucléique sans brin matrice, ou un variant d'un fragment fonctionnel d'une telle polymérase, ledit variant comprenant au moins une mutation d'un résidu à au moins une position sélectionnée dans le groupe consistant en T331, G332, G333, F334, R336, K338, H342, D343, V344, D345, F346, A397, D399 D434, V436, A446, L447, L448, G449, W450, G452, R454, Q455, F456, E457, R458, R461, N474, E491, D501, Y502, I503, P505, R508, N509 et A510, ou un résidu fonctionnellement équivalent, les positions indiquées étant déterminées par alignement avec SEQ ID N°1.

Dans un mode de réalisation particulier, le variant est capable de synthétiser un brin d'ADN ou un brin d'ARN.

La présente invention concerne notamment un variant d'une ADN polymérase de la famille polX et notamment d'une Pol IV de levure, Pol μ ou TdT sauvage, et comprenant la ou les mutations sélectionnées. Dans un exemple de réalisation particulier, le variant selon la présente invention est un variant de la TdT de séquence SEQ ID N°1 ou une séquence homologue qui

présente au moins 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% d'identité avec la séquence de la SEQ ID No 1, et porte la ou les mutations sélectionnées.

L'invention concerne également un acide nucléique codant un variant d'une ADN polymérase de la famille polX selon la présente invention, une cassette d'expression comprenant un acide nucléique selon la présente invention et un vecteur comprenant un acide nucléique ou une cassette d'expression selon la présente invention. L'acide nucléique codant pour le variant de la présente invention peut être celui de la forme mature ou de la forme précurseur de l'ADN polymérase selon la présente invention.

La présente invention concerne également l'utilisation d'un acide nucléique, d'une cassette d'expression ou d'un vecteur selon la présente invention pour transformer ou transfecter une cellule hôte. Elle concerne en outre une cellule hôte comprenant un acide nucléique, une cassette d'expression ou un vecteur codant une ADN polymérase de la famille polX selon la présente invention. Elle concerne l'utilisation d'un tel acide nucléique, d'une telle cassette d'expression, d'un tel vecteur ou d'une telle cellule hôte pour produire un variant d'une ADN polymérase de la famille polX selon la présente invention.

Elle concerne également un procédé de production d'un variant d'une ADN polymérase de la famille polX selon la présente invention comprenant la transformation ou la transfection d'une cellule hôte par un acide nucléique, une cassette d'expression ou un vecteur selon la présente invention, la mise en culture de la cellule hôte transformée/transfectée dans des conditions de culture permettant l'expression de l'acide nucléique codant ledit variant, et optionnellement la récolte du variant d'une ADN polymérase de la famille polX produit par la cellule hôte.

La cellule hôte peut être procaryote ou eucaryote. Notamment, la cellule hôte peut être un microorganisme, de préférence une bactérie, une levure ou un champignon. Dans un mode de réalisation, la cellule hôte est une bactérie, de préférence *E. coli*. Dans autre un mode de réalisation, la cellule hôte est une levure, de préférence *P. pastoris* ou *K. lactis*. Dans un autre mode de réalisation, la cellule hôte est une cellule de mammifère, de préférence une cellule COS7 ou CHO.

L'invention concerne également l'utilisation d'un variant d'une ADN polymérase de la famille polX selon la présente invention, pour synthétiser une molécule d'acide nucléique sans brin

matrice, à partir de nucléotides modifiés en 3'-OH. Bien entendu, le variant d'ADN polymérase de la famille polX selon la présente invention peut également être utilisé, dans le contexte de l'invention, pour synthétiser une molécule d'acide nucléique sans brin matrice, à partir de nucléotides non modifiés ou d'un mélange de nucléotides modifiés et non modifiés.

L'invention propose également un procédé de synthèse enzymatique d'une molécule d'acide nucléique sans brin matrice, selon lequel on met en contact un brin amorce avec au moins un nucléotide, préférentiellement un nucléotide modifié en 3'-OH, en présence d'un variant d'une ADN polymérase de la famille polX selon l'invention. La mise en œuvre du procédé peut notamment être réalisée en utilisant un variant purifié, un milieu de culture d'une cellule hôte transformée pour exprimer ledit variant, et/ou un extrait cellulaire d'une telle cellule hôte.

L'invention a également pour objet un kit pour la synthèse enzymatique d'une molécule d'acide nucléique sans brin matrice comprenant au moins un variant d'une ADN polymérase de la famille polX selon l'invention, des nucléotides, préférentiellement des nucléotides modifiés en 3'-OH, et optionnellement au moins un brin amorce, ou amorce nucléotidique, et/ou un tampon réactionnel.

Description des figures

Figure 1 : Gel SDS-PAGE de fractions d'un variant de TdT selon un exemple de réalisation de l'invention (M : Marqueur de poids moléculaire ; 1 : Centrifugat avant chargement ; 2 : Centrifugat après chargement ; 3 : Tampon de lavage après chargement ; 4 : Elution fraction 3 mL ; 5 : Elution fraction 30 mL ; 6 : Rassemblement pic d'élution ; 7 : Concentration) ;

Figure 2 : Alignement des séquences d'acides aminés des ADN polymérases Pol μ d'Homo sapiens (UniProtKB Q9NP87), Pol μ de Pan troglodytes (UniProtKB H2QUI0), Pol μ de Mus musculus (UniProtKB Q924W4), TdT de Canis lupus familiaris (UniProtKB F1P657), TdT de Mus musculus (UniProtKB Q3UZ80), TdT de Gallus gallus (UniProtKB P36195) et TdT d'Homo sapiens (UniProtKB P04053) obtenu au moyen du logiciel d'alignement en ligne Mutalin (http://multalin.toulouse.inra.fr/multalin/multalin.html) ;

Figure 3 : Comparaison de l'activité d'une TdT sauvage tronquée de séquence SEQ ID N°3 et de plusieurs variants de cette TdT tronquée comprenant différentes substitutions données par le

tableau 1, en présence d'une amorce préalablement marquée radioactivement en 5' et de nucléotides modifiés 3'-O-amino-2',3'-dideoxyadenosine-5'-triphosphate (gel ONH2) ou nucléotides modifiés 3'-biot-EDA-2',3'-dideoxyadenosine-5'-triphosphate (gel Biot-EDA) ; sur gel SDS-PAGE (No : pas d'enzyme présente ; wt : TdT sauvage tronquée de séquence SEQ ID N°3; DSi : Variants i définis dans le tableau 1) ;

Figure 4 : Etude de l'activité du variant DS124 selon l'invention (voir tableau 1), en présence d'une amorce préalablement marquée radioactivement en 5' et différents nucléotides modifiés avec 3'-O-amino-2',3'-dideoxyadenosine-5'-triphosphate sur gel SDS-PAGE ;

Figure 5 : Etude de l'activité des variants DS22, DS24, DS124, DS125, DS126, DS127 et DS128 en présence d'une amorce préalablement marquée radioactivement en 5' et différents nucléotides modifiés avec 3'-O-amino-2',3'-dideoxyadenosine-5'-triphosphate sur gel SDS-PAGE ;

Figure 6 : Synthèse d'un brin d'ADN de séquence : 5'- GTACGCTAGT-3' (SEQ ID N°15) à la suite de l'amorce de séquence 5'-AAAAAAAAAAGGGG-3' (SEQ ID N°14) au moyen d'un variant de la TDT selon l'invention présentant la combinaison de substitutions R336N - R454A - E457G (DS125).

Description détaillée de l'invention

Définitions

Les acides aminés sont représentés dans ce document par le code à une lettre ou à trois lettres selon la nomenclature suivante : A : Ala (alanine) ; R : Arg (arginine) ; N : Asn (asparagine) ; D : Asp (acide aspartique) ; C : Cys (cystéine) ; Q : Gln (glutamine) ; E : Glu (acide glutamique) ; G : Gly (glycine) ; H : His (histidine) ; I : Ile (isoleucine) ; L : Leu (leucine) ; K : Lys (lysine) ; M : Met (méthionine) ; F : Phe (phénylalanine) ; P : Pro (proline) ; S : Ser (sérine) ; T : Thr (thréonine) ; W : Trp (tryptophane) ; Y : Tyr (tyrosine) ; V : Val (valine).

Par « pourcentage d'identité » entre deux séquences d'acide nucléique ou d'acides aminés au sens de la présente invention, on entend désigner un pourcentage de nucléotides ou de résidus d'acides aminés identiques entre les deux séquences à comparer, obtenu après le meilleur alignement, ce pourcentage étant purement statistique et les différences entre les deux

séquences étant réparties au hasard et sur toute leur longueur. Le meilleur alignement ou alignement optimal est l'alignement pour lequel le pourcentage d'identité entre les deux séquences à comparer, comme calculé ci- après, est le plus élevé. Les comparaisons de séquences entre deux séquences d'acide nucléique ou d'acides aminés sont traditionnellement réalisées en comparant ces séquences après les avoir alignées de manière optimale, ladite comparaison étant réalisée par segment ou par fenêtre de comparaison pour identifier et comparer les régions locales de similarité de séquence. L'alignement optimal des séquences pour la comparaison peut être réalisé, outre manuellement, au moyen de l'algorithme d'homologie locale de Smith et Waterman (1981) (Ad. App. Math. 2 : 482), au moyen de l'algorithme d'homologie locale de Neddleman et Wunsch (1970) (J. Mol. Biol. 48 : 443), au moyen de la méthode de recherche de similarité de Pearson et Lipman (1988) (Proc. Natl. Acad. Sci. USA 85 : 2444), au moyen de logiciels informatiques utilisant ces algorithmes (GAP, BESTFIT, FASTA et TFASTA dans le Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), au moyen du logiciel d'alignement en ligne Mutalin (http://multalin.toulouse.inra.fr/multalin/multalin.html ; 1988, Nucl. Acids Res., 16 (22), 10881-10890). Le pourcentage d'identité entre deux séquences d'acide nucléique ou d'acides aminés est déterminé en comparant ces deux séquences alignées de manière optimale par fenêtre de comparaison dans laquelle la région de la séquence d'acide nucléique ou d'acides aminés à comparer peut comprendre des additions ou des délétions par rapport à la séquence de référence pour un alignement optimal entre ces deux séquences. Le pourcentage d'identité est calculé en déterminant le nombre de positions identiques pour lesquelles le nucléotide ou le résidu d'acide aminé est identique entre les deux séquences, en divisant ce nombre de positions identiques par le nombre total de positions dans la fenêtre de comparaison et en multipliant le résultat obtenu par 100 pour obtenir le pourcentage d'identité entre ces deux séquences.

