

FRANÇOIS VIGNEAULT

**IDENTIFICATION DES CONDITIONS OPTIMALES  
D'UTILISATION DE L'ADN POLYMÉRASE *VENT*  
EXO- DANS LA TECHNOLOGIE LMPCR.**

Mémoire présenté  
à la Faculté des études supérieures de l'Université Laval  
dans le cadre du programme de maîtrise en biologie cellulaire et moléculaire  
pour l'obtention du grade de maître ès sciences (M. Sc.)

FACULTÉ DE MÉDECINE  
UNIVERSITÉ LAVAL  
QUÉBEC

AVRIL 2004

© François Vigneault, 2004

## Résumé

L'étude des interactions ADN-protéines et de la structure de la chromatine dans les cellules vivantes est nécessaire à la compréhension des mécanismes de l'expression génique. La technique LMPCR (« ligation mediated polymerase chain reaction ») permet l'étude *in vivo* de ces interactions par une analyse plus réelle des événements qui se déroulent au sein même des cellules, comparativement aux études *in vitro*. Par contre, la qualité des résultats peut-être grandement influencée par l'ADN polymérase utilisée et nécessite donc l'optimisation de plusieurs paramètres. De ce fait, nous avons identifié les conditions optimales d'utilisation de l'ADN polymérase *Thermococcus litoralis* *exo-* (*Vent* *exo-*) dans la technique LMPCR telle que la quantité de polymérase et d'ADN. Nous avons montré que l'efficacité de *Vent* *exo-* à l'extension d'amorces et à l'amplification était similaire à celle de *Pyrococcus furiosus* *exo-* (*Pfu* *exo-*) et supérieure à *Thermus aquaticus* (*Taq*). De plus, nous avons observé que la thermostabilité de *Vent* *exo-* lui permettait de soutenir un plus grand nombre de cycles d'amplification que *Taq* facilitant ainsi la résolution de séquences riches en GC. D'autre part, nous avons montré que l'activité terminale transférase de *Vent* *exo-* était inhibée dans la plupart de nos conditions expérimentales et qu'il était donc possible de l'utiliser efficacement pour la production d'extrémités franches lors de l'étape d'extension d'amorces. Finalement, l'ADN polymérase *Vent* *exo-* s'avère être une alternative efficace pour les étapes d'extension d'amorces et d'amplification PCR dans la technologie LMPCR.

Francois Vigneault, étudiant

Dr Régén Drouin, directeur de recherche

## **Avant-Propos**

Les résultats présentés dans ce mémoire s'inscrivent dans le cadre d'une formation universitaire pour l'obtention du diplôme de maîtrise en science. Il est important d'exprimer mes remerciements à tous ceux et celles qui de façon directe ou indirecte m'ont aidé à réaliser ce projet. Ces travaux ont été réalisés au Laboratoire de Génétique Moléculaire de l'Unité de Recherche en Génétique Humaine et Moléculaire de l'Hôpital Saint-François d'Assise de Québec, sous la direction du Dr Régen Drouin.

Je tiens donc tout d'abord à remercier le Dr Régen Drouin qui m'a accueilli dans son équipe de recherche et qui a eu confiance dès le début en mes capacités. Il m'a grandement aidé à développer les qualités nécessaires pour progresser en recherche et m'a soutenu financièrement au cours de ces travaux.

Je voudrais ensuite remercier mes collègues de laboratoire Isabelle Paradis et Nathalie Bastien pour leur aide et soutien dans l'apprentissage des diverses techniques utilisées. Je remercie aussi tous les étudiants du laboratoire et plus particulièrement Stéphane Ouellet et Patrick Rochette pour leurs conseils et sans qui l'ambiance de travail n'aurait pas été la même.

Ce mémoire contient un article soumis pour publication au journal « Nucleic Acids Research » (section « methods » publié par voie électronique seulement) pour lequel j'ai effectué la totalité des expériences et j'ai rédigé l'article.

*À mes parents, grands-parents et à  
Catherine, je dédie ce travail*

# Table des matières

Résumé .....	i
Avant-Propos.....	ii
Table des matières.....	iv
Liste des tableaux .....	vi
Liste des figures .....	vi
Liste des abréviations.....	vii
Chapitre 1 Introduction .....	1
<b>1.1 La technologie LMPCR (<i>Ligation-Mediated Polymerase Chain Reaction</i>;     Réaction de polymérisation en chaîne permise par un adaptateur) .....</b>	<b>1</b>
1.1.1 Historique et rôle de la technologie LMPCR.....	1
1.1.2 Aperçu général.....	1
1.1.3 Étapes de la technique LMPCR .....	3
1.1.4 ADN polymérases pour LMPCR.....	5
1.1.4.1 Caractéristiques recherchées.....	5
1.1.4.2 <i>Sequenase</i> et <i>Taq</i> .....	5
1.1.4.3 <i>Vent</i> et <i>Pfu</i> <i>exo</i> -.....	5
1.1.5 Sélection des gènes : <i>p53</i> , <i>FMR1</i> et <i>C-Jun</i> .....	6
<b>1.2 <i>Thermococcus litoralis</i> <i>exo</i>- .....</b>	<b>7</b>
1.2.1 Historique.....	7
1.2.2 Thermostabilité .....	7
1.2.3 Activité exonucléase .....	8
1.2.4 Activité terminale transférase .....	8
1.2.5 Activité de déplacement de brin et de glissement de réplication.....	10
1.2.6 Incapacité d'incorporation des dUTP .....	11
<b>1.3 Les objectifs spécifiques de l'étude.....</b>	<b>12</b>
Chapitre 2 Optimal conditions and specific characteristics of <i>Vent</i> <i>exo</i> - DNA polymerase in ligation-mediated polymerase chain reaction protocols.....	15
<b>2.1 Résumé.....</b>	<b>16</b>
<b>2.2 Abstract.....</b>	<b>16</b>
<b>2.3 Introduction.....</b>	<b>16</b>
<b>2.4 Materials and methods .....</b>	<b>20</b>
2.4.1 Chemicals, enzymes and equipment.....	20
2.4.2 Synthetic Oligonucleotide Primers for LMPCR Analysis .....	20
2.4.3 Cell culture and DNA preparation .....	20
2.4.4 Ligation-mediated polymerase chain reaction (LMPCR).....	20
2.4.5 Phosphorimager quantification .....	20
<b>2.5 Results .....</b>	<b>21</b>
2.5.1 Primer extension step.....	22
2.5.2 PCR Amplification step .....	22

