

①⑨ RÉPUBLIQUE FRANÇAISE
INSTITUT NATIONAL
DE LA PROPRIÉTÉ INDUSTRIELLE
COURBEVOIE

①① N° de publication :
(à n'utiliser que pour les
commandes de reproduction)

3 052 462

②① N° d'enregistrement national : **16 55475**

⑤① Int Cl⁸ : **C 12 N 9/12** (2017.01), C 12 N 15/54, 15/63, C 12 P 19/34, C 12 Q 1/68

⑫

DEMANDE DE BREVET D'INVENTION

A1

②② Date de dépôt : 14.06.16.

③③ Priorité :

⑦① Demandeur(s) : *DNA SCRIPT Société par actions simplifiée — FR et INSTITUT PASTEUR — FR.*

⑦② Inventeur(s) : YBERT THOMAS et DELARUE MARC.

④③ Date de mise à la disposition du public de la demande : 15.12.17 Bulletin 17/50.

⑤⑥ Liste des documents cités dans le rapport de recherche préliminaire : *Se reporter à la fin du présent fascicule*

⑥⑥ Références à d'autres documents nationaux apparentés : vision demandée le 07/07/17 bénéficiant de la date de dépôt du 14/06/16 de la demande initiale n° 16 1655475.

⑦③ Titulaire(s) : DNA SCRIPT Société par actions simplifiée, INSTITUT PASTEUR.

○ Demande(s) d'extension :

⑦④ Mandataire(s) : CABINET BECKER ET ASSOCIES.

⑤④ VARIANTS D'UNE ADN POLYMERASE DE LA FAMILLE POLX.

⑤⑦ L'invention concerne des variants 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, comprenant au moins une mutation d'un résidu à au moins une position particulière, et des utilisations de ces variants, notamment pour la synthèse de molécules d'acide nucléiques comprenant des nucléotides modifiés en 3'-OH.

FR 3 052 462 - A1



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
5 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,
10 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 connaît cependant certaines limites. En effet, elle nécessite l'utilisation de multiples solvants et réactifs. En outre, elle ne permet d'obtenir que des
15 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
20 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
25 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-transférase 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 M330, T331, G332, G333, F334, K338, H342, D343, V344, D345, F346, A397, D399 D434, V436, A446, L447, L448, G449, W450, G452, 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 un autre 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.

- 5 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
- 10 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

15 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

20 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

25 d'Homo sapiens (UniProtKB P04053) obtenu au moyen du logiciel d'alignement en ligne Multalin (<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).

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 Pol μ 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 substitution, délétion et / ou insertion 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-méthyllysine, la N-éthylglycine, la N-méthylglycine, la N-éthylasparagine, 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 Pol β (UniProt P06746 chez humain ; Q8K409 chez la souris), la Pol σ , la Pol λ (UniProt Q9UGP5 chez humain ; Q9QUG2 et Q9QXE2 chez la souris) et la Pol μ (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 M330, T331, G332, G333, F334, K338, H342, D343, V344, D345, F346, A397, D399, D434, V436, A446, L447, L448, G449, W450, G452, 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.

- 5 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, D399, R461, et/ou R508), l'accroissement de l'accessibilité à la poche catalytique (en ciblant par exemple les positions R458, E455, A397, K338, et/ou N509), et/ou confère une plus grande flexibilité à la structure de l'enzyme lui permettant de recevoir
- 10 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 , Pol β , Pol λ 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
- 15 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
- 20 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
- 25 supprimés. L'annotation \emptyset désigne par contre une simple délétion ponctuelle du résidu considéré. Ainsi, la mutation D501 \emptyset 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,

D343, L447, L448, G449, W450, G452, Q455, E457 et R508, 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

5 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.

10 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

15 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 \emptyset

20 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

25 $LX_1YX_2X_3PX_4X_5RNA$ (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

X₃ représente un résidu choisi parmi E, Q, P, Y, L, K, G, N

X₄ représente un résidu choisi parmi W, S, V, E, R, Q, T, C, K, H

X₅ représente un résidu choisi parmi E, Q, D, H, L.

Préférentiellement, ledit variant présente au moins une délétion du résidu à la position X₁ et/ou
5 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 K338, H342, A397, S453, E457,
N474, D501, Y502, I503, R508 et N509, ou un résidu fonctionnellement équivalent,
10 préférentiellement une substitution d'un résidu à au moins une position sélectionnée dans le
groupe consistant en A397, E457, N474, D501, Y502 et I503, 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 K338A/C/G/S/T/N,
15 H342A/C/G/S/T/N, A397R/H/K/D/E, S453A/C/G/S/T, 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.

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.

20 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.

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
DS19	R336A - E457N - A397D
DS20	R336A - 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
DS58	R336A - E457S - A397D
DS59	R336A - 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
DS95	R336A - E457T - A397D
DS96	R336A - 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

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
5 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
10 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).

15 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 mutagenèse peuvent être guidées par des
20 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
25 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
5 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
10 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
15 la polyhédrine, 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
20 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
25 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
30 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
5 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-
10 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
15 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
20 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
25 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
30 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.

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.

- 5 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
10 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.

15 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
20 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)

25 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 boîtes kanamycine/chloramphénicol ont été isolées et notées Ec- SEQ ID N°3

Génération des variants

- 5 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>)

- 10 Le kit QuickChange II (Agilent) a été utilisé pour générer les plasmides des variants comportant la/les mutation(s) désirée(s). Le protocole de mutagenèse donné par le fabricant 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 boîtes kanamycine/chloramphénicol ont été isolées et notées Ec-DSx.

15 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 ensemer un erlen de 5 L contenant 2 L de milieu LB
- 20 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
- 25 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 NaCl). 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 recueilli 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 NaCl, 60 mM imidazol). A la fin de cette étape le tampon d'élution (20 mM Tris-HCl, pH 8.3, 0.5 M NaCl, 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 [NH₄]₂SO₄). 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)	<i>Homo sapiens</i>	494
H2QUI0	Pol μ (SEQ ID N°9)	<i>Pan troglodytes</i>	494
Q924W4	Pol μ (SEQ ID N°10)	<i>Mus musculus</i>	496
F1P657	TdT (SEQ ID N°11)	<i>Canis lupus familiaris</i>	509
Q3UZ80	TdT (SEQ ID N°1)	<i>Mus musculus</i>	510
P36195	TdT (SEQ ID N°12)	<i>Gallus gallus</i>	506
P04053	TdT (SEQ ID N°13)	<i>Homo sapiens</i>	509

Les alignements obtenus sont présentés à la figure 2.

10 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

15 Tableau 3 : Mélange réactionnel

Réactif	Concentration	volume
H ₂ O	-	15 μ L
Amorce	500 nM	2,5 μ L

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).

- 5 Le tampon 10x constitué de 250 mM Tris-HCl pH 7,2, 80 mM MgCl₂, 3,3 mM ZnSO₄ 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
 10 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
 15 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

- 20 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.

- 25 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.

- 5 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 :

- 10 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%.

- 15 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 :

- 20 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

5 tendant à augmenter de manière très importante l'encombrement stérique dudit nucléotide.

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
5 variant comprenant au moins une mutation d'un résidu à au moins une position sélectionnée dans le groupe consistant en M330, T331, G332, G333, F334, K338, H342, D343, V344, D345, F346, A397, D399, D434, V436, A446, L447, L448, G449, W450, G452, 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
10 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).
- 15 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
20 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, D343, L447, L448,
25 G449, W450, G452, Q455, E457, R461 et R508, 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 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),

5 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

10 (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

15 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 \emptyset

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

20 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.

8. Variant d'une ADN polymérase de la famille polX selon la revendication 7, ledit variant
25 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, R_1 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.