Les variants objets de la présente invention sont décrits en fonction de leurs mutations sur des résidus spécifiques, dont les positions sont déterminées par alignement avec, ou référence à, la séquence enzymatique SEQ ID N°1. Dans le contexte de l'invention, tout variant portant ces mêmes mutations sur des résidus fonctionnellement équivalents est également visé. Par « résidu fonctionnellement équivalent », on entend un résidu dans une séquence d'une ADN polymérase de la famille polX de séquence homologue à SEQ ID N°1 et présentant un rôle fonctionnel identique. Les résidus fonctionnellement équivalents sont identifiés en ayant recours à des alignements de séquences, par exemple au moyen du logiciel d'alignement en ligne Mutalin

(http://multalin.toulouse.inra.fr/multalin/multalin.html ; 1988, Nucl. Acids Res., 16 (22), 10881-10890). Après alignement, les résidus fonctionnellement équivalents se trouvent à des positions homologues sur les différentes séquences considérées. Les alignements de séquences et l'identification de résidus fonctionnellement équivalents peuvent se faire entre n'importe quelles ADN polymérases de la famille polX et leurs variants naturels, y compris inter-espèces. A titre d'exemple, le résidu L40 de la TdT humaine (UniProtKB P04053) est fonctionnellement équivalent au résidu M40 de la TdT de poulet (UniProtKB P36195) et au résidu V40 de la Polμ de Pan troglodytes (UniProtKB H2QUI0), lesdits résidus étant considérés après alignement des séquences (figure 2).

Par « fragment fonctionnel » est entendu un fragment d'ADN polymérase de la famille polX présentant l'activité ADN polymérase. Le fragment peut comprendre 100, 200, 300, 310, 320, 330, 340, 350, 360, 370, 380 ou plus acides aminés consécutifs d'une ADN polymérase de la famille polX. Préférentiellement le fragment comporte 380 acides aminés consécutifs d'une ADN polymérase de la famille polX consistant en le fragment catalytique de ladite enzyme.

Les termes «mutant» et «variant» peuvent être utilisés de manière interchangeable pour faire référence à des polypeptides dérivés d'ADN polymérases de la famille polX, ou dérivés de fragments fonctionnels de telles ADN polymérases, et notamment d'une TdT telle que la TdT murine selon la séquence SEQ ID N ° 1, et comprenant une altération, à savoir une substitution, une insertion et / ou délétion, à une ou plusieurs positions et ayant une activité d'ADN polymérases. Les variants peuvent être obtenus par diverses techniques bien connues dans l'art. En particulier, des exemples de techniques de modification de la séquence d'ADN codant pour la protéine de type sauvage, comprennent, mais sans s'y limiter, la mutagenèse dirigée, la mutagenèse aléatoire et la construction d'oligonucléotides synthétiques.

Le terme «modification» ou «mutation» tel qu'utilisé ici par rapport à une position ou un résidu d'acide aminé signifie que l'acide aminé dans la position considéré a été modifié par rapport à l'acide aminé de la protéine de type sauvage de référence. De telles modifications comprennent les substitutions, délétions et / ou insertions d'un ou plusieurs acides aminés, et notamment 1 à 5, 1 à 4, 1 à 3, 1 à 2 acides aminés, à une ou plusieurs positions, et notamment à 1, 2, 3, 4, 5 ou plus positions.

Le terme «substitution», en relation avec une position ou un résidu d'acides aminés, signifie que l'acide aminé dans la position particulière a été remplacé par un autre acide aminé que celui dans l'ADN polymérase sauvage ou parente. De préférence, le terme "substitution" désigne le remplacement d'un résidu d'acide aminé par un autre choisi parmi les 20 résidus naturels d'acides aminés standards, les résidus d'acide aminé d'origine naturelle rares (par exemple, l'hydroxyproline, l'hydroxylysine, l'allohydroxylysine, la 6-N-methylysine, la N-éthylglycine, la N-méthylglycine, la N-ethylasparagine, l'allo-isoleucine, la N-méthylisoleucine, la N-méthylvaline, la pyroglutamine, l'acide aminobutyrique, ornithine), et les résidus d'acide aminé non naturels rares, souvent fabriqués synthétiquement (par exemple, la norleucine, la norvaline et cyclohexyl-alanine). De préférence, le terme "substitution" désigne le remplacement d'un résidu d'acide aminé par un autre choisi parmi les 20 résidus d'acides aminés standards d'origine naturelle (G, P, A, V, L, I, M, C, F, Y, W, H, K, R, Q, N, E, D, S et T). La substitution peut être une substitution conservatrice ou non conservatrice. Les substitutions conservatrices se font au sein d'un même groupe d'acides aminés, parmi les aminés basiques (arginine, lysine et histidine), les acides aminés acides (acide glutamique et acide aspartique), les acides aminés polaires (glutamine et asparagine), les acides aminés hydrophobes (méthionine, leucine, isoleucine et valine), les acides aminés aromatiques (phénylalanine, tryptophane et tyrosine) et les petits acides aminés (glycine, alanine, serine et thréonine). Dans le présent document, la terminologie suivante est utilisée pour désigner une substitution: R454F indique que le résidu d'acide aminé en position 454 de la SEQ ID N °1 (arginine, R) est remplacé par une phénylalanine (F). N474S/T/N/Q signifie que le résidu d'acide aminé en position 474 (Asparagine, N) peut être remplacé par une sérine (S), une thréonine (T), une asparagine (N) ou une glutamine (Q). Le signe "+" indique une combinaison de substitutions.

L'invention a trait à des variants d'ADN polymérases de la famille polX (EC 2.7.7.7 ; Advances in Protein Chemistry, Vol. 71, 401-440) capables de synthétiser une molécule d'acide nucléique sans brin matrice, et notamment un brin d'ADN ou d'ARN. Les ADN polymérases de la famille polX comprennent notamment l'ADN polymérase Polβ (UniProt P06746 chez humain ; Q8K409 chez la souris), la Polσ, la Polλ (UniProt Q9UGP5 chez humain ; Q9QUG2 et Q9QXE2 chez la souris) et la Polμ (UniProt Q9NP87 chez humain ; Q9JIW4 chez la souris), la Pol4 (UniProt A7TER5 chez la levure *Vanderwaltozyma polyspora* ; P25615 chez la levure

Saccharomyces cerevisiae), et les Désoxyribonucléotidyl-transférase terminale ou TdT (EC 2.7.7.31 ; UniProt P04053 chez humain ; P09838 chez la souris).

L'invention a plus particulièrement pour objet un variant d'ADN polymérase de la famille polX capable de synthétiser une molécule d'acide nucléique sans brin matrice, ou un variant d'un fragment fonctionnel d'une telle polymérase, ledit variant comprenant au moins une mutation d'un résidu à au moins une position sélectionnée dans le groupe consistant en T331, G332, G333, F334, R336, K338, H342, D343, V344, D345, F346, A397, D399, D434, V436, A446, L447, L448, G449, W450, G452, R454, Q455, F456, E457, R458, R461, N474, E491, D501, Y502, I503, P505, R508, N509 et A510, ou un résidu fonctionnellement équivalent, les positions indiquées étant déterminées par alignement avec, ou référence à, la séquence SEQ ID N°1.

Dans un mode de réalisation, le variant est capable de un brin d'ADN et/ou un brin d'ARN.

Par « comprendre au moins une mutation » ou « comprenant au moins une mutation », on entend que le variant présente une ou plusieurs mutations telles qu'indiquées par rapport à la séquence polypeptidique SEQ ID N°1, mais qu'il peut présenter d'autres modifications, notamment des substitutions, des délétions ou des additions.

D'une manière générale, la mutation d'un ou plusieurs résidus aux positions ci-dessus permet l'élargissement de la poche catalytique (en ciblant par exemple les positions W450, D434, D435, H342, D343, T331, R336, D399, R461, et/ou R508), l'accroissement de l'accessibilité à la poche catalytique (en ciblant par exemple les positions R458, E455, R454, A397, K338, et/ou N509), et/ou confère une plus grande flexibilité à la structure de l'enzyme lui permettant de recevoir des nucléotides modifiés présentant un encombrement stérique important (en ciblant par exemple les positions V436, F346, V344, F334, M330, L448, E491, E457 et/ou N474).

Les variants objets de la présente invention peuvent être des variants de Pol IV, Pol μ, Polβ, Polλ ou de TdT, préférentiellement des variants de Pol IV, Pol μ, ou TdT. De manière alternative, les variants peuvent être des variants d'enzymes chimères, combinant par exemple des portions de séquences différentes d'au moins deux ADN polymérases de la famille polX.

Dans un mode de réalisation particulier, le variant présente au moins 60% d'identité avec la séquence selon SEQ ID N°1, préférentiellement au moins 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% et moins de 100% d'identité avec la séquence selon SEQ ID N°1.

Selon l'invention, la mutation peut consister en une substitution, une délétion ou une addition d'un ou plusieurs résidus d'acide aminé. Dans le cas de délétion, l'annotation X est utilisée, qui indique que le codon codant pour le résidu considéré est remplacé par un codon STOP, tous les acides aminés suivant ainsi que le résidu en question sont donc supprimés. Ainsi, la mutation D501X signifie que l'enzyme se termine au niveau du résidu précédent l'acide aspartique (D) à la position 501, c'est-à-dire la leucine (L) en position 500, tous les résidus au-delà ayant été supprimés. L'annotation Ø désigne par contre une simple délétion ponctuelle du résidu considéré. Ainsi, la mutation D501Ø signifie que l'acide aspartique (D) à la position 501 a été supprimé.

