

LA RÉACTION DE POLYMÉRISATION EN CHAÎNE (PCR)

PRINCIPE ET APPLICATIONS

TABLE DES MATIÈRES

LA RÉPLICATION DE L'ADN	4
Introduction	4
La réplication	4
1- Enchaînement des nucléotides	5
2- Changement de forme de l'ADN	5
3- Initiation de la réplication	6
4- Terminaison du brin d'ADN	8
5- Correction immédiate des erreurs.....	8
6- Chez les procaryotes	8
7- Chez les eucaryotes	9
La réplication chez les procaryotes	9
La réplication chez les eucaryotes	12
Les ADN polymérases eucaryotes	13
La polymérase alpha	13
La polymérase bêta	14
La polymérase gamma	14
La polymérase delta	14
La polymérase epsilon	14
La réplication des extrémités d'ADN linéaire	15
Correction lors de la réplication	15
Problème de la structure de la chromatine	16
Système de sauvegarde	16
Altérations d'origine physique	16

Altérations d'origine chimique	16
Systèmes multiples de sauvegarde	17
LA RÉACTION DE POLYMÉRISATION EN CHAÎNE (PCR)	18
But et principe	19
Le mélange réactionnel et les cycles de température	19
La dénaturation	19
L'hybridation	19
Les amorces	20
L'élongation	21
La Taq polymérase	22
Les conditions réactionnelles	22
Détection et analyse des produits PCR	24
Applications	25
Le clonage acellulaire	25
La RT-PCR (reverse-transcriptase PCR)	26
La PCR quantitative en temps réel	27
La PCR semi-quantitative ou compétitive	29
La PCR appliquée au diagnostic	30
Maladies génétiques	31
Maladies infectieuses	33
La PCR appliquée à l'identification	34

La réplication de l'ADN

Introduction

La molécule d'ADN doit sa pérennité et la transmission de l'information qu'elle contient au fil des divisions cellulaires à un mécanisme hautement complexe et très reproductible qu'est la réplication (ou duplication). Ce phénomène universel est très spécifique à la fois chez les organismes eucaryotes et procaryotes. Il assure une fidèle duplication du support de l'information génétique qu'est la molécule d'ADN.

Ce mécanisme hautement complexe implique une machinerie enzymatique très spécifique indispensable au processus de la vie des cellules. Cependant, au cours de la vie de la cellule, l'ADN est soumis à de nombreuses modifications comme des lésions et des mutations (deux phénomènes distincts) qui modifient la fidélité de réplication.

Pour pallier à ces perturbations, les cellules possèdent en complément des systèmes de réparation qui permettent la plupart du temps de rétablir l'information génétique originelle lors de la réplication.

Plusieurs chapitres vont distinguer les modes de la transmission de l'information génétique :

- Le mécanisme moléculaire de la réplication ;
- La réplication chez les procaryotes ;
- La réplication chez les eucaryotes ;
- Le système de réparation.

La réplication

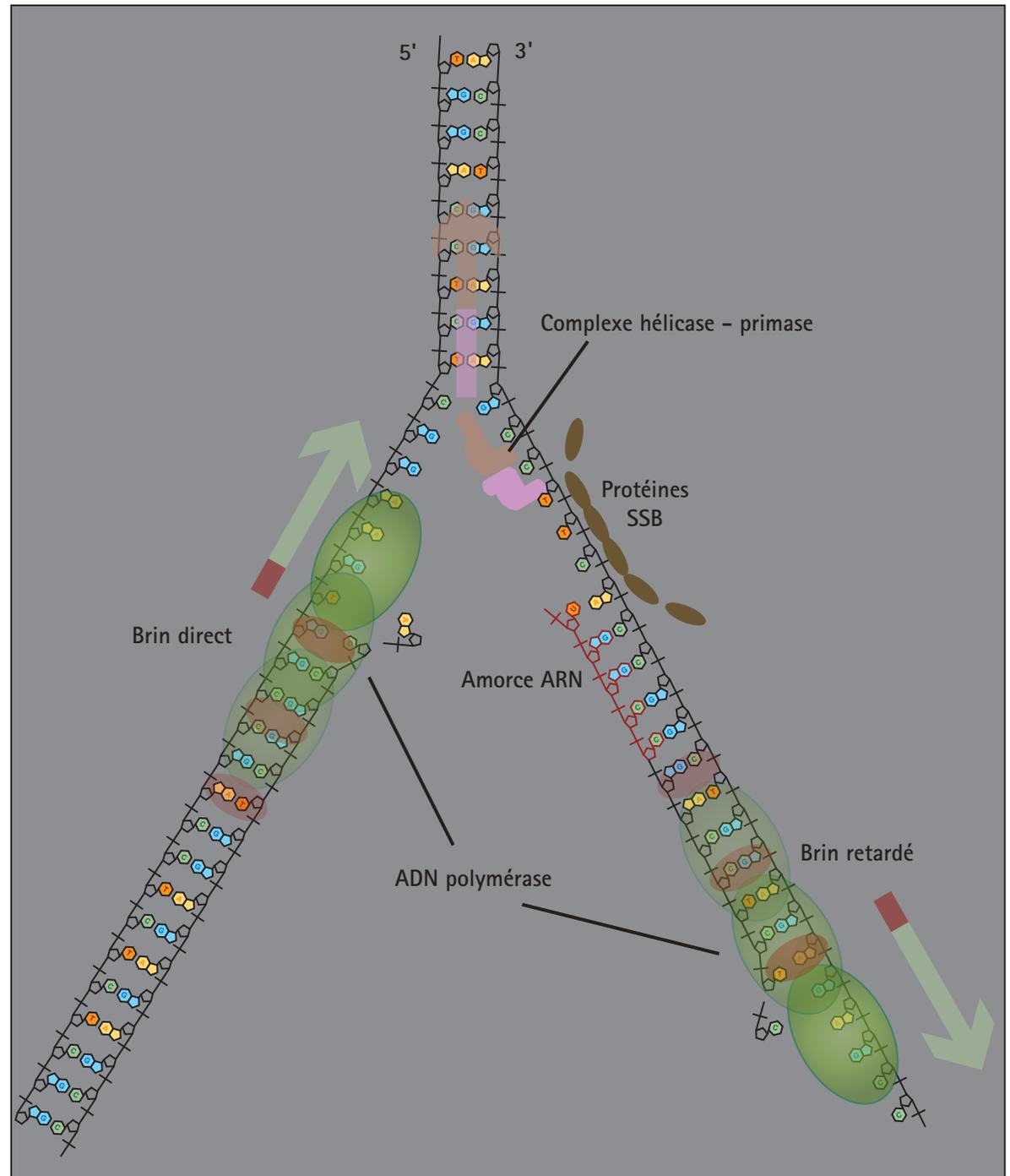
Le mécanisme de réplication est effectué selon un principe semi-conservatif qui assure une fidèle duplication de l'information génétique. Cette duplication semi-conservative repose sur le fait que chaque brin d'ADN sert de matrice pour la synthèse de chaque nouveau fragment et permet de conserver les séquences d'origine. Ce principe aboutit à deux molécules d'ADN identiques à la molécule initiale.

1- Enchaînement des nucléotides

La synthèse d'un brin d'ADN s'effectue par l'enchaînement de désoxyribonucléotides (dNTP). Les ADN polymérases n'effectuent jamais de synthèse ex nihilo, c'est-à-dire qu'elles ne peuvent réaliser une synthèse d'ADN en absence d'ADN servant de matrice (ADN parental). Ces enzymes catalysent la formation de liaisons phosphodiester qui impliquent le phosphate 5' (alpha) d'un dNTP libre et l'hydroxyle 3' libre d'un brin d'ADN en formation. Cet enchaînement s'effectue en respectant la règle de complémentarité des bases. Cette règle veut que le dNTP additionné soit le complémentaire du dNTP en vis-à-vis sur le brin matrice. Cela sous-entend que la polymérase est une enzyme de réplication qui a pour substrat un fragment d'ADN et catalyse la polymérisation de l'ADN monocaténaire complémentaire en formant des liaisons phosphodiester dans le respect de la règle de complémentarité des bases (A-T ; G-C).

2 - Changement de forme de l'ADN

La molécule d'ADN doit être accessible aux enzymes chargées de sa réplication. Or l'ADN en son état naturel est surenroulé. Par conséquent, cet état doit être modifié pour laisser l'accès aux enzymes responsables de la réplication. Les enzymes qui assurent ces modifications de l'état d'enroulement de l'ADN sont les topoisomérases. Elles déstructurent et restructurent la forme en isomérisant



l'ADN. On distingue les topoisomérases I et II avec, bien entendu, des fonctions spécifiques.

La topoisomérase de type I ne coupe que l'un des deux brins de l'ADN en se liant par une des tyrosines au phosphate en 5' libre de l'ADN coupé. La molécule d'ADN peut ensuite librement se dérouler et l'énergie de la liaison phosphodiester libérée est transférée lors de la ligation de l'ADN qui suit le déroulement, ce qui explique que l'énergie sous forme d'ATP n'est pas toujours nécessaire.

La topoisomérase de type II agit spécifiquement en coupant les deux brins de l'ADN. Contrairement à la précédente, elle est capable, chez les procaryotes, de créer ou de retirer des supertours en consommant de l'énergie. L'une des premières décrites est la gyrase d'E coli, enzyme très active car il lui est possible de faire ou défaire environ 100 tours par minute. Sensible à la novobicine, cette enzyme est moins bien connue chez les eucaryotes que chez les procaryotes. Les topoisomérases ne sont pas capables de produire les supertours observés lors de la formation des nucléosomes. En réalité elles jouent un rôle de relâchement dans les boucles pour favoriser la transcription.

Il est donc important de dérouler l'ADN et de le maintenir sous forme simple brin. Les deux brins sont séparés au moyen d'enzymes que sont les hélicases (ou déroulases) qui s'associent sur les brins d'ADN, coupent, déroulent et réassocient. Ces activités sont énergie-dépendantes et utilisent de l'ATP.

Il existe plusieurs types d'hélicases qui se fixent soit sur le brin 3'-5' soit sur le brin 5'-3' (hélicases II et III). Les brins d'ADN séparés sont stabilisés sous forme simple brin au moyen de protéines, les SSBP (single strand binding protein). Leur structure, sous forme de tétramère de 74 kDa, favorise la fixation sur la molécule d'ADN. L'association des tétramères sur le fragment d'ADN est régie par un phénomène coopératif où la fixation du premier tétramère favorise la liaison du deuxième etc..., ce qui finit par constituer un manchon rigide de SSB et empêche la réassociation des brins d'ADN.

3 - Initiation de la réplication

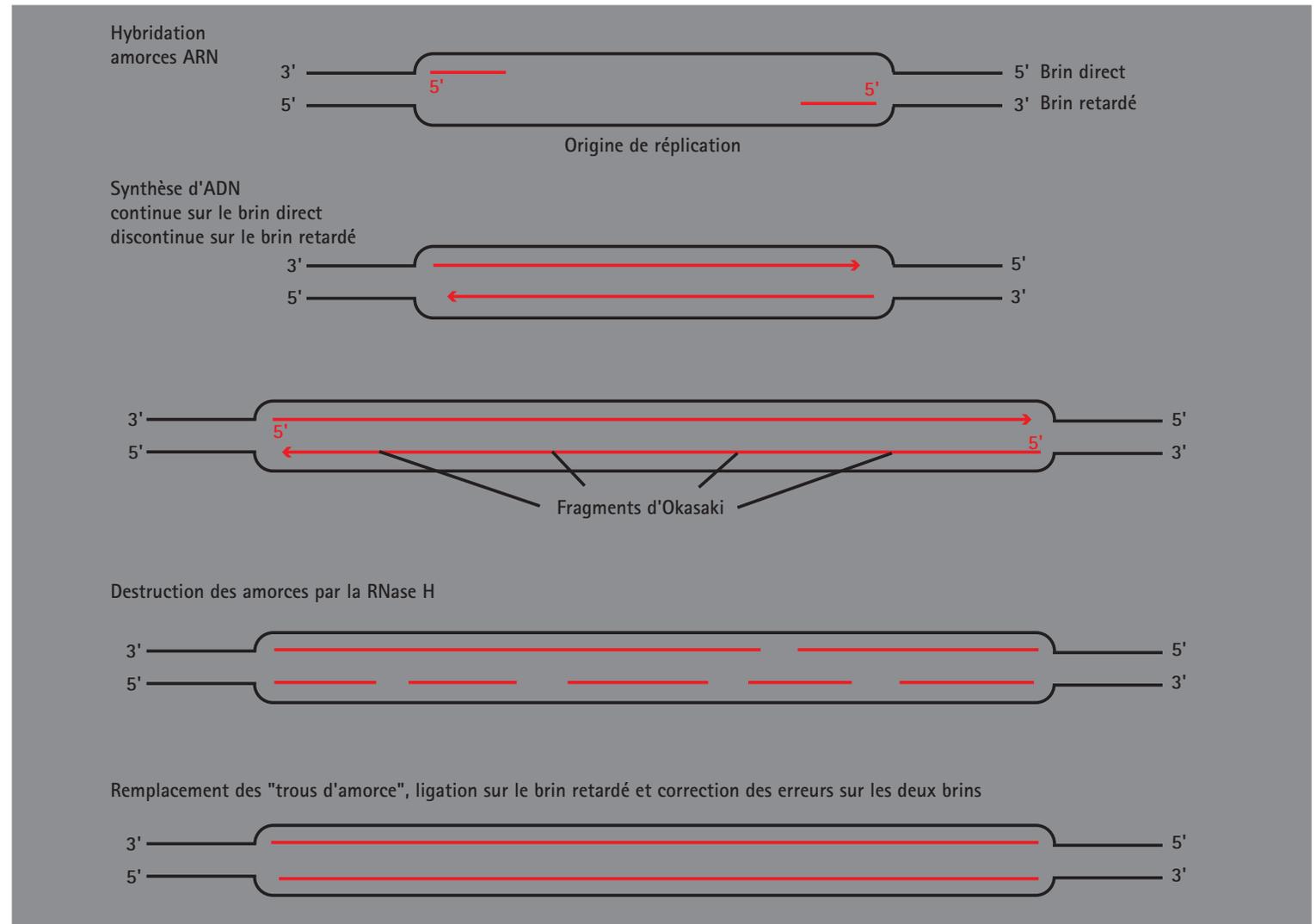
La réplication démarre à l'intérieur de la molécule d'ADN en un point qui est l'origine de la réplication et s'effectue dans les deux directions de l'orientation de l'ADN bicaténaire. Chez les virus et les bactéries, on distingue sur l'ADN une seule origine de réplication alors que, chez les eucaryotes, on distingue plusieurs origines. Cette multiplicité des origines de la réplication fait que la duplication de l'ADN eucaryote est bien plus rapide que celle de l'ADN procaryote, même si la vitesse de déplacement de la fourche de réplication est plus lente (inférieure à 5 000 pb /min pour les eucaryotes et de l'ordre de 100 000 pb/min pour les procaryotes).

Il est capital de noter que l'ADN polymérase doit l'amorce de son activité catalytique à une structure d'ADN obligatoirement double brin. L'initiation de la réplication nécessite cette structure double brin qui se manifeste par la synthèse et la liaison sur l'ADN monocaténaire d'une amorce d'ARN ou primer. La synthèse n'est effective qu'après l'élongation de cette amorce, ce qui fut démontré par Okasaki. Ce petit ARN est synthétisé par un complexe, le primosome. C'est cet élément qui détermine l'origine de l'amorce. Le primosome est constitué de trois types de protéines.

En premier lieu, une ADN polymérase dépendante appelée primase et codée par le gène dnaG.

On distingue ensuite deux protéines

qui s'associent à la primase pour former un complexe capable de synthétiser de l'ARN (les gènes responsables sont les dnaB et dnaC). Enfin, il existe des protéines spécifiques θ , η , θ' , η' qui assurent la reconnaissance du site où doit être synthétisée l'amorce, par conséquent leur fixation à l'ADN donne une structure qui permet la fixation des protéines issues de gènes dnaB et dnaC et de la primase. Chez les phages, ce mécanisme est bien décrit : lors du déroulement de l'ADN par les hélicases, les protéines SSB ne se fixent pas sur une petite portion de l'ADN. Du fait de



la présence complémentaire cette portion d'ADN va prendre une structure type hair-pin (épingle à cheveux). C'est cette structure qui favorise la reconnaissance du primosome et initie la synthèse.

Par conséquent, sur le brin matrice une séquence nucléotidique complémentaire courte (10 à 40 nucléotides) s'hybride et initie le mécanisme de réplication. Cette courte séquence de nucléotides indispensable est synthétisée par une ARN polymérase appelée primase.

La polymérase peut se détacher de l'ADN puis se réassocier à nouveau pour continuer la réplication. La polymérase est dite processive quand elle reste longtemps fixée à l'ADN.

4- Terminaison du brin d'ADN

Les amorces d'ARN qui initient la réplication sont dégradées par une enzyme spécifique qui hydrolyse l'ARN sur les hybrides ARN-ADN : la Rnase H. La lacune engendrée par cette action est comblée par l'action de l'ADN polymérase I. La soudure du brin d'ADN est effectuée par une enzyme appelée ligase qui génère des liaisons phosphodiester entre le 3'OH et les nucléosides triphosphates.

5- Correction immédiate des erreurs

Le taux d'erreur lors de la polymérisation est d'environ 1/10 000, cependant l'erreur observée sur l'ADN néo-synthétisé est de 1/108. Cette différence est expliquée par la reconnaissance des erreurs de l'ADN Polymérase et la réparation immédiate. Cette action de « contrôle-correction » dénommée proofreading est due à une activité 3'-exonucléase spécifique de ces polymérases. Le processus supposé est le suivant :

La polymérase en action entoure l'ADN, si la base ajoutée n'est pas adéquate la structure spatiale de l'ADN est modifiée ce qui permet de laisser agir l'activité exonucléasique 3'. Ce mécanisme extrait la base incorrectement incorporée ce qui permet le redémarrage de la polymérase avec l'incorporation du bon nucléotide.

6- Chez les procaryotes

Trois polymérases sont décrites : I, II et III, elles sont toutes trois indispensables pour la réplication. Elles possèdent toutes une activité de polymérisation en 5'>3' mais également une activité exonucléasique en 3'>5'.

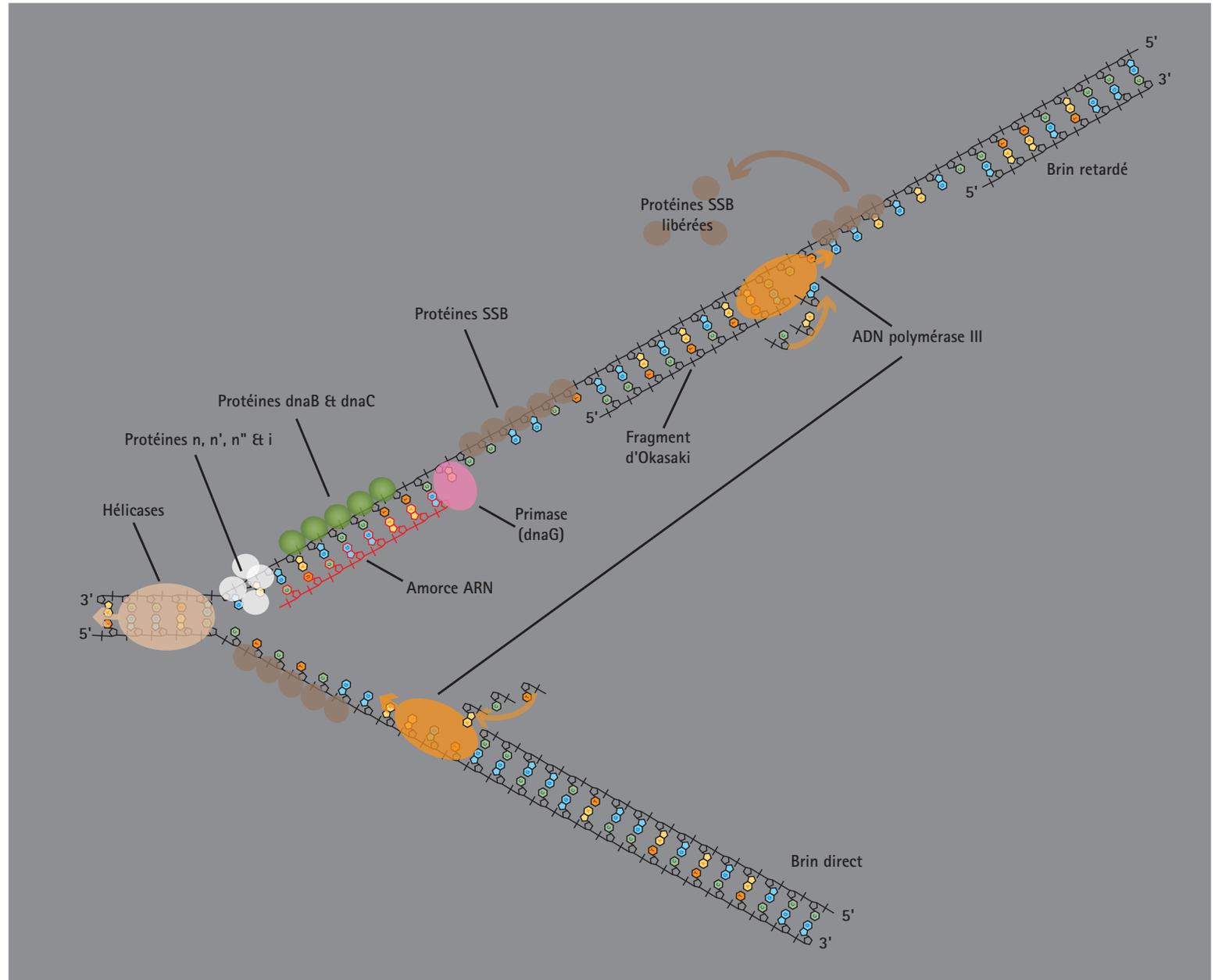
Pour l'ADN polymérase I, on distingue en plus une activité exonucléasique en 5'>3' ce qui correspond à la dégradation de l'ARN amorce au moment de la réplication.

7- Chez les eucaryotes supérieurs, les mammifères

On distingue 5 ADN polymérases nommées alpha, bêta, gamma, delta et epsilon. Elles assurent l'activité de polymérisation 5'>3', les formes gamma, delta et epsilon assurent aussi une activité exonucléase 3'>5'. Toutes les polymérases sont nucléaires seule la gamma est mitochondriale.

La réplication chez les procaryotes

Ce mécanisme complexe est initié par la synthèse préalable d'une amorce qui est à l'origine de la transcription des séquences d'ADN par la primase. Cette initiation nécessite de nombreuses protéines spécifiques qui constituent un complexe dénommé replisome.



Au moment de l'initiation, la protéine Dna A d'E. coli se fixe sur une séquence nucléotidique quasi constante sur les ADN procaryotes issus de diverses origines. Elle induit la séparation localisée et spécifique des deux brins d'ADN et définit l'origine de réplication.

Le déroulement très progressif de la molécule d'ADN s'effectue en présence d'énergie, sous forme d'ATP, au moyen de protéines spécifiques appelées hélicases (DnaB, PriA). La séparation des deux brins s'effectue dans les deux sens, ce qui entraîne des tensions largement suffisantes pour provoquer le surenroulement de la double hélice parentale non encore répliquée. Le surenroulement est amoindri par les topoisomérases. Elles opèrent en réalisant des coupures au niveau des liaisons phosphodiester sans pour autant laisser libre les extrémités 3' et 5' du brin coupé. Elles s'attachent ensuite à ces extrémités et effectuent soit des enroulements positifs (topoisomérase I) soit des enroulements négatifs (topoisomérase II) dans les entrecroisements des brins d'ADN. Les topoisomérases I et II équilibrent leur effet pour assurer un degré convenable de surenroulement de la molécule d'ADN.

Le démarrage de la réplication nécessite donc la séparation des deux brins parentaux ou matrice et la formation de deux fourches de réplication qui s'éloignent au gré de la réplication. Ainsi chaque initiation constitue une unité de réplication définie comme un réplicon.

La polymérisation s'opère uniquement dans le sens 5'>3', dans la mesure où les brins d'ADN sont antiparallèles la progression sera différente en fonction de l'orientation du brin.

Le brin avancé ou principal s'allongera de façon continue dans le sens 5'>3' ce qui correspond au déplacement de la fourche de réplication. L'autre brin, secondaire ou retardé, se forme de façon discontinue par petits fragments d'ADN dont la synthèse est systématiquement amorcée en 5'>3', donc en orientation inverse du déplacement de la fourche de réplication.

Ces fragments d'ADN intermédiaires, nommés fragments d'Okazaki, ont une taille comprise entre 1000 et 2000 chez les procaryotes (100 et 200 nucléotides chez les mammifères).