9. 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
5 sélectionnée dans le groupe consistant en K338, H342, A397, S453, 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 A397, E457, R461, N474, D501, Y502 et I503, ou un résidu fonctionnellement équivalent, les positions indiquées étant déterminées par alignement avec SEQ ID N°1.
10. Variant d'une ADN polymérase de la famille polX selon la revendication 9, dans lequel les substitutions sur les positions K338, H342, A397, S453, E457, R461, N474, D501, Y502, I503, R508 et N509 sont sélectionnées parmi le groupe consistant en K338A/C/G/S/T/N, H342A/C/G/S/T.N, A397R/H/K/D/E, S453A/C/G/S/T, 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..
- 15 11. 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.
12. Variant d'une ADN polymérase de la famille polX selon l'une des revendications
20 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.
13. Acide nucléique codant un variant d'une ADN polymérase de la famille polX selon l'une
25 des revendications 1 à 12.
14. Cassette d'expression d'un acide nucléique selon la revendication 13.
15. Vecteur comprenant un acide nucléique selon la revendication 13 ou une cassette d'expression selon la revendication 14.

16. Cellule hôte comprenant un acide nucléique selon la revendication 13 ou une cassette d'expression selon la revendication 14 ou un vecteur selon la revendication 15.
17. Utilisation d'un acide nucléique selon la revendication 13, d'une cassette d'expression selon la revendication 14, d'un vecteur selon la revendication 15 ou d'une cellule selon la revendication 16, pour produire un variant d'une ADN polymérase de la famille polX selon l'une quelconque des revendications 1-12.
18. Procédé de production d'un variant d'une ADN polymérase de la famille polX selon l'une quelconque des revendications 1-12, selon lequel on cultive une cellule hôte selon la revendication 16 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.
19. Utilisation d'un variant d'une ADN polymérase de la famille polX selon l'une quelconque des revendications 1-12, pour synthétiser une molécule d'acide nucléique sans brin matrice, à partir de nucléotides modifiés en 3'-OH.
20. Utilisation selon la revendication 19, pour synthétiser un brin d'ADN ou un brin d'ARN.
21. 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 1-12.
22. 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 1-12, des nucléotides, préférentiellement des nucléotides modifiés en 3'-OH, et optionnellement au moins une amorce nucléotidique.

1/2

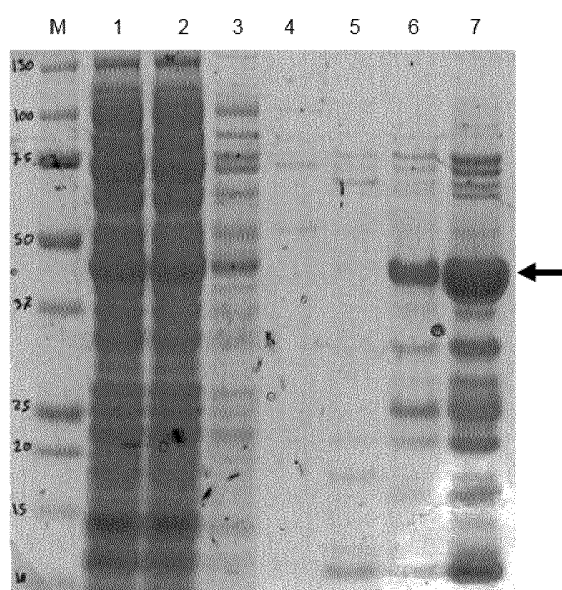


FIGURE 1

1 10 20 30 40 50 60 70 80 90

TD1|Q3U280|Q3U280_M
TD1|F1P657|F1P657_C
TD1|P4O533|TD1_HUMAN
TD1|P36195|TD1_CHICK
POLMU|Q9NB87|DPOCLM_B
POLMU|H2Q010|H2Q010_C
POLMU|Q924W4|Q924W4_C
consensus:50

100 110 120 130 140 150 160 170 180 190

TD1|Q3U280|Q3U280_M
TD1|F1P657|F1P657_C
TD1|P4O533|TD1_HUMAN
TD1|P36195|TD1_CHICK
POLMU|Q9NB87|DPOCLM_B
POLMU|H2Q010|H2Q010_C
POLMU|Q924W4|Q924W4_C
consensus:50

200 210 220 230 240 250 260 270 280 290

TD1|Q3U280|Q3U280_M
TD1|F1P657|F1P657_C
TD1|P4O533|TD1_HUMAN
TD1|P36195|TD1_CHICK
POLMU|Q9NB87|DPOCLM_B
POLMU|H2Q010|H2Q010_C
POLMU|Q924W4|Q924W4_C
consensus:50

300 310 320 330 340 350 360 370 380 390

TD1|Q3U280|Q3U280_M
TD1|F1P657|F1P657_C
TD1|P4O533|TD1_HUMAN
TD1|P36195|TD1_CHICK
POLMU|Q9NB87|DPOCLM_B
POLMU|H2Q010|H2Q010_C
POLMU|Q924W4|Q924W4_C
consensus:50

400 410 420 430 440 450 460 470 480 490

TD1|Q3U280|Q3U280_M
TD1|F1P657|F1P657_C
TD1|P4O533|TD1_HUMAN
TD1|P36195|TD1_CHICK
POLMU|Q9NB87|DPOCLM_B
POLMU|H2Q010|H2Q010_C
POLMU|Q924W4|Q924W4_C
consensus:50

500 510

TD1|Q3U280|Q3U280_M
TD1|F1P657|F1P657_C
TD1|P4O533|TD1_HUMAN
TD1|P36195|TD1_CHICK
POLMU|Q9NB87|DPOCLM_B
POLMU|H2Q010|H2Q010_C
POLMU|Q924W4|Q924W4_C
consensus:50