Préférentiellement, le variant selon l'invention comprend au moins une mutation d'un résidu à au moins une position sélectionnée dans le groupe consistant en T331, G332, G333, F334, R336, D343, L447, L448, G449, W450, G452, R454, Q455, E457 et R508, ou un résidu fonctionnellement équivalent, préférentiellement au moins une mutation d'un résidu à au moins une position sélectionnée dans le groupe consistant en R336, R454, E457, ou un résidu fonctionnellement équivalent, les positions indiquées étant déterminées par alignement avec SEQ ID N°1.

Dans un mode de réalisation particulier, ledit variant comprend au moins une mutation d'un résidu à au moins deux positions sélectionnées dans le groupe consistant en R336, R454 et E457, préférentiellement une mutation d'un résidu auxdites trois positions R336, R454 et E457, ou un résidu fonctionnellement équivalent, les positions indiquées étant déterminées par alignement avec SEQ ID N°1.

Dans un mode de réalisation particulier, le variant comprend en outre au moins une mutation d'un résidu dans au moins la région semi-conservée de séquence $X_1X_2GGFR_1R_2GKX_3X_4$ (SEQ ID N°4), dans laquelle

X_1 représente un résidu choisi parmi M, I, V, L
X_2 représente un résidu choisi parmi T, A, M, Q

X_3 représente un résidu choisi parmi M, K, E, Q, L, S, P, R, D
X_4 représente un résidu choisi parmi T, I, M, F, K, V, Y, E, Q, H, S, R, D.

Préférentiellement, ledit variant présente au moins une substitution d'un résidu à au moins une position R_1, R_2 et/ou K de la région semi-conservée de séquence SEQ ID N°4.

Dans un autre mode de réalisation particulier, le variant comprend en outre au moins une mutation d'un résidu dans au moins une région semi-conservée de séquence $X_1X_2LGX_3X_4GSR_1X_5X_6ER_2$ (SEQ ID N°5) dans laquelle

X_1 représente un résidu choisi parmi A, C, G, S
X_2 représente un résidu choisi parmi L, T, R
X_3 représente un résidu choisi parmi W, Y
X_4 représente un résidu choisi parmi T, S, I
X_5 représente un résidu choisi parmi Q, L, H, F, Y, N, E, D ou Ø
X_6 représente un résidu choisi parmi F, Y

Préférentiellement, ledit variant présente au moins une substitution d'un résidu à au moins une position S, R_1 et/ou E de la région semi-conservée de séquence SEQ ID N°5.

Dans un autre mode de réalisation particulier, le variant comprend en outre au moins une mutation d'un résidu dans au moins une région semi-conservée de séquence $LX_1YX_2X_3PX_4X_5RNA$ (SEQ ID N°6) dans laquelle

X_1 représente un résidu choisi parmi D, E, S, P, A, K
X_2 représente un résidu choisi parmi I, L, M, V, A, T
X_3 représente un résidu choisi parmi E, Q, P, Y, L, K, G, N
X_4 représente un résidu choisi parmi W, S, V, E, R, Q, T, C, K, H
X_5 représente un résidu choisi parmi E, Q, D, H, L.

Préférentiellement, ledit variant présente au moins une délétion du résidu à la position X_1 et/ou au moins une substitution au niveau des positions R et/ou N de la région semi-conservée de séquence SEQ ID N°6.

Dans un mode de réalisation particulier, le variant comprend une substitution d'un résidu à au moins une position sélectionnée dans le groupe consistant en R336, K338, H342, A397, S453, R454, E457, N474, D501, Y502, I503, R508 et N509, ou un résidu fonctionnellement équivalent, préférentiellement une substitution d'un résidu à au moins une position sélectionnée dans le groupe consistant en R336, A397, R454, E457, N474, D501, Y502 et I503, ou un résidu fonctionnellement équivalent, plus préférentiellement au moins une substitution d'un résidu à au moins une position sélectionnée dans le groupe consistant en R336, R454, E457, ou un résidu fonctionnellement équivalent, les positions indiquées étant déterminées par alignement avec SEQ ID N°1.

L'invention concerne de préférence un variant d'une ADN polymérase de la famille polX comprenant au moins une substitution parmi le groupe consistant en R336K/H/G/N/D, K338A/C/G/S/T/N, H342A/C/G/S/T/N, A397R/H/K/D/E, S453A/C/G/S/T, R454F/Y/W/A, E457G/N/S/T, N474S/T/N/Q, D501A/G/X, Y502A/G/X, I503A/G/X, R508A/C/G/S/T, N509A/C/G/S/T. Dans un mode de réalisation particulier, le variant comprend une substitution d'un résidu à au moins deux positions sélectionnées dans le groupe consistant en R336, R454, E457, ou un résidu fonctionnellement équivalent, préférentiellement une substitution d'un résidu auxdites trois positions, ou un résidu fonctionnellement équivalent, les positions indiquées étant déterminées par alignement avec SEQ ID N°1. Notamment, Les substitutions sont choisies parmi le groupe consistant en R336K/H/G/N/D, R454F/Y/W/A et E457N/D/G/S/T, préférentiellement parmi le groupe consistant en R336N/G, R454A et E457G/N/S/T.

Dans un mode de réalisation, le variant comprend au moins une substitution selon E457G/N/S/T.

Avantageusement, le variant comprend une combinaison de substitutions sélectionnées dans le groupe mentionné ci-dessus. La combinaison peut consister en 2, 3, 4, 5, 6, 7, 8, 9, 10 ou 11 substitutions sélectionnées dans ce groupe.

L'invention a plus particulièrement pour objet des variants d'une ADN polymérase de la famille polX capables de synthétiser une molécule d'acide nucléique, tel qu'un brin d'ADN ou d'ARN sans brin matrice, ou d'un fragment fonctionnel d'une telle polymérase, lesdits variants

comprenant au moins une combinaison de mutations décrites dans le tableau 1, les positions indiquées étant déterminées par alignement avec SEQ ID N°1.

Dans un mode de réalisation, le variant d'ADN polymérase de la famille polX comprend une combinaison de substitutions parmi R336G - E457N ; R336N - E457N ; R336N - R454A - E457N ; R336N - R454A - E457G ; R336N - E457G ; et R336G - R454A - E457N.

Tableau 1 : Exemples de combinaisons de mutations de variants d'ADN polymérase de la famille polX

	Combinaisons de mutations
DS1	R454F - E457N - A397D
DS2	R454F - E457N
DS3	R454Y - E457N - A397D
DS4	R454Y - E457N
DS5	R454W - E457N - A397D
DS6	R454W - E457N
DS7	R335A - E457N - A397D
DS8	R335A - E457N
DS9	R335G - E457N - A397D
DS10	R335G - E457N
DS11	R335N - E457N - A397D
DS12	R335N - E457N
DS13	R335D - E457N - A397D
DS14	R335D - E457N
DS15	R336K - E457N - A397D
DS16	R336K - E457N
DS17	R336H - E457N - A397D
DS18	R336H - E457N
DS21	R336G - E457N - A397D
DS22	R336G - E457N
DS23	R336N - E457N - A397D
DS24	R336N - E457N
DS25	R336D - E457N - A397D
DS26	R336D - E457N
DS27	R454A - E457N
DS28	R454A - E457A
DS29	R454A - E457G
DS30	R454A - E457D

DS31	E457N
DS32	E457D
DS33	R454A - E457N - A397D
DS34	R454A - E457N - A397K
DS35	R454A - E457N - N474S
DS36	R454A - E457D - A397D
DS37	D501X
DS38	D501X - E457N
DS39	D501X - E457N - A397D
DS40	R454F - E457S - A397D
DS41	R454F - E457S
DS42	R454Y - E457S - A397D
DS43	R454Y - E457S
DS44	R454W - E457S - A397D
DS45	R454W - E457S
DS46	R335A - E457S - A397D
DS47	R335A - E457S
DS48	R335G - E457S - A397D
DS49	R335G - E457S
DS50	R335N - E457S - A397D
DS51	R335N - E457S
DS52	R335D - E457S - A397D
DS53	R335D - E457S
DS54	R336K - E457S - A397D
DS55	R336K - E457S
DS56	R336H - E457S - A397D
DS57	R336H - E457S
DS60	R336G - E457S - A397D
DS61	R336G - E457S
DS62	R336N - E457S - A397D
DS63	R336N - E457S
DS64	R336D - E457S - A397D

DS65	R336D - E457S
DS66	R454A - E457S
DS70	E457S
DS72	R454A - E457S - A397D
DS73	R454A - E457S - A397K
DS74	R454A - E457S - N474S
DS75	D501X - E457S
DS76	D501X - E457S - A397D
DS77	R454F - E457T - A397D
DS78	R454F - E457T
DS79	R454Y - E457T - A397D
DS80	R454Y - E457T
DS81	R454W - E457T - A397D
DS82	R454W - E457T
DS83	R335A - E457T - A397D
DS84	R335A - E457T
DS85	R335G - E457T - A397D
DS86	R335G - E457T
DS87	R335N - E457T - A397D
DS88	R335N - E457T
DS89	R335D - E457T - A397D
DS90	R335D - E457T
DS91	R336K - E457T - A397D
DS92	R336K - E457T
DS93	R336H - E457T - A397D
DS94	R336H - E457T
DS97	R336G - E457T - A397D
DS98	R336G - E457T
DS99	R336N - E457T - A397D
DS100	R336N - E457T
DS101	R336D - E457T - A397D
DS102	R336D - E457T

DS103	R454A - E457T
DS104	E457T
DS105	R454A - E457T - A397D
DS106	R454A - E457T - A397K
DS107	R454A - E457T - N474S
DS108	D501X - E457T
DS109	D501X - E457T - A397D
DS110	D502X
DS111	D502X - E457N
DS112	D502X - E457TN - A397D
DS113	D502X - E457S
DS114	D502X - E457S - A397D
DS115	D502X - E457T
DS116	D502X - E457T - A397D
DS117	D503X
DS118	D503X - E457N
DS119	D503X - E457TN - A397D
DS120	D503X - E457S
DS121	D503X - E457S - A397D
DS122	D503X - E457T
DS123	D503X - E457T - A397D
DS124	R336N – R454A – E457N
DS125	R336N – R454A – E457G
DS126	R336N – E457G
DS127	R336G – R454A – E457N

Dans un mode de réalisation particulier, le variant est une construction chimérique d'ADN polymérases de la famille polX. Par « construction chimérique », on entend une enzyme chimère constituée par l'apport, et notamment la fusion ou la conjugaison, d'une ou plusieurs séquences déterminées d'une enzyme membre de la famille polX en remplacement d'une ou plusieurs séquences homologues dans le variant d'ADN polymérase considéré.