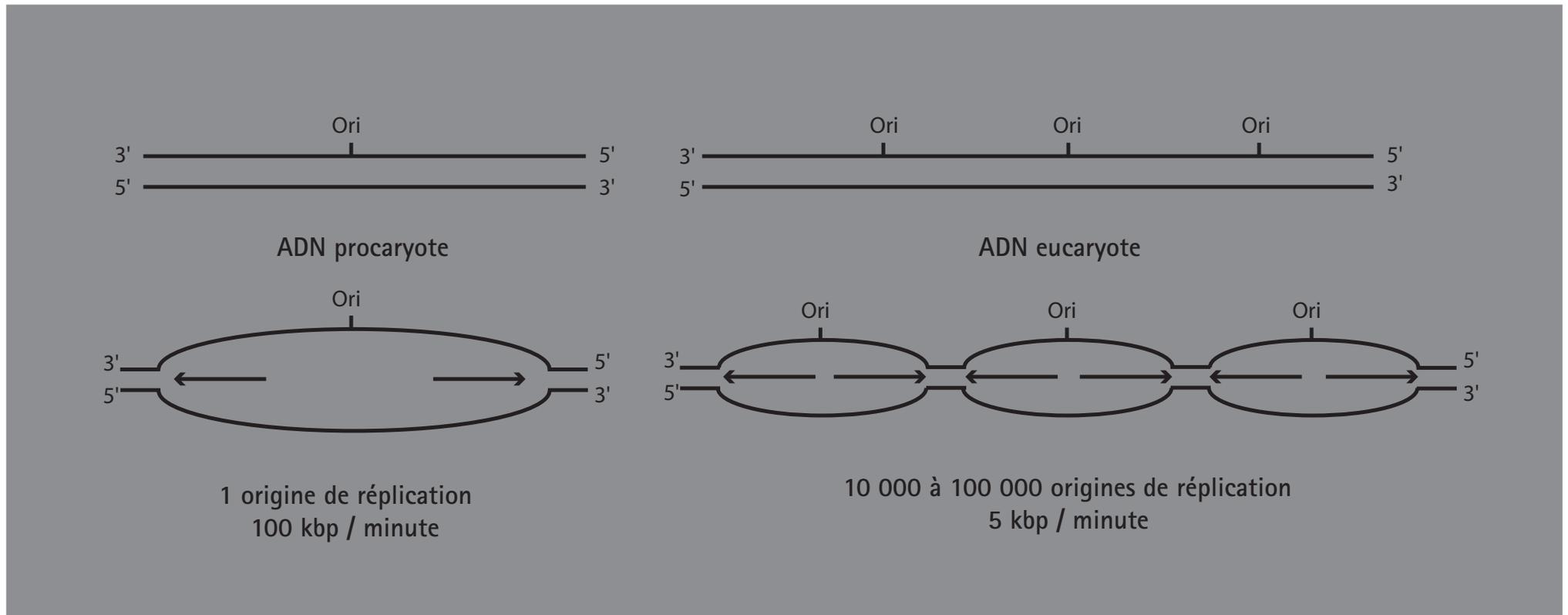
Le brin principal n'est amorcé qu'une seule fois et le brin retardé par contre est amorcé au début de chaque fragment d'Okazaki. En fonction de la procession du complexe de réplication, le primosome se déplace et synthétise dans le sens 3'>5' de petits oligonucléotides d'ARN qui sont

allongés par l'ADN polymérase III en petites séquences d'ADN.

L'ADN polymérase III synthétise un brin complémentaire à partir de l'extrémité 3'OH libre de l'amorce RNA en utilisant l'ADN comme matrice dans le sens 5'>3', ce qui disperse les protéines SSB fixées sur ce brin. Les deux brins sont par conséquent synthétisés en 5'>3' ; mais la synthèse ne peut pas être effectuée de façon continue sur le brin direct orienté 3'>5'.

L'ADN polymérase I fut la première polymérase mise en évidence. Elle a été isolée par Kornberg chez E. Coli en 1960. Elle était alors considérée comme l'enzyme de la réplication. En réalité une étude sur des mutants a montré qu'elle était en fait impliquée dans les mécanismes de réparation. La véritable réplicase est l'ADN polymérase III.

L'ADN polymérase III est constitué de 7 sous-unités codées par des gènes de structures différentes. Il semble que les polymérases soient



associées au niveau du point de réplication, chacune répliquant son brin, ce qui implique un repliement des brins sur eux-mêmes. L'enzyme, sur le brin matrice ou brin direct orienté 3'>5', synthétise en continu au fur et à mesure du déroulement des brins d'ADN. L'enzyme située sur l'autre brin (brin retardé) synthétise de l'ADN sous forme discontinue au moyen de petits fragments (1000 à 2000 pb) nommés les fragments d'Okasaki. L'ADN à répliquer est ensuite ouvert par les hélicases, une amorce est synthétisée sur les fragments 5'>3', les synthèses reprennent et ce processus se répète jusqu'à la réplication totale du brin d'ADN.

Lorsque l'ADN polymérase III arrive sur l'extrémité 5' de l'amorce précédente, elle est remplacée par l'ADN polymérase I qui agit en qualité d'exonucléase en dégradant l'ARN de 5' en 3'. Son activité polymérase remplace ensuite l'amorce d'ARN par de l'ADN. L'ensemble de ces fragments d'ADN sur le brin retardé sont associés par une enzyme qui rétablit la liaison phosphodiester : l'ADN ligase.

La formation de la première fourche de réplication est définie par le brin principal qui se forme à l'origine de la réplication. Sur le brin retardé, un fragment d'Okazaki peut dépasser l'origine de réplication, il devient par conséquent le brin principal de la deuxième fourche de réplication.

De part et d'autre de la fourche de réplication les deux fourches évoluent en sens inverse avec pour chacune un brin principal et un brin retardé. Il apparaît donc que le brin d'ADN est dupliqué de façon continue sur le brin 5'>3' et discontinue sur le brin 3'>5'.

Dès que les simples brins se forment, ils sont très rapidement liés aux protéines spécifiques SSBP (single strand binding protein, protéines sans activité enzymatique décrite) qui stabilisent les deux brins d'ADN sous forme monocaténaire indispensable à leur fonction de matrice. Les SSBP possèdent également un rôle protecteur de l'ADN monocaténaire contre les nucléases.

Bilan : La croissance des brins est différente en fonction de l'orientation. Elle est continue sur un brin et discontinue sur l'autre, c'est fonction de l'orientation des brins.

La réplication chez les eucaryotes

La réplication chez les eucaryotes est sensiblement identique à celle des procaryotes, la majeure distinction se fait essentiellement sur la fréquence, en effet le temps de régénération chez les eucaryotes est nettement plus long (cycles cellulaires et taille des génomes).

De plus l'organisation de l'ADN chez les eucaryotes est nettement plus complexe compte tenu des protéines (histones) qui s'y associent pour former la chromatine. Cette complexité motive les recherches avancées sur les mécanismes de contrôle de la réplication.

La réplication est bidirectionnelle et s'effectue généralement à partir de plusieurs origines. Le brin principal est synthétisé par un complexe composé par l'ADN polymérase et la deltaprotéine pCNA (proliferating cell nuclear antigen). Sur le brin secondaire, les amorces d'ARN et les fragments d'Okazaki sont l'œuvre d'une primase et de l'ADN polymérase alpha ; l'élimination des amorces est assurée par la RNase H et ce sont les polymérases alpha et bêta qui assurent le remplacement de l'ADN.

Les variations entre mécanismes de réplication procaryote et eucaryote concernent essentiellement les polymérases.

Les ADN polymérases eucaryotes

L'étude de ces enzymes a été complexe dans la mesure où elles sont peu abondantes, très sensibles à la protéolyse et leur fonctionnement nécessite des protéines associées.

La connaissance des mécanismes de réplication a été essentiellement permise par la mise au point d'un système de réplication in vitro au moyen du génome du virus SV 40 et le clonage de l'ADNc de quelques polymérases identifiées. Cependant les études n'ont permis d'éclaircir que le rôle des polymérases delta et epsilon, rôle qui n'est pas identique chez les levures et chez les mammifères.

La polymérase alpha (la primase)

Elle est constituée de 4 sous-unités dont les masses moléculaires sont respectivement de 165, 70, 58 et 48 kDa. La plus grosse sous-unité possède l'activité catalytique et a été clonée chez l'homme et la levure ; elle est constituée d'un polypeptide 165 kDa. Son gène est localisé sur le bras court du chromosome X, en Xp 21.3 – Xp22.1. Dans les cellules, cette sous-unité présente plusieurs masses moléculaires différentes comprises entre 120 et 180 kDa. Les formes les plus légères correspondent à des formes issues de modifications post-traductionnelles. Les sous-unités de 45 et 58 kDa possèdent une activité de type primase. Le rôle de la sous-unité de 70 kDa est encore mal connu. Cette polymérase a été nommée polymérase - alpha-primase.

Cette enzyme ne possède pas d'activité exonucléasique 3'>5' (excepté chez la drosophile), par conséquent elle n'est pas capable d'assurer des épreuves de correction (proofreading). Elle ne peut donc pas assurer à elle seule la réplication de l'ADN chez les eucaryotes. Son rôle sert à

synthétiser les amorces indispensables à la réplication sur le brin retardé par le biais de son activité de type primase.

La polymérase bêta

Elle est constituée de 335 acides aminés (39 kDa). Sa localisation est essentiellement nucléaire et elle est codée par un gène situé sur le bras court du chromosome 8. Cette polymérase possède un double rôle qui consiste à assurer à la fois la synthèse et la réparation de l'ADN amorcé. Elle possède aussi une fonction excision qui supprime l'ARN amorce sur le brin retardé. On peut lui attribuer un rôle similaire à celui de l'ADN polymérase des procaryotes.

La polymerase gamma

C'est une polymérase codée par un gène nucléaire, mais son activité est mitochondriale. Elle présente une structure variable en fonction des organismes : homotétramère de 4 x 47 kDa observé chez le poulet ; hétérodimère chez la drosophile (125 et 35 kDa) ; monomère chez la levure (143,5 kDa). La particularité de cette enzyme est de posséder une activité 5'-exonucléasique.

La polymérase delta

Observée initialement dans les cellules de la moelle osseuse de lapin et le thymus de veau, elle a été par la suite purifiée à partir des cellules HELA; la forme humaine est présente sous forme de dimère (125 kDa et 48 kDa). Son activité est potentialisée par une protéine de 36 kDa, PCNA (Proliferating cell Nuclear antigen). Elle possède une activité 3'>5'. Il apparaît que son rôle est différent en fonction des organismes :

- Synthèse du brin direct chez les virus ;
- Synthèse du brin retardé chez les levures ;
- Synthèse de deux brins chez les eucaryotes en jonction avec la polymérase epsilon.

La polymérase epsilon

Elle aussi fut purifiée à partir des cellules HELA, deux formes ont été trouvées ;

- Une première comprenant une sous-unité catalytique de 125-140 kDa et une autre de de 40 kDa ;
- Une deuxième, plus complexe, comprenant une sous-unité catalytique de 215–230 kDa et une série de sous-unités de 30-70 kDa.

Elle est impliquée aussi bien dans le mécanisme de réplication que de réparation de l'ADN.

Chez l'homme, l'enzyme est constituée de deux sous-unités liées par un site de protéolyse conduisant après coupure à des formes de 140 à 250 kDa avec chacune une activité catalytique.

Ces deux formes sont retrouvées dans les cellules ; il a été observé in vitro que la protéine n'est active que si les protéines accessoires telles PCNA ou RFA et RFC sont présentes. Cette enzyme possède une activité 3'>5' exonucléasique. Elle ne nécessite pas le facteur PCNA pour être active.

RF-A et RF-C sont des protéines (replication factors) avec des spécificités propres.

RFA-A est une protéine constituée de 3 sous-unités (11, 34 et 70kDa) qui se fixe sur l'ADN simple brin. Ce facteur semble être impliqué dans les mécanismes de recombinaison.

RF-C Se fixe au niveau des zones de synthèse des amorces. Ce facteur de réplication possède une fonction ATPasique ADN dépendante.

La réplication des extrémités d'ADN linéaire

Lors de la réplication d'ADN circulaire, le remplacement des amorces ARN par de l'ADN est parfaitement maîtrisé par les mécanismes de réparation. Dans le cas des molécules d'ADN linéaire le retrait de l'ARN aux extrémités des brins directs laisse un espace qui ne peut être comblé dans la mesure où les polymérases ne fonctionnent que dans le sens 3'>5'. Chez les eucaryotes, ce problème est levé par la présence des séquences télomériques. Chez les eucaryotes, la synthèse et la réplication du télomère sont généralement catalysés par la télomérase, une transcriptase inverse spécialisée qui maintient la longueur du télomère en équilibrant son raccourcissement et son élongation nets. La télomérase est une ribonucléoprotéine pour laquelle une partie de la composante ARN fournit la matrice utile à la synthèse des séquences télomériques. Elle est donc responsable de la réplication des extrémités d'ADN linéaire.

Correction lors de la réplication

L'activité de réplication, bien qu'effectuée selon un principe de copiage fidèle qui utilise une matrice, n'est pas infaillible et des erreurs de réplication surviennent. Ainsi a-t-il été observé un taux d'erreur par nucléotides de l'ordre 1 sur 1 milliard. Une aussi grande fidélité de réplication est due non seulement à la spécificité d'appariement des bases AT et GC mais aussi à la propriété de correction d'épreuve de l'ADN Pol III. La correction a

pour rôle d'éliminer le nucléotide mal apparié donc mal intégré au moment de la réplication. L'ADN pol III assure cette fonction grâce à son activité d'exonucléase 3'>5'. L'élimination du nucléotide mal apparié libère une extrémité 3'OH qui peut être utilisée par l'ADN polymérase pour insérer le bon nucléotide. Le taux d'erreur dans l'ADN est ainsi contrôlé par la même enzyme grâce à ses activités de polymérisation dans le sens 5'>3' et de dépolymérisation dans le sens 3'>5'.

Se pose le problème de la structure de la chromatine

Les modifications de la topologie de l'ADN après réplication ne posent pas de problèmes compte-tenu des nombreuses topoisomérases qui sont présentes. Cependant se pose la question de la structuration nucléosomique. Il apparaît que la molécule d'ADN après réplication s'organise très rapidement en nucléosomes (observations effectuées grâce au marquage à la thymidine tritiée).

Les nouveaux nucléosomes formés contiennent essentiellement des histones nouvellement synthétisées. Aucun mélange n'a été observé entre les anciennes et les nouvelles histones au sein d'un même nucléosome.

Le système de sauvegarde assure le maintien de l'intégrité de l'ADN

Les agressions physiques et chimiques de l'environnement altèrent les différentes molécules cibles de toutes les cellules. Les altérations de la molécule d'ADN sont les plus graves puisqu'elles peuvent être pérennisées au cours des réplifications. À ce titre, il existe des systèmes qui assurent le maintien de l'intégrité de l'ADN face à ces agressions.

Les altérations d'origine physique

Les rayons cosmiques et la radioactivité correspondent à des rayonnements très énergétiques qui peuvent directement produire des lésions de type modification de bases ou rupture de brins. Cependant l'agression par ces rayonnements peut être aussi indirecte car ils génèrent l'apparition d'ions superoxydes (ou radicaux libres), agents très réactifs. Les rayons solaires, moins énergétiques, induisent principalement des dimérisations de thymines.

Altérations d'origine chimique

Elles sont très variées et peuvent résulter du simple métabolisme de la cellule. Il a été observé que l'agitation thermique et les ions H⁺ peuvent supprimer 10000 bases puriques par jour et par cellule chez l'homme.

Ces lésions peuvent être directes (dépurination modification de bases, désamination, oxydation, création de liaisons covalentes entre les deux brins). Elles peuvent être aussi indirectes (substances intercalantes comme le bromure d'éthidium).

Toutes ces altérations se produisent principalement entre les mitoses et entraînent des mutations de type délétion et insertion.

Des systèmes multiples assurent la sauvegarde du contenu informatif de l'ADN

Dans le cadre de prévention, la cellule possède des systèmes de protection contre les agents susceptibles de provoquer des altérations. Si l'on considère l'exemple des ions superoxydes, ils sont directement détruits par la superoxyde dismutase (SOD). Les protons H⁺ sont directement pris en charge par des systèmes de régulation de l'équilibre acido-basique. Les oxydations sont réduites par les différents systèmes de réduction.

La fidélité de la réplication n'est pas absolue, cependant la correction immédiate des erreurs est assurée par un système assez fidèle. L'activité proofreading des polymérases (correction) laisse passer quelques erreurs. En général, après polymérisation, on observe environ un taux d'erreur de 10^{-10} à 10^{-11} . Il se trouve que le produit formé tout de suite après le passage de la polymérase possède encore un taux d'erreur voisin de 10^{-8} . Il existe un système qui a pour fonction de reconnaître le mauvais appariement, de détecter quel est le nouveau brin, de localiser l'erreur et de sectionner immédiatement après la mauvaise base.

La Réaction de polymérisation en chaîne (PCR)

La PCR (*Polymerase Chain Reaction*) fut inventée par K. Mullis en 1983 et brevetée en 1985. Son principe repose sur l'utilisation de l'ADN polymérase, il s'agit d'une répllication in vitro de séquences spécifiques d'ADN. Cette méthode permet de générer à des dizaines de milliards d'exemplaires un fragment d'ADN particulier (la séquence d'intérêt, ADN d'intérêt ou ADN cible) à partir d'un extrait d'ADN (ADN matriciel). En effet, si la séquence d'intérêt est présente dans l'extrait d'ADN, il est possible de sélectivement la répliquer (on parle d'amplification) en très grande quantité. La puissance de la PCR repose sur le fait que la quantité d'ADN matriciel n'est pas, en théorie, un facteur limitant. On peut donc amplifier des séquences nucléotidiques à partir de quantités infinitésimales d'extrait d'ADN.

La PCR est donc une technique de purification ou de clonage. L'ADN extrait à partir d'un organisme ou d'un échantillon contenant des ADN d'origines diverses n'est pas directement analysable. Il contient une masse trop importante de séquences nucléotidiques. Il convient donc d'isoler, de purifier la ou les séquences qui présentent un intérêt, qu'il s'agisse de la séquence d'un gène ou de séquences non codantes (introns, transposons, mini ou microsatellites...). À partir d'une telle masse de séquences que constitue l'ADN matriciel, la PCR peut donc sélectionner une ou plusieurs séquences déterminées et les amplifier par répllication à des dizaines de milliards de copies. La réaction terminée, la quantité extrêmement faible d'ADN matriciel contenue dans l'échantillon PCR n'aura pas varié. En revanche, la quantité de la ou des séquences amplifiées (l'ADN d'intérêt) sera très grande. La PCR permet donc d'amplifier un signal à partir d'un bruit de fond, il s'agit donc bien d'une méthode de clonage moléculaire, et cloner revient à purifier.

Les applications de la PCR sont multiples. C'est une technique désormais incontournable en biologie cellulaire et moléculaire. Elle permet notamment en quelques heures le « clonage acellulaire » d'un fragment d'ADN grâce à un système automatisé, alors qu'il faut plusieurs jours avec les techniques standard de clonage moléculaire. D'autre part, la PCR est largement utilisée à des fins de diagnostic pour détecter la présence d'une séquence d'ADN spécifique de tel ou tel organisme dans un fluide biologique. Elle est aussi employée pour réaliser des empreintes génétiques, qu'il s'agisse de l'identification génétique d'une personne dans le cadre d'une enquête judiciaire, ou de l'identification de variétés animales, végétales ou microbiennes destinée à des tests de qualité alimentaire, de diagnostic ou de sélection variétale. La PCR est encore indispensable à la réalisation d'un séquençage ou d'une mutagenèse dirigée. Enfin, il existe des variantes de la PCR : real-time PCR, PCR compétitive, PCR in situ, RT-PCR...

But et principe

La PCR permet donc d'obtenir par réplification in vitro de multiples copies d'un fragment d'ADN à partir d'un extrait. L'ADN matriciel peut tout autant être de l'ADN génomique que de l'ADN complémentaire obtenu par RT-PCR à partir d'un extrait d'ARN messagers (ARN poly-A), ou encore de l'ADN mitochondrial.

Le mélange réactionnel et les cycles de température

La réaction de polymérisation en chaîne est réalisée dans un mélange réactionnel qui comprend l'extrait d'ADN (ADN matriciel), la Taq polymérase, les amorces et les quatre désoxyribonucléosides triphosphates (dNTP) en excès dans une solution tampon. Les tubes contenant le mélange réactionnel sont soumis à des cycles de température réitérés plusieurs dizaines de fois dans le bloc chauffant d'un thermocycleur (appareil qui comporte une enceinte où l'on dépose les tubes échantillons et dans laquelle la température peut varier, très rapidement et précisément, de 0 à 100°C par effet Peltier). L'appareil permet la programmation de la durée et de la succession des cycles de paliers de température. Chaque cycle comprend trois périodes de quelques dizaines de secondes.

La dénaturation : 94°C

La première période s'effectue à une température de 94°C, dite température de dénaturation. À cette température, l'ADN matriciel, qui sert de matrice au cours de la réplification, est dénaturé : les liaisons hydrogène ne peuvent pas se maintenir à une température supérieure à 80°C et les ADN double-brin se dénaturent en ADN simple-brin (ADN monocaténares).

L'hybridation : 40 – 70°C

La deuxième période s'effectue à une température généralement comprise entre 40 et 70°C, dite température d'hybridation des amorces. La diminution de la température permet aux liaisons hydrogène de se reformer et donc aux brins complémentaires de s'hybrider. Les amorces, courtes séquences monocaténares complémentaires de régions qui flanquent l'ADN à amplifier, s'hybrident plus facilement que les longs brins d'ADN matriciel. Plus la température d'hybridation est élevée, plus l'hybridation est sélective, plus elle est spécifique.

Les amorces (primers)

Pour parvenir à amplifier sélectivement des séquences nucléotidiques à partir d'un extrait d'ADN par PCR, il est indispensable de disposer d'au moins une paire d'oligonucléotides. Ces oligonucléotides, qui vont servir d'amorces pour la réplication, sont synthétisés par voie chimique et doivent montrer la meilleure complémentarité possible avec les deux extrémités de la séquence d'intérêt que l'on souhaite amplifier. L'une des amorces est conçue pour reconnaître par complémentarité une séquence située en amont du brin 5'-3' du fragment d'ADN d'intérêt ; l'autre pour reconnaître, toujours par complémentarité, une séquence située en amont du brin complémentaire (3'-5') du même fragment d'ADN. Les amorces sont des ADN monocaténaires dont l'hybridation sur les séquences flanquant la séquence d'intérêt permettra sa réplication de façon sélective. La taille des amorces est généralement comprise entre 10 et 30 nucléotides afin de garantir une hybridation suffisamment spécifique sur les séquences d'intérêt de l'ADN matriciel.

D'autres critères, importants mais secondaires par rapport aux précédents, doivent être pris en compte pour la conception et l'emploi des amorces. La séquence des amorces doit comporter la plus grande proportion possible de guanines et de cytosines. En effet, la liaison entre guanines et cytosines est assurée par trois liaisons hydrogène au lieu de deux entre les adénines et les thymines, ce qui se traduit

par une meilleure stabilité de la liaison au moment de l'hybridation des amorces. Plus les amorces sont riches en cytosines et guanines, plus leur température d'hybridation (T_m) peut être élevée (elle est en général comprise entre 40 et 65°C, mais peut atteindre 70°C) et donc plus la spécificité

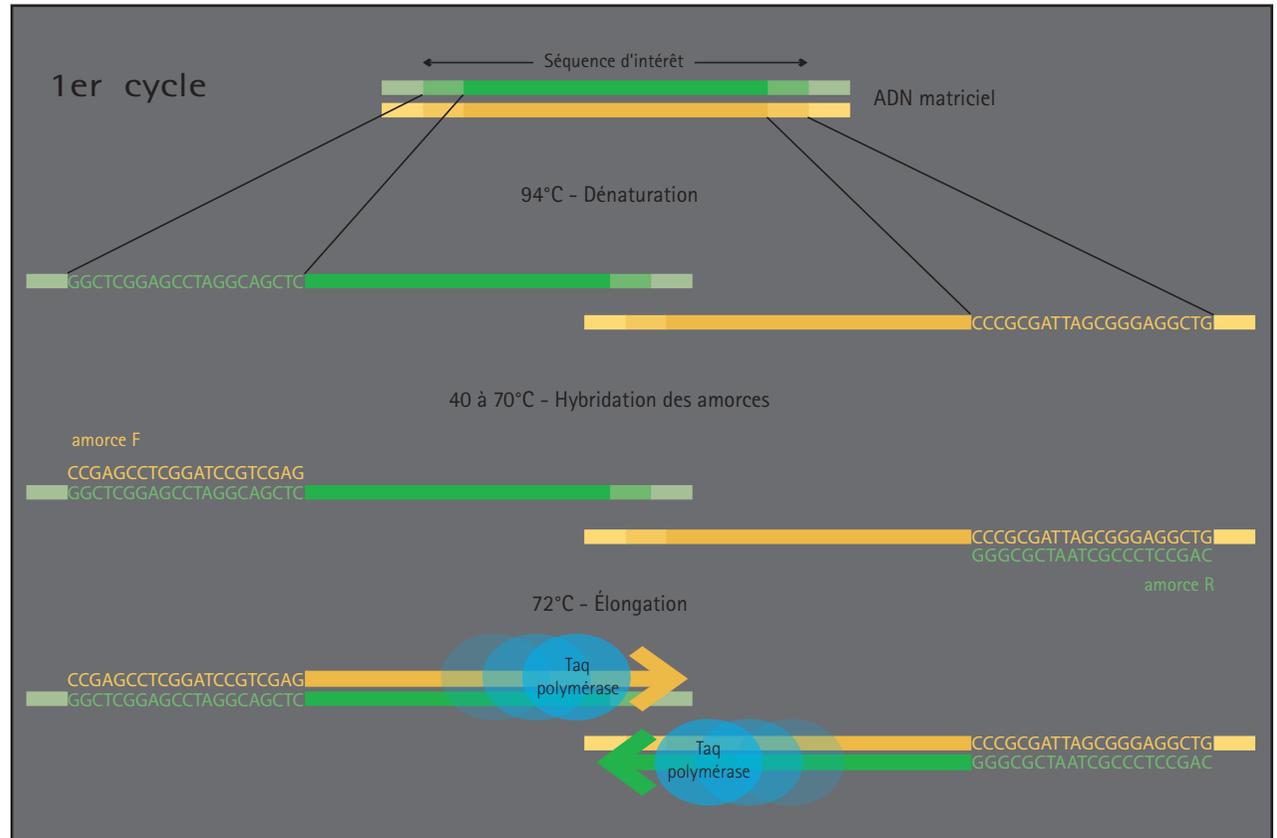


Figure 1: 1er cycle d'amplification par PCR

de l'hybridation est grande. À des températures d'hybridation basses, les amorces s'hybrident de façon beaucoup moins sélective ce qui peut se traduire par l'amplification de séquences d'ADN non souhaitées (amplification non spécifique).