FIGURE 2

Incorporation ONH2

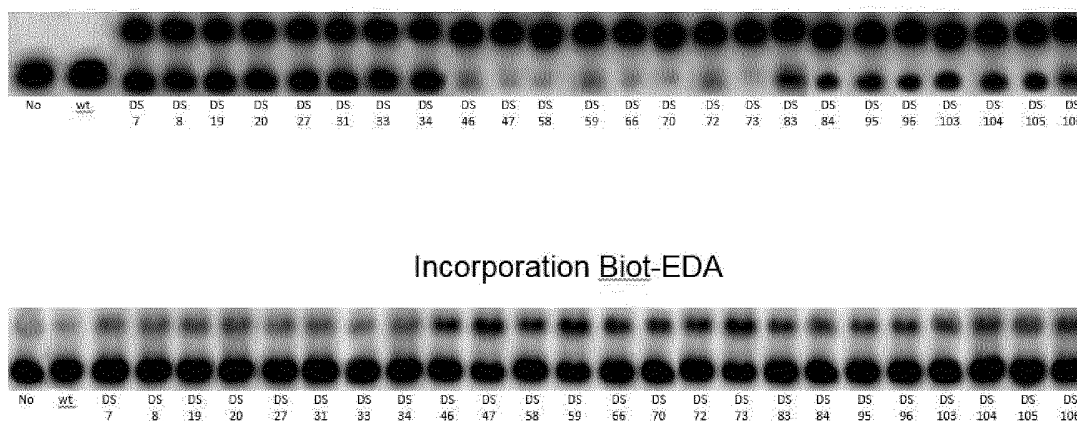


FIGURE 3

B2265FR00 séquence listing.txt

SEQUENCE LISTING

<110> DNA Script
Institut Pasteur

<120> Variants d'une ADN polymérase de la famille PolX

<130> B2265

<160> 14

<170> PatentIn version 3.5

<210> 1

<211> 510

<212> PRT

<213> artificial sequence

<220>

<223> TdT de souris

<400> 1

Met Asp Pro Leu Gln Ala Val His Leu Gly Pro Arg Lys Lys Arg Pro
1 5 10 15

Arg Gln Leu Gly Thr Pro Val Ala Ser Thr Pro Tyr Asp Ile Arg Phe
 20 25 30

Arg Asp Leu Val Leu Phe Ile Leu Glu Lys Lys Met Gly Thr Thr Arg
 35 40 45

Arg Ala Phe Leu Met Glu Leu Ala Arg Arg Lys Gly Phe Arg Val Glu
 50 55 60

Asn Glu Leu Ser Asp Ser Val Thr His Ile Val Ala Glu Asn Asn Ser
65 70 75 80

Gly Ser Asp Val Leu Glu Trp Leu Gln Leu Gln Asn Ile Lys Ala Ser
 85 90 95

Ser Glu Leu Glu Leu Leu Asp Ile Ser Trp Leu Ile Glu Cys Met Gly
 100 105 110

B2265FR00 séquence listing.txt

Ala Gly Lys Pro Val Glu Met Met Gly Arg His Gln Leu Val Val Asn
115 120 125

Arg Asn Ser Ser Pro Ser Pro Val Pro Gly Ser Gln Asn Val Pro Ala
130 135 140

Pro Ala Val Lys Lys Ile Ser Gln Tyr Ala Cys Gln Arg Arg Thr Thr
145 150 155 160

Leu Asn Asn Tyr Asn Gln Leu Phe Thr Asp Ala Leu Asp Ile Leu Ala
165 170 175

Glu Asn Asp Glu Leu Arg Glu Asn Glu Gly Ser Cys Leu Ala Phe Met
180 185 190

Arg Ala Ser Ser Val Leu Lys Ser Leu Pro Phe Pro Ile Thr Ser Met
195 200 205

Lys Asp Thr Glu Gly Ile Pro Cys Leu Gly Asp Lys Val Lys Ser Ile
210 215 220

Ile Glu Gly Ile Ile Glu Asp Gly Glu Ser Ser Glu Ala Lys Ala Val
225 230 235 240

Leu Asn Asp Glu Arg Tyr Lys Ser Phe Lys Leu Phe Thr Ser Val Phe
245 250 255

Gly Val Gly Leu Lys Thr Ala Glu Lys Trp Phe Arg Met Gly Phe Arg
260 265 270

Thr Leu Ser Lys Ile Gln Ser Asp Lys Ser Leu Arg Phe Thr Gln Met
275 280 285

Gln Lys Ala Gly Phe Leu Tyr Tyr Glu Asp Leu Val Ser Cys Val Asn
290 295 300

Arg Pro Glu Ala Glu Ala Val Ser Met Leu Val Lys Glu Ala Val Val

B2265FR00 séquence listing.txt

305 310 315 320
 Thr Phe Leu Pro Asp Ala Leu Val Thr Met Thr Gly Gly Phe Arg Arg
 325 330 335