2.5.3 Optimal MgSO <sub>4</sub> concentration and effect of salt on PCR efficiency .....	24
2.5.4 Number of cycles in the PCR amplification step.....	26
2.5.5 Highly GC-rich sequences .....	27
2.5.6 Polymerases efficiency at primer extension and amplification steps .....	27
<b>2.6 Discussion .....</b>	<b>29</b>
2.6.1 Primer extension step.....	30
2.6.2 PCR amplification step .....	30
2.6.3 Salt conditions.....	32
2.6.4 Thermostability of <i>Vent</i> exo- .....	33
2.6.5 Highly GC-rich sequence.....	34
2.6.6 Combining polymerases at primer extension and amplification steps.....	35
<b>2.7 Conclusion .....</b>	<b>37</b>
<b>2.8 Acknowledgement.....</b>	<b>39</b>
<b>2.9 Legend of figures.....</b>	<b>40</b>
<b>2.10 Figures 1 à 8 et tables 1 de l'article .....</b>	<b>45</b>
 Conclusion.....	 55
Références .....	60

# Liste des tableaux

## Tableau de l'article

Table 1 : Oligonucléotides sélectionnés pour l'étude .....	45
--	----

# Liste des figures

## Figures de l'introduction

<b>Figure 1</b> : Aperçu général du procédé LMPCR utilisant le traitement au DMS.....	2
<b>Figure 2</b> : Illustration des différentes étapes de la technique LMPCR. ....	4
<b>Figure 3</b> : Activité terminale désoxynucléotidyle transférase de l'ADN polymérase <i>Taq.</i> ...	9
<b>Figure 4</b> : Illustration du mécanisme de glissement de réplication.....	10
<b>Figure 5</b> : Illustration du mécanisme de déplacement de brin. ....	11

## Figures de l'article

<b>Figure 1</b> : Illustration des différentes étapes de la technique LMPCR .....	45
<b>Figure 2</b> : Étude de dosage de <i>Vent</i> exo- en fonction de la concentration en ADN à l'extension d'amorces .....	47
<b>Figure 3</b> : Étude de dosage de <i>Vent</i> exo- à l'amplification .....	48
<b>Figure 4</b> : Étude de concentration en ADN à l'amplification .....	49
<b>Figure 5</b> : Étude de concentration en MgSO <sub>4</sub> .....	50
<b>Figure 6</b> : Effet du nombre de cycles sur la thermostabilité des polymérases .....	51
<b>Figure 7 A et B</b> : Efficacité de <i>Vent</i> exo- à résoudre la répétition CGG du gène.....	52
<b>Figure 7 C et D</b> : Efficacité de <i>Vent</i> exo- à résoudre la répétition CGG du gène.....	53
<b>Figure 8</b> : Combinaison des polymérases au cours du processus LMPCR.....	54

## Liste des abréviations

ADN : acide désoxyribonucléique

DMS : sulfate de diméthyle

Exo- : sans activité exonucléase 3'-5'

KCl : chlorure de potassium

LMPCR : réaction de polymérisation en chaîne permise par un adaptateur « ligation-mediated polymerase chain reaction »

MgSO<sub>4</sub> : sulfate de magnésium

NaCl : chlorure de sodium

pb : paires de bases

PCR : réaction de polymérisation en chaîne

*Pfu* exo- : ADN polymérase *Pyrococcus furiosus* exo-

*Sequenase* : ADN polymérase T7-modifiée

*Taq* : ADN polymérase *Thermus aquaticus*

TDPCR : réaction de polymérisation en chaîne permise par une activité terminale transférase « terminal-transferase dependent PCR »

*Vent* : ADN polymérase *Thermococcus litoralis*

*Vent* exo- : ADN polymérase *Thermococcus litoralis* exo-

Tdt : terminale désoxynucléotidyle transférase



# Chapitre 1 Introduction

## 1.1 La technologie LMPCR (*Ligation-Mediated Polymerase Chain Reaction*; Réaction de polymérisation en chaîne permise par un adaptateur)

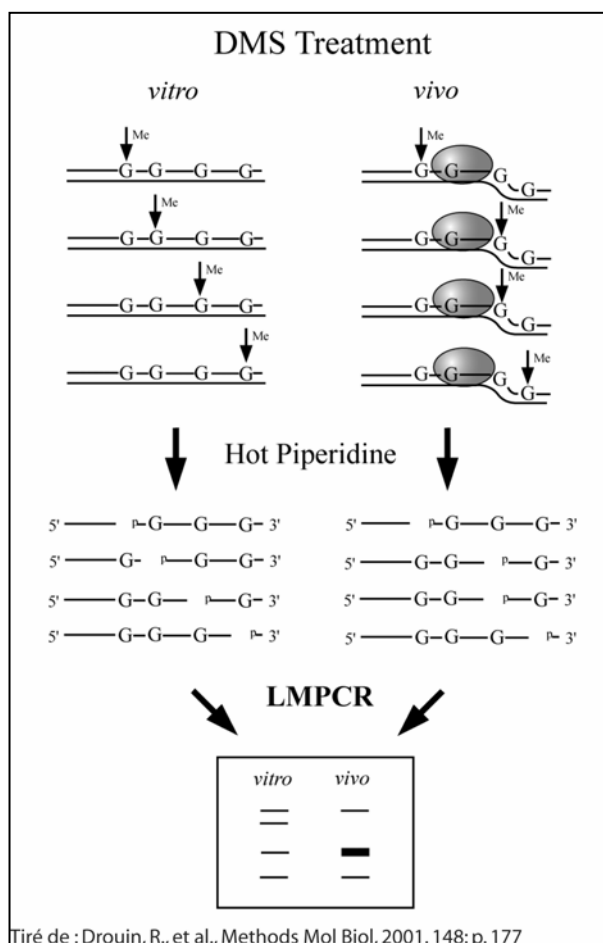
### 1.1.1 Historique et rôle de la technologie LMPCR