Les pourcentages en guanines et cytosines de chaque membre du couple d'amorces doivent être aussi proches que possible car la proportion de guanine et cytosine détermine la température d'hybridation. Si ces proportions diffèrent significativement d'une amorce à l'autre, il est très malaisé de définir une température d'hybridation convenable et qui permette une hybridation équilibrée des amorces. Il ne faut pas non plus que les amorces présentent de fortes complémentarités de séquence sans quoi elles s'associeraient entre-elles et n'effectueraient pas correctement l'amorçage de la réplication.

Il faut encore noter que certaines modifications chimiques du milieu réactionnel permettent d'optimiser la spécificité d'hybridation des amorces. L'ajout de détergent ou de glycérol, par exemple, inhibe la formation des liaisons hydrogène, rend plus difficile l'hybridation des amorces et donc favorise sa spécificité.

L'élongation : 72°C

La troisième période s'effectue à une température de 72°C, dite température d'élongation. À 72°C, la Taq polymérase se lie aux ADN monocaténaire amorcés et catalyse la réplication en utilisant les désoxyribonucléosides triphosphates présents dans le mélange réactionnel. Les régions de l'ADN matriciel en aval des amorces sont ainsi sélectivement synthétisées.

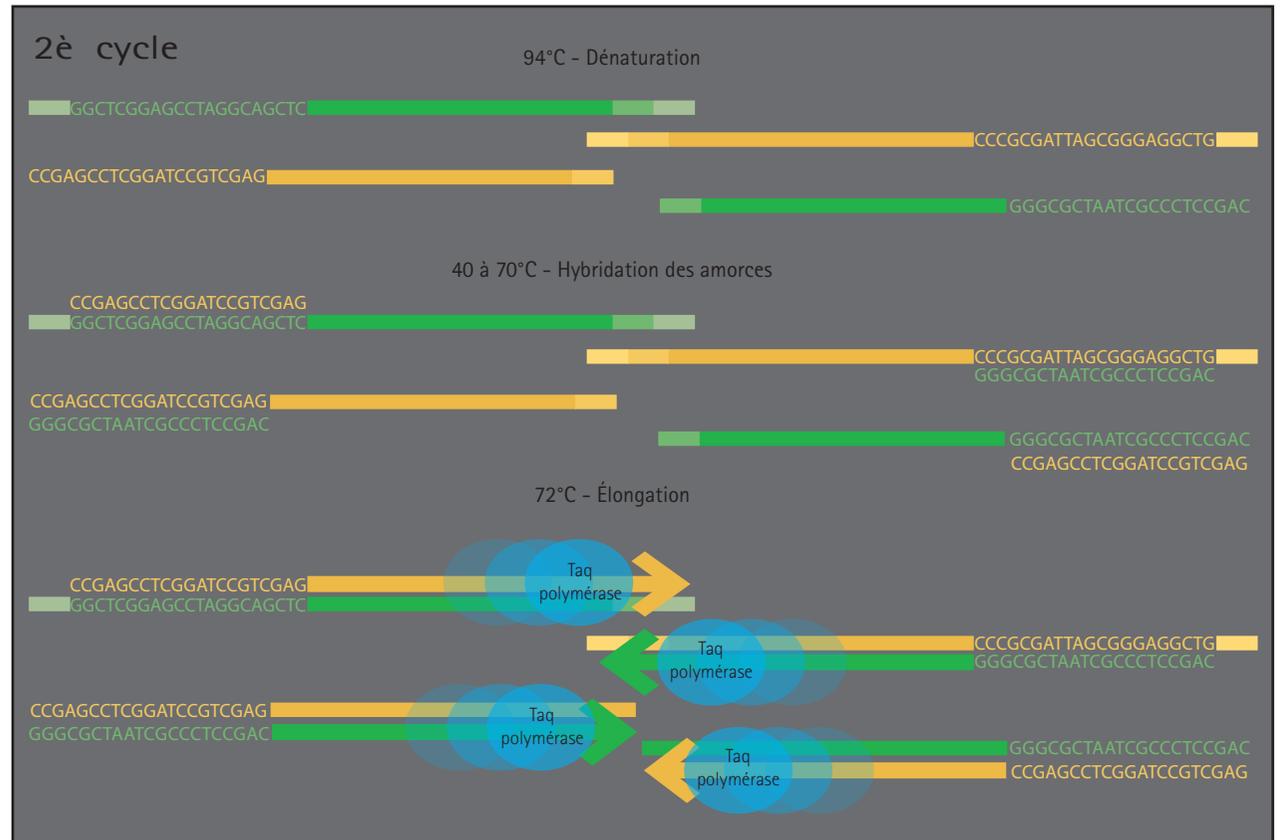


Figure 2: 2ème cycle d'amplification par PCR

Au cycle suivant, les fragments synthétisés au cycle précédent servent à leur tour de matrice et au bout de quelques cycles, l'espèce prédominante correspond à la séquence d'ADN comprise entre les régions où les amorces s'hybrident. Il faut compter 20 à 40 cycles pour synthétiser une quantité analysable d'ADN (environ 0,1 microgramme). Chaque cycle voit théoriquement doubler la quantité d'ADN présente au cours du cycle précédent.

Il est recommandé de rajouter un cycle final d'élongation à 72°C, notamment lorsque la séquence d'intérêt est de grande taille (supérieure à 1 kilobase), à raison de 2 minutes par kilobase.

La PCR permet d'amplifier des séquences dont la taille est inférieure à 6 kilobases.

La Taq polymérase

L'ADN polymérase permet la réplication. On utilise une ADN polymérase purifiée ou clonée à partir d'une bactérie extrêmophile, *Thermus aquaticus*, qui vit dans les sources chaudes et résiste à des températures supérieures à 100°C. Cette polymérase (Taq polymérase) possède la caractéristique remarquable de résister à des températures de l'ordre de 100°C, lesquelles sont généralement suffisantes pour dénaturer la plupart des protéines. *Thermus aquaticus* trouve sa température de confort à 72°C, température optimum pour l'activité de sa polymérase.

Les conditions réactionnelles

Les volumes de milieu réactionnel varient entre 10 et 100 µl. Il existe une multitude de formules de milieux



Figure 3: 3ème cycle d'amplification par PCR

réactionnels. Il est toutefois possible de définir une formule standard qui convient à la plupart des réactions de polymérisation en chaîne. Cette formule, à peu de chose près, a été choisie par la plupart des fabricants et fournisseurs qui, du reste, délivre une solution tampon prête à l'emploi avec la Taq polymérase. Concentrée 10 fois, sa formule est à peu près la suivante : 100 mM Tris-HCl, pH 9,0 ; 15 mM MgCl₂, 500 mM KCl.

Il est possible de rajouter des détergents (Tween 20, Triton X-100...) ou du glycérol afin d'augmenter les conditions de stringence qui rendent plus difficile et donc plus sélective l'hybridation des amorces. Cette démarche est généralement employée pour réduire le niveau d'amplifications non spécifiques dues à l'hybridation des amorces sur des séquences sans rapport avec la séquence d'intérêt. On peut aussi réduire la concentration en KCl jusqu'à l'éliminer ou encore augmenter la concentration en MgCl₂. En effet, certains couples d'amorces fonctionnent mieux avec des solutions enrichies en magnésium. D'autre part, lorsque sont utilisées de fortes concentrations en dNTP, il convient d'augmenter la concentration en magnésium à cause des interactions stœchiométriques entre magnésium et dNTP qui réduisent la quantité de magnésium libre dans le milieu réactionnel.

Les dNTP (desoxyribonucléotides) fournissent à la fois l'énergie et les nucléotides nécessaires à la synthèse de l'ADN lors de la polymérisation en chaîne. Ils sont incorporés au milieu réactionnel en excès, soit environ 200 µM final.

Selon le volume réactionnel choisi, la concentration en amorce peut varier entre 10 et 50 pmol par échantillon.

L'ADN matriciel peut provenir de n'importe quel organisme et même de matériels biologiques complexes qui comprennent des ADN de différents organismes. Mais pour garantir la réussite d'une PCR encore faut-il que l'ADN matriciel ne soit pas trop dégradé. Ce critère est évidemment d'autant plus crucial que la taille de la séquence d'intérêt est grande. Il est aussi important que l'extrait d'ADN ne soit pas contaminé par des inhibiteurs de la réaction de polymérisation en chaîne (détergents, EDTA, phénol, protéines...). La quantité d'ADN matriciel dans le milieu réactionnel pour initier

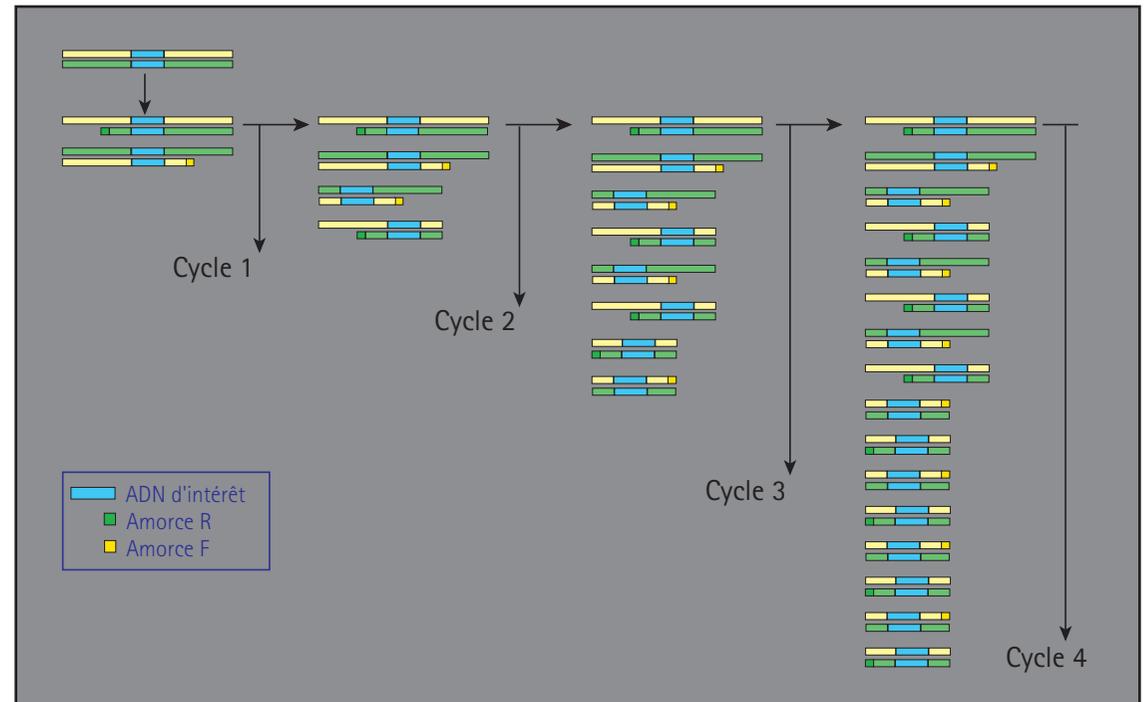


Figure 4: Hétérogénéité des fragments amplifiés par PCR

la réaction d'amplification peut se réduire à une copie unique. La quantité maximale ne peut en aucun cas excéder 2 µg. En général, les quantités utilisées se situent dans une gamme comprise entre 10 et 500 ng d'ADN matriciel.

La quantité de Taq polymérase par échantillon est généralement comprise entre 1 et 3 unités.

Le choix de la durée des cycles de température et du nombre de cycles dépend de la taille de la séquence d'intérêt ainsi que de la taille et de la complémentarité des amorces. Il convient de réduire les durées au minimum non seulement pour un gain de temps mais aussi pour prévenir le risque d'amplification non spécifique. Pour la dénaturation et l'hybridation des amorces, 30 secondes suffisent généralement. Pour l'élongation, il faut compter 1 minute par kilobase d'ADN d'intérêt et 2 minutes par kilobase pour le cycle final d'élongation. Le nombre de cycles, généralement compris entre 20 et 40, est inversement proportionnel à l'abondance en ADN matriciel.

Détection et analyse des produits PCR

Le produit d'une PCR est constitué d'un ou de plusieurs fragments d'ADN (la ou les séquences d'intérêt). La détection et l'analyse des produits peuvent être très rapidement réalisées par électrophorèse sur gel d'agarose (ou d'acrylamide). L'ADN est révélé par une coloration au bromure d'éthidium. Ainsi, les produits sont-ils visibles instantanément par transillumination aux ultraviolets (280 – 320 nm).

Des produits de très petite taille sont souvent visibles très près du front de migration sous forme de bandes plus ou moins diffuses. Ils correspondent à des dimères d'amorces et parfois aux amorces elles-mêmes. Selon les conditions réactionnelles, il arrive que des fragments non spécifiques d'ADN soient amplifiés en quantité plus ou moins abondante, formant des bandes nettes ou des « traînées » (smear).

Sur des systèmes automatisés, on utilise aujourd'hui un analyseur de fragment. Cet appareillage utilise le principe de l'électrophorèse capillaire. La détection des fragments est réalisée par une diode laser. Cela n'est possible que si la PCR est réalisée avec des amorces couplées à des fluorochromes.

Applications

Le clonage acellulaire

C'est l'une des applications les plus remarquables de la PCR. Elle permet d'isoler, c'est-à-dire de purifier un gène sans recourir aux méthodes traditionnelles de clonage moléculaire qui consistent à insérer une banque d'ADN dans un vecteur plasmidique qui est ensuite employé à transformer une souche bactérienne dont les clones, après sélection, sont criblés. La réalisation est beaucoup plus rapide et beaucoup moins aléatoire en utilisant la PCR.

On parle de clonage acellulaire, lorsqu'on utilise la PCR, car il est inutile d'employer un système cellulaire (bactérie, levure, cellule animale ou végétale) pour amplifier le clone. La réalisation d'un clonage moléculaire par PCR dépend de deux critères majeurs : le choix de l'extrait d'ADN (l'ADN matriciel) et des amorces. Il est en effet indispensable de disposer de données plus ou moins fiables sur la séquence du gène que l'on souhaite cloner et/ou des séquences qui le flanquent afin de synthétiser les jeux d'amorces nécessaires à son amplification en tout ou partie.

D'autre part, s'agit-il encore de réaliser la PCR sur l'ADN matriciel adéquat. On peut choisir l'ADN génomique qui comporte la séquence totale du génome et donc l'ensemble des gènes de l'espèce. Dans ce cas, les gènes comprennent tant les exons que les introns et leur amplification aboutit au clonage de la séquence complète du gène et même, selon les amorces que l'on a choisies, des régions régulatrices.

Mais on peut aussi choisir d'extraire les ARN messagers (ARNm), c'est-à-dire les seules séquences codantes du gène — les transcrits. Étant donné que les ARN sont instables, les ARN messagers sont transformés en ADN complémentaires (ADNc) par RT-PCR (voir infra), une variante de la PCR qui utilise la transcriptase inverse et permet de changer les séquences ARN en ADN. C'est sur cette banque d'ADNc que l'on réalise alors la PCR pour cloner le gène d'intérêt. Dans ce cas, la donne est plus complexe. La présence du transcrit d'un gène dans l'extrait dépend du type cellulaire, du tissu ou de l'organe à partir duquel on a réalisé l'extraction d'ARNm. En effet, la transcription est spécifique du type cellulaire. Pis, l'expression d'un gène est souvent régulée par des facteurs physiologiques, environnementaux... Autant dire donc que le gène d'intérêt n'est pas forcément transcrit et que la banque d'ADNc peut ne pas le contenir. Enfin faut-il encore signaler que la transcription est elle-même régulée et se voit souvent accompagnée d'un épissage alternatif. Ce phénomène conduit à l'élimination d'exons au moment de l'excision des introns et conduit à l'expression de différentes protéines à partir du même gène. Il en découle que selon le type cellulaire et les profils de régulation, on peut ne pas avoir affaire au même transcrit. Il est néanmoins très intéressant de cloner un transcrit puisque sa séquence nucléotidique correspond à la séquence aminoacide issue de la traduction. D'autre part, avec un ADNc il est plus facile de réaliser l'expression du gène et donc l'évaluation

fonctionnelle de la ou des protéines qui lui correspondent dans un modèle cellulaire d'expression.

Très fréquemment le clonage par PCR est pratiqué parallèlement sur ADN génomique (banque génomique) et différentes banques d'ADNc de sorte à déterminer la séquence complète du gène, son profil d'expression, les modalités de régulation d'épissage...

Plusieurs cas de figure se présentent :

1/ La séquence du gène est connue ;

2/ La séquence du gène n'est pas connue, mais il appartient à une famille dont plusieurs membres ont déjà été clonés et séquencés, et dont on peut déduire par comparaison des séquences très conservées au sein de la famille ;

3/ La séquence n'est pas connue, mais le gène a déjà été cloné et séquencé chez d'autres espèces d'où l'on peut déduire des séquences homologues entre espèces ;

4/ La séquence du gène n'est pas connue, mais la protéine qui lui correspond a été purifiée et séquencée et l'on peut déduire une séquence nucléotidique dégénérée à partir de la séquence aminoacide (génétique inverse), ce cas peut se combiner avec les deux précédents ;

5/ Le gène et la protéine qui lui correspond sont inconnus.

Le cas (1) ne présente pas de difficulté, la connaissance des séquences permet de synthétiser les jeux d'amorces nécessaires à l'amplification. Dans les cas (2) et (3), la difficulté augmente à cause des incertitudes en matière de séquence. Toutefois, les homologies sont assez fiables pour permettre de produire des jeux d'amorces performants à condition bien sûr que les espèces ne soient pas trop éloignées, ou que la famille du gène d'intérêt ne soit pas trop hétéromorphe. Le cas (4) peut poser des problèmes importants notamment dans la mesure où les séquences protéiques à partir desquelles les amorces sont déduites contiennent des aminoacides très dégénérés. Le cas (5) est quasiment désespéré.

La RT-PCR (Reverse Transcriptase PCR)

Comme on l'a vu au chapitre précédent, il peut s'avérer pertinent d'extraire les ARNm pour générer ensuite des copies d'ADNc. Cette réaction est catalysée par la transcriptase inverse des rétrovirus (reverse transcriptase) qui synthétise une chaîne d'ADN à partir d'une matrice d'ARN.

Dans un premier temps, les ARN totaux sont extraits. Les ARNm sont isolés à partir des ARN totaux par chromatographie d'affinité grâce à des oligodT (oligonucléotide polyT), car les ARN messagers se caractérisent par une séquence polyA en 3'. Puis, les ARNm sont soumis à la transcriptase inverse qui va générer une copie d'ADN (ADNc) de chaque ARNm. À l'issue de la transcription inverse, les ARNm sont hydrolysés (traitement alcalin, RNase ou température). Les étapes suivantes sont réalisées dans l'enceinte du thermocycleur. Les ADNc monocaténares

sont alors répliqués par l'ADN polymérase au cours d'un premier cycle de température. D'autres cycles sont réitérés afin d'amplifier les ADNc bicaténaires en grande quantité.

Dans un phénotype cellulaire donné, on estime que 10 à 15 000 gènes sont exprimés chez l'homme et la plupart des mammifères. Certains transcrits cellulaires sont exprimés à quelques centaines voire quelques milliers de copies par cellule, mais la majorité des transcrits représente un faible nombre de copies. Les profils d'expression des transcrits connaissent des variations qualitatives ou quantitatives qui reflètent la dynamique biologique de la cellule. L'identification des variations d'expression de gènes dans un contexte physiologique ou pathologique donné peut donc apporter des informations précieuses concernant la fonction des gènes et l'influence de facteurs de modulation de leur expression, qu'ils soient physiologiques ou d'origine environnementale. L'analyse des variations d'expression de gènes impliqués dans une pathologie peut orienter vers de nouvelles cibles thérapeutiques ou de diagnostic. Enfin, d'un point de vue fondamental, étudier le profil d'expression de gènes permet d'avancer dans la compréhension des mécanismes de physiologie cellulaire.

La PCR quantitative en temps réel (Quantitative real-time PCR)

Mise au point au milieu des années quatre-vingt-dix, la PCR quantitative permet de déterminer le taux d'ADN ou d'ARN spécifiques dans un

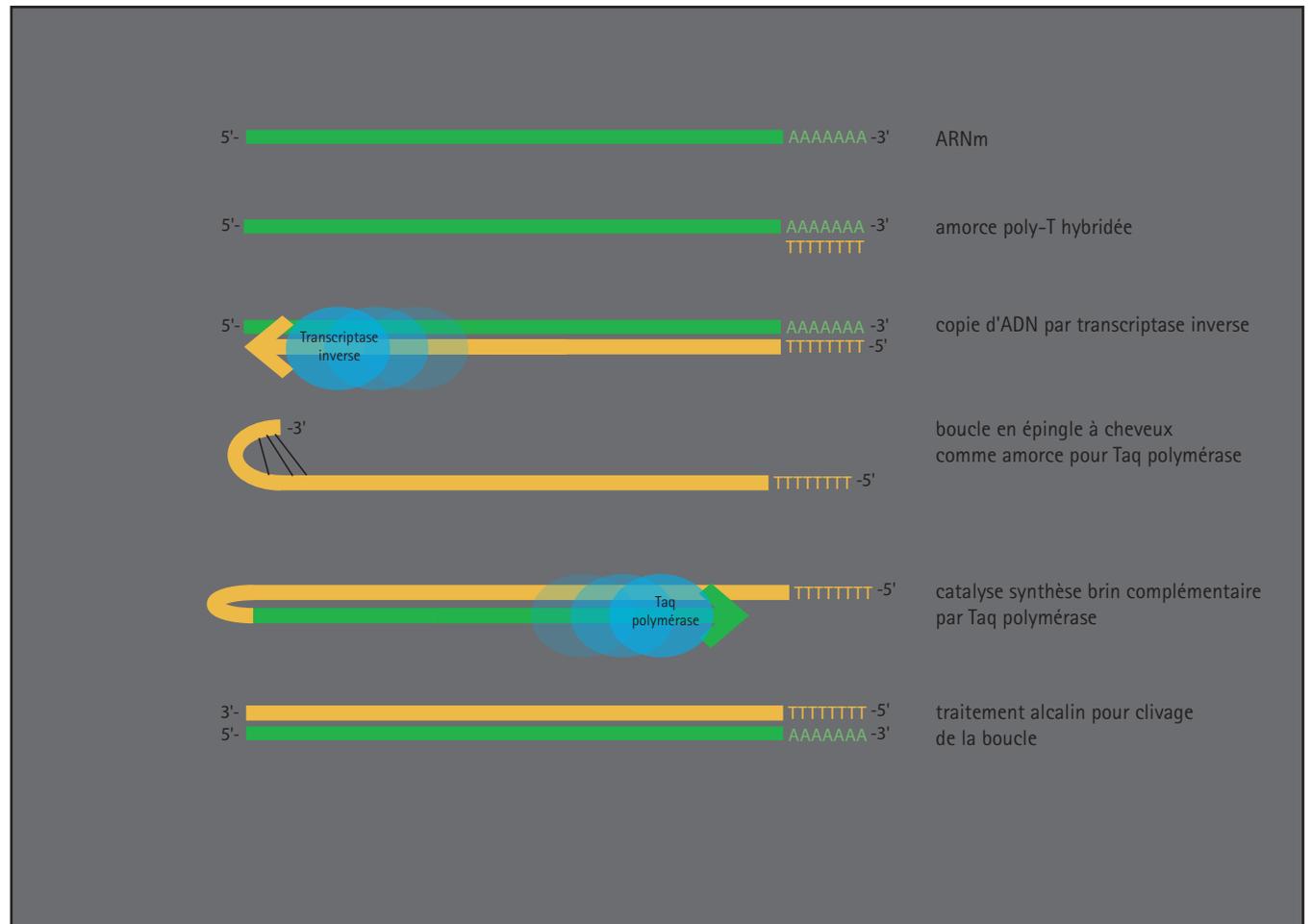


Figure 5 : Le principe de la RT-PCR

échantillon biologique. La méthode est basée sur la détection d'un signal fluorescent qui est produit de façon proportionnelle à l'amplification du produit PCR, cycle après cycle. Elle nécessite un thermocycleur couplé à un système de lecture optique qui mesure une émission de fluorescence.

Une sonde nucléotidique est synthétisée de telle sorte qu'elle puisse s'hybrider sélectivement à l'ADN d'intérêt, entre les séquences où les amorces s'hybrident. La sonde est marquée sur l'extrémité 5' par un fluorochrome signal (par exemple, la 6-carboxyfluoresceine), et sur l'extrémité 3' par un fluorochrome extincteur (quencher) (par exemple, la 6-carboxy-tétraméthyl-rhodamine).

Cette sonde doit montrer une température d'hybridation (T_m) supérieure à celle des amorces afin qu'elle s'hybride à 100 % pendant la phase d'élongation (paramètre critique).

Tant que les deux fluorochromes restent présents au niveau de la sonde, l'extincteur empêche la fluorescence du signal. En fait, la proximité de l'extincteur et du signal induit une absence d'émission de fluorescence. Or, pendant la phase d'élongation, la Taq polymérase, qui possède une activité nucléase 5'-3' intrinsèque, dégrade la sonde et donc libère le fluorochrome signal. Le taux de fluorescence alors libéré est proportionnel à la quantité de produits PCR générée à chaque cycle.