Ainsi, l'invention propose un variant de la TdT de séquence SEQ ID N°1 comprenant, outre une ou plusieurs mutations ponctuelles à l'une et/ou l'autre des positions ci-dessus, une substitution des résidus compris entre les positions C378 à L406, ou les résidus fonctionnellement équivalents, par les résidus H363 à C390 de la polymérase Polμ de séquence SEQ ID N°2, ou les résidus fonctionnellement équivalents.

De manière alternative ou additionnelle, les variants objets de la présente invention peuvent présenter une délétion d'un ou plusieurs résidus successifs d'acides aminés au niveau de la partie N-terminale. Ces délétions peuvent notamment cibler un ou des domaines enzymatiques impliqués dans la liaison avec d'autres protéines et/ou impliqués dans la localisation cellulaire. Par exemple la séquence polypeptidique de la TdT comprend en N-terminal un domaine BRCT d'interaction avec d'autres protéines telles que Ku70/80 et un domaine de localisation au niveau du noyau (NLS).

Dans un mode de réalisation particulier de la présente invention, le variant est un variant de la TdT de séquence SEQ ID N°1 présentant, outre une ou plusieurs des mutations décrites ci-dessus, une suppression des résidus 1-129 correspondant à l'extrémité N-terminale de la TdT sauvage.

Dans certains cas particuliers, les stratégies de mutagénèse peuvent être guidées par des informations connues telles que les séquences de variants naturels, la comparaison de séquence avec des protéines liées, des propriétés physiques, l'étude d'une structure tridimensionnelle ou de simulations informatique impliquant plusieurs entités.

La présente invention concerne un acide nucléique codant un variant d'une ADN polymérase de la famille polX capable de synthétiser une molécule d'acide nucléique sans brin matrice selon la présente invention. La présente invention concerne également une cassette d'expression d'un acide nucléique selon la présente invention. Elle concerne en outre un vecteur comprenant un acide nucléique ou une cassette d'expression selon la présente invention. Le vecteur peut être sélectionné parmi un plasmide et un vecteur viral.

L'acide nucléique codant pour le variant d'ADN polymérase peut être de l'ADN (ADNc ou ADNg), de l'ARN, un mélange des deux. Il peut être sous forme simple chaîne ou en duplexe ou un mélange des deux. Il peut comprendre des nucléotides modifiés, comprenant par exemple

une liaison modifiée, une base purique ou pyrimidique modifiée, ou un sucre modifié. Il peut être préparé par toutes méthodes connues de l'homme du métier, dont la synthèse chimique, la recombinaison, la mutagenèse, etc...

La cassette d'expression comprend tous les éléments nécessaires à l'expression du variant d'une ADN polymérase de la famille polX capable de synthétiser une molécule d'acide nucléique sans brin matrice selon la présente invention, notamment les éléments nécessaires à la transcription et à la traduction dans la cellule hôte. La cellule hôte peut être procaryote ou eucaryote. En particulier, la cassette d'expression comprend un promoteur et un terminateur, facultativement un amplificateur. Le promoteur peut être procaryote ou eucaryote. Des exemples de promoteurs procaryotes préférés sont les suivants : LacI, LacZ, pLacT, ptac, pARA, pBAD, les promoteurs d'ARN polymérase de bactériophage T3 or T7, le promoteur de la polyhèdrine, le promoteur PR ou PL du phage lambda. Des exemples de promoteurs eucaryotes préférés sont les suivants : promoteur précoce du CMV, promoteur de la thymidine kinase de HSV, promoteur précoce ou tardif de SV40, le promoteur de la métallothionéine-L de souris, et les régions LTR de certains rétrovirus. De manière générale, pour le choix d'un promoteur adapté, l'homme du métier pourra avantageusement se référer à l'ouvrage de Sambrook et al. (1989) ou encore aux techniques décrites par Fuller et al. (1996; Immunology in Current Protocols in Molecular Biology).

La présente invention concerne un vecteur portant un acide nucléique ou une cassette d'expression codant pour un variant d'une ADN polymérase de la famille polX capable de synthétiser une molécule d'acide nucléique sans brin matrice selon la présente invention. Le vecteur est de préférence un vecteur d'expression, c'est-à-dire qu'il comprend les éléments nécessaires à l'expression du variant dans la cellule hôte. La cellule hôte peut être un procaryote, par exemple *E. coli*, ou un eucaryote. L'eucaryote peut être un eucaryote inférieur comme une levure (par exemple, *P. Pastoris* ou *K. lactis*) ou un champignon (par exemple du genre *Aspergillus*) ou un eucaryote supérieur comme une cellule d'insecte (Sf9 ou Sf21 par exemple), de mammifère ou de plante. La cellule peut être une cellule mammifère, par exemple COS (lignée cellulaire de singe vert) (par exemple, COS 1 (ATCC CRL-1650), COS 7 (ATCC CRL-1651), CHO (US 4,889,803 ; US 5,047,335, CHO-K1 (ATCC CCL-61)), des cellules de souris et des cellules humaines. Dans un mode de réalisation particulier, la cellule est non-humaine et non-embryonnaire. Le vecteur peut être un plasmide, un phage, un phagemide, un

cosmide, un virus, un YAC, un BAC, un plasmide pTi d'*Agrobacterium*, etc... Le vecteur peut comprendre de préférence un ou plusieurs éléments sélectionnés parmi une origine de réplication, un site de clonage multiple et un gène de sélection. Dans un mode de réalisation préféré, le vecteur est un plasmide. Des exemples non-exhaustifs de vecteurs procaryotes sont les suivants : pQE70, pQE60, pQE-9 (Qiagen), pbs, pD10, phagescript, psiX174, pbluescript SK, pbsks, pNH8A, pNH16A, pNH18A, pNH46A (Stratagene); ptrc99a, pKK223-3, pKK233-3, pDR540, pBR322, et pRIT5 (Pharmacia), pET (Novagen). Des exemples non-exhaustifs de vecteurs eucaryotes sont les suivants : pWLNEO, pSV2CAT, pPICZ, pcDNA3.1 (+) Hyg (Invitrogen), pOG44, pXT1, pSG (Stratagene); pSVK3, pBPV, pCI-neo (Stratagene), pMSG, pSVL (Pharmacia); et pQE-30 (QLAexpress). Les vecteurs viraux peuvent être de manière non-exhaustive des adénovirus, des AAV, des HSV, des lentivirus, etc... De préférence, le vecteur d'expression est un plasmide ou un vecteur viral.

La séquence codant le variant selon la présente invention peut comprendre ou ne pas comprendre de peptide signal. Dans le cas où elle n'en comprend pas, une méthionine peut être éventuellement ajoutée à l'extrémité N-terminale. Dans une autre alternative, un peptide signal hétérologue peut être introduit. Ce peptide signal hétérologue peut être dérivé d'un procaryote tel que *E. coli* ou d'un eucaryote, notamment une cellule mammifère, d'insecte ou d'une levure.

La présente invention concerne l'utilisation d'un polynucléotide, d'une cassette d'expression ou d'un vecteur selon la présente invention pour transformer ou transfecter une cellule. La présente invention concerne une cellule hôte comprenant un acide nucléique, une cassette d'expression ou un vecteur codant un variant d'une ADN polymérase de la famille polX capable de synthétiser une molécule d'acide nucléique sans brin matrice et son utilisation pour produire un variant d'une ADN polymérase de la famille polX capable de synthétiser une molécule d'acide nucléique sans brin matrice recombinant selon la présente invention. Le terme « cellule hôte » englobe les cellules filles résultant de la culture ou de la croissance de cette cellule. Dans un mode de réalisation particulier, la cellule est non-humaine et non-embryonnaire. Elle concerne également une méthode de production d'un variant d'une ADN polymérase de la famille polX capable de synthétiser une molécule d'acide nucléique sans brin matrice recombinant selon la présente invention comprenant la transformation ou transfection d'une cellule par un polynucléotide, une cassette d'expression ou un vecteur selon la présente invention ; la mise en culture de la cellule transfectée/transformée ; et la récolte du variant d'une

ADN polymérase de la famille polX capable de synthétiser une molécule d'acide nucléique sans brin matrice produit par la cellule. Dans un mode de réalisation alternatif, la méthode de production d'un variant d'une ADN polymérase de la famille polX capable de synthétiser une molécule d'acide nucléique sans brin matrice recombinant selon la présente invention comprenant la fourniture d'une cellule comprenant un polynucléotide, une cassette d'expression ou un vecteur selon la présente invention ; la mise en culture de la cellule transfectée/transformée ; et la récolte du variant d'une ADN polymérase de la famille polX capable de synthétiser une molécule d'acide nucléique sans brin matrice produit par la cellule. En particulier, la cellule peut être transformée/transfectée de manière transitoire ou stable par l'acide nucléique codant le variant. Cet acide nucléique peut être contenu dans la cellule sous forme d'épisome ou sous forme chromosomique. Les méthodes de production de protéines recombinantes sont bien connues par l'homme du métier. Par exemple, on peut citer les modes spécifiques décrits dans US 5,004,689, EP 446 582, Wang et al. (Sci. Sin. B 24:1076-1084, 1994 et Nature 295, page 503) pour une production dans *E. coli*, et JAMES et al. (Protein Science (1996), 5:331-340) pour une production en cellules mammifères.

Les variants d'ADN polymérase selon la présente invention sont particuliers intéressants pour la synthèse d'acides nucléiques sans brin matrice. Plus particulièrement, les variants selon l'invention présentent une poche catalytique accrue particulièrement adaptée pour la synthèse d'acide nucléique au moyen de nucléotides modifiés présentant un encombrement plus important que les nucléotides naturels. Les variants selon l'invention peuvent notamment permettre d'incorporer dans un brin d'acide nucléique des nucléotides modifiés tels que ceux décrits dans la demande WO2016/034807.