LMPCR est une des rares techniques disponibles pour l'analyse *in vivo* de l'ADN. Il s'agit en fait d'une technique de séquençage génomique qui permet la cartographie des cassures monocaténares de l'ADN au niveau de résolution des nucléotides. Elle a été développée conjointement par Mueller & Wold et Pfeifer et coll. en 1989, pour l'étude d'empreintes protéiques ("*footprinting*") sur l'ADN dans les cellules vivantes [1, 2]. Les études *in vivo* possèdent l'avantage de fournir une représentation plus adéquate de la structure de la chromatine au cœur des cellules, alors que les études *in vitro* peuvent perturber significativement l'environnement intracellulaire. Depuis son développement, la technologie LMPCR a été utilisée avec succès pour la cartographie des interactions ADN-protéines ayant lieu dans les cellules vivantes et a ainsi participé à l'avancement de notre compréhension des mécanismes de régulation génique [3-8]. De plus, la technologie LMPCR a été utilisée avec succès dans les études d'analyse des patrons de méthylation des cytosines [2, 9, 10], le positionnement des nucléosomes [3, 11, 12] ainsi que dans les études de réparations de l'ADN [13-15]. Finalement, certaines études ont porté sur l'optimisation de la technique via l'utilisation de différentes ADN polymérases par exemple [16-18].

### 1.1.2 Aperçu général

La technologie LMPCR consiste à évaluer la réactivité locale de l'ADN de cellules vivantes face à divers agents modifiants, comparée à celle de l'ADN purifié. Deux étapes caractérisent cette technique soit : (1) le traitement d'ADN purifié et d'ADN cellulaire à l'aide d'agents modifiants comme le DMS, les UVC et la DNaseI et (2) la visualisation des modifications nucléotidiques sur gel de séquençage (**Figure 1**). Une comparaison peut alors être faite entre la fréquence des cassures provenant de l'ADN cellulaire versus l'ADN purifié. En examinant la différence entre les patrons de cassures *in vivo* et *in vitro*, on peut

ainsi déterminer la présence de facteurs protéiques liés directement à l'ADN. Il faut noter que les modifications induites doivent directement causer ou être potentiellement converties en cassures monocaténaire et générer une extrémité 5' phosphorylée.



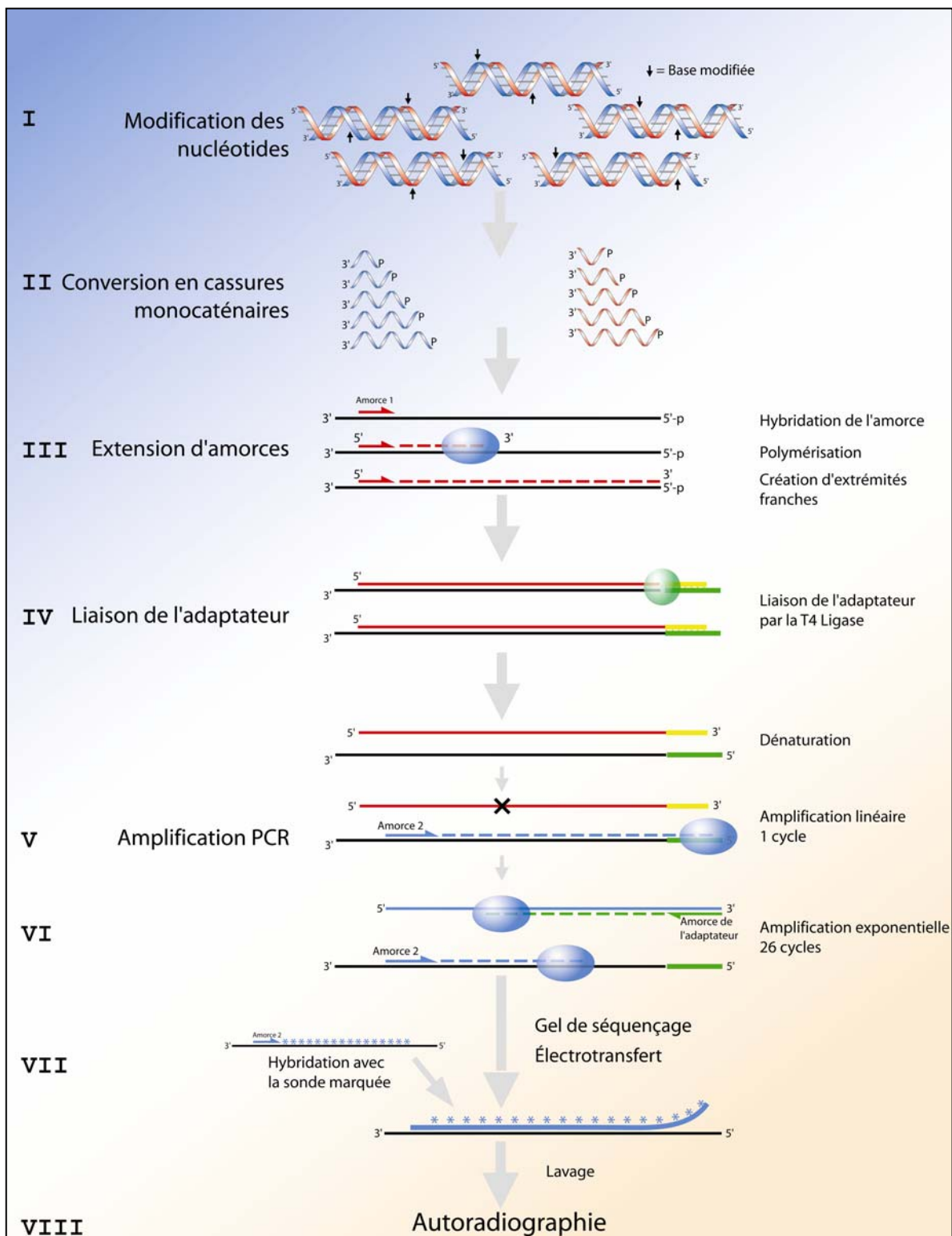
**Figure 1:** Aperçu général du procédé LMPCR utilisant le traitement au DMS. Le DMS induit la méthylation des guanines qui sont alors converties en cassures monocaténaire par un traitement à la pipéridine chaude. Les fragments d'ADN sont alors soumis aux étapes de la technique LMPCR pour ensuite être visualisés sur gel de séquençage. La comparaison des patrons de méthylation entre l'ADN purifié (*vitro*) et l'ADN cellulaire nous permet de déterminer la présence de facteurs liés à l'ADN.