Le thermocycleur est conçu de telle sorte que chaque échantillon (la PCR est réalisée

généralement dans des plaques 96 puits) soit connecté à un système optique. Celui-ci comprend un émetteur laser connecté à une fibre optique.

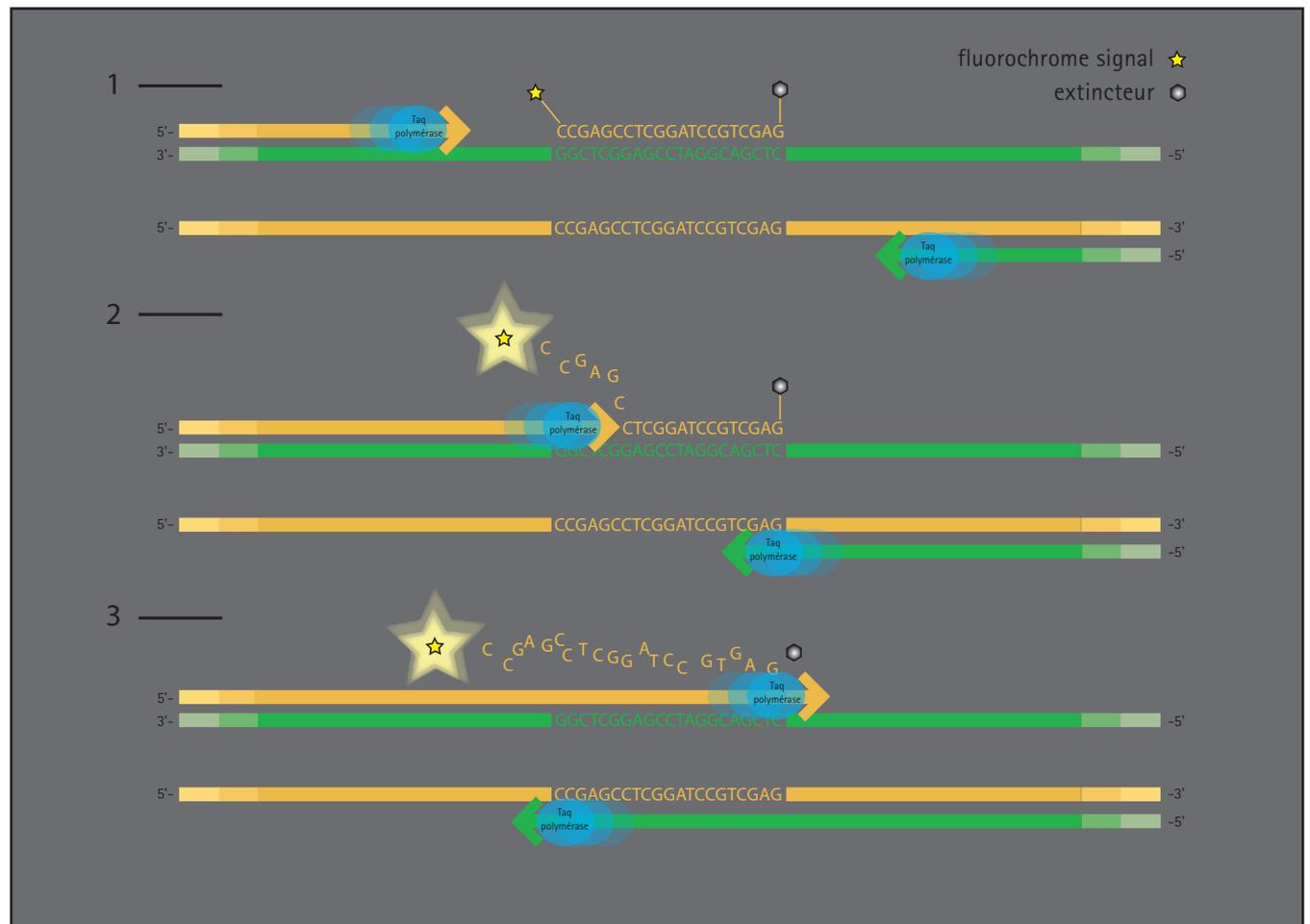


Figure 6 : Le principe de la PCR quantitative en temps réel

Le laser, par l'intermédiaire de la fibre optique, excite le fluorochrome au sein du mélange réactionnel PCR. La fluorescence émise est retransmise, toujours par le biais de la fibre optique, à une caméra numérique reliée à un ordinateur. Un logiciel analyse et thésaurise ensuite les données. La PCR quantitative est une méthode de haute spécificité et sensibilité. Elle s'avère très opportune pour d'innombrables applications. Une PCR conventionnelle n'apporte que des données qualitatives (présence ou absence de l'ADN d'intérêt, purification de cet ADN). La PCR quantitative, comme son nom l'indique, permet de connaître en outre précisément la quantité de l'ADN d'intérêt (ou de l'ARN, puisqu'il est possible de conduire une RT-PCR quantitative avec le même appareillage). Elle est en effet très souvent utilisée à cette fin, par exemple afin de déterminer la charge virale, notamment dans les cas d'hépatite C ou de SIDA. L'une des plus remarquables et utiles applications est l'analyse de l'expression d'un gène grâce à la mesure quantitative des transcrits.

La PCR semi-quantitative ou compétitive

Il s'agit dans la plupart des cas de RT-PCR. Dans le cas de la PCR quantitative, le taux d'ARN ou d'ADN d'intérêt est mesuré en tant que quantité absolue. Dans les cas de PCR semi-quantitative ou de PCR compétitive, il s'agit de mesurer des quantités relatives grâce à des standards qui correspondent à des ARN ou plus rarement à des ADN. Ces standards peuvent être internes ou externes. Les standards externes peuvent être homologues ou hétérologues. Le standard est un ARN (plus rarement un ADN) qui est présent dans l'extrait d'ARN (standard interne) ou qui est rajouté en quantité connue dans le mélange réactionnel (standard externe). Le standard est amplifié en même temps que l'ARN d'intérêt. Il y a donc compétition entre l'amplification du standard et celle de l'ADN d'intérêt. Plus la quantité de standard est importante (il s'agit néanmoins que l'amplification soit toujours en phase exponentielle) moins l'ARN d'intérêt sera amplifié et donc plus sa quantité sera faible *in fine*. Bien sûr, la méthode d'analyse de l'échantillon PCR doit permettre de discriminer le standard par rapport à l'ARN d'intérêt d'une part et d'autre part d'évaluer la quantité relative d'ADN d'intérêt par comparaison avec la quantité de standard qui est connue.

Les standards internes sont des ARN endogènes, correspondant aux ARN de gènes dont l'expression est présumée constante (actine, bêta2-microglobuline...) et qui sont présents au sein de la population d'ARN matrices lors de la transcription inverse. Ces standards présentent un désavantage majeur : ils nécessitent l'emploi d'amorces différentes de celles qui sont utilisées pour l'ARN d'intérêt. Les cinétiques d'amplification sont donc sensiblement différentes et il est très difficile voire impossible de garantir une expression constante entre différents échantillons.

Les standards d'ARN externes homologues sont des ARN synthétiques qui partagent les mêmes sites d'hybridation des amorces que l'ARN d'intérêt et qui possèdent la même séquence globale, à une légère mutation, délétion ou insertion près qui vont permettre l'identification et la quantification de celui-ci par rapport au signal rendu par l'ARN d'intérêt. Ces standards permettent d'une part d'apprécier la variabilité introduite

au niveau de la RT et, d'autre part, présentent globalement la même efficacité d'amplification que l'ARN d'intérêt que ce soit au niveau de la RT ou de la PCR.

Les standards d'ARN externes hétérologues sont des ARN exogènes et leur taux peut donc être contrôlé. Ils présentent toutefois, à la différence des standards externes homologues une efficacité d'amplification différente comparé à celle de l'ARN d'intérêt.

Dans le cas de la RT-PCR quantitative (PCR semi-quantitative), le standard consiste en une solution titrée d'ADN de séquence identique à celle de l'ADN d'intérêt à quantifier. Une série de dilution est réalisée, chacune étant utilisée pour une amplification. Il s'agit ensuite de définir le nombre idéal de cycles pour se placer dans la phase exponentielle de la réaction tout en s'assurant d'une amplification efficace. Ensuite, chaque dilution d'ADN standard ainsi que l'ADN extrait de l'échantillon à quantifier sont soumis en parallèle à la réaction de PCR. Une courbe étalon est établie avec les dilutions de standards [signal = f (concentration)]. Connaissant la valeur du signal mesuré sur l'échantillon à quantifier, le nombre de copies correspondant peut être extrapolé à partir de la courbe.

Dans le cas de la PCR compétitive, une série de dilutions d'ARN standard externe homologue synthétique est co-amplifiée avec des quantités équivalentes d'ARN total (et donc une quantité équivalente du gène natif). Le standard entre en compétition avec l'ARN d'intérêt vis-à-vis de la polymérase et des amorces. Plus la concentration en standard augmente, plus le signal du gène d'intérêt diminue. Ici, la PCR n'a pas besoin d'être réalisée en phase exponentielle et les résultats présentent une correcte reproductibilité. Cependant la méthode est lourde et ne permet pas de gérer beaucoup d'échantillons simultanément.

La PCR appliquée au diagnostic

La PCR représente un fabuleux outil de diagnostic. Elle est déjà très utilisée dans la détection de maladies génétiques. L'amplification de tout ou partie d'un gène responsable d'une maladie génétique permet de révéler la ou les mutations délétères, leurs positions, leurs tailles et leurs natures. On peut ainsi détecter des délétions, des inversions, des insertions et même des mutations ponctuelles, soit grâce à une analyse directe des produits PCR par électrophorèse, soit en combinant la PCR à d'autres techniques.

Mais la PCR peut encore être employée pour détecter des maladies infectieuses (virales, bactériennes, parasitaires...), comme c'est déjà le cas pour le SIDA, l'hépatite C, ou les infections à chlamydia. Même si d'autres outils de diagnostic sont efficaces pour détecter ces maladies, la PCR présente l'énorme avantage de produire des résultats très fiables et rapides à partir d'échantillons biologiques infimes dans lesquels la présence du pathogène n'est pas toujours décelable selon les autres techniques.

Maladies génétiques

Dans le cadre des maladies génétiques, il s'agit de détecter une mutation sur la séquence d'un gène. Plusieurs cas de figure se présentent. Les plus simples concernent les insertions et délétions. Dans ces cas, la mutation se manifeste par la modification de la taille du gène ou d'une partie du gène. Dans la mesure où la mutation est connue, décrite, il suffit donc d'amplifier tout ou partie du gène. S'agissant d'une insertion, le produit PCR issu de l'ADN d'un malade est plus long que celui qui est issu d'une personne saine. Une délétion présente un résultat contraire. L'analyse des produits PCR par électrophorèse, et donc l'évaluation de leur taille, conduit directement au diagnostic.

La détection d'inversions et de mutations ponctuelles s'avère plus délicate. La différence de taille entre ADN sain et malade est nulle dans le cas d'une inversion, quasi-nulle dans le cas d'une mutation ponctuelle. On ne peut donc plus retenir le critère de taille des produits PCR pour aboutir au résultat. Il faut donc recourir à des techniques complémentaires à la PCR. Trois approches peuvent être retenues, le *southern blot*, le *RFLP* (*Restriction Fragment Length Polymorphism*) ou la détection de *mismatch*.

Le *southern blot* consiste à hybrider sur le produit PCR une sonde oligonucléotidique marquée, grâce à un isotope radioactif ou un fluorochrome, dont la séquence est complémentaire et donc spécifique de celle qui correspond à la mutation. Cette stratégie est

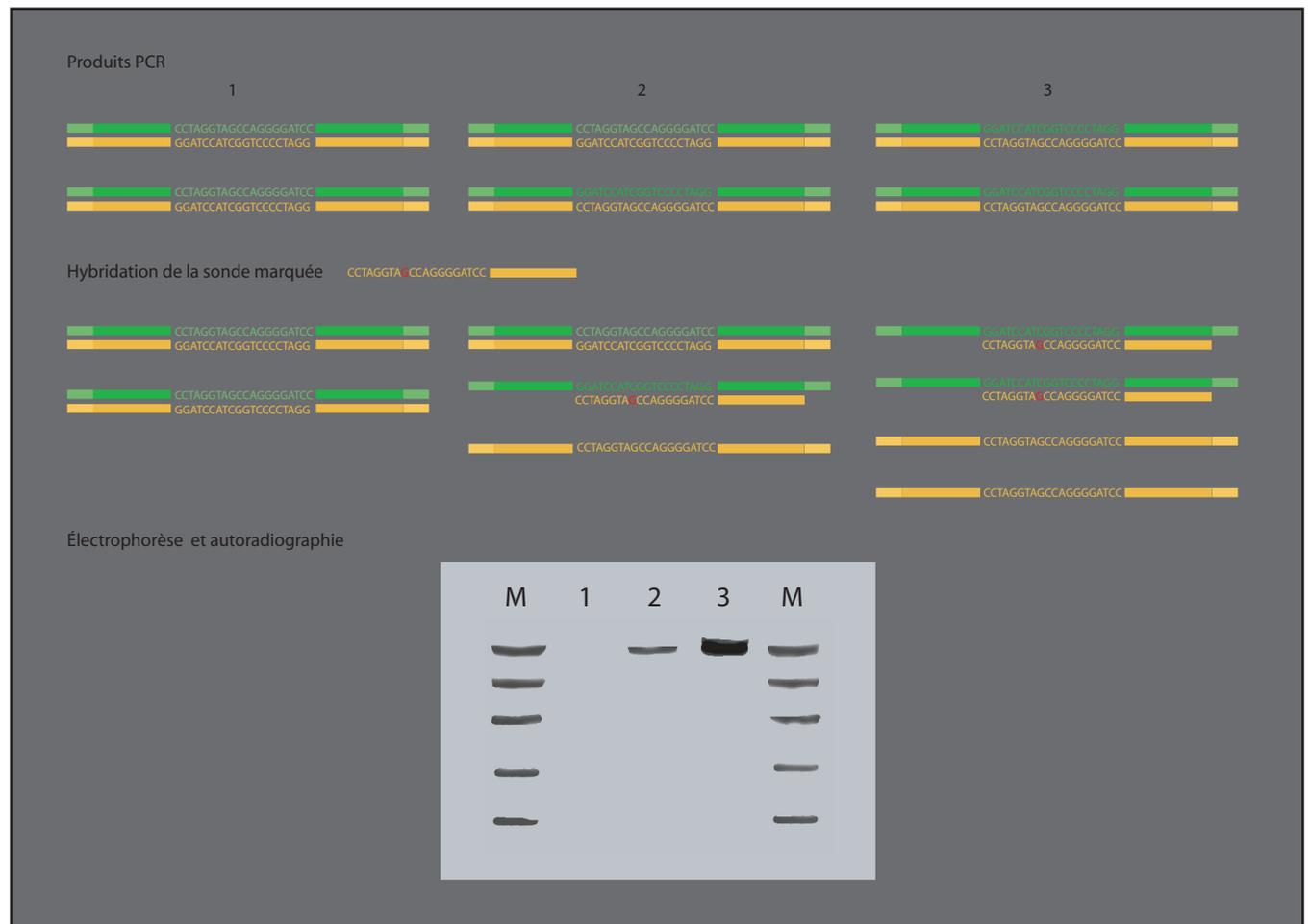


Figure 7 : Southern blot sur produits d'amplification par PCR

bien adaptée aux cas d'inversion.

Le RFLP permet de détecter les inversions comme les mutations ponctuelles. Il met en jeu une enzyme de restriction capable d'hydrolyser le

produit PCR au niveau de la séquence où se situe la mutation. Cette approche n'est possible que si un site de restriction est effectivement présent sur cette séquence, qu'il s'agisse de l'allèle muté ou de l'allèle sauvage. L'enzyme de restriction hydrolyse donc soit le produit PCR issu de l'ADN sain soit celui qui est issu de l'ADN malade. À partir de ces produits PCR, on obtient ainsi soit un soit deux fragments d'ADN qui sont ensuite révélés par électrophorèse.

La détection de *mismatch* est, comme le RFLP, adaptée aux inversions et aux mutations ponctuelles. Le produit PCR issu de l'ADN du patient (ADN échantillon) est mélangé au produit PCR issu de l'ADN d'une personne saine (ADN de référence). Ce mélange est ensuite dénaturé par la température puis réhybridé. Si l'ADN échantillon est muté, les appariements entre ADN échantillon et ADN de référence seront incomplets au niveau de la mutation. Les mésappariements (*mismatch*) concernent une seule paire de base dans le cas d'une mutation

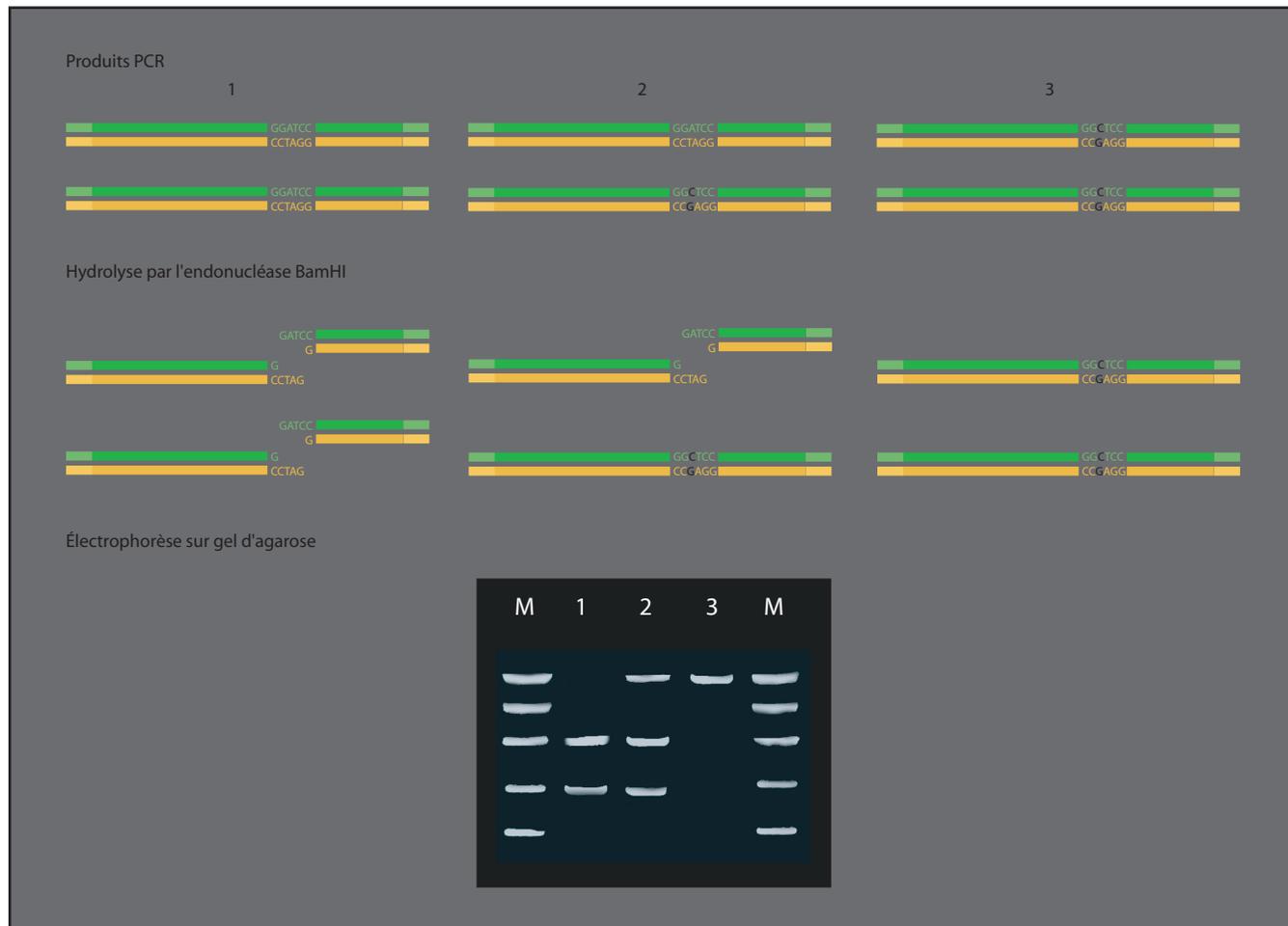


Figure 8 : RFLP sur produits d'amplification par PCR

ponctuelle, plusieurs paires de bases dans le cas d'une inversion. Ces mésappariements sont ensuite dégradés grâce à la nucléase S1, enzyme qui ne dégrade que les ADN monocaténares. Une autre solution consiste à cliver les mésappariements par voie chimique (tetroxyde d'osmium,

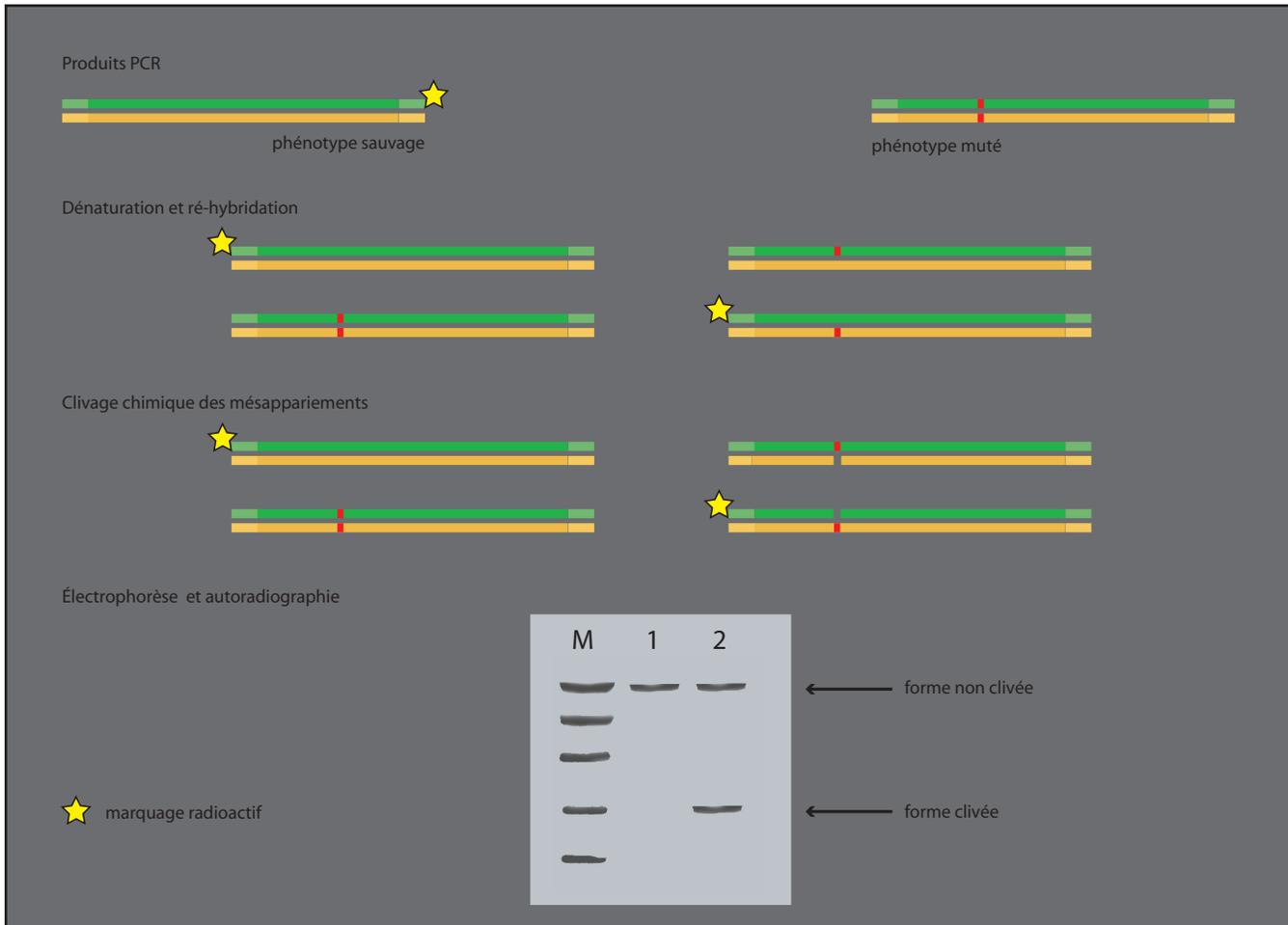


Figure 9 : Détection de mésappariements sur produits d'amplification par PCR

puis pipéridine), mais elle convient plutôt aux mutations ponctuelles. En résumé, la mutation induit un mésappariement au niveau duquel un clivage, enzymatique ou chimique, conduit à la génération de deux fragments à partir d'un produit PCR unique. Ces fragments sont analysés par électrophorèse.

Maladies infectieuses

La contamination par des virus ou des microorganismes (bactéries, parasites...) se traduit obligatoirement par la présence de leur matériel génétique dans tout ou partie de l'organisme infecté. La PCR est donc un outil d'autant plus performant pour détecter la présence d'un pathogène dans un échantillon biologique que sa sensibilité et sa spécificité sont très grandes.