 Gly Lys Met Thr Gly His Asp Val Asp Phe Leu Ile Thr Ser Pro Glu
 340 345 350

 Ala Thr Glu Asp Glu Glu Gln Gln Leu Leu His Lys Val Thr Asp Phe
 355 360 365

 Trp Lys Gln Gln Gly Leu Leu Leu Tyr Cys Asp Ile Leu Glu Ser Thr
 370 375 380

 Phe Glu Lys Phe Lys Gln Pro Ser Arg Lys Val Asp Ala Leu Asp His
 385 390 395 400

 Phe Gln Lys Cys Phe Leu Ile Leu Lys Leu Asp His Gly Arg Val His
 405 410 415

 Ser Glu Lys Ser Gly Gln Gln Glu Gly Lys Gly Trp Lys Ala Ile Arg
 420 425 430

 Val Asp Leu Val Met Cys Pro Tyr Asp Arg Arg Ala Phe Ala Leu Leu
 435 440 445

 Gly Trp Thr Gly Ser Arg Gln Phe Glu Arg Asp Leu Arg Arg Tyr Ala
 450 455 460

 Thr His Glu Arg Lys Met Met Leu Asp Asn His Ala Leu Tyr Asp Arg
 465 470 475 480

 Thr Lys Arg Val Phe Leu Glu Ala Glu Ser Glu Glu Glu Ile Phe Ala
 485 490 495

 His Leu Gly Leu Asp Tyr Ile Glu Pro Trp Glu Arg Asn Ala
 500 505 510

B2265FR00 séquence listing.txt

<210> 2

<211> 494

<212> PRT

<213> artificial sequence

<220>

<223> Pol μ humaine

<400> 2

Met Leu Pro Lys Arg Arg Arg Ala Arg Val Gly Ser Pro Ser Gly Asp
 1 5 10 15

Ala Ala Ser Ser Thr Pro Pro Ser Thr Arg Phe Pro Gly Val Ala Ile
 20 25 30

Tyr Leu Val Glu Pro Arg Met Gly Arg Ser Arg Arg Ala Phe Leu Thr
 35 40 45

Gly Leu Ala Arg Ser Lys Gly Phe Arg Val Leu Asp Ala Cys Ser Ser
 50 55 60

Glu Ala Thr His Val Val Met Glu Glu Thr Ser Ala Glu Glu Ala Val
 65 70 75 80

Ser Trp Gln Glu Arg Arg Met Ala Ala Ala Pro Pro Gly Cys Thr Pro
 85 90 95

Pro Ala Leu Leu Asp Ile Ser Trp Leu Thr Glu Ser Leu Gly Ala Gly
 100 105 110

Gln Pro Val Pro Val Glu Cys Arg His Arg Leu Glu Val Ala Gly Pro
 115 120 125

Arg Lys Gly Pro Leu Ser Pro Ala Trp Met Pro Ala Tyr Ala Cys Gln
 130 135 140

Arg Pro Thr Pro Leu Thr His His Asn Thr Gly Leu Ser Glu Ala Leu
 145 150 155 160

B2265FR00 séquence listing.txt

Glu Ile Leu Ala Glu Ala Ala Gly Phe Glu Gly Ser Glu Gly Arg Leu
 165 170 175

Leu Thr Phe Cys Arg Ala Ala Ser Val Leu Lys Ala Leu Pro Ser Pro
 180 185 190

Val Thr Thr Leu Ser Gln Leu Gln Gly Leu Pro His Phe Gly Glu His
 195 200 205

Ser Ser Arg Val Val Gln Glu Leu Leu Glu His Gly Val Cys Glu Glu
 210 215 220

Val Glu Arg Val Arg Arg Ser Glu Arg Tyr Gln Thr Met Lys Leu Phe
 225 230 235 240

Thr Gln Ile Phe Gly Val Gly Val Lys Thr Ala Asp Arg Trp Tyr Arg
 245 250 255

Glu Gly Leu Arg Thr Leu Asp Asp Leu Arg Glu Gln Pro Gln Lys Leu
 260 265 270

Thr Gln Gln Gln Lys Ala Gly Leu Gln His His Gln Asp Leu Ser Thr
 275 280 285

Pro Val Leu Arg Ser Asp Val Asp Ala Leu Gln Gln Val Val Glu Glu
 290 295 300

Ala Val Gly Gln Ala Leu Pro Gly Ala Thr Val Thr Leu Thr Gly Gly
 305 310 315 320

Phe Arg Arg Gly Lys Leu Gln Gly His Asp Val Asp Phe Leu Ile Thr
 325 330 335

His Pro Lys Glu Gly Gln Glu Ala Gly Leu Leu Pro Arg Val Met Cys
 340 345 350

Arg Leu Gln Asp Gln Gly Leu Ile Leu Tyr His Gln His Gln His Ser

B2265FR00 séquence listing.txt

355 360 365

Cys Cys Glu Ser Pro Thr Arg Leu Ala Gln Gln Ser His Met Asp Ala
370 375 380

Phe Glu Arg Ser Phe Cys Ile Phe Arg Leu Pro Gln Pro Pro Gly Ala
385 390 395 400

Ala Val Gly Gly Ser Thr Arg Pro Cys Pro Ser Trp Lys Ala Val Arg
405 410 415

Val Asp Leu Val Val Ala Pro Val Ser Gln Phe Pro Phe Ala Leu Leu
420 425 430

Gly Trp Thr Gly Ser Lys Leu Phe Gln Arg Glu Leu Arg Arg Phe Ser
435 440 445

Arg Lys Glu Lys Gly Leu Trp Leu Asn Ser His Gly Leu Phe Asp Pro
450 455 460

Glu Gln Lys Thr Phe Phe Gln Ala Ala Ser Glu Glu Asp Ile Phe Arg
465 470 475 480

His Leu Gly Leu Glu Tyr Leu Pro Pro Glu Gln Arg Asn Ala
485 490

<210> 3

<211> 401

<212> PRT

<213> artificial sequence

<220>

<223> TdT de souris tronquée

<400> 3

Thr Met Gly Ser Ser His His His His His Ser Ser Gly Leu Val
1 5 10 15

Pro Arg Gly Ser His Met Ser Pro Ser Pro Val Pro Gly Ser Gln Asn

B2265FR00 séquence listing.txt

20 25 30

Val Pro Ala Pro Ala Val Lys Lys Ile Ser Gln Tyr Ala Cys Gln Arg
35 40 45

Arg Thr Thr Leu Asn Asn Tyr Asn Gln Leu Phe Thr Asp Ala Leu Asp
50 55 60

Ile Leu Ala Glu Asn Asp Glu Leu Arg Glu Asn Glu Gly Ser Cys Leu
65 70 75 80

Ala Phe Met Arg Ala Ser Ser Val Leu Lys Ser Leu Pro Phe Pro Ile
85 90 95

Thr Ser Met Lys Asp Thr Glu Gly Ile Pro Cys Leu Gly Asp Lys Val
100 105 110

Lys Ser Ile Ile Glu Gly Ile Ile Glu Asp Gly Glu Ser Ser Glu Ala
115 120 125

Lys Ala Val Leu Asn Asp Glu Arg Tyr Lys Ser Phe Lys Leu Phe Thr
130 135 140

Ser Val Phe Gly Val Gly Leu Lys Thr Ala Glu Lys Trp Phe Arg Met
145 150 155 160

Gly Phe Arg Thr Leu Ser Lys Ile Gln Ser Asp Lys Ser Leu Arg Phe
165 170 175

Thr Gln Met Gln Lys Ala Gly Phe Leu Tyr Tyr Glu Asp Leu Val Ser
180 185 190

Cys Val Asn Arg Pro Glu Ala Glu Ala Val Ser Met Leu Val Lys Glu
195 200 205

Ala Val Val Thr Phe Leu Pro Asp Ala Leu Val Thr Met Thr Gly Gly
210 215 220

B2265FR00 séquence listing.txt

Phe Arg Arg Gly Lys Met Thr Gly His Asp Val Asp Phe Leu Ile Thr
 225 230 235 240

Ser Pro Glu Ala Thr Glu Asp Glu Glu Gln Gln Leu Leu His Lys Val
 245 250 255

Thr Asp Phe Trp Lys Gln Gln Gly Leu Leu Leu Tyr Cys Asp Ile Leu
 260 265 270

Glu Ser Thr Phe Glu Lys Phe Lys Gln Pro Ser Arg Lys Val Asp Ala
 275 280 285

Leu Asp His Phe Gln Lys Cys Phe Leu Ile Leu Lys Leu Asp His Gly
 290 295 300

Arg Val His Ser Glu Lys Ser Gly Gln Gln Glu Gly Lys Gly Trp Lys
 305 310 315 320

Ala Ile Arg Val Asp Leu Val Met Cys Pro Tyr Asp Arg Arg Ala Phe
 325 330 335

Ala Leu Leu Gly Trp Thr Gly Ser Arg Gln Phe Glu Arg Asp Leu Arg
 340 345 350

Arg Tyr Ala Thr His Glu Arg Lys Met Met Leu Asp Asn His Ala Leu
 355 360 365

Tyr Asp Arg Thr Lys Arg Val Phe Leu Glu Ala Glu Ser Glu Glu Glu
 370 375 380

Ile Phe Ala His Leu Gly Leu Asp Tyr Ile Glu Pro Trp Glu Arg Asn
 385 390 395 400

Ala

<210> 4

B2265FR00 séquence listing.txt

<211> 11
 <212> PRT
 <213> artificial sequence

<220>
 <223> séquence semi-conservée

<220>
 <221> MISC_FEATURE
 <222> (1)..(1)
 <223> X=M, I, V, L

<220>
 <221> MISC_FEATURE
 <222> (2)..(2)
 <223> X=T, A, M, Q

<220>
 <221> MISC_FEATURE
 <222> (10)..(10)
 <223> X=M, K, E, Q, L, S, P, R, D

<220>
 <221> MISC_FEATURE
 <222> (11)..(11)
 <223> X=T, I, M, F, K, V, Y, E, Q, H, S, R, D

<400> 4

Xaa Xaa Gly Gly Phe Arg Arg Gly Lys Xaa Xaa
 1 5 10

<210> 5
 <211> 13
 <212> PRT
 <213> artificial sequence

<220>
 <223> région semi-conservée

<220>
 <221> MISC_FEATURE
 <222> (1)..(1)
 <223> X=A, C, G, S

<220>
 <221> MISC_FEATURE

B2265FR00 séquence listing.txt

<222> (2)..(2)

<223> X=L, T, R

<220>

<221> MISC_FEATURE

<222> (5)..(5)

<223> X=W, Y

<220>

<221> MISC_FEATURE

<222> (6)..(6)

<223> X=T, S, I

<220>

<221> MISC_FEATURE

<222> (10)..(10)

<223> X=Q, L, H, F, Y, N, E, D,

<220>

<221> MISC_FEATURE

<222> (11)..(11)

<223> X=F, Y

<400> 5

Xaa Xaa Leu Gly Xaa Xaa Gly Ser Arg Xaa Xaa Glu Arg

1 5 10

<210> 6

<211> 11

<212> PRT

<213> artificial sequence

<220>

<223> région semi-conservée

<220>

<221> MISC_FEATURE

<222> (2)..(2)

<223> X=D, E, S, P, A, K

<220>

<221> MISC_FEATURE

<222> (4)..(4)

<223> X=I, L, M, V, A, T

<220>

<221> MISC_FEATURE

B2265FR00 séquence listing.txt

<222> (5)..(5)

<223> X=E, Q, P, Y, L, K, G, N

<220>

<221> MISC_FEATURE

<222> (7)..(7)

<223> X=W, S, V, E, R, Q, T, C, K, H

<220>

<221> MISC_FEATURE

<222> (8)..(8)