La caractéristique unique de la technologie LMPCR, ce qui lui donne sa très grande sensibilité et son nom, provient de la liaison d'un adaptateur, un "linker" bicaténaire

asymétrique, à l'extrémité 5' phosphorylée de chaque fragment d'ADN avec une extrémité franche. L'adaptateur procure alors une séquence commune et connue. Par conséquent, tous les fragments d'ADN contenant un adaptateur à leur extrémité 5' peuvent être amplifiés de façon exponentielle et ainsi rendre possible la cartographie des cassures monocaténares et indirectement des nucléotides endommagés.

### 1.1.3 Étapes de la technique LMPCR

La **Figure 2** illustre les différentes étapes de la technologie LMPCR. L'étape I consiste en la conversion des dommages nucléotidiques en cassures monocaténares. À l'étape II, l'ADN est dénaturé par la chaleur afin d'obtenir une population de fragments monocaténares d'ADN de différentes longueurs contenant un phosphate en 5'. À l'étape III, à l'aide d'une amorce spécifique du gène ou de la séquence à l'étude, une extension d'amorce est effectuée avec une ADN polymérase n'ayant pas d'activité terminale transférase, pour ainsi obtenir des extrémités franches. L'étape IV consiste à la liaison de l'adaptateur asymétrique à l'extrémité franche des fragments. À l'étape V, après dénaturation, une deuxième amorce adjacente à la première est utilisée afin d'augmenter la spécificité, pour ainsi réaliser une amplification linéaire du fragment isolé ainsi que de son extrémité connue de l'adaptateur. Étant donné qu'il n'y a pas d'amorce s'hybridant au brin complémentaire, ce dernier ne sera pas amplifié au cours des étapes suivantes. Par la suite, à l'étape VI, les fragments d'ADN sont amplifiés de façon exponentielle, et ce, à l'aide de la deuxième amorce et de l'amorce correspondante à la séquence du long oligonucléotide de l'adaptateur. Les fragments d'ADN sont alors séparés selon leur longueur par électrophorèse sur gel de séquençage. Après cette séparation, les fragments d'ADN sont transférés sur une membrane de nylon par électrotransfert et fixés par les rayons ultraviolets de type C (pic à 254 nm). Cette membrane est hybridée avec une sonde radioactive préparée préalablement par PCR et spécifique à la séquence étudiée (étape VII). Ensuite, la membrane est lavée et le signal est détecté par autoradiographie (étape VIII). Ce même signal peut aussi être quantifié par exposition de la membrane à une plaque d'analyse du phosphore radioactif et lu par un système d'imagerie.



**Figure 2** : Illustration des différentes étapes de la technique LMPCR.

## 1.1.4 ADN polymérase pour LMPCR

### 1.1.4.1 Caractéristiques recherchées

Au cours de la procédure LMPCR, l'utilisation d'ADN polymérase est requise pour deux étapes soit l'extension d'amorces et l'amplification PCR. Lors de l'extension d'amorces, la polymérase idéale serait thermostable et très efficace, ne posséderait aucune activité terminale transférase, pourrait polymériser efficacement des fragments d'une longueur pouvant aller jusqu'à 750 pb même si l'ADN comporte des séquences riches en GC ou des structures secondaires particulières.

Pour l'étape d'amplification PCR, l'ADN polymérase idéale serait thermostable et très efficace et serait capable d'amplifier sans discrimination un mélange de fragments de diverses longueurs (entre 50 et 750 pb). De plus, elle doit pouvoir résoudre des séquences avec des contenus différents en GC (de 5% à 95%) et finalement être capable de polymériser à travers des structures secondaires particulières de l'ADN, par exemple, les structures en épingle à cheveux.

### 1.1.4.2 *Sequenase* et *Taq*

L'utilisation de la *Sequenase* à l'extension d'amorces suivi de l'ADN polymérase *Taq* à l'amplification a été originalement utilisée dans les protocoles LMPCR. Rapidement, il a été montré que la thermostabilité de la *Sequenase* était limitative pour l'extension d'amorces de séquences riches en GC [16]. De plus, il a été rapporté que cette polymérase possède une activité terminale transférase dépendante de la séquence, qui amène ainsi la production d'extrémités protubérantes qui inhibent toutes les liaisons subséquentes de l'adaptateur [18].

### 1.1.4.3 *Vent* et *Pfu* exo-

Garrity et Wold ont montré que l'ADN polymérase *Vent* pouvait être utilisé au cours de l'extension d'amorces et de l'amplification PCR avec plus d'efficacité qu'avec la *Sequenase* et la *Taq* respectivement [16]. Par la suite, le développement de nouvelles polymérases a entraîné de nouvelles études d'optimisation. De ce fait, les conditions

optimales d'utilisation de l'ADN polymérase *Pfu* *exo-* pour la technique LMPCR ont été établies. Il a été démontré que l'absence d'activité Tdt de *Pfu* *exo-* permettait une extension d'amorces plus efficace qu'avec *Sequenase*, ainsi qu'une augmentation de l'efficacité de l'amplification par rapport à l'ADN polymérase *Taq* ce qui en a fait la polymérase la plus efficace d'utilisation pour la technique LMPCR jusqu'à présent [18].