La performance du diagnostic par PCR repose essentiellement sur un critère : le choix des amorces susceptibles d'amplifier de façon très

sélective une séquence de l'ADN du virus ou du microorganisme. L'ADN matriciel, d'autre part, doit être extrait à partir d'un tissu dans lequel le microorganisme est présent. Il suffit donc d'amplifier une séquence spécifique du pathogène à partir d'un prélèvement réalisé sur le patient et d'analyser le produit PCR par électrophorèse. La taille du fragment d'ADN amplifié, qui doit être conforme à la taille attendue, garantit la fiabilité du résultat et donc du diagnostic.

Dans le cas du dépistage du SIDA (VIH), par exemple, les tests de routine reposent sur la méthode ELISA qui consiste à détecter les anticorps

anti-VIH ou encore des antigènes viraux dans le sérum des patients par une technique de type *immuno-assay*. Cette méthode, assez fiable et peu coûteuse, présente néanmoins quelques inconvénients. Les faux positifs sont assez fréquents à cause d'immuno-réactivités croisées. Les échantillons positifs font donc l'objet d'un test de contrôle par une autre technique de routine, le *western blot*.

Reste le problème des séropositifs qui ne sont pas porteurs du virus, comme les enfants dont la mère est malade du SIDA. Le sang de ces nouveaux-nés contient généralement des anticorps anti-VIH d'origine maternelle et ils sont donc séropositifs. En revanche, ils ne sont pas forcément porteurs du virus. Dans ce type de cas, le diagnostic par PCR s'avère pertinent. La méthode consiste à amplifier une séquence spécifique du provirus à partir d'un extrait de lymphocytes. Le même principe est employé pour la détection du toxoplasme chez les nouveaux-nés dont la mère est porteuse.

Il est bien sûr possible de diagnostiquer le SIDA par RT-PCR en recherchant l'ARN viral dans le sérum des patients. Des méthodes quantitatives ou semi-quantitatives sont au point qui permettent en outre d'évaluer la charge virale.

La PCR appliquée à l'identification

La PCR est remarquablement efficace pour identifier des espèces, des variétés ou des individus par empreinte génétique. Cette application repose sur les connaissances acquises en matière de structure des génomes. Il s'agit tout simplement d'amplifier des séquences nucléotidiques qui sont spécifiques soit d'espèce, soit de variété, soit d'individu. Chez les eucaryotes notamment, ces séquences sont très nombreuses et offrent une vaste palette qui permet des identifications de manière très précise et très sélective. En effet, les génomes d'organismes eucaryotes comportent, à la différence des procaryotes, des séquences codantes et des séquences non codantes. Les séquences codantes correspondent aux gènes et sont donc traduites en protéines. Les séquences non codantes, qui ne sont donc pas traduites, représentent une large proportion de l'ADN génomique des eucaryotes (jusqu'à 98 %).

Les séquences codantes sont très homologues chez les individus d'une même espèce. En effet, l'espèce est caractérisée par des caractères communs qui sont garantis par ses gènes. Les différences phénotypiques entre les individus qui la composent reposent sur les variations alléliques et les différents allèles d'un même gène montrent des différences de séquence qui sont infimes (de l'ordre d'une base sur mille). D'une espèce à l'autre, en fonction de la distance phylogénétique qui les sépare, les séquences des gènes qui codent pour la même fonction présentent des homologies très fortes, d'autant plus fortes que la fonction du gène est essentielle à l'embryogénèse ou au métabolisme. En conséquence, les séquences codantes présentent peu d'intérêt en matière d'identification.

En revanche, les séquences non codantes sont très polymorphes entre espèces comme entre individus d'une même espèce. Elles présentent ainsi un large choix de marqueurs génétiques qui permettent d'établir des tests d'identification redoutablement discriminants. Parmi ces marqueurs, on

trouve notamment les minisatellites (ou VNTR, *variable number of tandem repeats* – nombre variable de répétitions en tandem) et les microsatellites (ou STR, *short tandem repeats* – courtes répétitions en tandem). Les VNTR et STR sont des polymorphismes de répétition composés de séquences qui se répètent en tandem. Ces séquences répétées mesurent de dix à quarante paires de bases pour les VNTR, de une à cinq paires de bases pour les STR. D'un individu à l'autre, la séquence répétée d'un VNTR ou d'un STR est identique mais le nombre de répétitions, et donc la taille du VNTR ou du STR peuvent être très variables (on parle d'allèles). D'autre part, il existe une grande variété de VNTR et de STR sur les génomes eucaryotes.

La mise en évidence du polymorphisme d'un STR ou d'un VNTR se fait par PCR à l'aide d'amorces qui s'hybrident aux séquences non polymorphes flanquantes. Le ou les produits d'amplification sont ensuite soit analysés par électrophorèse, soit subissent une analyse de fragments à l'aide d'un séquenceur capillaire.

Il est aujourd'hui possible d'amplifier simultanément plusieurs STR ou VNTR en utilisant plusieurs couples d'amorces. La variété des produits d'amplification obtenus permet d'aboutir à des empreintes qui sont spécifiques des individus. D'autre part, la puissance de la PCR permet d'amplifier les micro et minisatellites à partir de très peu d'ADN.

L'identification par empreinte génétique s'est beaucoup banalisée ces dernières années dans le cadre d'enquêtes judiciaires. Mais ces techniques sont tout aussi performantes chez d'autres espèces que l'homme et permettent non seulement d'identifier des individus mais aussi des variétés ou des

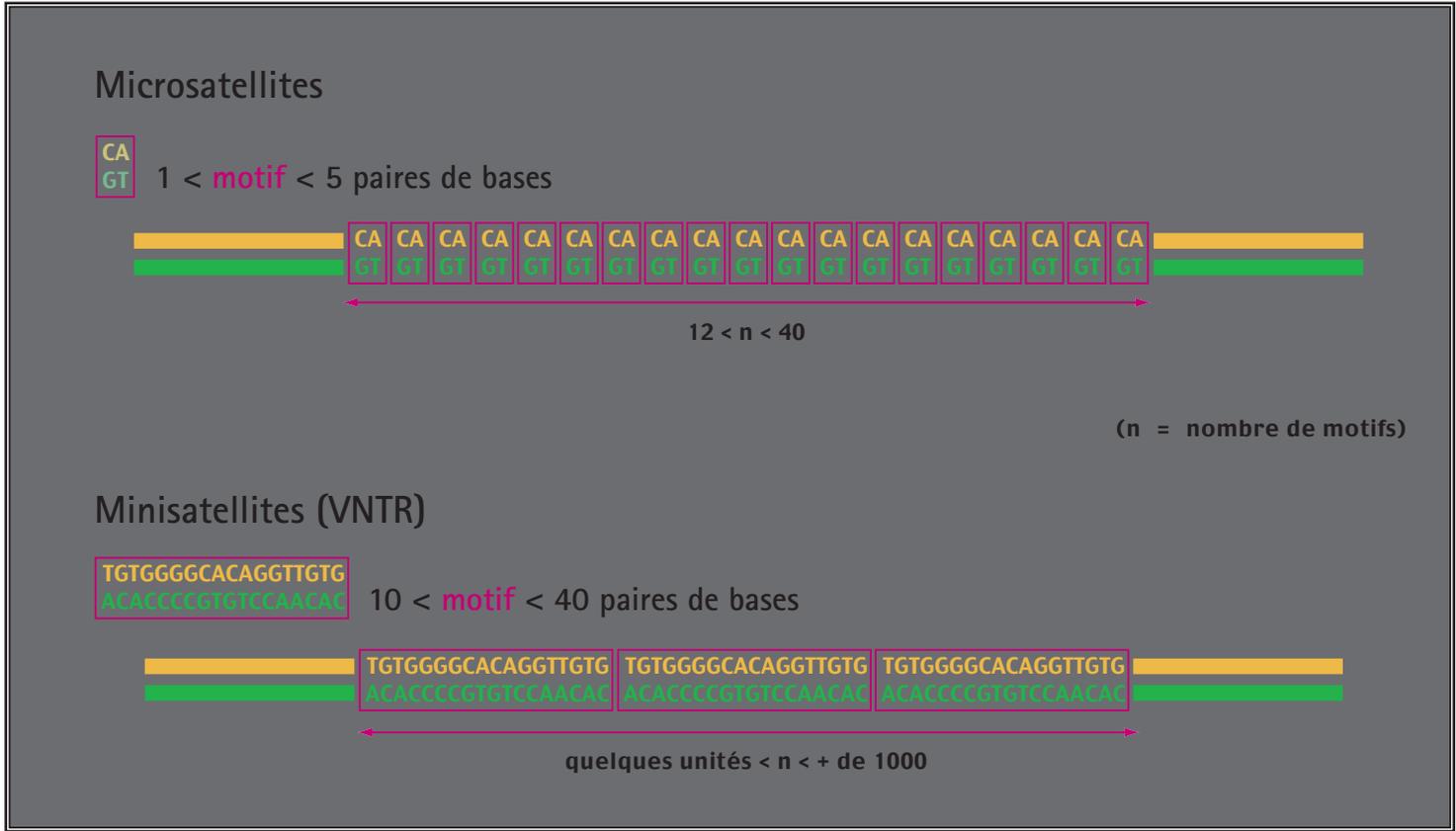


Figure 10 : Micro et minisatellites

espèces. Le type d'identification dépend simplement du choix des marqueurs. De même, à des fins d'identification variétale peut-on communément procéder selon des protocoles dérivés de la PCR. Deux approches méritent d'être mentionnées, la RAPD et l'AFLP. La RAPD (*Random Amplification of Polymorphic DNA*) est une PCR à fin d'identification variétale qui emploie des couples d'amorces aléatoires de taille réduite (une dizaine de bases). Ces amorces vont s'hybrider de manière aléatoire certes, mais la PCR aboutit généralement à un profil d'amplification par électrophorèse qui est spécifique de la variété à partir de laquelle l'ADN matriciel est issu. L'AFLP

(*Amplification of Fragment Length Polymorphism*) est une méthode beaucoup plus performante. Elle consiste en premier lieu à hydrolyser l'ADN génomique avec une ou deux endonucléases de restriction.

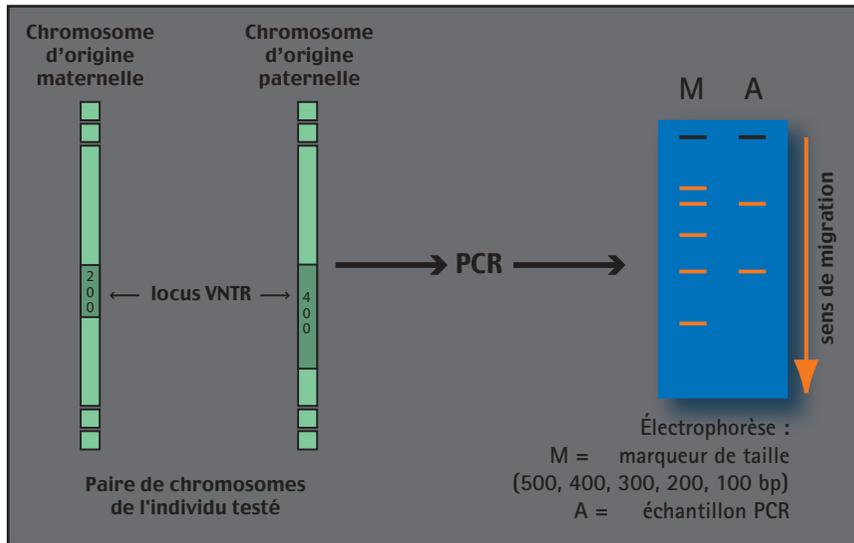


Figure 11 : Principe du test de filiation

Ensuite, on procède à la ligation d'adaptateurs (des séquences définies d'ADN d'une quinzaine de nucléotides) au niveau des extrémités cohésives générées par les enzymes de restriction. Enfin, le produit de la ligation est amplifié par PCR avec un couple d'amorces qui s'hybride au niveau des adaptateurs. L'AFLP donne un résultat comparable à la RAPD. Toutefois, l'AFLP montre des résultats

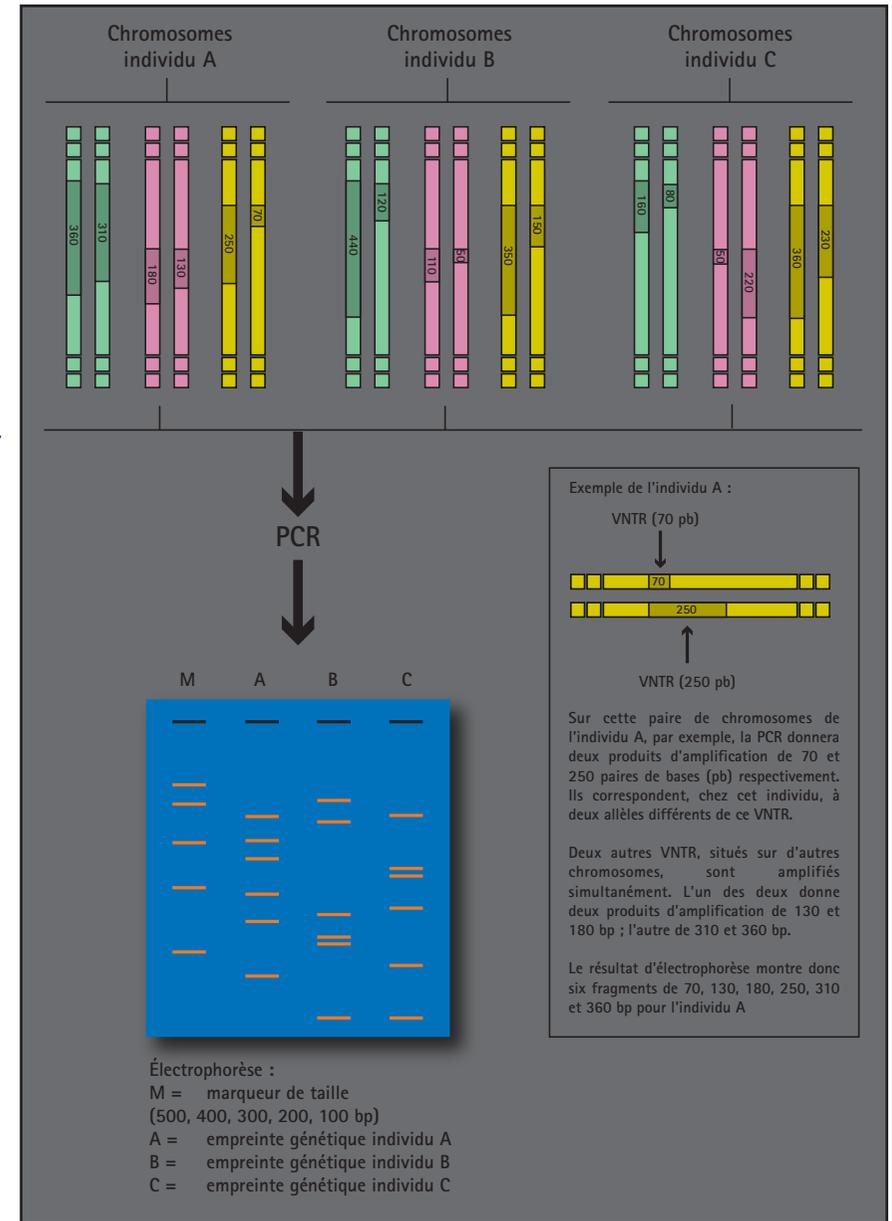


Figure 12 : Principe de l'identification humaine par empreinte génétique

plus propres et très reproductibles. Il s'agit de la méthode la plus performante à ce jour appliquée à l'identification variétale.

La PCR et RT-PCR quantitative temps-réel

Matthieu Bagory

2006

Résumé

La présente étude s'inscrit dans le cadre d'un travail personnel en immunotechniques. Il s'agit d'un état de l'art des techniques de quantification temps-réel par PCR et RT-PCR. Les principes, la technologie ainsi que des applications seront étudiés.

1 Introduction

Depuis la découverte de la *Polymerase Chain Reaction*¹ par Kary Mullis en 1986 et qui lui valu le prix Nobel de chimie en 1993, les techniques de PCR sont rapidement apparues comme des outils indispensables en biologie moléculaire. La quantification en temps réelle, et avec elle un élargissement des applications à de très nombreux domaines tels que le diagnostic clinique et l'agroalimentaire, fut certes proposée en 1992 par Higuchi R., a connu un récent et spectaculaire développement ces dernières années [1].

2 Principes

2.1 La *Polymerase Chain Reaction*

La PCR consiste en la répllication à la chaîne d'une séquence d'ADN par des enzymes, les ADN polyméras.

La PCR repose sur deux principes [1] :

1. Les « ADN polyméras ADN dépendantes thermostables » ont des propriétés de synthèse enzymatique et d'initiation « ADN double brins spécifique »
2. L'hybridation et la deshybridation des brins complémentaires d'ADN est fonction de la température

En contrôlant la température il est ainsi possible de contrôler l'activité enzymatique des ADN polyméras.

La PCR est basée sur un cycle de dénaturation - hybridation - élongation, chacune de ces étapes étant « pilotée » par une température différente.

1. La dénaturation (95° C pendant 10 à 15 minutes) consiste à déshybrider l'ADN², casser et dénaturer les structures et enzymes secondaires, homogénéiser le milieu réactionnel par agitation thermique, et activer les ADN polyméras.
2. L'hybridation (56 à 64° C pendant 2 à 60 secondes) : des amorces sens et anti-sens « s'accrochent » aux ADN matrices en raison d'une température thermodynamiquement favorable.
3. L'élongation (72° C pendant 4 à 120 secondes) : l'ADN complémentaire est synthétisé par les polyméras à partir des dNTPs³ libres dans le milieu réactionnel.

2.2 La *Reverse Transcriptase*, ou *transcriptase inverse*

La RT est une enzyme qui transcrit l'information génétique de l'ARN en ADN (appelé aussi ADNc). Elle est notamment très utilisée par les rétrovirus⁴ et les rétrotransposons⁵ [1]. Cette propriété est très utile pour quantifier des ANR, parfois présents en infime quantité, en utilisant l'ADNc produit par la RT pour effectuer une PCR.

2.3 Le quenching par FRET

Le *Fluorescence resonance energy transfer*⁶ est un phénomène par lequel l'énergie émise par un fluorochrome appelé *reporter* est absorbée par un *quencher* qui la réémettra à son tour dans une longueur d'onde différente ou sous forme de chaleur. Pour que cette absorption fonctionne il faut que *reporter* et *quencher* soient proches, et que leur spectre respectivement d'émission et de réception se chevauchent, mais non l'inverse. Ainsi, en concevant des sondes pour lesquels *reporter* et *quencher* se trouvent très proches, puis éloignés après la réaction que l'on souhaite observer, la fluorescence du *reporter* sera un marqueur direct de la réaction.

²Séparer les brins de l'ADN en double hélice

³Bases (adénine, cytosine, guanine, thymine) triphosphatées et désoxydées

⁴Virus à ARN

⁵Séquences d'ADN capables de se déplacer dans le génome de l'hôte

⁶Aussi appelé *Förster Resonance Energy Transfer*

¹Réaction de Polymérisation en Chaîne

2.4 Les stratégies de quantifications

La quantification des acides nucléiques peut se faire soit de façon « absolue » après établissement d'un étalonnage, soit de façon relative par rapport à un gène de référence [2].

1. Quantification absolue par étalonnage avec un standard : les cinétiques sont mesurées pour différentes concentrations du gène ou transcrit cible. On en déduit une courbe du Ct en fonction de la quantité d'acides nucléique et une mesure du coefficient d'efficacité E (qui doit être le plus proche possible de 2). La courbe d'étalonnage ainsi constituée permet de déduire la concentration à partir d'une mesure du Ct . La principale difficulté réside dans la capacité à titrer⁷ avec précision et reproductibilité les solutions d'étalonnage. Etant donnée cette limitation, la quantification absolue est le plus souvent réservée à des applications bien particulières, telles que la virologie [2].
2. Quantification relative : un gène de référence est quantifié en même temps que le gène étudié, la concentration de ce dernier étant ensuite normalisée par rapport à celle connue du gène de référence.

2.5 Principes mathématiques de la PCR quantitative temps réel

La PCR est une réaction en chaîne dans laquelle les produits servent de matrice pour la réaction suivante suivant une amplification théoriquement quadratique :

$$N = N_0 \cdot 2^n \quad (1)$$

Où N_0 est le nombre initial de molécules, n le nombre de cycle et N le nombre de molécules amplifiées.

Cette amplification théorique dépend dans la réalité d'une efficacité d'amplification E , ou proportion moyenne de molécules produites à chaque cycle ($E = 2$ idéalement et $E = 1$ si aucune molécule n'est produite). Expérimentalement, E varie entre 1,78 et 1,97 [2]

$$N = N_0 \cdot E^n \quad (2)$$

Soit sous forme logarithmique :

$$\log N = \log N_0 + n \cdot \log E \quad (3)$$

⁷Par dilution limite puis analyse statistique

A partir d'une courbe expérimentale du nombre de molécules amplifiées N en fonction du nombre de cycle, il est possible d'en déduire la valeur de N_0 .

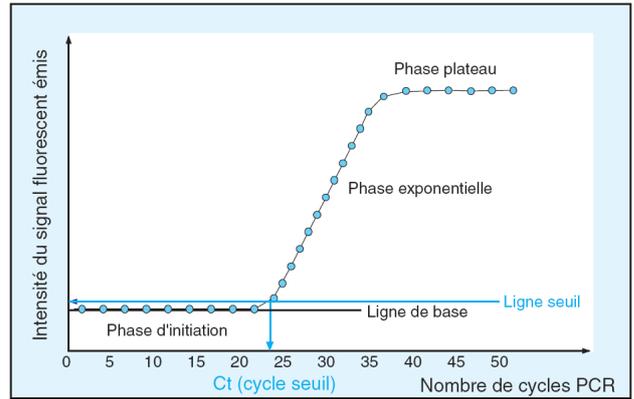


FIG. 1 – Profil d'une courbe PCR temps réel. Le nombre de copies est fonction du nombre de cycles. Plusieurs phases sont à distinguer : une phase en delà de laquelle aucun signal n'est détecté, une phase linéaire et une phase plateau où la PCR atteint un maximum [2]

La principale difficulté réside dans la détermination du cycle PCR au delà duquel la mesure est significative et ne correspond plus à du bruit. On parle de cycle seuil Ct , et il s'agit en quelque sorte de la sensibilité de l'instrumentation de détection.

La quantification temps-réel est plus fiable que la PCR conventionnelle, dans la mesure où l'ensemble du profil d'amplification est connu, permettant d'identifier rapidement des réactions dont l'efficacité d'amplification dévie, par exemple en présence d'un inhibiteur de la polymérase [3].

3 La fluorométrie temps-réel

Le principe du temps-réel en PCR repose sur la visualisation de la phase exponentielle de la réaction par une fluorochrome spécifique à l'ADN ou par une sonde oligonucléotique fluorescente spécifique à une séquence.

3.1 L'instrumentation

Les cycles de PCR quantitative temps-réel sont effectués dans des tubes fermés, limitant ainsi le risque de contamination, et sous un contrôle très précis de la température dans un thermocycleur. Les performances de l'instrumentation dépendent fortement :

1. du rapport signal-sur-bruit, fonction de la technologie de détection employé et de la qualité des optiques
2. de la répétabilité de la mesure, fonction de l'homogénéité en température des échantillons.

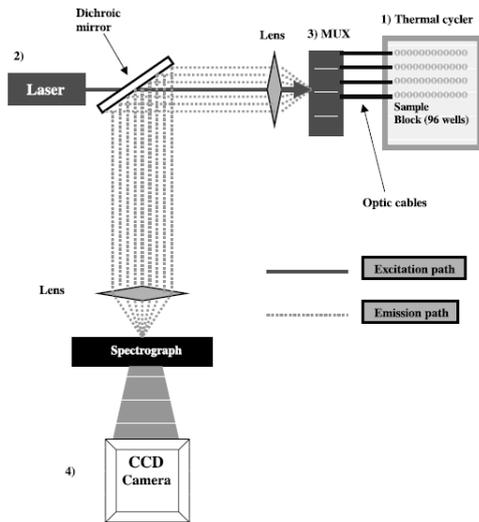


FIG. 2 – Principe de base d'un système PCR temps-réel [4], composé 1) d'un thermocycleur 2) d'un laser de fluorescence 3) d'un multiplexeur pour répartir l'excitation sur l'ensemble des échantillons 4) d'une caméra CCD pour mesurer la fluorescence émise

3.2 Les agents intercalants

Les agents intercalants sont des fluorochromes spécifiques à l'ADN double brin nouvellement synthétisé, le plus utilisé étant le SYBRTMGreen I. Sous une excitation ultra-violet, ce dernier émet un signal de fluorescence, dont la mesure quantifie l'ADN double brin, qu'il soit d'intérêt ou non. C'est là la principale limitation de cette technique : sa non spécificité, ainsi que sa dépendance à la « masse » d'ADN. Pourtant, cette méthode reste très sensible et reproductible, peu cher, et aide grandement dans les études de faisabilité et de calibration.

3.3 les sondes d'hydrolyse ou sondes TaqManTM

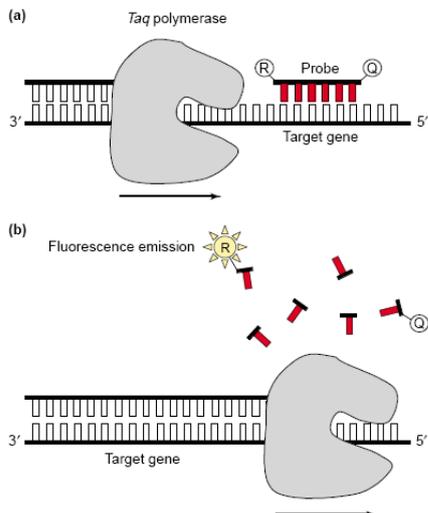


FIG. 3 – Principe d'une sonde d'hydrolyse [5]

Il s'agit d'une sonde fluorogénique spécifique à la séquence que l'on souhaite quantifier. Lors de la phase d'hybridation la sonde vient s'accrocher à la séquence, mais ne génère aucune fluorescence du fait du phénomène de *quenching*. Au cours de la phase d'élongation, l'activité hydrolytique de l'enzyme Taq polymérase vient cliver *reporter* et *quencher*, générant un signal détectable dans la longueur d'onde d'émission du *reporter*.