<223> X=E, Q, D, H, L

<400> 6

Leu Xaa Tyr Xaa Xaa Pro Xaa Xaa Arg Asn Ala

1 5 10

<210> 7

<211> 20

<212> DNA

<213> artificial sequence

<220>

<223> amorce

<400> 7

taatacgact cactataggg

20

<210> 8

<211> 19

<212> DNA

<213> artificial sequence

<220>

<223> amorce

<400> 8

gctagttatt gctcagcgg

19

<210> 9

<211> 494

<212> PRT

<213> Pan troglodytes

<400> 9

B2265FR00 séquence listing.txt

Met Leu Pro Lys Arg Arg Arg Ala Arg Val Gly Ser Pro Ser Gly Asp
 1 5 10 15

Ala Ala Ser Ser Thr Pro Pro Ser Thr Arg Phe Pro Gly Val Ala Ile
 20 25 30

Tyr Leu Val Glu Pro Arg Met Gly Arg Ser Arg Arg Ala Phe Leu Thr
 35 40 45

Arg Leu Thr Arg Ser Lys Gly Phe Arg Val Leu Asp Ala Cys Ser Ser
 50 55 60

Glu Ala Thr His Val Val Met Glu Glu Thr Ser Ala Glu Glu Ala Val
 65 70 75 80

Ser Trp Gln Glu Arg Arg Met Ala Ala Ala Pro Pro Gly Cys Thr Pro
 85 90 95

Pro Ala Leu Leu Asp Ile Ser Trp Leu Thr Glu Ser Leu Gly Ala Gly
 100 105 110

Gln Pro Val Pro Val Glu Cys Arg His Arg Leu Glu Val Ala Gly Pro
 115 120 125

Arg Lys Gly Pro Leu Ser Pro Ala Trp Met Pro Ala Tyr Val Cys Gln
 130 135 140

Arg Pro Thr Pro Leu Thr His His Asn Thr Gly Leu Ser Glu Ala Leu
 145 150 155 160

Glu Thr Leu Ala Glu Ala Ala Gly Phe Glu Gly Ser Glu Gly Arg Leu
 165 170 175

Leu Thr Phe Cys Arg Ala Ala Ser Val Leu Lys Ala Leu Pro Ser Pro
 180 185 190

Val Thr Thr Leu Ser Gln Leu Gln Gly Leu Pro His Phe Gly Glu His
 195 200 205

B2265FR00 séquence listing.txt

Ser Ser Arg Val Val Gln Glu Leu Leu Glu His Gly Val Cys Glu Glu
210 215 220

Val Glu Arg Val Gln Arg Ser Glu Arg Tyr Gln Thr Met Lys Leu Phe
225 230 235 240

Thr Gln Ile Phe Gly Val Gly Val Lys Thr Ala Asp Arg Trp Tyr Arg
245 250 255

Glu Gly Leu Arg Thr Leu Asp Asp Leu Arg Glu Gln Pro Gln Lys Leu
260 265 270

Thr Gln Gln Gln Lys Ala Gly Leu Gln His His Gln Asp Leu Ser Thr
275 280 285

Pro Val Leu Arg Ser Asp Val Asp Ala Leu Gln Gln Val Val Glu Glu
290 295 300

Ala Val Gly Gln Ala Leu Pro Gly Ala Thr Val Thr Leu Thr Gly Gly
305 310 315 320

Phe Arg Arg Gly Lys Leu Gln Gly His Asp Val Asp Phe Leu Ile Thr
325 330 335

His Pro Lys Glu Gly Gln Glu Ala Gly Leu Leu Pro Arg Val Met Cys
340 345 350

Arg Leu Gln Asp Gln Gly Leu Ile Leu Tyr His Gln His Gln His Ser
355 360 365

Cys Trp Glu Ser Pro Thr Arg Leu Ala Gln Gln Ser His Met Asp Ala
370 375 380

Phe Glu Arg Ser Phe Cys Ile Phe Arg Leu Pro Gln Pro Pro Gly Ala
385 390 395 400

B2265FR00 séquence listing.txt

Ala Val Gly Gly Ser Thr Arg Pro Cys Pro Ser Trp Lys Ala Val Arg
 405 410 415

Val Asp Leu Val Val Ala Pro Val Ser Gln Phe Pro Phe Ala Leu Leu
 420 425 430

Gly Trp Thr Gly Ser Lys Leu Phe Gln Arg Glu Leu Arg Arg Phe Ser
 435 440 445

Arg Lys Glu Lys Gly Leu Trp Leu Asn Ser His Gly Leu Phe Asp Pro
 450 455 460

Glu Gln Lys Thr Phe Phe Gln Ala Ala Ser Glu Glu Asp Ile Phe Arg
 465 470 475 480

His Leu Gly Leu Glu Tyr Leu Pro Pro Glu Gln Arg Asn Ala
 485 490

<210> 10

<211> 496

<212> PRT

<213> Mus musculus

<400> 10

Met Leu Pro Lys Arg Arg Arg Val Arg Ala Gly Ser Pro His Ser Ala
 1 5 10 15

Val Ala Ser Ser Thr Pro Pro Ser Val Val Arg Phe Pro Asp Val Ala
 20 25 30

Ile Tyr Leu Ala Glu Pro Arg Met Gly Arg Ser Arg Arg Ala Phe Leu
 35 40 45

Thr Arg Leu Ala Arg Ser Lys Gly Phe Arg Val Leu Asp Ala Tyr Ser
 50 55 60

Ser Lys Val Thr His Val Val Met Glu Gly Thr Ser Ala Lys Glu Ala
 65 70 75 80

B2265FR00 séquence listing.txt

Ile Cys Trp Gln Lys Asn Met Asp Ala Leu Pro Thr Gly Cys Pro Gln
 85 90 95

Pro Ala Leu Leu Asp Ile Ser Trp Phe Thr Glu Ser Met Ala Ala Gly
 100 105 110

Gln Pro Val Arg Glu Glu Gly Arg His His Leu Glu Val Ala Glu Pro
 115 120 125

Arg Lys Glu Pro Pro Val Ser Ala Ser Met Pro Ala Tyr Ala Cys Gln
 130 135 140

Arg Pro Ser Pro Leu Thr His His Asn Thr Leu Leu Ser Glu Ala Leu
 145 150 155 160

Glu Thr Leu Ala Glu Ala Ala Gly Phe Glu Ala Asn Glu Gly Arg Leu
 165 170 175

Leu Ser Phe Ser Arg Ala Asp Ser Val Leu Lys Ser Leu Pro Cys Pro
 180 185 190

Val Ala Ser Leu Ser Gln Leu His Gly Leu Pro Tyr Phe Gly Glu His
 195 200 205

Ser Thr Arg Val Ile Gln Glu Leu Leu Glu His Gly Thr Cys Glu Glu
 210 215 220

Val Lys Gln Val Arg Cys Ser Glu Arg Tyr Gln Thr Met Lys Leu Phe
 225 230 235 240

Thr Gln Val Phe Gly Val Gly Val Lys Thr Ala Asn Arg Trp Tyr Gln
 245 250 255

Glu Gly Leu Arg Thr Leu Asp Glu Leu Arg Glu Gln Pro Gln Arg Leu
 260 265 270

Thr Gln Gln Gln Lys Ala Gly Leu Gln Tyr Tyr Gln Asp Leu Ser Thr

B2265FR00 séquence listing.txt

275 280 285

Pro Val Arg Arg Ala Asp Ala Glu Ala Leu Gln Gln Leu Ile Glu Ala
290 295 300

Ala Val Arg Gln Thr Leu Pro Gly Ala Thr Val Thr Leu Thr Gly Gly
305 310 315 320

Phe Arg Arg Gly Lys Leu Gln Gly His Asp Val Asp Phe Leu Ile Thr
325 330 335

His Pro Glu Glu Gly Gln Glu Val Gly Leu Leu Pro Lys Val Met Ser
340 345 350

Cys Leu Gln Ser Gln Gly Leu Val Leu Tyr His Gln Tyr His Arg Ser
355 360 365

His Leu Ala Asp Ser Ala His Asn Leu Arg Gln Arg Ser Ser Thr Met
370 375 380

Asp Ala Phe Glu Arg Ser Phe Cys Ile Leu Gly Leu Pro Gln Pro Gln
385 390 395 400

Gln Ala Ala Leu Ala Gly Ala Leu Pro Pro Cys Pro Thr Trp Lys Ala
405 410 415

Val Arg Val Asp Leu Val Val Thr Pro Ser Ser Gln Phe Pro Phe Ala
420 425 430

Leu Leu Gly Trp Thr Gly Ser Gln Phe Phe Glu Arg Glu Leu Arg Arg
435 440 445