Toutes ces expériences ont montré l'importance du ratio molécules d'ADN / molécules de polymérase ainsi que la nécessité d'établir les conditions optimales d'utilisation de la polymérase choisie. Plusieurs laboratoires emploient l'ADN polymérase *Vent* *exo-* ou *Pfu* *exo-*, mais aucune comparaison n'a été effectuée à ce jour pour identifier laquelle de ces deux enzymes possédait la plus grande efficacité dans l'accomplissement des étapes d'extension d'amorces et d'amplification PCR

### **1.1.5 Sélection des gènes : *p53*, *FMRI* et *C-Jun***

Dans cette étude, nous avons décidé d'utiliser des amorces couvrant une partie de l'exon 5 du gène *p53* et la région promotrice des gènes *FMRI* et *C-Jun*, du fait que ces régions ont été cartographiées à maintes reprises avec succès dans diverses études par la technologie LMPCR. Les régions sélectionnées sont principalement composées de séquences très riches en GC, ce qui favorise la mise en évidence de l'efficacité individuelle des ADN polymérases testées. De plus, une de nos expériences porte sur la capacité des polymérases à résoudre des structures complexes de l'ADN. L'exon 1 du gène *FMRI* est composé d'une succession de triplets CGG dont le nombre varie de 6 à 250 chez l'homme; cette suite de triplets peut former une structure secondaire inhibant la polymérisation [19]. L'utilisation de ces trois segments de gènes nous permettra d'identifier les conditions générales qui prévaudront par la suite lors de l'utilisation de la *Vent* *exo-* pour la majorité des gènes qui seront traité dans l'avenir.

## 1.2 *Thermococcus litoralis* exo-

### 1.2.1 Historique

L'ADN polymérase *Vent* a été isolée et clonée à partir de la bactérie *Thermococcus litoralis* [20]. Cette hyperthermophile a préalablement été isolée à partir d'une cheminée thermale des fonds sous-marins [21], communément appelé « thermal *Vent* » en anglais, d'où son appellation courante d'ADN polymérase *Vent*. Un mutant exonucléase déficient a ensuite été produit (ADN polymérase *Vent* exo-) et les caractéristiques spécifiques de ces deux polymérases ont alors été déterminées [22]. Depuis, l'ADN polymérase *Vent* a été grandement utilisée dans des expériences de séquençage, mais aussi en LMPCR où elle s'est avérée plus efficace que les polymérases utilisées préalablement, soit la *Sequenase* et la *Taq* [16]. Par contre, aucune étude d'optimisation des conditions d'utilisation de la *Vent* exo-, avec la technique LMPCR n'a été publiée, et de ce fait, nous avons décidé d'examiner les caractéristiques intrinsèques de cette polymérase.

### 1.2.2 Thermostabilité

L'avènement des polymérases résistantes à la chaleur, comme l'ADN polymérase *Taq*, a été une étape clé dans le développement de la réaction de polymérisation en chaîne [23]. La technologie LMPCR utilise la réaction de polymérisation en chaîne pour amplifier les fragments d'ADN. Lorsqu'il y a présence des séquences très riches en GC, l'amplification PCR de ces fragments doit inclure une très haute température de dénaturation pendant de nombreux cycles. Une polymérase très thermostable résiste à ce genre de température durant de longues périodes, permettant ainsi de cartographier des séquences riches en GC. La *Vent* exo- ainsi que la *Vent* se caractérisent par une demi-vie similaire de 2 heures à 100°C, où *Taq* ne montre plus aucune activité après ce temps d'incubation. À 95°C, la demi-vie de la *Taq* est de 1.6 h, comparativement à 8 h pour *Vent* exo- et la *Vent* [22]. L'ADN polymérase *Vent* exo- montre donc un meilleur potentiel à résister aux hautes températures requises pour l'amplification PCR de séquences riches en GC, ce qui a d'ailleurs été démontré au préalable pour la *Vent* [16].