La cinétique d'incorporation des variants d'ADN polymérase, et notamment des variants de la TdT selon l'invention, présentant les mutations ou les combinaisons de mutations spécifiques décrites ci-dessus, est largement améliorée comparativement à la cinétique d'incorporation d'un ADN polymérase sauvage. Ces variants peuvent avantageusement être utilisés dans le cadre d'une méthode de synthèse d'ADN enzymatique à hautes performances.

L'invention a donc également pour objet une utilisation d'un variant d'une ADN polymérase de la famille polX selon la présente invention, pour synthétiser une molécule d'acide nucléique

sans brin matrice, à partir de nucléotides modifiés en 3'-OH, et notamment ceux décrits dans la demande WO2016034807.

L'invention a également pour objet un procédé de synthèse enzymatique d'une molécule d'acide nucléique sans brin matrice, selon lequel on met en contact un brin amorce avec au moins un nucléotide, préférentiellement un nucléotide modifié en 3'-OH, en présence d'un variant d'une ADN polymérase de la famille polX selon l'invention.

De manière avantageuse, les variants selon l'invention peuvent être utilisés pour mettre en œuvre le procédé de synthèse décrit dans la demande WO2015/159023.

L'invention a également pour objet un kit pour la synthèse enzymatique d'une molécule d'acide nucléique sans brin matrice comprenant au moins un variant d'une ADN polymérase de la famille polX selon l'invention, des nucléotides, préférentiellement des nucléotides modifiés en 3'-OH, et optionnellement au moins une amorce nucléotidique.

Toutes les références citées dans cette description sont incorporées par référence dans la présente demande. D'autres caractéristiques et avantages de l'invention apparaîtront mieux à la lecture des exemples suivants donnés bien entendu à titre illustratif et non limitatif.

Exemples

Exemple 1 - Génération, production et purification de variants d'ADN polymérase de la famille polX selon l'invention

Génération des souches productrices

Le gène tronqué de la TdT de souris a été généré à partir du plasmide pET28b dont la construction est décrite dans [Boulé et al., 1998, *Mol. Biotechnol.*, **10**, 199-208]. La séquence correspondante SEQ ID N°3 (correspondant à SEQ ID N°1 tronquée des 120 premiers acides aminés) a été amplifiée en utilisant les amorces suivantes:

- T7-pro: TAATACGACTCACTATAGGG (SEQ ID N°7)
- T7-ter: GCTAGTTATTGCTCAGCGG (SEQ ID N°8)

selon des techniques d'amplification PCR et de biologie moléculaire usuelles. Elle a été clonée dans un plasmide pET32 pour donner le vecteur pET32- SEQ ID N°3.

Le plasmide pET32- SEQ ID N°3 a d'abord été séquencé puis transformé dans les souches *E. coli* commerciales BL21 (DE3) (Novagen). Les colonies capables de pousser sur boites kanamycine/chloramphénicol ont été isolées et notées Ec- SEQ ID N°3

Génération des variants

Le vecteur pET32- SEQ ID N°3 a été utilisé comme vecteur de départ. Des amorces comportant la mutation ponctuelle (ou dans certains cas les mutations ponctuelles si celles-ci sont suffisamment proches) ont été générés à partir de l'outil en ligne d'Agilent : (http://www.genomics.agilent.com/primerDesignProgram.jsp)

Le kit QuickChange II (Agilent) a été utilisé pour générer les plasmides des variants comportant la/les mutation(s) désirée(s). Le protocole de mutagénèse donner par le fabriquant a été respecté scrupuleusement afin d'obtenir un plasmide pET32-DSi (i est le numéro du variant considéré donné par le tableau 1). A la fin de la procédure, le plasmide pET32-DSx a été d'abord séquencé puis transformé dans les souches *E. coli* commerciales BL21 (DE3) (Novagen). Les colonies capables de pousser sur boites kanamycine/chloramphénicol ont été isolées et notées Ec-DSx.

Production

Les cellules Ec- SEQ ID N°3 et Ec-DSx ont été mises en préculture dans un erlen de 250 mL contenant 50 mL de milieu LB auquel ont été ajoutées des quantités appropriées de kanamycine et chloramphénicol. La culture a été incubée à 37°C sous agitation durant toute une nuit. La préculture a ensuite été utilisée pour ensemencer un erlen de 5 L contenant 2 L de milieu LB additionné des quantités appropriées de kanamycine et chloramphénicol. La densité optique (DO) de départ était de 0,01. La culture a été incubée à 37°C sous agitation. La DO a été régulièrement mesurée jusqu'à atteindre une valeur comprise entre 0,6 et 0,9. Une fois cette valeur atteinte, 1 mL d'isopropyl β-D-1-thiogalactopyranoside 1 M a été ajouté au milieu de culture. La culture a été de nouveau incubée à 37°C jusqu'au lendemain. Les cellules ont alors été récoltées par centrifugation sans excéder les 5 000 rpm. Les différents culots obtenus ont été rassemblés dans un culot unique au cours de lavage avec du tampon de lyse (20 mM Tris-HCl, pH 8.3, 0.5 M NaCI). Le culot de cellules a été congelé à -20°C. Il peut être conservé ainsi durant plusieurs mois.

Extraction

Le culot de cellule congelé lors de l'étape précédente a été décongelé dans un bain-marie porté entre 25 et 37°C. Une fois entièrement décongelé il a été re-suspendu dans environ 100 mL de tampon de lyse. Une attention particulière a été portée à la re-suspension qui doit aboutir à une solution très homogène et notamment à l'absence totale d'agrégats. Ainsi re-suspendue, les cellules ont été lysées à l'aide d'une presse de French sous une pression de 14 000 psi. Le lysat recueillit a été centrifugé à haute vitesse, 10 000 g durant 1h à 1h30. Le centrifugat a été filtré sur filtre 0,2 µM et recueilli dans un tube de volume suffisant.

Purification

La TdT a été purifiée sur colonne d'affinité. Des colonnes 5 mL His-Trap Crude (GE Life Sciences) ont été utilisées avec des pompes péristaltiques (Peristaltic Pump - MINIPULS® Evolution, Gilson). Dans un premier temps la colonne a été équilibrée à l'aide de 2 à 3 CV (volume de colonne) de tampon de lyse. Le centrifugat de l'étape précédente a alors été chargé sur la colonne à une vitesse comprise entre 0,5 et 5 mL/min environ. Une fois la totalité de centrifugat chargée, la colonne a été lavée à l'aide de 3 CV de tampon de lyse puis 3 CV de tampon de lavage (20 mM Tris-HCl, pH 8.3, 0.5 M NaCI, 60 mM imidazol). A la fin de cette étape le tampon d'élution (20 mM Tris-HCl, pH 8.3, 0.5 M NaCI, 1 M imidazol) a été injecté dans la colonne à environ 0,5 à 1 ml/min pour un volume total de 3 CV. Durant toute la phase d'élution la sortie de colonne a été collectée par fractions de 1 mL. Ces fractions ont été analysées par SDS-PAGE afin de déterminer quelles sont les fractions contenant le pic d'élution. Une fois déterminées, celles-ci ont été rassemblées dans une fraction unique et dialysée contre du tampon de dialyse (20 mM Tris-HCl, pH 6.8, 200 mM NaCl, 50 mM MgOAc, 100 mM $[NH_4]_2SO_4$. La TdT a alors été concentrée (Filtres à centrifuger Amicon Ultra-30, Merk Millipore) jusqu'à une concentration finale de 5 à 15 mg/mL. La TdT concentrée a été congelée à -20°C pour le stockage à long terme après addition de 50% glycérol. Durant l'intégralité de la phase de purification, des aliquotes des différents échantillons ont été prélevés (environ 5 µL) pour une analyse en gel SDS-PAGE dont les résultats sont présentés par la figure 1.

Exemple 2 – Alignement de séquences entre différentes polymérases de la famille des Pol X susceptibles d'être utilisées pour la création de variants selon l'invention

Différentes ADN polymérases de la famille des Pol X ont été alignées en utilisant le logiciel d'alignement en ligne Mutalin (http://multalin.toulouse.inra.fr/multalin/multalin.html, accédé le 04/04/2016).

Tableau 2 : Séquences alignées

Identifiant	ADN polymérase	Espèce	Longueur
Q9NP87	Pol μ (SEQ ID N°2)	*Homo sapiens*	494
H2QUI0	Pol μ (SEQ ID N°9)	*Pan troglodytes*	494
Q924W4	Pol μ (SEQ ID N°10)	*Mus musculus*	496
F1P657	TdT (SEQ ID N°11)	*Canis lupus familiaris*	509
Q3UZ80	TdT (SEQ ID N°1)	*Mus musculus*	510
P36195	TdT (SEQ ID N°12)	*Gallus gallus*	506
P04053	TdT (SEQ ID N°13)	*Homo sapiens*	509

Les alignements obtenus sont présentés à la figure 2.

Exemple 3 – Etude d'activité des variants en présence de substrats non naturels

L'activité de différents variants selon l'invention a été déterminée par l'essai suivant. Les résultats ont été comparés à ceux obtenus avec l'enzyme naturelle dont chacun des variants est issu.

Test d'activité

Tableau 3 : Mélange réactionnel

Réactif	Concentration	volume
H_2O	-	15 µL
Amorce	500 nM	2,5 µL
Tampon	10x	2,5 µL
Nucléotide modifié	250 µM	2,5 µL
Enzyme	20 µM	2,5 µL

L'amorce utilisée, de séquence 5'-AAAAAAAAAAGGGG-3' (SEQ ID N°14), a été préalablement marquée radioactivement en 5' au moyen d'un protocole standard de marquage impliquant l'enzyme PNK (NEB) et l'usage d'ATP radioactif (PerkinElmer).

Le tampon 10x constitué de 250 mM Tris-HCl pH 7,2, 80 mM $MgCl_2$, 3,3 mM $ZnSO_4$ a été utilisé.