Phe Ser Arg Gln Glu Lys Gly Leu Trp Leu Asn Ser His Gly Leu Phe
450 455 460

Asp Pro Glu Gln Lys Arg Val Phe His Ala Thr Ser Glu Glu Asp Val
465 470 475 480

B2265FR00 séquence listing.txt

Phe Arg Leu Leu Gly Leu Lys Tyr Leu Pro Pro Glu Gln Arg Asn Ala
 485 490 495

<210> 11

<211> 509

<212> PRT

<213> Canis lupus

<400> 11

Met Asp Pro Leu Gln Met Ala His Ser Gly Pro Arg Lys Lys Arg Pro
 1 5 10 15

Arg Gln Met Gly Ala Pro Met Val Ser Pro Pro His Asn Ile Lys Phe
 20 25 30

Gln Asp Leu Val Leu Tyr Ile Leu Glu Lys Lys Met Gly Thr Thr Arg
 35 40 45

Arg Ala Phe Leu Met Glu Leu Ala Arg Arg Lys Gly Phe Arg Val Asp
 50 55 60

Asn Glu Phe Ser Asp Ser Ile Thr His Ile Val Ala Glu Asn Asn Ser
 65 70 75 80

Gly Ser Asp Val Leu Glu Trp Leu Gln Val Gln Asn Ile Lys Ala Ser
 85 90 95

Ser Gln Leu Glu Leu Leu Asp Ile Ser Trp Leu Ile Glu Ser Met Gly
 100 105 110

Ala Gly Lys Pro Val Glu Met Thr Gly Lys His Gln Leu Met Arg Arg
 115 120 125

Asp Tyr Thr Ala Ser Pro Asn Pro Glu Leu Gln Lys Thr Leu Pro Val
 130 135 140

Ala Val Lys Lys Ile Ser Gln Tyr Ala Cys Gln Arg Arg Thr Thr Leu
 145 150 155 160

B2265FR00 séquence listing.txt

Asn Asn Tyr Asn Asn Val Phe Thr Asp Ala Phe Glu Val Leu Ala Glu
 165 170 175

Asn Tyr Glu Phe Arg Glu Asn Glu Val Phe Ser Leu Thr Phe Met Arg
 180 185 190

Ala Ala Ser Val Leu Lys Ser Leu Pro Phe Thr Ile Ile Ser Met Lys
 195 200 205

Asp Thr Glu Gly Ile Pro Cys Leu Gly Asp Gln Val Lys Cys Ile Ile
 210 215 220

Glu Glu Ile Ile Glu Asp Gly Glu Ser Ser Glu Val Lys Ala Val Leu
 225 230 235 240

Asn Asp Glu Arg Tyr Gln Ser Phe Lys Leu Phe Thr Ser Val Phe Gly
 245 250 255

Val Gly Leu Lys Thr Ser Glu Lys Trp Phe Arg Met Gly Phe Arg Thr
 260 265 270

Leu Ser Lys Ile Lys Ser Asp Lys Ser Leu Lys Phe Thr Pro Met Gln
 275 280 285

Lys Ala Gly Phe Leu Tyr Tyr Glu Asp Leu Val Ser Cys Val Thr Arg
 290 295 300

Ala Glu Ala Glu Ala Val Gly Val Leu Val Lys Glu Ala Val Gly Ala
 305 310 315 320

Phe Leu Pro Asp Ala Phe Val Thr Met Thr Gly Gly Phe Arg Arg Gly
 325 330 335

Lys Lys Met Gly His Asp Val Asp Phe Leu Ile Thr Ser Pro Gly Ser
 340 345 350

B2265FR00 séquence listing.txt

Thr Asp Glu Asp Glu Glu Gln Leu Leu Pro Lys Val Ile Asn Leu Trp
 355 360 365

Glu Arg Lys Gly Leu Leu Leu Tyr Cys Asp Leu Val Glu Ser Thr Phe
 370 375 380

Glu Lys Leu Lys Leu Pro Ser Arg Lys Val Asp Ala Leu Asp His Phe
 385 390 395 400

Gln Lys Cys Phe Leu Ile Leu Lys Leu His His Gln Arg Val Asp Gly
 405 410 415

Gly Lys Cys Ser Gln Gln Glu Gly Lys Thr Trp Lys Ala Ile Arg Val
 420 425 430

Asp Leu Val Met Cys Pro Tyr Glu Arg Arg Ala Phe Ala Leu Leu Gly
 435 440 445

Trp Thr Gly Ser Arg Gln Phe Glu Arg Asp Leu Arg Arg Tyr Ala Ser
 450 455 460

His Glu Arg Lys Met Ile Leu Asp Asn His Ala Leu Tyr Asp Lys Thr
 465 470 475 480

Lys Lys Ile Phe Leu Lys Ala Glu Ser Glu Glu Glu Ile Phe Ala His
 485 490 495

Leu Gly Leu Asp Tyr Ile Glu Pro Trp Glu Arg Asn Ala
 500 505

<210> 12

<211> 506

<212> PRT

<213> Gallus gallus

<400> 12

Met Glu Arg Ile Arg Pro Pro Thr Val Val Ser Gln Arg Lys Arg Gln
 1 5 10 15

B2265FR00 séquence listing.txt

Lys Gly Met Tyr Ser Pro Lys Leu Ser Cys Gly Tyr Glu Ile Lys Phe
 20 25 30

Asn Lys Leu Val Ile Phe Ile Met Gln Arg Lys Met Gly Met Thr Arg
 35 40 45

Arg Thr Phe Leu Met Glu Leu Ala Arg Ser Lys Gly Phe Arg Val Glu
 50 55 60

Ser Glu Leu Ser Asp Ser Val Thr His Ile Val Ala Glu Asn Asn Ser
 65 70 75 80

Tyr Pro Glu Val Leu Asp Trp Leu Lys Gly Gln Ala Val Gly Asp Ser
 85 90 95

Ser Arg Phe Glu Ile Leu Asp Ile Ser Trp Leu Thr Ala Cys Met Glu
 100 105 110

Met Gly Arg Pro Val Asp Leu Glu Lys Lys Tyr His Leu Val Glu Gln
 115 120 125

Ala Gly Gln Tyr Pro Thr Leu Lys Thr Pro Glu Ser Glu Val Ser Ser
 130 135 140

Phe Thr Ala Ser Lys Val Ser Gln Tyr Ser Cys Gln Arg Lys Thr Thr
 145 150 155 160

Leu Asn Asn Cys Asn Lys Lys Phe Thr Asp Ala Phe Glu Ile Met Ala
 165 170 175

Glu Asn Tyr Glu Phe Lys Glu Asn Glu Ile Phe Cys Leu Glu Phe Leu
 180 185 190

Arg Ala Ala Ser Val Leu Lys Ser Leu Pro Phe Pro Val Thr Arg Met
 195 200 205

Lys Asp Ile Gln Gly Leu Pro Cys Met Gly Asp Arg Val Arg Asp Val

B2265FR00 séquence listing.txt

210 215 220

Ile Glu Glu Ile Ile Glu Glu Gly Glu Ser Ser Arg Ala Lys Asp Val
225 230 235 240

Leu Asn Asp Glu Arg Tyr Lys Ser Phe Lys Glu Phe Thr Ser Val Phe
 245 250 255

Gly Val Gly Val Lys Thr Ser Glu Lys Trp Phe Arg Met Gly Leu Arg
 260 265 270

Thr Val Glu Glu Val Lys Ala Asp Lys Thr Leu Lys Leu Ser Lys Met
 275 280 285

Gln Arg Ala Gly Phe Leu Tyr Tyr Glu Asp Leu Val Ser Cys Val Ser
 290 295 300

Lys Ala Glu Ala Asp Ala Val Ser Ser Ile Val Lys Asn Thr Val Cys
305 310 315 320

Thr Phe Leu Pro Asp Ala Leu Val Thr Ile Thr Gly Gly Phe Arg Arg
 325 330 335

Gly Lys Lys Ile Gly His Asp Ile Asp Phe Leu Ile Thr Ser Pro Gly
 340 345 350

Gln Arg Glu Asp Asp Glu Leu Leu His Lys Gly Leu Leu Leu Tyr Cys
 355 360 365

Asp Ile Ile Glu Ser Thr Phe Val Lys Glu Gln Ile Pro Ser Arg His
 370 375 380

Val Asp Ala Met Asp His Phe Gln Lys Cys Phe Ala Ile Leu Lys Leu
385 390 395 400

Tyr Gln Pro Arg Val Asp Asn Ser Ser Tyr Asn Met Ser Lys Lys Cys
 405 410 415

B2265FR00 séquence listing.txt

Asp Met Ala Glu Val Lys Asp Trp Lys Ala Ile Arg Val Asp Leu Val
 420 425 430

Ile Thr Pro Phe Glu Gln Tyr Ala Tyr Ala Leu Leu Gly Trp Thr Gly