3.4 Les sondes FRET en tandem ou sondes LightCyclerTM

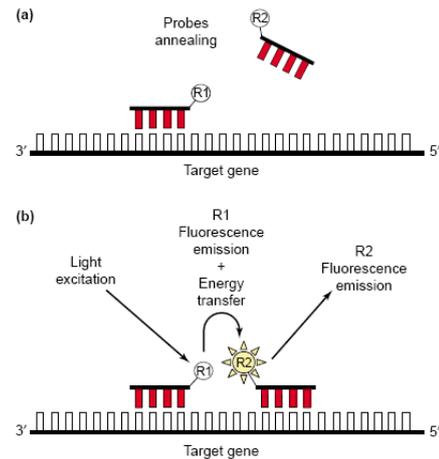


FIG. 4 – Principe d'une sonde FRET en tandem [5]

Il s'agit aussi d'une sonde fluorogénique spécifique à une séquence. Deux sondes sont conçues : une première contenant un *reporter*, une deuxième un *quencher* réémettant dans une autre longueur d'onde, et dont les deux sites d'accroche sur la séquence sont proches. Lors de la phase d'hybridation, en excitant dans la longueur d'onde du *reporter*, le signal recueilli dans la la longueur d'onde du *quencher* correspond à la PCR.

3.5 Les sondes d'hybridation phare ou sondes beaconTM

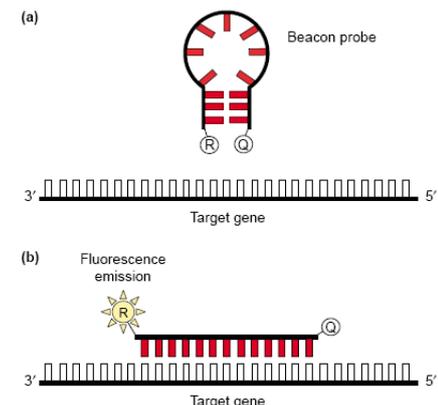


FIG. 5 – Principe d'une sonde d'hybridation phare [5]

Il s'agit d'une sonde fluorogénique spécifique à une séquence, et dont la conformation tridimensionnelle à l'état non lié fait qu'un *reporter* et un *quencher* non fluorescent⁸ se trouvent très près l'un de l'autre, n'émettant ainsi aucun signal. Lorsque la sonde se lie à la séquence, le changement de conformation spatial induit sépare les deux fluorophores et le *reporter* peut générer un signal mesurable dans sa longueur d'onde d'émission. La principale difficulté de ces sondes réside dans le design crucial de la boucle.

3.6 Historique et technologies des thermocycleurs

Le premier thermocycleur disponible, le ABI 7700, fut commercialisé par Applied Biosystems en 1997 et reste encore aujourd'hui le *best-seller* des automates de PCR. Il s'agit d'un block cylindrique de 96 puits couplé à un fluomètre. Peu de temps après, Roche Diagnostics arriva sur le marché avec le LightCycler, à l'origine développé par Idaho Technologies[3].

Il existe aujourd'hui de très nombreux compagnies et automates. Ils peuvent être classés en deux catégories [2] :

1. Les systèmes haut débit, pouvant analyser au moins 96 PCR simultanément : ABI Prism™7000, 7700 et 7900 (Applied Biosystems), iCycler™(Bio-rad), MX-4000™(Stratagene) et le DNA Engine Opticon™(MJ Research). A l'exception de l'ABI Prism™7700 qui comprend une résistance électrique et un groupe froid, ces systèmes un thermocycleur 96 puits chauffés par des modules à effet Peltier et/ou effet Joule, pour des vitesses de l'ordre de 1 à 3° C par seconde.
2. Les systèmes flexibles : LightCycler™(Roche Diagnostics), SmartCycler™(Eurogentec) et Rotor-Gene™(Corbett Research). Le contrôle de la température se fait par de l'air homogénéisé puis pulsé ou par rotation d'un carroussel contenant des capillaires où ont lieu les PCR, pour des vitesse pouvant atteindre jusqu'à 20° C par secondes

4 Quelles applications pour la PCR et la RT-PCR quantitative temps-réel ?

Les technologies de PCR et de RT-PCR s'inscrivent dans le cadre plus large de la problématique

⁸Aussi appelé *dark quencher*, par exemple les *black hole quencher*™, couvrant un très large spectre d'absorption et restituant l'énergie sous forme de chaleur. Le principal avantage est une grande efficacité de *quenching*, limitant ainsi le bruit de fond

de l'identification et de la quantification des acides nucléiques.

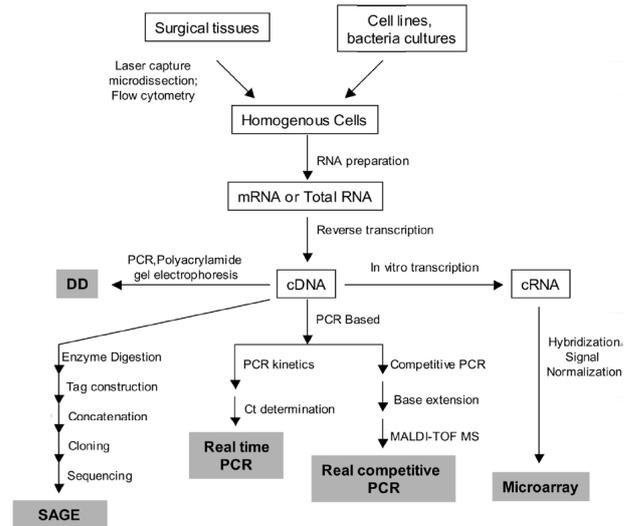


FIG. 6 – Techniques de quantification des acides nucléiques[6]

La détection et la quantification d'ADN par PCR et d'ARN par RT-PCR trouve de très nombreuses applications. En Microbiologie clinique, ces techniques offrent une détection et une quantification d'agents pathogènes bactériens et fongiques bien plus sensible et rapide que les diagnostics traditionnels, dont la durée pouvait s'exprimer en jours voir en semaines. En virologie, pouvoir déterminer par quantification absolue la titration virale du SIDA ou des hépatites est d'un grand intérêt dans les phases précoces de la maladie et dans le suivi des traitements, notamment antiviraux [2]. Des limitations sont cependant à prendre en compte :

1. La grande variabilité génétique des virus, nécessitant donc de développer des essais spécifiques à chaque sous-classe virale.
2. Le risque de faux positif par contamination extérieure.
3. Le risque de vrai négatif par inhibition de l'ADN polymérase. L'analyse en parallèle d'un contrôle interne positif permettrait d'y remédier.

En oncologie clinique, les applications concernent essentiellement [2] :

1. La recherche de transcrits de fusion, principalement dans les cas de leucémies, avant ou après une atogreffe, pour évaluer la réponse à un traitement, pour prédire une rechute ou évaluer la maladie résiduelle minimale.
2. La quantification d'oncogènes, l'amplification génique faisant partie des critères d'indication de thérapies ciblées, comme l'Herceptin® pour le cancer du sein métastatique.

Dans le domaine de l'expression génique, les techniques de RT-PCR sont largement utilisées pour [2] :

1. Estimer le niveau d'expression d'un gène.
2. Analyser l'expression des variants d'épissage génique.
3. Valider les résultats des puces à ADN

5 Conclusion

La PCR et la RT-PCR quantitative peuvent se résumer en quatre mots [2] :

1. Spécificité
2. Sensibilité
3. Bonne reproductibilité
4. linéarité pour une dynamique de huit ordres de grandeur

Ajouter à cela le fait d'être rapide, automatisable et en tube fermé, et vous obtenez une technologie d'un immense intérêt et bénéfique pour la médecine et la biologie.

Références

- [1] <http://fr.wikipedia.org/>. Réaction en chaîne par polymérase. *Wikipedia France*, 2006.
- [2] C. Tse and J. Capeau. Quantification des acides nucléiques par PCR quantitative en temps réel. *Annale de Biologie Clinique*, 61 :279–293, 2003.
- [3] Jochen Wilhem and Alfred Pingoud. Real-time Polymerase Chain Reaction. *ChemBioChem*, 4 :1120–1128, 2003.
- [4] Yong-lee Ong and Alexandra Irvine. Quantitative real-time PCR : a critique of method and practical considerations. *Hematology*, 7(1) :59–67, 2002.
- [5] Simone Mocellin, Carlos R. Rossi, Pierluigi Poliati, Donato Nitti, and Francesco M. Marincola. Quantitative real-time PCR : a powerful ally in cancer research. *Trends in Molecular Medicine*, 9(5), May 2003.
- [6] Chumming Ding and Charles R. Cantor. Quantitative analysis of nucleic acids - the last few years of progress. *Biochemistry and Molecular Biology*, 37(1) :1–10, 2004.
- [7] David G. Ginzinger. Gene quantification using real-time quantitative PCR : an emerging technology hits the mainstream. *Experimental Hematology*, 30 :503–512, 2002.
- [8] H. Cavé, C. Acquaviva, I. Bièche, D. Brault, F. de Fraipont, F. Fina, S. Loric, L. Maison-neuve, Namour F., and S. Tuffery. La RT-PCR en diagnostique clinique. *Annale de Biologie Clinique*, 61 :635–644, 2003.
- [9] S. A. Bustin, V. Benes, T. Nolan, and Pfaffl M. W. Quantitative real-time RT-PCR - a perspective. *Journal of Molecular Endocrinology*, 34 :597–601, 2005.
- [10] Annapaula Guilietti, Lut Overberg, Dirk Valckx, Brigitte Decallone, Roger Bouillon, and Chantal Mathieu. An Overview of real-time quantitative PCR : Applications to quantify cytokine gene expression. *Methods*, 25 :386–401, 2001.
- [11] Stephen A. Bustin and Tania Nolan. Pitfalls of quantitative real-time Reverse-Transcription Polymerase Chain Reaction. *Methods*, 25 :386–401, 2001.
- [12] Stephen Bustin and Mueller Reinhold. Real-time reverse transcription PCR (qRT-PCR) and its potential use in clinical diagnosis. *Clinical Science*, 109 :365–379, 2005.
- [13] Phillip S. Bernard and Wittwer Carl T. Real-time PCR technology for cancer diagnostics. *Clinical Chemistry*, 48(8) :1178–1185, 2002.

La PCR en temps réel: principes et applications

Elyse Poitras et Alain Houde*

La technologie de PCR en temps réel devient de plus en plus populaire dans différents secteurs d'activité. Cette technologie est basée sur la détection et la quantification d'un reporter fluorescent dont l'émission est directement proportionnelle à la quantité d'amplicons générés pendant la réaction de PCR. Étant donné qu'elle utilise généralement des systèmes en tubes fermés et que la quantification ne requiert aucune manipulation post-amplification, les problèmes de contamination post-PCR par les amplicons sont significativement réduits. Le processus complet est automatisé du début à la fin rendant cette technologie très performante pour des applications d'analyses à grande échelle. Cet article présente une description des principes à la base de la PCR en temps réel, des différentes technologies de détection des amplicons et des exemples d'applications courantes.

Historique

Russell Higuchi fut l'un des premiers à faire l'analyse des cinétiques de la PCR (polymerase Chain Reaction) en élaborant un système qui détectait le produit de la PCR au fur et à mesure qu'il s'accumulait. Ce système en « temps réel » utilisait le bromure d'éthidium comme agent intercalant dans chacune des réactions d'amplification et un thermocycleur modifié pour stimuler l'émission des échantillons par rayonnements UV. L'émission de la fluorescence était détectée à l'aide d'une caméra CCD (charge-coupled device). Une augmentation de l'émission de la fluorescence était observée lorsque le bromure d'éthidium se fixait à l'ADN double brin produit au cours de l'amplification. En traçant l'augmentation de l'émission de fluorescence en fonction du nombre de cycles, le système produit des courbes d'amplification exhibant un schéma plus complet du processus de la PCR que la simple détermination d'amplicons (produits d'amplification) accumulés en fin d'amplification (Higuchi et al, 1992).

Principe

Depuis son invention, la PCR est devenue la technique la plus utilisée pour la détection de l'ADN et de l'ARN. À partir d'une simple copie d'une séquence particulière d'acides nucléiques, cette séquence peut être spécifiquement amplifiée et détectée. Sa nature exponentielle rend cette technique attrayante pour des analyses quantitatives. Théoriquement, il existe une relation quantitative entre la quantité de la séquence cible

de départ et la quantité du produit amplifié à n'importe quel cycle. En pratique, il n'est pas rare que les réactions de PCR en replica donnent des taux différents d'amplicons. Le développement de la PCR quantitative en temps réel a éliminé les variabilités traditionnelles associées à la PCR quantitative et permet la quantification du produit de la PCR de façon fiable et routinière.

Au début de la réaction PCR, les réactifs sont en excès mais en concentration assez faible afin d'éviter que la renaturation des amplicons n'entre en compétition avec l'hybridation des amorces (primers). L'amplification est alors réalisée de façon constante à un taux exponentiel à l'aide d'une ADN polymérase thermostable. Après la phase exponentielle, la réaction d'amplification entre dans une phase linéaire où le taux d'amplification devient extrêmement variable, même au niveau de replica d'un même échantillon, à cause d'une compétition entre la renaturation des amplicons et l'hybridation des amorces. Suit ensuite une phase plateau où le taux d'amplification décroît à près de zéro générant très peu d'amplicons.

Afin de recueillir des données quantitatives avec précision, chacun des échantillons doit être analysé dans sa phase exponentielle d'amplification qui est la phase la plus reproductible de la réaction de PCR. La PCR en temps réel fait donc le suivi de la fluorescence émise pendant la réaction avec un indicateur de la production des amplicons durant chaque cycle, à l'opposé de la PCR quantitative conventionnelle où les amplicons ne sont détectés qu'à la toute fin du processus.

La technologie de la PCR en temps réel est basée sur la détection et la quantification d'un «reporter» fluorescent. L'augmentation du signal fluorescent est directement proportionnelle à la quantité d'amplicons générés durant la réaction de PCR. En observant la quantité de fluorescence émise à chaque cycle, il devient possible de suivre la réaction PCR durant sa phase exponentielle où la première augmentation significative dans la quantité d'amplicons est en corrélation directe avec la quantité initiale de la matrice originale cible (template). Plusieurs instruments de PCR en temps réel sont présentement sur le marché. Ces appareils utilisent généralement un système en tubes fermés et la quantification ne requiert aucune manipulation post-amplification, ce qui minimise ou élimine les problèmes de contamination par les amplicons suite à la réaction de PCR et réduit le temps d'analyse (Bustin, 2000). Le processus complet est donc automatisé du début à la fin rendant ainsi cette technologie intéressante pour des applications d'analyses à grande échelle (high-throughput) (Martell et al, 1999).

Technologies de détection

Tous les systèmes de PCR en temps réel reposent donc sur la détection et la quantification d'un émetteur fluorescent pendant le processus d'amplification et l'augmentation du signal d'émission fluorescente est directement proportionnelle à la quantité d'amplicons produits durant la réaction. Il existe deux principes généraux pour la détection quantitative des amplicons : les agents se liant à l'ADN double brin (ex. SYBR Green I) et les sondes fluorescentes. Pour cette dernière catégorie, il existe présentement quatre technologies principales : hydrolyse de sondes (Taqman assay), hybridation de 2 sondes (HybProbes), balises moléculaires (Molecular Beacons) et amorces scorpion (Scorpion primers). Selon Wittwer et al. (1997), ces différentes technologies de détection auraient une sensibilité équivalente. Cependant, ces technologies présentent des différences au niveau de la spécificité (Bustin, 2000).

1) Agents se liant à l'ADN double brin (Double-stranded DNA binding dyes: Lightcycler assay)

Les molécules qui se lient à l'ADN double brin peuvent être divisées en deux classes: les agents intercalants comme le bromure d'éthidium (Higuchi et al, 1992), le YO-PRO-1 (Ishiguro et al, 1995; Tseng et al, 1997), le SYBR Green I (Morrison et al, 1998) et les agents se fixant au sillon mineur (minor groove binders) comme le Hoeschst 33258 (Searle et Embrey, 1990; Nielsen, 1991). Leur émission fluorescente augmente lorsque qu'ils sont liés à l'ADN double brin. Pour être utilisés dans une réaction de PCR en temps réel, ces agents doivent rencontrer deux exigences : augmenter en fluorescence lorsque lié à l'ADN double brin et ne pas inhiber la réaction de PCR. Le SYBR Green I, dont le mécanisme de liaison n'est pas bien défini, est l'agent le plus fréquemment utilisé. Ses avantages sont

qu'il est économique, facile à utiliser et possède plus de sensibilité que le bromure d'éthidium sans inhiber la réaction d'amplification.

Lors de la réaction d'amplification par PCR, le colorant libre en solution exhibe peu de fluorescence. Durant l'étape d'élongation, une augmentation de la fluorescence est associée à la quantité de colorant se fixant à l'ADN double brin naissant. Lorsque suivi en temps réel, l'augmentation du signal de fluorescence est observée pendant l'étape de polymérisation et l'émission fluorescente décroît complètement lorsque l'ADN est dénaturé à l'étape suivante. Conséquemment, l'émission de fluorescence est mesurée à la fin de chaque étape d'élongation pour chacun des cycles par un système de lecture intégré à l'appareil de PCR en temps réel qui permet de suivre l'augmentation de la quantité d'ADN amplifié durant la réaction (Bustin, 2000) (figure 1).

La technologie basée sur le SYBR Green I ne nécessite aucune sonde fluorescente mais sa spécificité repose entièrement sur ses amorces (Bustin, 2000). Elle ne requiert donc aucune expertise particulière pour le design des sondes fluorescentes et n'est pas affectée par des mutations dans l'ADN cible qui influencent l'hybridation des sondes spécifiques (Mackay et al, 2002). Étant donné que le SYBR Green I se fixe à n'importe quelle molécule d'ADN double brin, cette technologie présente une certaine versatilité puisque le même agent peut être utilisé pour détecter plus d'un produit d'amplification dans la même séquence réactionnelle.

Cette technologie présente aussi certains désavantages: 1) étant donné que l'ADN double brin total émet des signaux, il devient impossible en cours de réaction de s'assurer de la spécificité des amplicons ou de discriminer les différents amplicons dans le cas de multiplexage; 2) le mauvais appariement (mis-priming), générant souvent des bandes d'ADN superflues observables sur gel d'électrophorèse, peut conduire à des faux positifs ou une surestimation de la quantification; 3) l'émission de fluorescence peut être biaisée par la masse moléculaire de l'ADN amplifié par un amplicon plus long qui fixera davantage de molécules fluorescentes par rapport à un amplicon plus court dans la même réaction (Bustin, 2000).

2) Hydrolyse de sondes (Hydrolysis probes: Taqman assay)

La technologie *Taqman* est basée sur l'activité 5'-exonucléasique de la Taq polymérase pour hydrolyser une sonde hybridée à sa séquence cible sur l'amplicon durant l'étape d'hybridation/extension de la PCR. Un fluorochrome émetteur (reporter) (ex. FAM : 6-carboxy-fluorocein) est fixé à l'extrémité 5' de la sonde d'hybridation et son émission est inhibée par un second fluorochrome suppresseur (quencher) présent à l'extrémité 3' (ex. TAMRA : 6-carboxy-tetramethyl-rhodamine). Lorsque stimulé, le fluorochrome émetteur transfère son énergie au fluorochrome suppresseur voisin par le principe

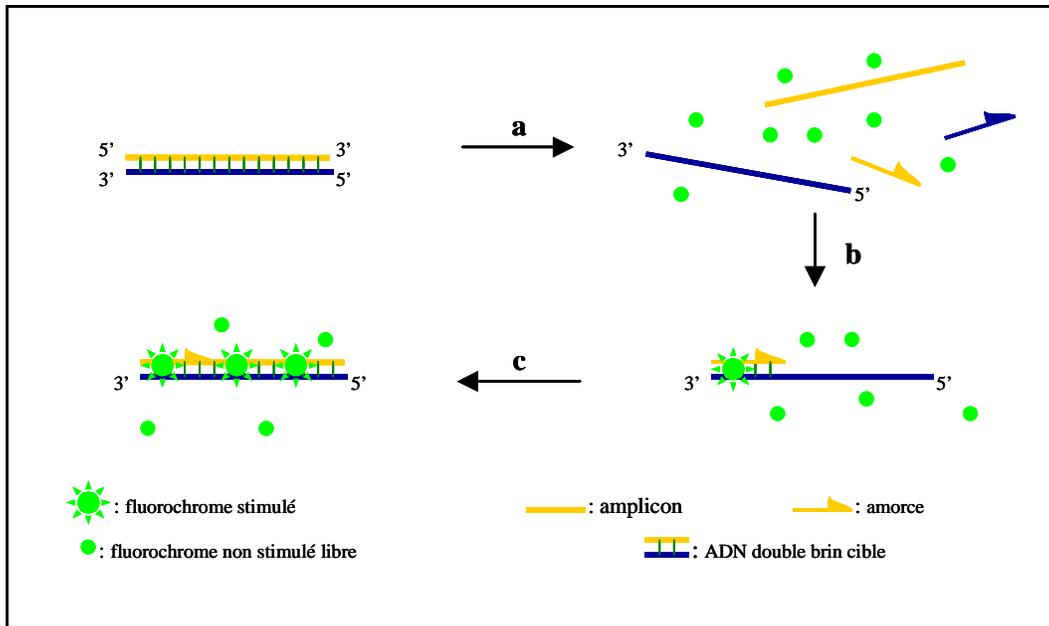


Figure 1 : *Agents se liant à l'ADN double brin (Double-stranded DNA binding dyes: Lightcycler assay)*. (a) Durant la dénaturation, le SYBR Green I libre exhibe peu de fluorescence. (b) À la température d'appariement, quelques molécules se lient au double brin d'ADN naissant résultant en une émission de fluorescence lors de l'excitation. (c) Durant la phase de polymérisation, de plus en plus de molécules se lient au brin naissant et l'accroissement de la fluorescence peut-être suivi en temps réel.

FRET (fluorescence resonance energy transfer) qui dissipe cette énergie sous forme de chaleur plutôt que d'émettre de la fluorescence (Mackay et al, 2002). Étant donné que l'activité 5'-exonucléasique de la Taq polymérase est spécifique à l'ADN double brin, les sondes libres en solution demeurent intactes et aucune fluorescence n'est émise. Lors de l'étape d'hybridation, la sonde et les amorces se fixent à leurs séquences complémentaires respectives. À l'étape suivante, la Taq polymérase débute l'élongation du nouveau brin d'ADN à partir de l'amorce jusqu'à ce qu'elle rencontre sur son passage la sonde hybridée qu'elle déplace et hydrolyse avec son activité 5'-exonucléasique. Le reporter est alors libéré de l'environnement du suppresseur permettant ainsi l'émission de fluorescence qui augmente à chaque cycle proportionnellement au taux d'hydrolyse de la sonde (figure 2A).

Comme la Taq polymérase hydrolysera la sonde seulement lorsque celle-ci est hybridée à sa séquence complémentaire, les conditions de température de l'étape de polymérisation doivent être ajustées de façon à permettre à la sonde de rester hybridée durant cette étape. La majorité des sondes ont une température de dissociation (T_m) autour de 70°C ou de 5 à 10°C plus élevée que les amorces. Par conséquent, la technologie *Taqman* utilise une étape combinée d'hybridation et de polymérisation à 60-62°C assurant l'hybridation et la stabilité de la sonde durant l'extension. Ceci permet aussi une activité 5'-exonucléasique maximale de la Taq polymérase mais, l'efficacité de l'activité de

polymérisation de l'enzyme sera légèrement réduite à cette température suboptimale. Pour de longs amplicons, une étape d'hybridation/polymérisation plus longue ou encore une augmentation de la concentration en Mn^{2+} ou Mg^{2+} pourrait s'avérer nécessaire pour stabiliser l'hybridation de la sonde à sa séquence cible.

Les principes à respecter dans le design des sondes *Taqman* sont aussi applicables aux autres sondes linéaires et comprennent comme règles générales : 1) une longueur de 20-40 nucléotides, 2) un contenu en G-C variant de 40-60%, 3) aucun patron de séquence répétée, 4) aucune séquence permettant une hybridation ou un chevauchement avec les amorces, 5) un A, un C ou un T à l'extrémité 5' parce qu'un G supprime la fluorescence de l'émetteur même après clivage et, 6) un T_m de 5 à 10°C plus élevé que les amorces afin de s'assurer qu'elles s'hybrident avant les amorces et qu'elles demeureront hybridées pendant l'étape combinée d'hybridation et de polymérisation (Bustin, 2000; Mackay et al, 2002).

Les sondes fluorescentes possèdent comme avantage par rapport aux agents se liant à l'ADN une spécificité accrue et une meilleure capacité de multiplexage. La spécificité d'hybridation entre la sonde fluorescente et sa séquence d'ADN cible réduit significativement l'émission de fluorescence non spécifique due à des mauvais appariements ou des dimères d'amorces (primer-dimers). Des réactions multiplexes peuvent être élaborées en utilisant des fluorochromes émetteurs distincts liés à des sondes différentes dans une même réaction de PCR.