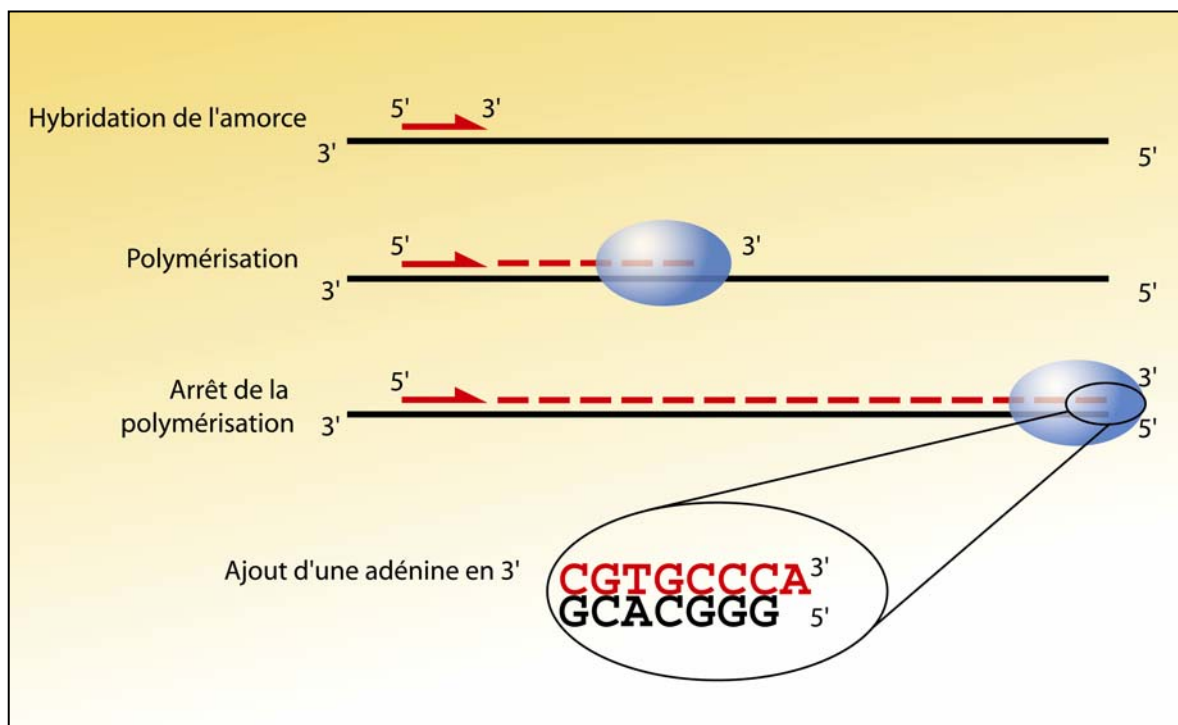
### 1.2.3 Activité exonucléase

L'ADN polymérase *Vent* possède une activité exonucléase 3'-5' qui augmente la fidélité de la réplication de l'ADN par cette enzyme [24]. L'introduction d'une double mutation causant la substitution de 2 acides aminés dans le domaine exonucléase a aboli cette activité, sans toutefois modifier les paramètres cinétiques de polymérisation et produisant ainsi la polymérase *Vent* *exo-* [22]. Lors de la procédure LMPCR, l'utilisation d'une polymérase possédant une activité exonucléase lors de l'extension d'amorces pourrait nuire à la formation d'extrémités franches et ainsi inhiber toute liaison subséquente de l'adaptateur. L'utilisation de la *Pfu* *exo-* pour l'extension d'amorces s'est avérée très efficace dans la formation d'extrémités franches [18], et de ce fait, nous avons émis l'hypothèse qu'il en serait de même pour l'ADN polymérase *Vent* *exo-*. Garrity et Wold ont montré que *Vent* était plus efficace que *Sequenase* pour l'extension d'amorces [16], mais la comparaison entre *Vent* et *Vent* *exo-* n'avait pas été faite à ce jour.

### 1.2.4 Activité terminale transférase

La formation d'extrémités franches lors de l'extension d'amorce peut aussi être influencée par la présence d'activité terminale désoxynucléotidyle transférase (Tdt). Cette activité se caractérise par l'addition d'un nucléotide en 3' au-delà du patron du brin répliqué, comme il a été observé pour l'ADN polymérase *Taq* (Figure 3) [25, 26].





**Figure 3** : Activité terminale désoxynucléotidyle transférase de l'ADN polymérase *Taq*.

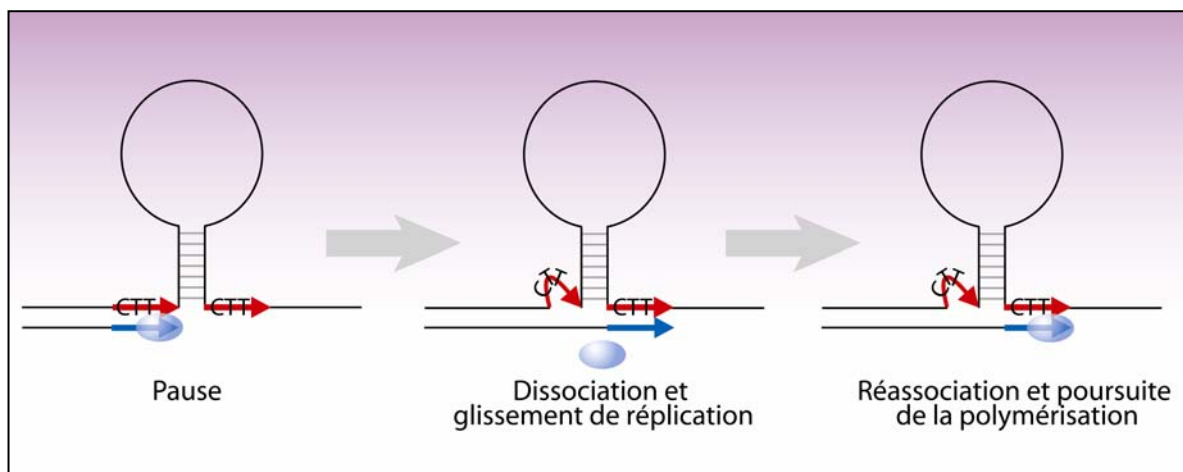
La création d'extrémités protubérantes par *Taq* inhiberait donc toute liaison de l'adaptateur asymétrique utilisé lors de la technique LMPCR, et c'est pourquoi cette polymérase n'est pas utilisée à l'étape d'extension d'amorces. Les études sur *Vent* ont démontré que 90% des fragments polymérisés par *Vent* possèdent des extrémités franches [22], ce qui explique les excellents résultats de liaison d'adaptateur obtenus avec LMPCR par Garrity et Wold [16]. Dans le cas de *Vent* *exo-*, il a été rapporté que 30 % des fragments étaient caractérisés par l'addition d'un nucléotide en 3' et que cela inhiberait la liaison efficace d'adaptateurs à extrémités franches [22]. Donc selon ces résultats, *Vent* *exo-* ne serait pas un bon choix pour LMPCR, car 30% des fragments seraient perdus au cours de la procédure LMPCR. Par contre, l'absence d'activité exonucléase étant un facteur critique pour l'étape d'extension d'amorce, tel que démontré par les études sur la *Pfu* *exo-* [18], nous avons cru bon d'évaluer le potentiel de la *Vent* *exo-* lorsque soumise aux conditions rigoureuses de la technique LMPCR.

De plus, si une polymérase possède une activité Tdt, les fragments amplifiés par cette enzyme devraient subir un retard de migration sur le gel de séquençage lorsque comparés à

des fragments n'ayant pas subi l'addition d'un nucléotide supplémentaire en 3'. Dans le cas de fragments amplifiés avec *Taq*, toutes les bandes seront retardées, mais dans le cas de *Vent* exo-, on devrait s'attendre à ce que 30 % des bandes soit retardées. Ce retard de migration de 30% des molécules devrait se traduire par la présence de doubles bandes.