 435 440 445

Ser Arg Gln Phe Gly Arg Asp Leu Arg Arg Tyr Ala Thr His Glu Arg
 450 455 460

Lys Met Met Leu Asp Asn His Ala Leu Tyr Asp Lys Arg Lys Arg Val
 465 470 475 480

Phe Leu Lys Ala Gly Ser Glu Glu Glu Ile Phe Ala His Leu Gly Leu
 485 490 495

Asp Tyr Val Glu Pro Trp Glu Arg Asn Ala
 500 505

<210> 13

<211> 509

<212> PRT

<213> Homo sapiens

<400> 13

Met Asp Pro Pro Arg Ala Ser His Leu Ser Pro Arg Lys Lys Arg Pro
 1 5 10 15

Arg Gln Thr Gly Ala Leu Met Ala Ser Ser Pro Gln Asp Ile Lys Phe
 20 25 30

Gln Asp Leu Val Val Phe Ile Leu Glu Lys Lys Met Gly Thr Thr Arg
 35 40 45

Arg Ala Phe Leu Met Glu Leu Ala Arg Arg Lys Gly Phe Arg Val Glu
 50 55 60

Asn Glu Leu Ser Asp Ser Val Thr His Ile Val Ala Glu Asn Asn Ser
 65 70 75 80

B2265FR00 séquence listing.txt

Gly Ser Asp Val Leu Glu Trp Leu Gln Ala Gln Lys Val Gln Val Ser
 85 90 95

Ser Gln Pro Glu Leu Leu Asp Val Ser Trp Leu Ile Glu Cys Ile Arg
 100 105 110

Ala Gly Lys Pro Val Glu Met Thr Gly Lys His Gln Leu Val Val Arg
 115 120 125

Arg Asp Tyr Ser Asp Ser Thr Asn Pro Gly Pro Pro Lys Thr Pro Pro
 130 135 140

Ile Ala Val Gln Lys Ile Ser Gln Tyr Ala Cys Gln Arg Arg Thr Thr
 145 150 155 160

Leu Asn Asn Cys Asn Gln Ile Phe Thr Asp Ala Phe Asp Ile Leu Ala
 165 170 175

Glu Asn Cys Glu Phe Arg Glu Asn Glu Asp Ser Cys Val Thr Phe Met
 180 185 190

Arg Ala Ala Ser Val Leu Lys Ser Leu Pro Phe Thr Ile Ile Ser Met
 195 200 205

Lys Asp Thr Glu Gly Ile Pro Cys Leu Gly Ser Lys Val Lys Gly Ile
 210 215 220

Ile Glu Glu Ile Ile Glu Asp Gly Glu Ser Ser Glu Val Lys Ala Val
 225 230 235 240

Leu Asn Asp Glu Arg Tyr Gln Ser Phe Lys Leu Phe Thr Ser Val Phe
 245 250 255

Gly Val Gly Leu Lys Thr Ser Glu Lys Trp Phe Arg Met Gly Phe Arg
 260 265 270

B2265FR00 séquence listing.txt

Thr Leu Ser Lys Val Arg Ser Asp Lys Ser Leu Lys Phe Thr Arg Met
 275 280 285

Gln Lys Ala Gly Phe Leu Tyr Tyr Glu Asp Leu Val Ser Cys Val Thr
 290 295 300

Arg Ala Glu Ala Glu Ala Val Ser Val Leu Val Lys Glu Ala Val Trp
 305 310 315 320

Ala Phe Leu Pro Asp Ala Phe Val Thr Met Thr Gly Gly Phe Arg Arg
 325 330 335

Gly Lys Lys Met Gly His Asp Val Asp Phe Leu Ile Thr Ser Pro Gly
 340 345 350

Ser Thr Glu Asp Glu Glu Gln Leu Leu Gln Lys Val Met Asn Leu Trp
 355 360 365

Glu Lys Lys Gly Leu Leu Leu Tyr Tyr Asp Leu Val Glu Ser Thr Phe
 370 375 380

Glu Lys Leu Arg Leu Pro Ser Arg Lys Val Asp Ala Leu Asp His Phe
 385 390 395 400

Gln Lys Cys Phe Leu Ile Phe Lys Leu Pro Arg Gln Arg Val Asp Ser
 405 410 415

Asp Gln Ser Ser Trp Gln Glu Gly Lys Thr Trp Lys Ala Ile Arg Val
 420 425 430

Asp Leu Val Leu Cys Pro Tyr Glu Arg Arg Ala Phe Ala Leu Leu Gly
 435 440 445

Trp Thr Gly Ser Arg Gln Phe Glu Arg Asp Leu Arg Arg Tyr Ala Thr
 450 455 460

His Glu Arg Lys Met Ile Leu Asp Asn His Ala Leu Tyr Asp Lys Thr
 465 470 475 480

B2265FR00 séquence listing.txt

Lys Arg Ile Phe Leu Lys Ala Glu Ser Glu Glu Glu Ile Phe Ala His
485 490 495

Leu Gly Leu Asp Tyr Ile Glu Pro Trp Glu Arg Asn Ala
500 505

<210> 14

<211> 14

<212> DNA

<213> artificial sequence

<220>

<223> amorce

<400> 14

aaaaaaaaa gggg

14



RAPPORT DE RECHERCHE PRÉLIMINAIRE PARTIEL

établi sur la base des dernières revendications
déposées avant le commencement de la recherche

voir FEUILLE(S) SUPPLÉMENTAIRE(S)

N° d'enregistrement
national

FA 829050
FR 1655475

DOCUMENTS CONSIDÉRÉS COMME PERTINENTS		Revendications concernées	Classement attribué à l'invention par l'INPI
Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes		
X	WO 2016/064880 A1 (MOLECULAR ASSEMBLIES INC [US]) 28 avril 2016 (2016-04-28) * page 1, ligne 22 - page 2, ligne 19 * * page 5, ligne 21 - ligne 24; figure 4 * * page 7, ligne 4 - page 10 * * page 16; tableau 9 * * page 22 - page 24; tableaux *	1-5,7, 13-22	C12N9/12 C12N15/54 C12N15/63 C12P19/34 C12Q1/68 C12N9/1264 C12P19/34 C12Y207/07031
A	Baoli Yangs ET AL: "Mutational Analysis of Residues in the Nucleotide Binding Domain of Human Terminal Deoxynucleotidyl Transferase*", 22 avril 1994 (1994-04-22), pages 11859-11868, XP055345779, Extrait de l'Internet: URL: http://www.jbc.org/content/269/16/11859.full.pdf [extrait le 2017-02-14] * le document en entier *	1-5,7, 13-22	
A	M. DELARUE ET AL: "Crystal structures of a template-independent DNA polymerase: murine terminal deoxynucleotidyltransferase", EMBO JOURNAL, vol. 21, no. 3, 1 février 2002 (2002-02-01), pages 427-439, XP002767368, DOI: 10.1093/emboj/21.3.427 * page 428, colonne de droite; figure 1 * * page 430, colonne de gauche - page 431; figure 2 * * page 432, colonne de droite *	1-5,7, 13-22	DOMAINES TECHNIQUES RECHERCHÉS (IPC) C12N C12P
		-/--	
Date d'achèvement de la recherche		Examineur	
17 février 2017		Le Cornec, Nadine	
<p>CATÉGORIE DES DOCUMENTS CITES</p> <p>X : particulièrement pertinent à lui seul Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie A : arrière-plan technologique O : divulgation non-écrite P : document intercalaire</p> <p>T : théorie ou principe à la base de l'invention E : document de brevet bénéficiant d'une date antérieure à la date de dépôt et qui n'a été publié qu'à cette date de dépôt ou qu'à une date postérieure. D : cité dans la demande L : cité pour d'autres raisons & : membre de la même famille, document correspondant</p>			