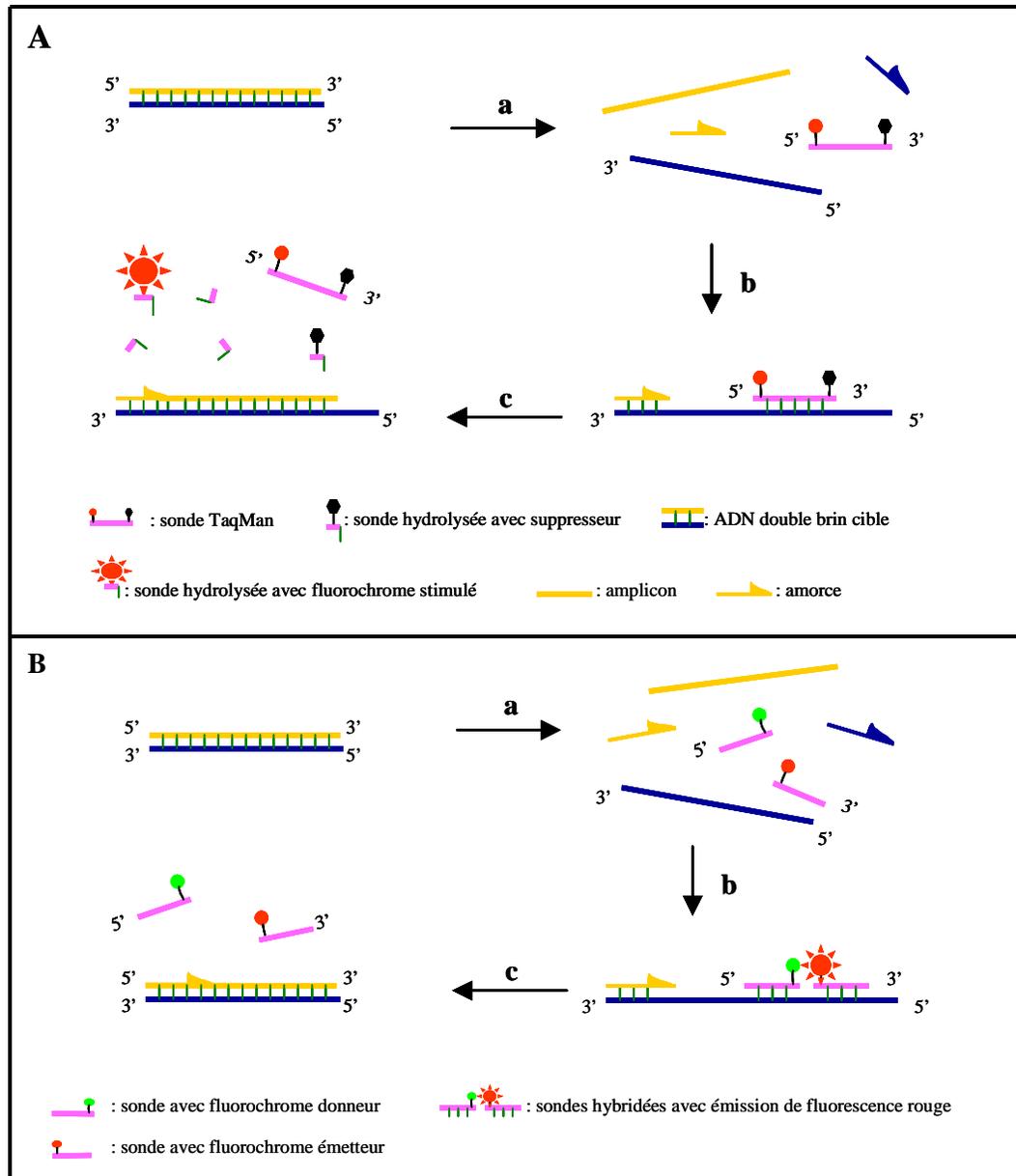


Figure 2 : **A: Hydrolyse de sondes (Hydrolysis probes: Taqman assay)** (a) Durant l'étape de dénaturation, la sonde est libre en solution. (b) À la température d'appariement, la sonde et les amorces s'hybrident à leurs séquences cibles respectives et la proximité des fluorochromes permet l'inhibition de la fluorescence. La polymérisation débute. (c) La polymérase déplace et hydrolyse la sonde. Le fluorochrome émetteur est libéré de l'environnement du suppresseur permettant ainsi l'émission de la fluorescence. **B: Hybridation de 2 sondes (Hybridization probes)** (a) Durant l'étape de dénaturation, les deux sondes demeurent séparées et en solution. (b) À la température d'appariement, les sondes s'hybrident à leurs séquences cibles respectives et la proximité des fluorochromes permet l'émission de fluorescence rouge par le principe FRET. (c) Les sondes retournent libres en solution.

La technologie *Taqman* est toutefois moins efficace et moins flexible que d'autres technologies en temps réel pour la détection de mutations spécifiques (Bustin, 2000; Mackay et al, 2002).

3) Hybridation de 2 sondes (Hybridization probes)

Cette technologie est aussi connue sous le nom de

HybProbes et repose sur l'utilisation de deux sondes linéaires complémentaires à une séquence cible pour maximiser la spécificité du signal (Wittwer et al, 1997). Une première sonde, bloquée à son extrémité 3' afin de prévenir son extension durant l'étape d'élongation, transporte en 3' un fluorochrome donneur (FITC) qui produit une lumière fluorescente verte lorsque excité

par une source de lumière. Son spectre d'émission est plus large que celui du fluorochrome accepteur (Red 640 ou Red 705) attaché à l'extrémité 5' d'une seconde sonde. En solution, les deux sondes sont libres et séparées. Étant donné que le transfert d'énergie par le principe FRET dépend de la distance entre les deux fluorochromes, il en résulte alors seulement un bruit de fond de fluorescence verte émis par le fluorochrome donneur. Pendant l'étape d'hybridation, les deux sondes se fixent à leurs séquences cibles respectives localisées à moins de 10 nucléotides (Mackay et al, 2002) dans un arrangement en tête-à-queue. La proximité des deux fluorochromes permet le transfert énergétique de la fluorescéine verte par le principe FRET au fluorochrome accepteur rouge et provoque son émission fluorescente. Pendant l'étape de polymérisation, les deux sondes retournent de façon indépendante en solution ce qui supprime l'émission de fluorescence rouge (figure 2B). L'accroissement de la fluorescence rouge est proportionnel à la quantité d'ADN synthétisé durant la réaction PCR et commence à diminuer lorsque la quantité d'amplicons produits devient suffisamment importante pour provoquer une compétition de l'ADN amplifié avec l'hybridation simultanée des 2 sondes sur l'ADN cible (Mackay et al, 2002).

Cette technologie possède donc une grande spécificité et permet aussi une grande flexibilité dans le design des sondes. De plus, comme les sondes ne sont pas hydrolysées, elles sont réutilisées à chacun des cycles (Bustin, 2000). En plus du fait que la séquence cible doit être localisée vers l'extrémité 3' de l'amplicon et que le T_m des deux sondes doit être similaire, les principes généraux dans le design des sondes *Taqman* sont applicables pour cette technologie.

4) Balises moléculaires (sonde d'hybridation en épingle à cheveux: *Molecular Beacons*)

Une balise moléculaire est une sonde d'hybridation d'ADN en forme d'épingle à cheveux. La portion de la sonde qui compose la boucle est complémentaire à la séquence cible d'ADN. Le tronc de la balise moléculaire est formé de deux bras avec des séquences complémentaires (Tyagi et Kramer, 1996). Un émetteur fluorescent (FAM, TAMRA, TET, ROX) est fixé à l'extrémité d'un des bras et un suppresseur (quencher) (4-(4'-diméthylamino-phenylazo)-benzene: DABCYL) est fixé à l'extrémité de l'autre bras. Le suppresseur est un chromophore non-fluorescent qui dissipe en chaleur l'énergie du fluorochrome émetteur par le principe FRET. Les sondes libres adoptent en solution une structure en épingle à cheveux et le tronc maintient les bras ensemble pour une suppression efficace de l'émission de fluorescence. Pendant l'étape d'hybridation, lorsque la sonde rencontre une séquence qui lui est complémentaire, elle adopte une conformation transitoire qui force le tronc à se séparer libérant ainsi les 2 bras. La sonde s'hybride alors préférentiellement à sa séquence complémentaire cible sur la matrice. Dans cette conformation, le fluorochrome

émetteur est éloigné de son suppresseur restaurant ainsi l'émission de fluorescence qui peut être détectée alors que les autres balises moléculaires demeurent en structure d'épingle à cheveux (position fermée). Si la séquence d'ADN cible ne correspond pas parfaitement (mismatch) à la séquence de la balise moléculaire aucune hybridation et, par conséquent, aucune émission fluorescente ne survient. Ceci est principalement dû aux propriétés thermodynamiques de la structure de la balise moléculaire qui favorisent surtout la formation de l'épingle à cheveux sauf en cas d'hybridation parfaite des séquences (Bustin, 2000) (figure 3A). La spécificité de cette technologie est telle qu'elle permet de détecter les différences de séquences au nucléotide près, ce qui peut être parfois difficile à réaliser avec les technologies de sondes linéaires (Giesendorf et al, 1998; Marras et al, 1999). Elle constitue donc une technologie efficace pour la détection et le criblage à grande échelle des SNPs (single nucleotide polymorphisms) (Tyagi et al, 1998; Mhlanga et Malmberg, 2001). Les sondes sont dessinées pour demeurer intactes durant la réaction d'amplification et doivent se réhybrider à la cible à chaque cycle pour mesurer le signal constituant ainsi un autre avantage par rapport aux sondes *Taqman* hydrolysées à chaque cycle.

Le désavantage le plus important associé à l'utilisation des balises moléculaires est le design des sondes d'hybridation. Un design optimal du tronc des balises moléculaires est crucial. Avec un design non adéquat, le tronc pourrait adopter une conformation différente qui éloignerait le fluorochrome émetteur de l'environnement immédiat du suppresseur résultant ainsi en une population de sondes mal supprimées et un bruit de fond important. En contrepartie, si les forces d'hybridation du tronc (dépendantes de la longueur et de la composition des séquences complémentaires des bras) sont trop importantes, elles peuvent interférer avec l'hybridation de la balise moléculaire à sa séquence complémentaire cible et l'émission de fluorescence. Un profil de dénaturation thermique précis de chacune des balises moléculaires doit être établi pour déterminer les caractéristiques de dissociation (melting) des bras.

Plusieurs variantes de cette technologie comme les Sunrise primers (maintenant sous appellation commerciale Amplifluor™ hairpin primers) (Nazarenko et al, 1997; Myakishev et al, 2001), les cyclicons (Kandimalla et Agrawal, 2000) et les amorces scorpion (Mhlanga et Malmberg, 2001) ont été proposées pour la détection d'acides nucléiques spécifiques dans des solutions homogènes.

5) Amorces scorpion (*Scorpion primer: self-fluorescing amplicon*)

Les amorces scorpion représentent une variante de la technologie des balises moléculaires. Les fluorochromes et la sonde (partie balise moléculaire) sont intégrés à même l'amplicon de façon irréversible durant l'amplification par

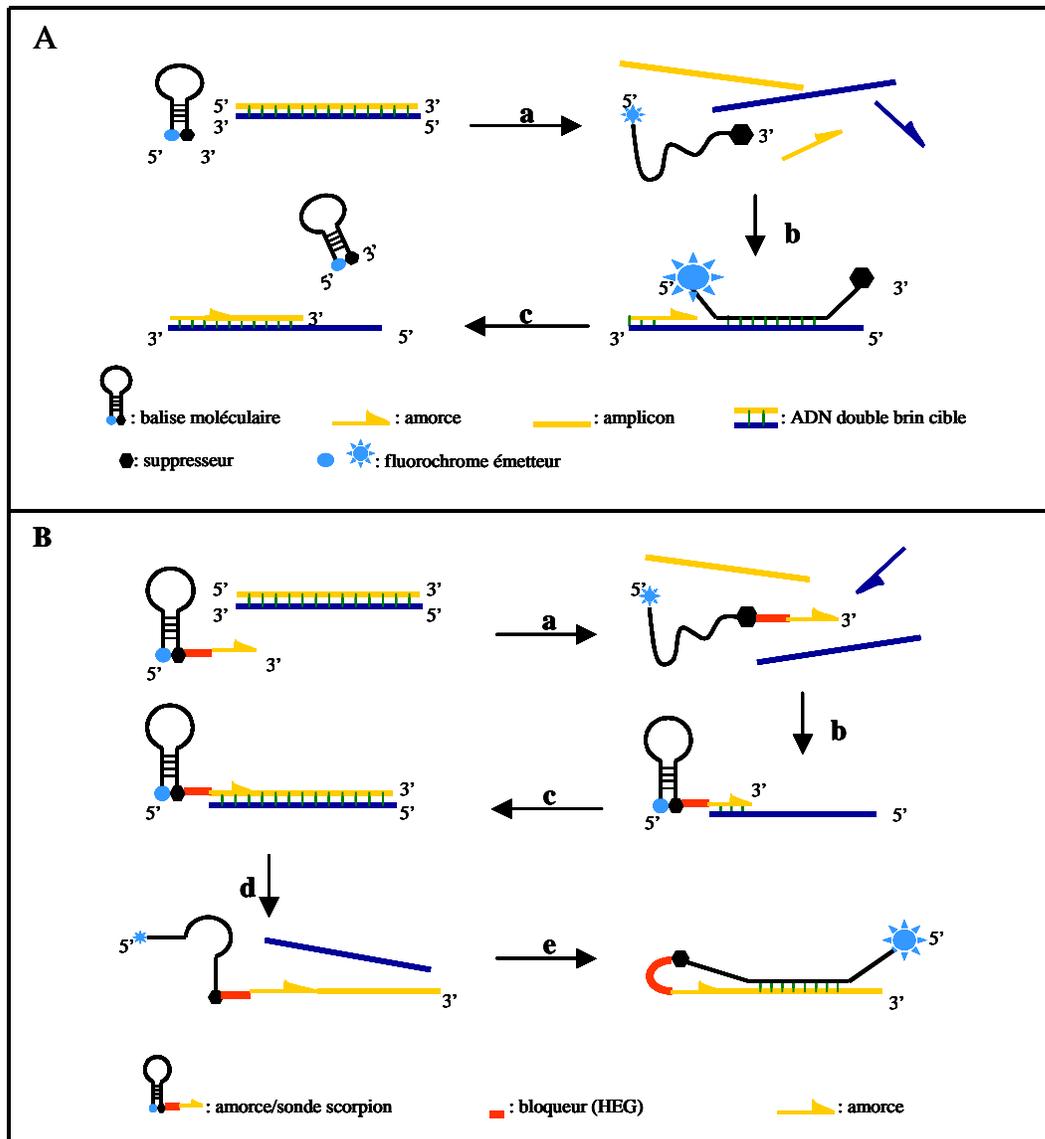


Figure 3 : **A: Balises moléculaires (Molecular Beacons).** (a) Durant l'étape de dénaturation, la balise moléculaire est sous forme relaxée et libre en solution mais la proximité des fluorochromes permet l'inhibition de la fluorescence. (b) Lorsque la sonde s'hybride à sa séquence cible, le fluorochrome émetteur est suffisamment éloigné de son supprimeur pour permettre l'émission de fluorescence. (c) À l'étape de polymérisation, la balise moléculaire retourne en solution sous forme d'épingle à cheveux. **B: Amorce scorpion (Scorpion primer).** (a) Durant l'étape de dénaturation, la balise moléculaire est sous forme relaxée et libre en solution mais la proximité des fluorochromes permet l'inhibition de la fluorescence. (b) L'amorce scorpion se fixe à sa séquence complémentaire cible. (c) Polymérisation du brin complémentaire. (d) Dénaturation des brins d'ADN. (e) Hybridation de la séquence complémentaire de la partie balise moléculaire à sa séquence cible permettant l'émission de fluorescence.

PCR. Le design d'une amorce/sonde scorpion est pratiquement similaire à celui des balises moléculaires. L'ajout d'une molécule d'hexéthylène glycol (HEG), aussi appelé bloqueur (stopper), est nécessaire pour empêcher la réplication de la balise moléculaire par l'ADN polymérase

pendant la réaction de PCR. Le HEG est donc situé après le fluorochrome supprimeur et est suivi d'une région amorce. Le fluorochrome émetteur porté en 5' de la région balise moléculaire peut être le FAM ou le ROX et le supprimeur est normalement le rouge de méthyl. La région amorce du

scorpion permet donc d'intégrer la balise moléculaire dans le nouvel amplicon pendant la réaction de PCR. La boucle de l'épingle à cheveux est dessinée dans le but de permettre l'hybridation de la sonde à sa séquence complémentaire cible située sur l'amplicon (figure 3B). Cette hybridation force l'épingle à cheveux à changer de conformation permettant ainsi l'émission de fluorescence (Whitcombe et al, 1999; Thelwell et al, 2000; Mackay et al, 2002). Thelwell et al (2000) rapportent que la technologie amorce/sonde scorpion est généralement plus efficace que les technologies *Taqman* et balises moléculaires, particulièrement dans un programme de PCR possédant des cycles très courts.

Cycle seuil (Threshold cycle)

Le concept du « cycle seuil » est à la base d'une quantification précise et reproductible pour les techniques fluorescentes en PCR. Les valeurs de fluorescence sont enregistrées au cours de chaque cycle et représentent la quantité d'amplicons produits en un point précis dans la réaction (figure 4). Plus il y a de matrices (template) à amplifier au départ de la réaction PCR, moins élevé sera le nombre de cycle requis pour atteindre un point où le signal d'émission de fluorescence sera statistiquement et significativement plus élevé que le bruit de fond (Gibson et al, 1996). Ce point est défini comme étant le cycle seuil (Ct) et apparaîtra toujours au cours de la phase exponentielle d'amplification. Par conséquent, la quantification n'est pas affectée par l'épuisement d'un des réactifs comme lors de la phase plateau ce qui explique pourquoi le système en temps réel est si reproductible. La valeur du Ct peut être traduite en un résultat quantitatif en la comparant avec les valeurs du Ct générées avec des matrices de quantification connues (Bustin, 2000).

Applications

La PCR en temps réel, à cause de sa capacité à produire des résultats rapides, spécifiques et quantitatifs, trouve de plus en plus d'applications dans différents domaines. Bien que cette liste ne soit pas exhaustive, voici les exemples d'application les plus courants.

La PCR en temps réel permet maintenant de réaliser des études plus fines d'expression génétique à partir de tissus ou de lignées cellulaires comme pour l'analyse quantitative de l'expression de différents gènes dans des études de modulation du cycle cellulaire avec des tissus exposés à des carcinogènes non génotoxiques (cic TNO BIBRA), les analyses de promoteurs dans des lignées cellulaires avec le gène reporter de la chloramphénicol acétyl transférase (Jeyaseelan et al, 2001), les études de modifications d'expression génétique dans les expériences drogues/réponses (drug/response), l'étude des ARNm variants résultant d'épissage alternatif (alternative splicing) (Vandenbroucke et al, 2001), etc. Elle peut aussi permettre de standardiser la quantité d'ADN de départ et d'évaluer la quantité d'ARNm en fin de réaction dans des études de

transcription *in vitro* (Liu et al, 2002).

La PCR en temps réel s'avère aussi un outil puissant pour des analyses de mutations comme les SNPs et des études de génotypage à grande échelle comme pour le gène du récepteur d'œstrogène (Täpp et al, 2000).

De plus en plus des tests utilisant la technologie des balises moléculaires sont dessinés pour la détection et la quantification rapide d'agents pathogènes viraux, bactériens et parasitaires. Vet et al (1999) ont déjà proposé un système multiplexe qui permet la détection et la quantification simultanées des rétrovirus HIV-1, HIV-2, des virus lymphotrophiques-T humains de type I et de type II avec un seuil de détection de 10 copies de génome et Poddar (1999) pour la détection d'adénovirus. Des tests de détection ont aussi été mis au point pour *Salmonella* avec une sensibilité de 2 CFU par réaction de PCR en temps réel (Chen et al, 2000) et de 1 CFU/ml après 6 h d'enrichissement pour *Escherichia coli* O157:H7 à partir d'échantillons de lait cru et de jus de pomme (Fortin et al, 2001). Chez les parasites, un test de PCR en temps réel a été développé pour *Toxoplasma gondii* (Lin et al, 2000). Plusieurs laboratoires travaillent présentement à la mise au point de nombreux autres tests diagnostiques pour différents agents pathogènes.

Traditionnellement, l'analyse du nombre de copies d'un plasmide pour évaluer la stabilité génétique des collections de cellules était effectuée par transfert de Southern (Southern blot). La PCR en temps réel permet maintenant de réaliser ces analyses de façon beaucoup plus rapide et plus précise. Elle peut aussi être utilisée dans l'évaluation de la quantité d'ADN chromosomique contaminant ou résiduel bactérien ou de mammifères dans la production de protéines recombinantes.

L'analyse de la bio-distribution d'un vecteur est un élément important dans l'évaluation de l'efficacité de protocoles de thérapie génique. La PCR en temps réel simplifie les études d'évaluation de la présence/absence d'ADN ou d'ARNm cible dans les différents tissus et permet de déterminer les conditions d'équilibre des ARNm exprimés (steady-state expressed mRNA).

Conclusion

Selon la littérature, la technique la plus fréquemment utilisée jusqu'à maintenant est la technologie des sondes d'hydrolyse (*Taqman assay*) bien que sa popularité soit principalement due à sa maturité commerciale. Cette technologie s'est avérée plus efficace que celle des agents intercalants au niveau de la spécificité et du multiplexage. Le développement des technologies de détection à base de balises moléculaires amène encore davantage de spécificité et de précision entre autre pour la détection des SNPs et des analyses de génotypage comparativement aux technologies basées sur les sondes linaires. La PCR en temps réel s'avère très intéressante pour des applications d'analyses à grande échelle étant donné que le processus complet est

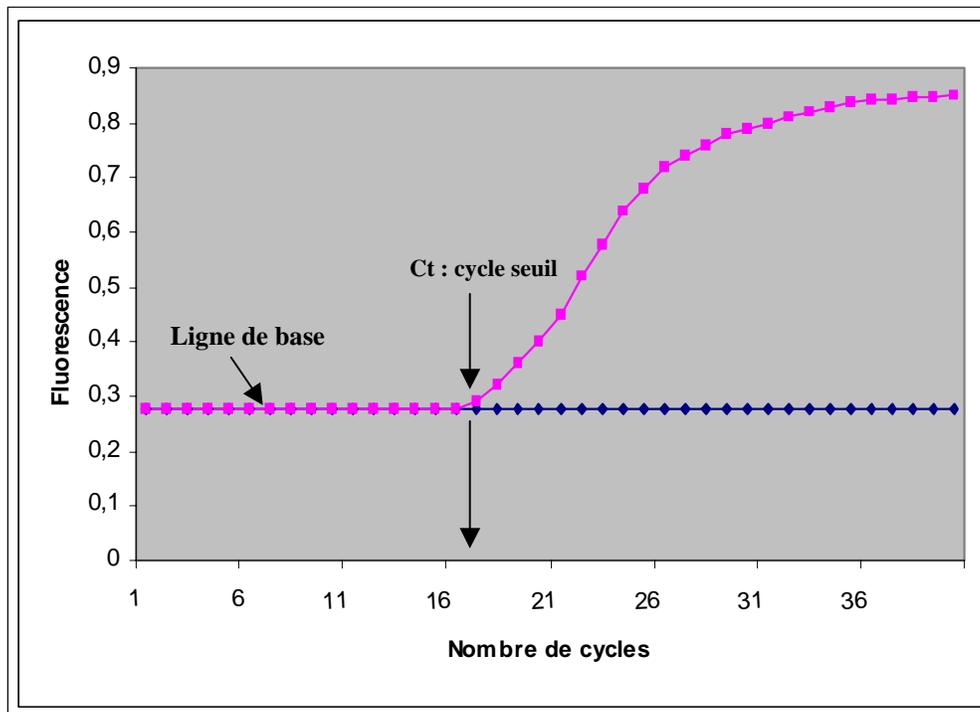


Figure 4 : Modèle graphique de la PCR en temps réel où l'intensité de la fluorescence est exprimée en fonction du nombre de cycles. L'intensité de la fluorescence à chaque cycle est proportionnelle à la concentration d'amplicons, le cycle seuil (Ct) représente le nombre de cycles requis où le signal d'émission de fluorescence est statistiquement et significativement plus élevé que la ligne de base.

automatisé du début à la fin. Comme cette technique comprend généralement des systèmes en tubes fermés et que la quantification ne requiert aucune manipulation post-amplification, les problèmes de contamination post-PCR par les amplicons sont fortement réduits.

Des analyses d'expression de gènes, de réponse cellulaire à différentes drogues (drug/response), de détection de mutations, de génotypage, de détection et quantification d'agents pathogènes, de quantification d'ADN et d'ARN, d'essais d'expression et de distribution pour la thérapie

génique, de quantification du nombre de copies dans une collection de cellules et, finalement, d'évaluation d'ADN résiduel sont, présentement, quelques exemples d'applications de plus en plus répandus de la PCR en temps réel.

La PCR en temps réel est donc un outil puissant et les développements récents dans les différentes technologies de détection laissent entrevoir des applications plus innovatrices les unes que les autres.

Références

Bustin, S. A. 2000. Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *Journal of Molecular Endocrinology* # 25: 169-193.

Chen, W., Martinez, G. and Mulchandani, A. 2000. Molecular beacons: A real time polymerase chain reaction assay for detecting *Salmonella*. *Analytical Biochemistry* # 280: 166-172.