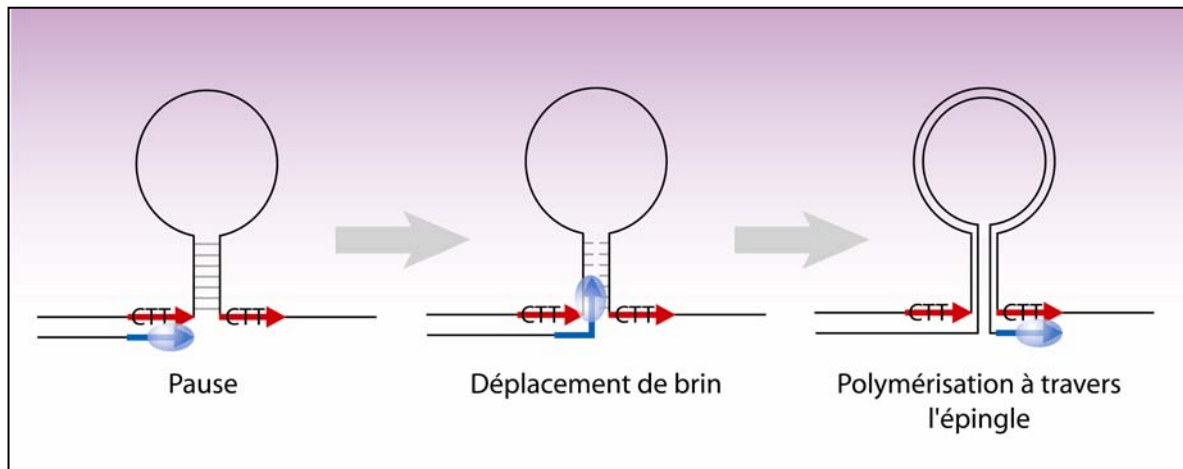
### 1.2.5 Activité de déplacement de brin et de glissement de réplication

La présence de structures secondaires particulières de l'ADN, comme les épingle à cheveux, peut nuire à la réplication efficace pour la majorité des ADN polymérase. Les polymérase incapables de résoudre ce genre de structure sont victimes de glissements de réplication (« Replication Slippage »). Ce phénomène se caractérise par l'arrêt de la polymérisation lorsque l'épingle à cheveux est rencontrée, suivi d'une dissociation de la polymérase. Il s'ensuit un réappariement du brin copié avec la répétition finale de l'épingle à cheveux induisant ainsi un glissement entre les brins d'ADN, puis une réassociation de la polymérase qui reprend alors la polymérisation (Figure 4). Le tout résulte en une délétion de la structure secondaire particulière de l'ADN, et du même coup, d'une portion du fragment à amplifier [27]. Les ADN polymérase *Taq*, *Pyra* exo-, *Pfu* et *Vent* sont des exemples de polymérase qui peuvent subir le glissement de réplication [28].



**Figure 4 :** Illustration du mécanisme de glissement de réplication.

Une polymérase sera par contre capable de résoudre ce genre de structure de l'ADN si elle est dotée d'une activité de déplacement de brin (« strand-displacement activity »). Dans ce cas, lorsque la polymérase rencontre une épingle à cheveux, elle possède le potentiel de déplacer les brins hybridés formant cette structure et ainsi répliquer entièrement le brin complémentaire. C'est le cas des ADN polymérases *Bst*, *Vent* et *Vent* *exo*- (Figure 5)[28].



**Figure 5** : Illustration du mécanisme de déplacement de brin.

En fait, dans le cas de *Vent* l'activité de déplacement de brin dépend de la température, absente à 50°C, mais activée à 65°C et toujours présente peu importe la température dans le cas de *Vent* *exo*-. De plus, la concentration en ions magnésium influence cette activité pour les deux polymérases; des concentrations supérieures à 4 mM en  $MgSO_4$  inhibent l'activité de déplacement de brin et induisent le glissement de réplication pour *Vent* et *Vent* *exo*- [28]. L'utilisation de l'ADN polymérase *Vent* *exo*- à haute température et à basse concentration en  $MgSO_4$  serait donc préférable à celle de *Pfu* *exo*- pour la cartographie par LMPCR des séquences pouvant comporter des structures secondaires particulières de l'ADN.

### 1.2.6 Incapacité d'incorporation des dUTP

Un autre facteur notable pouvant affecter l'amplification efficace des fragments d'ADN dans le procédé LMPCR, est l'incapacité de certaines polymérases à incorporer les dUTP. Il a été montré qu'au fil des cycles d'amplification, il y avait une accumulation de dUTP provenant de la désamination des dCTP, principalement causée par les hautes températures

des cycles PCR [29]. La présence d'uraciles dans l'ADN inhibe la capacité de polymériser des ADN polymérase archaéal comme *Pfu* et *Vent* ainsi que leurs dérivés [30]. Ces polymérase reconnaissent les résidus d'uraciles présents dans le brin à répliquer, causant par la suite un arrêt de la polymérisation pour ainsi éviter l'incorporation inadéquate d'adénines [31]. Il a été montré qu'en ajoutant des dUTPase au mélange réactionnel d'amplification, l'accumulation de dUTP était inhibée, ce qui favorise grandement l'utilisation d'ADN polymérase archaéal pour ainsi obtenir une plus grande efficacité d'amplification [29]. Finalement, lorsque l'on effectue la technique LMPCR sur des séquences riches en GC, on s'attend donc à favoriser l'accumulation des dUTP au cours de l'étape d'amplification, ce qui pourrait nuire au rendement lorsque *Vent* *exo*- et *Pfu* *exo*- sont utilisées.

### **1.3 Les objectifs spécifiques de l'étude**

Le développement de la technologie LMPCR passe inévitablement par l'optimisation des nombreuses étapes qui la compose. Comme nous l'avons mentionné plus tôt, l'un des critères qui influence le plus ce procédé est l'ADN polymérase utilisé au cours des étapes d'extension d'amorces et d'amplification par PCR. La gamme d'ADN polymérase disponible ne cesse de croître et l'identification de celle qui sera la plus efficace pour la technique LMPCR demeure une nécessité.