Les nucléotides modifiés utilisés sont des 3'-O-amino-2',3'-dideoxynucléotides-5'-triphosphate (ONH2, Firebird Biosciences) ou des 3'-biot-EDA-2',3'-dideoxynucléotides-5'-triphosphate (Biot-EDA, Jena Biosciences), tels que 3'-O-amino-2',3'-dideoxyadenosine-5'-triphosphate ou 3'-biot-EDA-2',3'-dideoxyadenosine-5'-triphosphate par exemple. Le groupement 3'-O-amino est un groupement plus volumineux fixé sur l'extrémité 3'-OH. Le groupement 3'-biot-EDA est un groupement extrêmement volumineux et non flexible fixé sur l'extrémité 3'-OH.

Les performances d'incorporation d'un nucléotide modifié données par les variants listés dans le table 1, par rapport à la TdT naturelle (SEQ ID N°3) ont été évaluées en réalisant des tests d'activité simultanés, pour lesquels seuls l'enzyme varie.

Les réactifs ont été ajoutés dans l'ordre donné dans le tableau 3 ci-dessus puis incubés à 37°C pendant 90min. La réaction a alors été stoppée par l'ajout de bleu formamide (formamide 100%, 1 à 5 mg de bleu de bromophénol; Simga)

Gel et radiographie

Un gel dénaturant de polyacrylamide 16% (Biorad) a été utilisé pour l'analyse du test d'activité précédent. Le gel a été préalablement coulé et laissé à polymériser. Il a ensuite été monté sur une cuve à électrophorèse de dimensions appropriées remplie de tampon TBE (Sigma). Les différents échantillons ont été directement chargés sur le gel sans pré-traitement.

Le gel a alors été soumis à une différence de potentiel de 500 à 2 000 V durant 3 à 6 heures. Une fois la migration satisfaisante, le gel a été désincarcé puis transféré dans une cassette d'incubation. Un écran phosphore (Amersham) a été utilisé pendant 10 à 60 min pour la révélation effectuée à l'aide d'un instrument Typhoon (GE Life Sciences) préalablement paramétré avec un mode de détection adéquat.

Résultats

Les résultats comparatifs des deux enzymes utilisées sont présentés par la figure 3.

Plus précisément, sur le premier gel (Incorporation ONH2), la TdT naturelle (colonne wt) est incapable d'incorporer les nucléotides modifiés 3'-O-amino-2',3'-dideoxyadenosine-5'-triphosphate comme le montre la comparaison avec le contrôle négatif (colonne No).

Parmi les différents variants, 3 groupes différents peuvent être observés :

Un premier groupe de variants (colonnes DS7 à DS34) est capables d'une incorporation à 50% environ.

Un second groupe de variants (colonnes DS46 à DS73) est capables d'une incorporation à plus de 95%, parfois plus de 98%.

Un troisième groupe de variants (colonnes DS83 à DS106) est capable d'une incorporation entre 60 et 80%.

Sur le second gel (Incorporation Biot-EDA), la TdT naturelle (colonne wt) est également incapable d'incorporer les nucléotides modifiés 3'-biot-EDA-2',3'-dideoxyadenosine-5'-triphosphate comme le montre la comparaison avec le contrôle négatif (colonne No).

Parmi les différents variants, 3 groupes différents peuvent être observés :

Un premier groupe de variants (colonnes DS7 à DS34) est capable d'une incorporation entre 5 et 10% environ.

Un second groupe de variants (colonnes DS46 à DS73) est capable d'une incorporation à plus de 30%, parfois plus de 40%.

Un troisième groupe de variants (colonnes DS83 à DS106) est capable d'une incorporation entre 10 et 25%.

Ces résultats confirment que les variants de TdT selon l'invention sont tous capables d'utiliser comme substrat des nucléotides modifiés, notamment en 3'-OH, contrairement à l'enzyme sauvage. De manière particulièrement avantageuse, certains variants présentent des taux d'incorporation très élevés et cela même en présence de nucléotides portant des modifications tendant à augmenter de manière très importante l'encombrement stérique dudit nucléotide.

Exemple 4 – Etude de la cinétique des variants selon l'invention

Un mutant présentant la combinaison de substitutions R336N - R454A - E457N (DS124) a été généré et produit suivants l'exemple 1 précédent.

Test d'activité

Dans le test d'activité les enzymes sont mises en présence de nucléotides modifiés ONH2 et incubé à 37°C suivant différent temps. Les réactions sont arrêtées afin d'observer la cinétique d'incorporation de DS124 et la comparer à la cinétique de l'enzyme naturelle WT (SEQ ID N°3).

Tableau 4 : Mélange réactionnel

Réactif	*Concentration*	*Volume*
H_2O	-	15 µL
Tampon	10x	2,5 µL
Nucléotides	2.5 µM	2,5 µL
Enzyme	80 µM	2,5 µL
Amorce	1 µM	2,5 µL

L'amorce et le tampon utilisés sont conformes à l'exemple 3.

Les nucléotides modifiés utilisés sont des 3'-O-amino-2',3'-dideoxynucléotides-5'-triphosphate (ONH2, Firebird Biosciences) : 3'-O-amino-2',3'-dideoxyguanosine-5'-triphosphate, 3'-O-amino-2',3'-dideoxycytidine-5'-triphosphate et 3'-O-amino-2',3'-dideoxythymidine-5'-triphosphate. Le groupement 3'-O-amino est un groupement plus volumineux fixé sur l'extrémité 3'-OH.

Les performances d'incorporation du mélange de nucléotides par l'enzyme DS124 ont été évaluées en réalisant des tests d'activité pour lesquels des premix contenant tous les réactifs (ajouté dans l'ordre du tableau 4) mis à part l'amorce, sont préparés. Ils sont distribués dans différent puits réactionnels. Au temps initial t=0 l'amorce est ajoutée dans tous les puits de façon simultanée. Aux différents temps t=2min, t=5min, t=10min, t=15min, t=30min et t=90min la réaction est stoppée par l'ajout de bleu formamide (formamide 100%, 1 à 5 mg de bleu de bromophénol; Simga).

Gel et radiographie

L'analyse du test d'activité est réalisée par migration des différents échantillons dans un gel de polyacrylamide suivant le protocole décrit dans l'exemple 3.

Résultats

Les résultats comparatifs des deux enzymes (DS124 et WT) sont présentés par la figure 4.

Plus précisément, sur ce gel le contrôle négatif (colonne No) donne la taille attendue de l'amorce utilisée lorsqu'elle n'a pas été allongée, c'est-à-dire lorsqu'il n'y a pas eu d'incorporation de nucléotides. La TdT naturelle (colonne WT) n'est pas capable d'incorporer les nucléotides modifiés (ici ONH2-dGTP) : on observe une bande au même niveau que celle de la colonne No.

Pour tous les nucléotides testés et pour tous les temps allant de 90min (ici utilisé comme un contrôle positif) à 2min, soit une réduction du temps d'incubation d'un facteur 45, le variant DS124 est capable d'incorporer les nucléotides modifiés avec une efficacité apparente de 100%.

Ces résultats confirment que les variants de la TdT selon l'invention sont capables de performances d'incorporation bien supérieures à celles de la TdT naturelle, tant en termes d'efficacité d'incorporation que de rapidité d'incorporation. La cinétique des variants de la TdT selon l'invention est largement améliorée par les mutations ou combinaisons de mutations spécifiques décrites par la présente invention.

Exemple 5 – Etude de la spécificité des variants selon l'invention

Les mutants présentant une combinaison de substitution selon le tableau 5 ci-dessous ont été générés et produits suivant l'exemple 1.

Tableau 5 : Liste des variants enzymatiques utilisés

#	Combinaisons de mutations
DS124	R336N - R454A - E457N
DS24	R336N - E457N
DS125	R336N - R454A - E457G
DS126	R336N - E457G
DS127	R336G - R454A - E457N
DS22	R336G - E457N
DS128	R336A - R454A - E457G
WT	SEQ ID N°3

Test d'activité

Dans ce test d'activité les différents variants ont été mis en présence d'un mélange de nucléotides naturels et de nucléotides modifiés hautement concentré. La concentration de l'enzyme est également augmentée afin de raccourcir le temps d'incubation et de parvenir à une addition quantitative (cf. exemple 4).

L'activité de différents variants générés a été déterminée par l'essai suivant :

Chaque variant est testé selon deux conditions : (1) en l'absence de nucléotides (remplacés par H_2O) ou (2) en présence du mélange de nucléotides. Les résultats des différents variants sont comparés entre eux. Un échantillon de contrôle a été ajouté, il ne contenait ni nucléotide ni enzyme (remplacés par H_2O).

Tableau 6 : Mélange réactionnel

Réactif	*Concentration*	*Volume*
H_2O	-	15 µL
Amorce	1 µM	2,5 µL
Tampon	10x	2,5 µL
Mélange Nucléotides (10 :90)	2.5 µM	2,5 µL
Enzyme	80 µM	2,5 µL

L'amorce et le tampon utilisés sont identiques à l'exemple 3.

Lorsqu'il est présent, le mélange de nucléotides est constitué de nucléotides naturels 2′-deoxynucleotide 5′-triphosphate (Nuc, Sigma-Aldrich) tels que des 2′-deoxyguanosine 5′-triphosphate (dGTP) et de nucléotides modifiés 3'-O-amino-2',3'-dideoxynucléotides-5'-triphosphate (ONH2, Firebird Biosciences) tels que 3'-O-amino-2',3'-dideoxyguanosine-5'-triphosphate par exemple. Le groupement 3'-O-amino est un groupement plus volumineux fixé sur l'extrémité 3'-OH. Le mélange est constitué à 90% de nucléotides modifiés ONH2-dGTP et à 10% de nucléotides naturels dGTP.

Les performances d'incorporation du mélange de nucléotides par les variants listés dans le tableau 5 entre eux ont été évaluées en réalisant des tests d'activité simultanés, pour lesquels seule l'enzyme varie.

Les réactifs ont été ajoutés dans l'ordre donné dans le tableau 6 ci-dessus, puis incubés à 37°C pendant 15min. La réaction a alors été stoppée par l'ajout de bleu formamide (formamide 100%, 1 à 5 mg de bleu de bromophénol; Simga).

Gel et radiographie

L'analyse du test d'activité a été réalisée par migration des différents échantillons dans un gel de polyacrylamide suivant le protocole décris dans l'exemple 3.