**RAPPORT DE RECHERCHE
PRÉLIMINAIRE PARTIEL**
établi sur la base des dernières revendications
déposées avant le commencement de la recherche
voir FEUILLE(S) SUPPLÉMENTAIRE(S)

N° d'enregistrement
national

FA 829050
FR 1655475

DOCUMENTS CONSIDÉRÉS COMME PERTINENTS		Revendications concernées	Classement attribué à l'invention par l'INPI
Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes		
X	<p>DATABASE Protein [Online]</p> <p>26 juin 2015 (2015-06-26), "predicted: DNA nucleotidylexotransferase isoform X1", XP002767369, extrait de NCBI Database accession no. XP_012873149 * le document en entier * -----</p>	1-5,7,13	<p>DOMAINES TECHNIQUES RECHERCHÉS (IPC)</p>
Date d'achèvement de la recherche		Examineur	
17 février 2017		Le Cornec, Nadine	
<p>CATÉGORIE DES DOCUMENTS CITES</p> <p>X : particulièrement pertinent à lui seul Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie A : arrière-plan technologique O : divulgation non-écrite P : document intercalaire</p> <p>T : théorie ou principe à la base de l'invention E : document de brevet bénéficiant d'une date antérieure à la date de dépôt et qui n'a été publié qu'à cette date de dépôt ou qu'à une date postérieure. D : cité dans la demande L : cité pour d'autres raisons & : membre de la même famille, document correspondant</p>			

**ABSENCE D'UNITÉ D'INVENTION
FEUILLE SUPPLÉMENTAIRE B**

Numéro de la demande

FA 829050
FR 1655475

La division de la recherche estime que la présente demande de brevet ne satisfait pas à l'exigence relative à l'unité d'invention et concerne plusieurs inventions ou pluralités d'inventions, à savoir :

1. revendications: 1-5, 7, 13-22(toutes en partie)

Variant d'une ADN polymérase de la famille polX capable de synthétiser une molécule d'acide nucléique sans brin matrice comprenant au moins une mutation d'un résidu à la position M330, position déterminée par rapport à la séquence ID no.1. acide nucléique, cassette d'expression, vecteur, cellule hôte, utilisation du variant , procédé de production, procédé de synthèse enzymatique , kit pour la synthèse enzymatique

2-35. revendications: 1-22(en partie)

la même chose que pour le sujet no.1 mais pour les résidus aux positions suivantes: T331, G332, G333, F334, K338, H342, D343, V344, D345, F346, A397, D399, D434, V436, A446, L447, L448, G449, W450, G452, Q455, F456, E457, R458, R461, N474, E491, D501, Y502, I503, P505, R508, N509, A510

La première invention a été recherchée.

D1 poursuit le même but que celui de la présente demande. Il cherche à obtenir une désoxyribonucleotidyl transférase terminale (TdT) qui peut fonctionner en l'absence de brin matrice et pouvant utiliser des nucléotides modifiés.

La TdT est une ADN polymérase appartenant à la famille des ADN polymérases polX. D1 (page 7 ligne 5 -page 9, page 16 SEQ ID no.9) divulgue la structure tridimensionnelle de cette enzyme notamment de son site actif et à partir de sa structure, explore les différentes substitutions à des positions spécifiques qui ont un rôle à jouer dans l'interaction avec le nucléotide modifié en 3'-OH et aussi qui ont le moins d'interférence possible avec la modification en position 3'-OH car elle occupe un plus grand volume au sein de l'enzyme. Ces modifications stériques sont analysées, déterminées grâce à une modélisation informatique afin d'obtenir une enzyme active qui puisse incorporer des nucléotides modifiés en 3'-OH.

Les positions au niveau et aux alentours des motifs GGFR (332-336) et TGS (451-454) ainsi que les résidus à l'extrémité C-terminale (508-510) sont dévoilés comme positions préférées.

Le concept de muter une ADN polymérase de la famille polX, plus particulièrement la désoxyribonucleotidyl transférase terminale (TdT) capable de synthétiser une molécule d'acide nucléique sans brin matrice et capable d'incorporer des nucléotides modifiés en 3'-OH n'est pas nouveau au vu de D1 .

De plus, D2 divulgue aussi des mutants de la TdT humaine ayant une substitution aux positions R336, D343 et D345. Ces mutants sont compris dans la portée de la revendication 1.

Au vu de D1, le problème est la mise à disposition d'autres enzymes TdT mutées pouvant incorporer des nucléotides modifiées en position 3'-OH.

**ABSENCE D'UNITÉ D'INVENTION
FEUILLE SUPPLÉMENTAIRE B**

Numéro de la demande

FA 829050

FR 1655475

La division de la recherche estime que la présente demande de brevet ne satisfait pas à l'exigence relative à l'unité d'invention et concerne plusieurs inventions ou pluralités d'inventions, à savoir :

Les solutions au problème posé correspondent aux différentes mutations de la TdT aux positions indiquées dans la revendication 1:

Solution 1:

Variant d'une ADN polymérase de la famille polX (TdT) ayant au moins une mutation à la position M330, position déterminée par alignement avec la séquence ID no.1.

Solutions 2 à 35:

Même chose que pour la solution 1 mais pour les résidus aux positions suivantes T331, G332, G333, F334, K338, H342, D343, V344, D345, F346, A397, D399, D434, V436, A446, L447, L448, G449, W450, G452, Q455, F456, E457, R458, R461, N474, E491, D501, Y502, I503, P505, R508, N509, A510

Au vu de la différence entre les différents mutants et vu qu'on ne peut distinguer aucune autre caractéristique, qui, à la lumière de l'art antérieur pourrait être considérée comme caractéristique technique particulière, il n'y a pas de concept inventif commun reliant les multiples inventions revendiquées dans la présente demande. C'est pourquoi il y a un manque d'unité et les différentes inventions, n'appartenant pas à un concept inventif commun sont formulées comme les différentes solutions ci-dessus.

ANNEXE AU RAPPORT DE RECHERCHE PRÉLIMINAIRE**RELATIF A LA DEMANDE DE BREVET FRANÇAIS NO. FR 1655475 FA 829050**

La présente annexe indique les membres de la famille de brevets relatifs aux documents brevets cités dans le rapport de recherche préliminaire visé ci-dessus.

Les dits membres sont contenus au fichier informatique de l'Office européen des brevets à la date du **17-02-2017**

Les renseignements fournis sont donnés à titre indicatif et n'engagent pas la responsabilité de l'Office européen des brevets, ni de l'Administration française

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
WO 2016064880 A1	28-04-2016	US 2016108382 A1	21-04-2016
		WO 2016064880 A1	28-04-2016