Fortin, N. Y., Mulchandani, A. and Chen, W. 2001. Use of real time polymerase chain reaction and molecular beacons for the detection of *Escherichia coli* O157:H7. *Analytical Biochemistry* # 289: 281-288.

Gibson, U. E., Heid, C. A. and Williams, P. M. 1996. A novel method for real time quantitative RT-PCR. *Genome Research* # 6: 995-1001.

Giesendorf, B. A., Vet, J. A., Tyagi, S., Mensink, E. J., Trijbels, F. J. and Blom, H. J. 1998. Molecular beacons: a new approach for semiautomated mutation analysis. *Clinical Chemistry* # 44: 482-486.

Higuchi, R., Dollinger, G., Walsh, P. S. and Griffith, R. 1992. Simultaneous amplification and detection of specific DNA sequences. *Biotechnology* # 10: 413-417.

- Ishiguro, T., Saitoh, J., Yawata, H., Yamagishi, H., Iwasaki, S. and Mitoma, Y. 1995. Homogeneous quantitative assay of Hepatitis C virus RNA by polymerase chain reaction in the presence of a fluorescent intercalater. *Analytical Biochemistry* # 229: 207-213.
- Jeyaseelan, K., Ma, D. and Armugam A. 2001. Real-time detection of gene promotor activity: quantification of toxin gene transcription. *Nucleic Acids Research* # 29 No.12 E58.
- Kandimalla, E. R. and Agrawal, S. 2000. "Cyclicons" as hybridation-based fluorescent primer-probes: synthesis, properties and application in real-time PCR. *Bioorganic and Medicinal Chemistry* # 8: 1911-1916.
- Lin, M.-H., Chen, T.-C., Kuo, T.-T., Tseng, C. and Tseng, C.-P. 2000. Real-time PCR for quantitative detection of *Toxoplasma gondii*. *Journal of Clinical Microbiology* # 38: 4121-4125.
- Liu, J., Feldman, P. and Chung, T. D. Y. 2002. Real-time monitoring *in vitro* transcription using molecular beacons. *Analytical Biochemistry* # 300: 40-45.
- Mackay, I. M., Arden, K. E. and Nitsche A. 2002. Real time PCR in virology. *Nucleic Acids Research* # 30: 1292-1305.
- Marras, S. A., Kramer, F. R. and Tyagi, S. 1999. Multiplex detection of single-nucleotide variations using molecular beacons. *Genetic Analysis* # 14: 151-156.
- Martell, M. Gomez, J., Esteban, J. I., Sauleda, S., Quer, J., Cabot, B., Esteban, R. and Guardia, J. 1999. High-throughput real-time reverse transcription-PCR quantitation of Hepatitis C virus RNA. *Journal of Clinical Microbiology* # 37: 327-332.
- Mhlanga, M. M. and Malmberg, L. 2001. Using molecular beacons to detect single nucleotide polymorphisms with real-time PCR. *Methods* # 25: 463-471.
- Morrison, T. M., Weis, J. J. and Wittwer, C. T. 1998. Quantification of low-copy transcripts by continuous SYBR green I monitoring during amplification. *Biotechniques* # 24: 952-962.
- Myakishev, M. V., Khripin Y., Hu, S. and Hamer D. H. 2001. High-throughput SNP Genotyping by allele-specific PCR with universal energy-transfer-labeled primers. *Genome Research* # 11: 163-169.
- Nazarenko, I. A., Bhatnagaer, S. K. and Hohman, R.J. 1997. A closed tube format for amplification and detection of DNA based on energy transfer. *Nucleic Acids Research* # 25: 2516-2521.
- Nielsen, P. E. 1991. Sequence-selective DNA recognition by synthetic ligands. *Bioconjugate Chemistry* # 2:1-12.
- Poddar, S. K. 1999. Detection of adenovirus using PCR and molecular beacon. *Journal of Virological Methods* # 82: 19-26.
- Searle, M. S. and Embrey, K. E. 1990. Sequence-specific interaction of Hoescht 33258 with the minor groove of an adenine-tract DNA duplex studied in solution by 1H NMR spectroscopy. *Nucleic Acids Research* # 18: 3753-3762.
- Täpp, I., Malmberg, I., Rennel, E., Wik, M. and Syvänen, A.-C. 2000. Homogenous scoring of single-nucleotide polymorphisms: comparison of the 5'-nuclease Taq Man assay and molecular beacon probes. *Biotechniques* # 28: 732-738.
- Thelwell, N., Millington, S., Solinas, A., Booth, J. and Brown, T. 2000. Mode of action and application of scorpion primers to mutation detection. *Nucleic Acids Research* # 28: 3752-3761.
- Tseng, S. Y., Macool, D., Elliot, V., Tice, G., Jackson, R., Barbour, M. and Amorese, D. 1997. An homogeneous fluorescence polymerase chain reaction assay to identify *Salmonella*. *Analytical Biochemistry* # 245: 207-212.
- Tyagi, S. and Kramer, F.R. 1996. Molecular beacons: Probes that fluoresce upon hybridization. *Nature Biotechnology* # 14: 303-308.
- Tyagi, S., Bratu, D. P. and Kramer, F. R. 1998. Multicolor molecular beacons for allele discrimination. *Nature Biotechnology* # 16: 49-53.
- Vandenbroucke, I. I., Vandesompele, J., De Paepe, A. and Messiaen, L. 2001. Quantification of splice variant using real-time PCR. *Nucleic Acids Research* # 29 No.13 E68.
- Vet, J. A. M., Majithia, A. R., Marras, S. A. E., Tyagi, S., Dube, S., Poiesz, B. J. and Kramer, F. R. 1999. Multiplex detection of four pathogenic retroviruses using molecular beacons. *Proceedings of the National Academy of Science* # 96: 6394-6399.
- Whitcombe, D., Theaker, J., Guy, S. P., Brown, T. and Little, S. 1999. Detection of PCR products using self probing amplicons and fluorescence. *Nature Biotechnology* # 17: 804-807.
- Wittwer, C. T., Herrmann, M. G., Moss, A. A. and Rasmussen, R. P. 1997. Continuous fluorescence monitoring of rapid cycle DNA amplification. *Biotechniques* # 22: 130-138.

Summary:

Real-Time PCR technology is getting more and more popular for different sectors of activity. This technology is based on the detection and quantification of a fluorescent reporter whose emission is directly

proportional to the quantity of amplicons produced during the PCR reaction. Because it is generally performed with closed tube systems and that quantification does not require any post-amplification

handling, the risks of post-PCR contamination by amplicons are significantly reduced. The complete process is automated from the beginning to the end, rendering this technology very performing for high-

throughput analyses. This paper presents a description of the Real-Time PCR based principles, different amplicons detection technologies and examples of current applications.



ÉTAT DE L'ART SUR LA PCR

Support Théorique

PCR Réaction de Polymérisation en Chaîne

1983 : Invention de la technique PCR par K.Mullis.

Objectif : synthèse d'un grand nombre de fragments d'ADN identiques.

Condition requises :

- Connaître la séquence du fragment d'intérêt ou de ses séquences flanquantes en amont et en aval ;
- Synthétiser les oligonucléotides (ou amorces) complémentaires à ces séquences.

Etapas de la PCR :

- 1/ Dénaturation de l'ADN ;
- 2/ Hybridation des amorces ;
- 3/ Elongation par la Taq pol.

Principe de la PCR : n cycles = 2^{n-1} molécules d'ADN théoriques.

OPTIMISATIONS

Paramètres importants :

- Choix de l'enzyme et concentration ;
 - Concentration en magnésium ;
 - Concentration en désoxyribonucléotides = DNTP ;
 - Amorces ;
 - Températures d'hybridation et cycles d'amplification ;
 - Matrice (ADN) ;
 - Additifs et co-solvants (éthanol, urée, DMS, SDS,... = inhibiteurs de Taq polymérase) ;
 - Préparation des réactifs.
1. CHOIX de l'ENZYME

Taq polymérase :

- Résistante à de fortes températures ;
- Activité optimale à des températures de l'ordre de 70°C ;
- Enzyme la plus adaptée pour la PCR ;
- Pas d'activité de correction d'erreur.

Concentration d'utilisation : 1 à 5 unités par échantillon

- Si la concentration est TROP FAIBLE on observe une FAIBLE efficacité d'amplification ;
- Si la concentration est TROP ELEVEE l'amplification est NON SPÉCIFIQUE.

2. CONCENTRATION en MAGNESIUM

- La Taq pol nécessite du magnésium libre (Mg^{2+}) et ceci en plus du Mg^{2+} piégé par l'ADN matrice, les amorces et surtout les DNTP ;

- Mg^{2+} est un catalyseur de la fonction enzymatique de polymérisation de la plupart des ADN pol.

Concentration d'utilisation : 1,5 mM mais généralement le Mg^{2+} est déjà à la concentration voulue dans le tampon de la Taq pol.

3. DNTP

- Les DNTP chélatent le Mg^{2+} ;
- Augmenter la concentration en DNTP n'entraîne pas toujours une meilleure efficacité de la PCR ;

Concentration d'utilisation : 200 μ M suffisent pour synthétiser 12,5 μ g d'ADN (incorporation de 50 % des DNTP) ;

- Si la concentration est ÉLEVÉE, on observe une AUGMENTATION DU TAUX D'ERREUR ;
- Si la concentration est TROP ÉLEVÉE, il peut y avoir une INHIBITION de l'enzyme.

4. CHOIX des AMORCES

- Primordial pour l'efficacité de la PCR ;
- Longueur : 15 à 30 nucléotides (pour les PCR classiques) ;
- Site unique d'hybridation.

Séquence :

- Pourcentage de G+C doit être compris idéalement entre 40 et 60% ;
- Extrémité en 3' (point de départ de la polymérisation) riche en G et C ;
- Même particularité en 5' (fixation de l'amorce) ;

Notion de Tm :

Temperature of melting (température de fusion)

- Température à laquelle la moitié des hybrides formés sont dissociés ;
- Le Tm doit être idéalement compris entre 45°C et 70°C ;
- Si inférieur à 55°C, diminution de la spécificité d'hybridation des amorces d'où risque d'amplification non spécifique ;
- Tm forward = Tm reverse ;

Calcul des Tm :

(soit N le nombre de nucléotides)

Pour N inférieur ou égale à 20 : $Tm = 4 \times (G+C) + 2 \times (A+T)$

Pour N supérieur à 20 : $Tm = 81,5 + 16,6 \times \log Na^+ + 0,41(\% G+C) - 600/N$ (dans les milieux classiques, concentration en Na^+ est de 0,053 moles/l)

- Si homologie de séquences dans le choix des amorces = hybridation des amorces entre elles ;
- S'il existe une région palindromique au sein de l'amorce = repliement ou formation d'une boucle.

Concentration d'utilisation :

- La plus basse possible pour donner une sensibilité et une spécificité maximale : 5 à 25 pmol par amorce ;
- Si la concentration est trop FAIBLE, cela peut entraîner un MAUVAIS rendement de la PCR ;
- Si la concentration est trop ELEVÉE, on observe une AUGMENTATION de dimères d'amorces.

5. BILAN :

LA SPECIFICITE dépend de trois facteurs :

- Choix des amorces ;
- Températures d'hybridation ;
- Concentration des réactifs.

AUGMENTATION DE LA SPECIFICITE d'une réaction PCR

- Diminuer le Mg^{2+} ;
- Jouer sur un équilibre fin entre les concentrations en amorce, enzyme et ADN ;
- Augmenter la taille des amorces ;
- Augmenter la température d'hybridation.

OPTIMISATION DES ETAPES

1. DENATURATION :

Avant de commencer les cycles, il est préférable d'ajouter une étape de dénaturation initiale de 2 à 5 minutes qui permet d'améliorer les étapes de dénaturation qui suivent.

Cruciale pour la séparation des deux brins

- Normalement entre 94°C - 96°C ;
- Entre 45s et 1min pour que le tube atteigne la température ;
- Prévoir un temps plus long pour de l'ADN génomique .

Si problème PCR : vérifier que le tube soit à la bonne température.

2. HYBRIDATION

- Etape la plus délicate à optimiser ;
- Détermination de la température ;
- Température optimale plus élevée que celle prédite : choisir une température de quelques degrés supérieure (ou égale) au T_m calculé ;
- Si perte de sensibilité : baisser la température d'hybridation.

3. ELONGATION

Etape la plus simple à optimiser, les conditions sont fournies par le fabriquant.

- 72°C température optimale pour la plupart des enzymes ;
- Durée variable selon la longueur de l'ADN à amplifier ;
- L'activité de la Taq est d'environ 1000 pb/min (selon la Taq).

PROBLEMES RENCONTRES ET SOLUTIONS EVENTUELLES

1. Produits PCR avec un « smear » important

- Augmenter la température d'hybridation ;
- Diminuer le temps de polymérisation ;
- Faire varier la concentration en $MgCl_2$ sans modifier la concentration en DNTP ;
- Utiliser moins d'ADN ;
- Combiner un paramètre après l'autre.

2. Produits PCR non spécifiques

- Augmenter la température d'hybridation ;
- Augmenter le temps d'hybridation ;
- Utiliser moins d'amorces ;
- Utiliser moins d'ADN ;
- Augmenter la taille des amorces (pour élever la température d'hybridation) ;
- Ajouter du glycérol ou de la BSA qui joue sur la stringence ;
- Combiner un paramètre après l'autre.

3. Pas de produits PCR alors que votre protocole PCR donnait les résultats attendus avant ?

- Assurez vous de mettre tous les composants nécessaires à la réaction ;
- Changer la solution de DNTP (sensible aux congélations et décongélation), astuce : aliquoter en petit volume ;
- Même remarque pour les amorces et la Taq ;
- Si vous venez d'acheter un nouveau lot d'amorces, vérifier leurs séquences et concentrations ;
- Augmenter la quantité d'amorces ;
- Augmenter la quantité d'ADN ;
- Diminuer la température d'hybridation ;
- Combiner un paramètre après l'autre ;

4. Faible amplification PCR. Que faire pour augmenter le rendement ?

- Augmenter la quantité d'ADN ;
- Diminuer la température d'hybridation ;
- Augmenter la quantité d'amorces ;
- Diminuer graduellement le Tm jusqu'à la température la plus faible ;
- Changer la concentration en MgCl₂ : plus élevée si les produits PCR < 1000 pb ou plus faible si produits PCR > 1000 pb ;
- Ajouter des adjuvants : BSA (0,1 à 0,8 µg/µl au final), ou du glycérol utilisé entre 5 à 10 %.



Travaux dirigés applications de la PCR

Etudes de cas, responsable C. Siatka

Problématique 1

Détermination d'amorces pour réaliser une amplification spécifique de séquence.

Un fragment Eco RI-Bam HI de 1 kb contenant un gène de souris appelé BM est sous cloné dans un vecteur appelé pGEM2, aux sites Eco RI et Bam HI de pGEM2. Le plasmide obtenu lors de ce sous clonage est appelé pGEM-BM. La séquence partielle du plasmide pGEM2 au niveau des sites de clonage Eco RI et Bam HI est écrite ci-dessous.

5'- ATTTAGGTGACACTATAGAATACACGGAATTTCGAGCTCGCCCGG
GGATCCTCTAGAGTCGACCTGCAGCCCAAGCTTCCGGTCCCTAT
AGTGAGTCGT - 3'

On souhaite utiliser le fragment Eco RI-Bam HI de 1 kb comme sonde pour faire une analyse RFLP chez différentes lignées de souris. Pour préparer la sonde, on choisit de procéder par amplification spécifique de séquence (PCR) en incorporant des nucléotides marqués lors de l'amplification.

A partir de la séquence du site multiple de clonage de pGEM2 donnée ci-dessus, donnez la séquence de 2 amorces de 20 nucléotides permettant d'amplifier le fragment Eco RI-Bam HI (1 kb) de pBM1 cloné aux sites Eco RI et Bam HI de pGEM2.

On rappelle que les enzymes de restriction Eco RI et Bam HI reconnaissent les séquences suivantes:

Eco RI
G*AATTC

Bam HI
G*GATCC

Problématique 2

Détermination de la taille optimale des amorces

On détermine la séquence d'une amorce de n nucléotides.

a- Calculez la probabilité qu'une séquence de n nucléotides pris au hasard soit identique à la séquence de l'amorce.

b- Etant donné la taille L et la complexité C du génome d'un organisme, calculez le nombre moyen de fois que l'on peut trouver une séquence identique à la séquence de l'amorce choisie ci-dessus dans le génome de cet organisme. (on appelle complexité C d'un génome la longueur totale des séquences uniques (non répétées) composant ce génome; L et C sont exprimés en nombre de paires de bases).

c- Calculez la taille minimale d'une amorce qui ne s'hybriderait qu'à une position dans les génomes suivants:

organisme	taille du génome (L)	complexité du génome (C)
<i>E. coli</i>	4 e+6 pb	4 e+6 pb
levure	1,4 e+7 pb	1,4 e+7 pb
drosophile	1,1 e+8 pb	1,05 e+8 pb
mammifères	2,9 e+9 pb	1,8 e+9 pb

En pratique, il faut ajouter trois bases à la longueur théorique pour avoir une bonne sécurité car la valeur calculée est une valeur statistique.

Problématique 3

Amplification spécifique de séquence (PCR)

On souhaite amplifier par PCR une partie de la séquence (donnée ci-dessous) du gène de la β -6 tubuline du maïs.

Rappel: la taille du génome haploïde du maïs est de 5×10^9 pb

```
1  tttaagta ctgtgtgctt gttgcaggat ctgtaactaa ttectatgcg attctcttgt
61  tttagggcg aagatgaggg agatcctgca catccagggg gggcaatgtg gcaaccagat
121 tggcgccaag ttctgggagg tgggtgtgca tgaacatggc attgacccta ccgggcggta
181 cactggcaat tccgacctc agttggagcg tgtaatgtc tactacaatg aagcctcctg
241 cggacgcttt gttccccgcg ctgttctcat ggatcttgag cctgggacaa tggacagtgt
301 ccggaccgga ccctatgggc agatcttccg cctgacaac ttgtgtttg ggcaatctgg
361 tgctggtaac aattgggcta agggcacta caccgagggt gctgagctca ttgactctgt
421 tctggatggt gtgaggaagg aagctgagaa ctgtgactgc ttgcaaggat tccaagtatg
481 ccactccctt ggtggtggtg ctggatctgg tatgggtacg ctgttgatct caaagatcag
541 ggaagagtac cctgaccgca tgatgcttac atttcagtt tcccctcac cgaaagtatc
601 tgataccgtg gttgagccat acaatgccac ttttctgtc caccagttgg tcgagaatgc
661 tgatgagtgc atggttctcg ataacgaagc cctctatgac atctgcttca ggactcttaa
721 gctgaccacc ctagctttg gtgatctgaa ccatttgatc tctgcaacca tgagtggagt
781 cacctgtgct ctaaggttcc ctggtcagct gaactccgac ctgaggaagc tggcagtgaa
841 cctgataccc tccccctgc tcaacttct catggtcggc ttcgctcgcg tgacgtcccg
901 tggctcccag cagtaccggg cctcacagt ccccgagctc acgcagcaga tgtgggatgc
961 caagaacatg atgtgtgccc ctgaccctcg ccatgggctg tactcaccg cctcggccat
1021 gttccgctgg aagatgagca ccaaggaggt tgacgagcag atgatcaacg tccagaacaa
1081 gaactcgtcc tacttctgtg agtggatccc caacaacgtc aagtccagcg tgtcgcacat
1141 cccgccagg ggctgtcca tggcgtccac ctcatcggc aactcgacct ccatccagga
1201 gatgttccgg aggtgagcgc agcagttcac tgccatgttc aggaggaagg ctttcttga
1261 ctggtacacg ggcgagggca tggacgagat ggagttcacc gaggccgaga gcaacatgaa
1321 cgacctcgtg tggagtacc agcagtacca ggacgcgact gccgacgagg aggagtacga
1381 ggacgaggag gaggtgcagg ccgatgacat gtgaggggag ggctgttacc gtgtgaagcc
1441 ttgtgtccc tagggcaage ggacctgat gagttcgggt tcccttctg tgtgtgtgccc
1501 atctttctac tgctagcgtg cccacctcgc tggcccattc cgctgctgtt gacgtatgta
1561 ttttcttgt gctatggaac ctgctttg gtacgttact atctgctag tatgcttggc
1621 gtttgaggtt cctggcgtga attaagcct tccgtatgca gtgattggag ttggagaccg
1681 gctgcttctg ccaggcgaag caattgacag cgactgctgta tacactcaag
```

a- Déterminez la séquence de deux amorces utilisables pour amplifier de façon spécifique environ 1 kilobase de la séquence du gène de la tubuline.

b- Calculez la température de fusion (T_m) de chacune de ces deux amorces.

c- Proposez un protocole d'amplification (température et durée des différentes étapes).

d- Calculez le nombre de cycles théoriquement nécessaires pour obtenir $1 \mu\text{g}$ de ce fragment en partant soit de 10 ng, soit de 500 ng d'ADN génomique.

Problématique 4

Quantification de la présence d'OGM par PCR quantitative.

Pour déterminer le pourcentage de soja génétiquement modifié dans un lot de fèves, on utilise la technique de PCR quantitative en temps réel. L'analyse repose sur une double amplification. L'amplification d'un gène spécifique de l'espèce soja (gène codant la lectine) permet d'évaluer la quantité d'ADN total présent dans l'échantillon. L'amplification d'une partie du transgène (promoteur CaMV 35S) permet d'évaluer la quantité d'ADN correspondant à la présence d'OGM. La réglementation française impose l'obligation d'étiquetage des OGM ou des produits dérivés d'OGM dès lors que le pourcentage d'OGM est supérieur à 1%.

Pour limiter les biais de l'analyse, les deux amplifications sont réalisées au cours d'une même réaction de PCR (PCR duplex).

En même temps que les échantillons inconnus que l'on cherche à caractériser, on analyse une série d'échantillons connus qui contiennent 1% d'OGM, ce qui permet d'établir une courbe d'étalonnage. L'analyse de cette courbe d'étalonnage a donné les résultats suivants :

Quantité d'ADN total de soja dans l'échantillon (en picogramme)	200 000	100 000	50 000	25 000	5 000	2 500
Quantité d'ADN de soja transgénique dans l'échantillon (en picogramme)	2 000	1 000	500	250	50	0
Valeur de Ct mesurée pour l'amplification de la lectine	20,36	21,17	22,80	23,87	26,08	26,98
Valeur de Ct mesurée pour l'amplification du promoteur CaMV35S	26,08	27,12	28,62	29,36	31,46	-

a. Tracez séparément les deux droites d'étalonnage en exprimant le Ct (affiché en ordonnée) en fonction du logarithme de la quantité d'ADN (affiché en abscisse). Déterminez les pentes de chaque droite puis calculez l'efficacité de chaque PCR selon la formule :

$$\text{Efficacité de PCR} = [10^{|1/\text{pente}|} - 1] \times 100$$

Ces efficacités de PCR sont-elles satisfaisantes pour analyser valablement le pourcentage de soja transgénique contenu dans un lot ?

b. L'analyse de deux échantillons de soja a donné les résultats suivants

valeurs de Ct mesurées pour l'amplification de la lectine : 20,60 (éch. 1) et 23,87 (éch. 2)

valeurs de Ct mesurées pour l'amplification de l'OGM : 26,36 (éch. 1) et 26,33 (éch. 2)

A l'aide des droites d'étalonnage ou de leurs équations, calculez le pourcentage de soja transgénique contenu dans ces 2 échantillons de soja. Les produits correspondants doivent-ils être étiquetés ?

Problématique 5

Amplification spécifique d'un microsatellite.

On rappelle qu'un microsatellite est une séquence très courte fortement répétée dans le génome. Dans l'exemple ci-dessous, il s'agit de la répétition CA. Le nombre de répétition du motif CA à une même position dans l'ADN génomique est très variable selon les différents individus qui composent une population homogène. Il existe en général plusieurs endroits différents du génome qui contiennent un microsatellite d'un type particulier (le microsatellite de type CA par exemple). On distingue bien entendu ces différents endroits par la séquence de l'ADN entourant le microsatellite, séquence que l'on ne retrouve qu'une fois dans le génome.

a- En considérant l'exemple ci-dessous, proposez une méthode fondée sur la PCR pour différencier différents individus de la population. Expliquez (et schématisez) quel type de résultat on doit obtenir lorsqu'on applique la méthode que vous proposez.

b- Déterminez la séquence de deux amorces de 20 paires de bases utilisables pour amplifier de façon spécifique le fragment contenant ce microsatellite.

```
1 TTTTAAGTTA CTGTGTGCTT GTTGCAGGAT CTGTA ACTAA TTCCTATGCG ATTCTCTTGT
61 TTGTAGGGCG AAGATGAGGG AGATCCTGCA CATCCAGGGA GGGCAATGTG GCAACCAGAT
121 TGGCGCCAAG TTCTGGGAGG TGGTGTGCGA TGAACATGGC ATTGACCACA CACACACACA
181 CACACACACA CACACACACA CACACACACA CACACACACA CACACACACA CACACACACA
241 CACACACACA CACACACACA CACACACACA CACACACACA CACACACACA CACACACACA
301 CACACACACA CACACACACA CACACACACA CACACACACA CACACAACCG CCTCGGCCAT
361 TGCTGGTAAC AATTGGGCTA AGGGCCACTA CACCGAGGGT GCTGAGCTCA TTGACTCTGT
421 TCTGGATGTT GTGAGGAAGG AAGCTGAGAA CTGTGACTGC TTGCAAGGAT TCCAAGTATG
481 CCACTCCCTT GGTGGTGGTA CTGGATCTGG TATGGGTACG CTGTTGATCT CAAAGATCAG
```