Résultats

Les résultats comparatifs des enzymes utilisées sont présentés à la figure 5.

Plus précisément, sur ce gel le contrôle négatif (colonne No) donne la taille attendu de l'amorce utilisée lorsqu'elle n'a pas été allongée, c'est-à-dire lorsqu'il n'y a pas eu d'incorporation de nucléotides. Les échantillons suivants vont par paire, chaque paire correspondant à un même variant enzymatique testé dans les deux conditions : en l'absence et en présence de nucléotides (sous forme de mélange lorsqu'ils sont présents).

Parmi les différents variants testés, 3 groupes différents peuvent être observés :

Le premier groupe est le variant DS128, constituant un contrôle négatif. Ce variant présente des taux extrêmement faibles d'incorporation des nucléotides : entre 5% et 10% d'incorporation sont observés lorsque le mélange de nucléotides est présent ; ce qui correspond à la proportion de nucléotides naturels présents dans le mélange.

Le second groupe est constitué par les variants DS127 et DS22. Ces variants présentent des taux d'incorporation élevés des nucléotides: entre 50% et 60% d'incorporation sont observés lorsque le mélange de nucléotides est présent. Toujours dans ce cas, une bande de sur-addition correspondant à l'incorporation successive de deux nucléotides est observée pour ces deux variants. L'intensité de cette bande correspond à la proportion de nucléotides naturels présents dans le mélange de nucléotides.

Le dernier groupe est constitué par les variants DS124, DS24, DS125 et DS126. Ces variants présentent des taux d'incorporation extrêmement élevés des nucléotides: entre 80% et 100%, pour DS124 et DS125, lorsque le mélange de nucléotides est présent. Dans ce cas, aucune bande de sur-addition n'est présente. Dans le cas des variants DS24 et DS126 la proportion de non incorporation est semblable à la proportion de nucléotides naturels présents dans le mélange.

Ces résultats confirment que les variants de la TdT selon l'invention sont capables d'utiliser préférentiellement les nucléotides modifiés parmi un mélange de nucléotides modifiés et de nucléotides naturels. De manière particulièrement avantageuse ces variants présentent des taux d'incorporation extrêmement élevés des nucléotides modifiés et sont capables de discriminer

les nucléotides naturels de façon à ne pas les incorporer et ainsi améliorer grandement la qualité de l'ADN synthétiser en évitant les sur-additions.

Exemple 6 – Exemple de synthèse d'un brin d'ADN sans brin matrice

Un variant de TdT présentant la combinaison de substitutions R336N - R454A - E457G (DS125) a été généré et produit selon l'exemple 1.

Le variant DS125 est utilisé afin de synthétiser la séquence : 5'- GTACGCTAGT-3' (SEQ ID N°15) à la suite de l'amorce de séquence 5'-AAAAAAAAAAGGGG-3' (SEQ ID N°14). L'amorce a été préalablement marquée radioactivement en 5' au moyen d'un protocole standard de marquage impliquant l'enzyme PNK (NEB) et l'usage d'ATP radioactif (PerkinElmer).

L'amorce est fixée à un support solide par interaction avec un fragment de capture de séquence : 5'- CCTTTTTTTTTT -3' complémentaire (SEQ ID N°16). Le fragment de capture possède sur son extrémité 3' un groupement lui permettant de réagir de façon covalente avec un groupe réactionnel fixé sur une surface. Par exemple ce groupement peut être NH2, le groupe réactionnel N-hydroxysuccinimide et la surface une bille magnétique (Dynabeads, Thermofisher). L'interaction de l'amorce avec le fragment de capture est réalisée dans des conditions standard d'hybridation de fragment d'ADN.

Les nucléotides modifiés utilisés sont des 3'-O-amino-2',3'-dideoxynucléotides-5'-triphosphate (ONH2, Firebird Biosciences) tels que 3'-O-amino-2',3'-dideoxyguanosine-5'-triphosphate, 3'-O-amino-2',3'-dideoxycytidine-5'-triphosphate, 3'-O-amino-2',3'-dideoxythymidine-5'-triphosphate ou 3'-O-amino-2',3'-dideoxytadénosine-5'-triphosphate. Le groupement 3'-O-amino est un groupement plus volumineux fixé sur l'extrémité 3'-OH.

Synthèse

Tableau 7: Mélange réactionnel

Réactif	*Concentration*	*Volume*
H_2O	-	210 µL
Tampon	10x	70 µL
Nucléotides	2.5 µM	35 µL
Enzyme	80 µM	35 µL
Amorce sur support solide	1 µM	-

Le tampon 10x constitué de 250 mM Tris-HCl pH 7,2, 80 mM MgCl2, 3,3 mM ZnSO4 a été utilisé.

Le tampon de lavage L utilisé est constitué de Tris-HCl 25 mM à pH 7,2.

Le tampon de déprotection D utilisé est constitué d'acétate de sodium 50 mM pH 5,5 en présence de 10 mM MgCl2.

Avant le démarrage de la synthèse, les billes constituant le support solide sur lesquelles ont été hybridées les amorces pour une quantité total équivalente d'amorce de 35 pmol, ont été lavées plusieurs fois avec le tampon L. Après ces lavages, les billes ont été maintenues sur un aimant et le surnageant retiré dans sa totalité.

Plusieurs premix constitués des différents réactifs ajoutés dans l'ordre du tableau 7 ont été préparés. Chacun de ces premix contient des nucléotides différents suivant le tableau 8 ci-dessous.

Tableau 8 : Composition des premix

Numéro du premix	Nucléotide du premix
1	G
2	T
3	A
4	C
5	G
6	C
7	T
8	A
9	G
10	T

La synthèse commence lorsque le premix 1 est ajouté aux billes préalablement lavées et débarrassées de leur surnageant. Les étapes de synthèse selon le tableau 9 ci-dessous s'enchainent, afin de produire la nouvelle séquence 5'- GTACGCTAGT-3'.

Tableau 9 : Etape du procédé de synthèse d'un brin d'ADN sans brin matrice

Etapes	*Action*	*Volume*	*Durée*
Elongation 1	Ajout premix 1	350 µL	15min
Prélèvement 1	Prélèvement	5 µL	<1min
1er Lavage 1	Ajout tampon L	350 µL	5min
1er Déprotection 1	Ajout tampon D	350 µL	15min
2nd Déprotection 1	Ajout tampon D	350 µL	15min
2nd Lavage 1	Ajout tampon L	350 µL	5min
Elongation 2	Ajout premix 2	350 µL	15min
Prélèvement 2	Prélèvement	5 µL	<1min
1er Lavage 2	Ajout tampon L	350 µL	5min
1er Déprotection 2	Ajout tampon D	350 µL	15min
2nd Déprotection 2	Ajout tampon D	350 µL	15min
2nd Lavage 2	Ajout tampon L	350 µL	5min
Elongation 3	Ajout premix 3	350 µL	15min
Prélèvement 3	Prélèvement	5 µL	<1min
1er Lavage 3	Ajout tampon L	350 µL	5min
1er Déprotection 3	Ajout tampon D	350 µL	15min
2nd Déprotection 3	Ajout tampon D	350 µL	15min
2nd Lavage 3	Ajout tampon L	350 µL	5min
Elongation 4	Ajout premix 4	350 µL	15min
Prélèvement 4	Prélèvement	5 µL	<1min
1er Lavage 4	Ajout tampon L	350 µL	5min
1er Déprotection 4	Ajout tampon D	350 µL	15min
2nd Déprotection 4	Ajout tampon D	350 µL	15min
2nd Lavage 4	Ajout tampon L	350 µL	5min
Elongation 5	Ajout premix 5	350 µL	15min
Prélèvement 5	Prélèvement	5 µL	<1min
1er Lavage 5	Ajout tampon L	350 µL	5min
1er Déprotection 5	Ajout tampon D	350 µL	15min
2nd Déprotection 5	Ajout tampon D	350 µL	15min
2nd Lavage 5	Ajout tampon L	350 µL	5min
Elongation 6	Ajout premix 6	350 µL	15min
Prélèvement 6	Prélèvement	5 µL	<1min
1er Lavage 6	Ajout tampon L	350 µL	5min
1er Déprotection 6	Ajout tampon D	350 µL	15min
2nd Déprotection 6	Ajout tampon D	350 µL	15min
2nd Lavage 6	Ajout tampon L	350 µL	5min
Elongation 7	Ajout premix 7	350 µL	15min
Prélèvement 7	Prélèvement	5 µL	<1min
1er Lavage 7	Ajout tampon L	350 µL	5min
1er Déprotection 7	Ajout tampon D	350 µL	15min
2nd Déprotection 7	Ajout tampon D	350 µL	15min
2nd Lavage 7	Ajout tampon L	350 µL	5min
Elongation 8	Ajout premix 8	350 µL	15min
Prélèvement 8	Prélèvement	5 µL	<1min

1er Lavage 8	Ajout tampon L	350 µL	5min
1er Déprotection 8	Ajout tampon D	350 µL	15min
2nd Déprotection 8	Ajout tampon D	350 µL	15min
2nd Lavage 8	Ajout tampon L	350 µL	5min
Elongation 9	Ajout premix 9	350 µL	15min
Prélèvement 9	Prélèvement	5 µL	<1min
1er Lavage 9	Ajout tampon L	350 µL	5min
1er Déprotection 9	Ajout tampon D	350 µL	15min
2nd Déprotection 9	Ajout tampon D	350 µL	15min
2nd Lavage 9	Ajout tampon L	350 µL	5min
Elongation 10	Ajout premix 10	350 µL	15min
Prélèvement 10	Prélèvement	5 µL	<1min

Entre chaque étape, hormis celle de prélèvement, les billes sont collectées au moyen d'un aimant et le surnageant est retiré dans sa totalité.

Chaque prélèvement est ajouté à une solution de 15µL de bleu formamide (formamide 100%, 1 à 5 mg de bleu de bromophénol; Simga) afin de stopper la réaction et préparer l'analyse.

Gel et radiographie

L'analyse du test d'activité est réalisée par migration des différents échantillons dans un gel de polyacrylamide suivant le protocole décrit dans l'exemple 3.

Résultats

Les résultats de cette synthèse sont présentés à la figure 6.

La colonne 0 (No, pas de nucléotides) donne la taille attendue de l'amorce utilisée lorsqu'elle n'a pas été allongée, c'est-à-dire lorsqu'il n'y a pas eu d'incorporation de nucléotides.

Les colonnes 1 à 10 correspondent aux prélèvements 1 à 10 lors de la synthèse. Chaque incorporation de nucléotides a été réalisée par l'enzyme avec une performance maximale. Aucune étape de purification supplémentaire n'est réalisée.

Une expérience similaire de synthèse a été réalisée avec la TdT naturelle. Celle-ci étant incapable d'incorporer des nucléotides modifiés il n'a pas était possible de synthétiser la séquence désirée.

REVENDICATIONS

1. Variant d'une ADN polymérase de la famille polX capable de synthétiser une molécule d'acide nucléique sans brin matrice, ou d'un fragment fonctionnel d'une telle polymérase, ledit variant comprenant au moins une mutation d'un résidu à au moins une position sélectionnée dans le groupe consistant en E457, T331, G332, G333, F334, R336, K338, H342, D343, V344, D345, F346, A397, D399, D434, V436, A446, L447, L448, G449, W450, G452, R454, Q455, F456, R458, R461, N474, E491, D501, Y502, I503, P505, R508, N509 et A510, ou un résidu fonctionnellement équivalent, les positions indiquées étant déterminées par alignement avec SEQ ID N°1.

2. Variant d'une ADN polymérase de la famille polX selon la revendication 1, ledit variant étant capable de synthétiser un brin d'ADN et/ou un brin d'ARN.

3. Variant d'une ADN polymérase de la famille polX selon la revendication 1 ou 2, ledit variant étant un variant de Pol IV, Pol μ, ou de la désoxyribonucléotidyl-transférase terminale (TdT).

4. Variant d'une ADN polymérase de la famille polX selon l'une des revendications précédentes, et présentant au moins 60% d'identité avec la séquence selon SEQ ID N°1, préférentiellement au moins 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% d'identité avec la séquence selon SEQ ID N°1.

5. Variant d'une ADN polymérase de la famille polX selon l'une des revendications précédentes, dans lequel au moins une mutation consiste en une substitution, une délétion ou une addition d'un ou plusieurs résidus d'acide aminé.

6. Variant d'une ADN polymérase de la famille polX selon l'une des revendications précédentes, ledit variant comprenant au moins une mutation d'un résidu à au moins une position sélectionnée dans le groupe consistant en T331, G332, G333, F334, R336, D343, L447, L448, G449, W450, G452, R454, Q455, E457, R461 et R508, ou un résidu fonctionnellement équivalent, préférentiellement au moins une mutation d'un résidu à au moins une position sélectionnée dans le groupe consistant en R336, R454 et E457, ou un résidu fonctionnellement équivalent, les positions indiquées étant déterminées par alignement avec SEQ ID N°1.

7. Variant d'une ADN polymérase de la famille polX selon l'une des revendications précédentes, ledit variant comprenant au moins une mutation d'un résidu à au moins deux positions sélectionnées dans le groupe consistant en R336, R454 et E457, préférentiellement une mutation d'un résidu auxdites trois positions R336, R454 et E457.

8. Variant d'une ADN polymérase de la famille polX selon l'une des revendications précédentes, ledit variant présentant au moins une mutation d'un résidu dans au moins une région semi-conservée de séquence

(i) $X_1X_2GGFR_1R_2GKX_3X_4$ (SEQ ID N°4),

dans laquelle

X_1 représente un résidu choisi parmi M, I, V, L

X_2 représente un résidu choisi parmi T, A, M, Q

X_3 représente un résidu choisi parmi M, K, E, Q, L, S, P, R, D

X_4 représente un résidu choisi parmi T, I, M, F, K, V, Y, E, Q, H, S, R, D

(ii) $X_1X_2LGX_3X_4GSR_1X_5X_6ER_2$ (SEQ ID N°5)

dans laquelle

X_1 représente un résidu choisi parmi A, C, G, S

X_2 représente un résidu choisi parmi L, T, R

X_3 représente un résidu choisi parmi W, Y

X_4 représente un résidu choisi parmi T, S, I

X_5 représente un résidu choisi parmi Q, L, H, F, Y, N, E, D ou Ø

X_6 représente un résidu choisi parmi F, Y

(iii) $LX_1YX_2X_3PX_4X_5RNA$ (SEQ ID N°6)

X_1 représente un résidu choisi parmi D, E, S, P, A, K

X_2 représente un résidu choisi parmi I, L, M, V, A, T

X_3 représente un résidu choisi parmi E, Q, P, Y, L, K, G, N

X_4 représente un résidu choisi parmi W, S, V, E, R, Q, T, C, K, H

X_5 représente un résidu choisi parmi E, Q, D, H, L.

9. Variant d'une ADN polymérase de la famille polX selon la revendication 8, ledit variant présentant

- au moins une substitution d'un résidu à au moins une position R_1, R_2 et/ou K de la région semi-conservée de séquence SEQ ID N°4 ; et/ou
- au moins une substitution d'un résidu à au moins une position S, R1 et/ou E de la région semi-conservée de séquence SEQ ID N°5 ; et/ou
- une délétion du résidu à la position X1 et/ou au moins une substitution au niveau des positions R et/ou N de la région semi-conservée de séquence SEQ ID N°6.

10. Variant d'une ADN polymérase de la famille polX selon l'une des revendications précédentes, ledit variant comprenant une substitution d'un résidu à au moins une position sélectionnée dans le groupe consistant en R336, K338, H342, A397, S453, R454, E457, R461, N474, D501, Y502, I503, R508 et N509, ou un résidu fonctionnellement équivalent, préférentiellement une substitution d'un résidu à au moins une position sélectionnée dans le groupe consistant en R336, A397, R454, E457, R461, N474, D501, Y502 et I503, ou un résidu fonctionnellement équivalent, plus préférentiellement une substitution d'un résidu à au moins une position sélectionnée dans le groupe consistant en R336, R454 et E457, ou un résidu fonctionnellement équivalent, encore plus préférentiellement au moins une substitution sur le résidu à la position E457, ou un résidu fonctionnellement équivalent, les positions indiquées étant déterminées par alignement avec SEQ ID N°1.

11. Variant d'une ADN polymérase de la famille polX selon la revendication 10, dans lequel les substitutions sur les positions R336, K338, H342, A397, S453, R454, E457, R461, N474, D501, Y502, I503, R508 et N509 sont sélectionnées parmi le groupe consistant en R336K/H/N/G/D, K338A/C/G/S/T/N, H342A/C/G/S/T.N, A397R/H/K/D/E, S453A/C/G/S/T, R454F/Y/W/A, E457G/N/S/T, N474S/T/N/Q, D501A/G/X, Y502A/G/X, I503A/G/X, R508A/C/G/S/T, N509A/C/G/S/T.

12. Variant d'une ADN polymérase de la famille polX selon l'une des revendications précédentes dans laquelle le variant comprend ou présente une substitution, délétion, combinaisons de substitutions et/ou de délétions listées dans le tableau 1, les positions indiquées étant déterminées par alignement avec SEQ ID N°1.

13. Variant d'une ADN polymérase de la famille polX selon l'une des revendications précédentes, ledit variant étant un variant de la TdT de séquence SEQ ID N°1 et comprenant en outre une substitution des résidus compris entre les positions C378 à L406, ou les résidus

fonctionnellement équivalents par les résidus H363 à C390 de la polymérase Polμ de séquence SEQ ID N°2, ou les résidus fonctionnellement équivalents.

14. Variant d'une ADN polymérase de la famille polX selon l'une des revendications précédentes dans laquelle le variant comprend une combinaison de substitution choisie parmi R336G-E457N ; R336N-E457N ; R336N-R454A-E457N ; R336N-E457N ; R336N-R454A-E457G ; R336N-E457G ; R336G-R454A-E457N ; R336G-E457N, les positions indiquées étant déterminées par alignement avec SEQ ID N°1.

15. Acide nucléique codant un variant d'une ADN polymérase de la famille polX selon l'une des revendications 1 à 14.

16. Cassette d'expression d'un acide nucléique selon la revendication 15.

17. Vecteur comprenant un acide nucléique selon la revendication 15 ou une cassette d'expression selon la revendication 16.

18. Cellule hôte comprenant un acide nucléique selon la revendication 15 ou une cassette d'expression selon la revendication 16 ou un vecteur selon la revendication 17.

19. Utilisation d'un acide nucléique selon la revendication 15, d'une cassette d'expression selon la revendication 16, d'un vecteur selon la revendication 17 ou d'une cellule selon la revendication 18, pour produire un variant d'une ADN polymérase de la famille polX selon l'une quelconque des revendications 1-14.

20. Procédé de production d'un variant d'une ADN polymérase de la famille polX selon l'une quelconque des revendications 1-14, selon lequel on cultive une cellule hôte selon la revendication 18 dans des conditions de culture permettant l'expression de l'acide nucléique codant ledit variant, et optionnellement on récupère ledit variant ainsi exprimé à partir du milieu de culture ou desdites cellules hôtes.

21. Utilisation d'un variant d'une ADN polymérase de la famille polX selon l'une quelconque des revendications 1-14, pour synthétiser une molécule d'acide nucléique sans brin matrice, à partir de nucléotides modifiés en 3'-OH.

22. Utilisation selon la revendication 21, pour synthétiser un brin d'ADN ou un brin d'ARN.

23. Procédé de synthèse enzymatique d'une molécule d'acide nucléique sans brin matrice, selon lequel on met en contact un brin amorce avec au moins un nucléotide, préférentiellement un nucléotide modifié en 3'-OH, en présence d'un variant d'une ADN polymérase de la famille polX selon l'une quelconque des revendications 1-14.

24. Kit pour la synthèse enzymatique d'une molécule d'acide nucléique sans brin matrice comprenant au moins un variant d'une ADN polymérase de la famille polX selon l'une quelconque des revendications 1-14, des nucléotides, préférentiellement des nucléotides modifiés en 3'-OH, et optionnellement au moins une amorce nucléotidique.