Dans les pages précédentes, nous avons fait mention des caractéristiques que devrait posséder la polymérase idéale. Néanmoins, même avec une idée préconçue des critères que nous recherchons, il n'existe qu'une façon pour déterminer si une polymérase est adéquate; la soumettre au processus expérimental de la technique LMPCR. Pour ce faire, il faut d'abord identifier les conditions optimales d'utilisation de la polymérase choisie, pour ainsi comparer son efficacité avec celle des autres polymérase présentement utilisées. Dans cette étude, nous nous sommes engagés à vérifier l'efficacité de l'ADN polymérase *Vent* *exo*- aux étapes d'extension d'amorces et d'amplification PCR.

Le premier objectif consistait à déterminer la concentration optimale de *Vent* *exo*- à utiliser au cours de l'étape d'extension d'amorces, ainsi que la relation entre la quantité initiale d'ADN et la quantité de polymérase. Nous voulions vérifier du même coup l'étendue de

l'activité Tdt de l'ADN polymérase *Vent* *exo*-, ainsi que l'influence de cette activité sur la production d'extrémités franches, nécessaires à l'étape subséquente de liaison de l'adaptateur. Étant donné que les conditions optimales d'utilisation de *Pfu* *exo*- ont préalablement été déterminées et que son efficacité semblait être supérieure à toutes les autres polymérases testées à ce jour [18], il allait de soi que l'efficacité de *Vent* *exo*- devait être comparée avec celle de *Pfu* *exo*-. Dans un même temps, nous voulions vérifier l'efficacité de *Vent* *exo*- à résoudre différentes séquences et c'est pourquoi nos expérimentations ont été effectuées sur trois gènes possédant des caractéristiques particulières, soit les gènes *p53*, *FMRI* et *C-Jun*.

Notre second objectif portait principalement sur l'efficacité de *Vent* *exo*- lors de l'étape d'amplification PCR. Nous voulions déterminer la concentration optimale d'utilisation de *Vent* *exo*- à une quantité d'ADN donnée. Par la suite, nous avons identifié l'influence de la quantité d'ADN, sur l'efficacité de *Vent* *exo*- à amplifier de manière uniforme et efficace les fragments d'ADN présents. De plus, nous avons investigué quelle était la concentration idéale en  $MgSO_4$ , pour l'obtention d'une amplification efficace tout en demeurant à l'intérieur des concentrations nécessaires au bon fonctionnement de l'activité de déplacement de brin de *Vent* *exo*-. Finalement, nous avons étudié l'effet du nombre de cycle à l'amplification PCR sur la thermostabilité des ADN polymérases *Vent* *exo*-, *Pfu* *exo*- et *Taq*.

Notre troisième objectif était de vérifier, au niveau des étapes d'extension d'amorces et d'amplification, l'effet des cations, plus précisément le sodium et le potassium présents dans les tampons de réaction sur l'efficacité de polymérisation des ADN polymérases *Pfu* *exo*-, *Vent* *exo*- et *Taq*. Dans un même ordre d'idées, nous avons évalué l'efficacité de différentes polymérases combinées à l'extension et à l'amplification, dans la résolution de la structure particulière du gène *FMRI*, formée par une répétition de triplets CGG, représentant du même coup une séquence constituée à 100% de GC. Dans le but de favoriser la résolution de cette séquence, l'effet de l'ajout de 7-déaza-dGTP ainsi que de DMSO dans les tampons de réaction a été évalué pour les différentes polymérases testées.

Notre quatrième objectif consistait à évaluer quelle combinaison de polymérases était la plus efficace à l'extension d'amorces et à l'amplification. Du même coup, nous avons

vérifié la présence d'activité Tdt des polymérase *Pfu* *exo-*, *Vent* *exo-*, *Vent* et *Taq*; (1) soit dans l'efficacité à produire des extrémités franches à l'étape d'extension d'amorces, (2) soit par la présence d'un retard de migration sur gel provenant de l'utilisation de ces polymérase lors de l'étape d'amplification PCR.

## **Chapitre 2 Optimal conditions and specific characteristics of *Vent* exo- DNA polymerase in ligation-mediated polymerase chain reaction protocols.**

François VIGNEAULT<sup>1,2</sup> and Régen DROUIN<sup>1,2,\*</sup>

<sup>1</sup>Unité de Recherche en Génétique Humaine et Moléculaire, Research Center, Hôpital St-François d'Assise, Centre Hospitalier Universitaire de Québec, 10 de l'Espinay street, Québec, Québec, Canada.

<sup>2</sup>Present address: Department of Pediatrics, CHUS, 3001, 12 Ave North, Sherbrooke, Quebec, Canada, J1H 5N4

Key words: *Vent* exo- DNA polymerase; *Pfu* exo- DNA polymerase; DNA sequence context; Ligation-Mediated Polymerase Chain Reaction; PCR reaction buffer

\* To whom correspondence should be addressed:

Régen Drouin,  
Department of Pediatrics  
Centre Hospitalier Universitaire de Sherbrooke  
Hôpital Fleurimont,  
3001, 12 Avenue North  
Sherbrooke (Québec) J1H 5N4, Canada  
Tel.: (819) 820-6827  
Fax: (819) 564-5217  
Email: Regen.Drouin@USherbrooke.ca

## 2.1 Résumé

Nous avons développé une procédure optimisée pour la technologie LMPCR (“ligation mediated polymerase chain reaction”) en utilisant l’ADN polymérase *Thermococcus litoralis* *exo-* (*Vent* *exo-*). Le dosage optimal de *Vent* *exo-* aux étapes d’extension d’amorces et d’amplification PCR, ainsi que la quantité optimale d’ADN à utiliser ont été établis. Nous montrons que *Vent* *exo-* peut être efficacement utilisée dans la création d’extrémités franches requises pour l’étape subséquente de liaison de l’adaptateur. *Vent* *exo-* s’avère être plus efficace que *Pyrococcus furiosus* *exo-* (*Pfu* *exo-*) dans cette tâche lorsque utilisée avec des quantités d’ADN supérieures à 1,25 µg. Nous montrons que *Vent* *exo-* résout les séquences très riches en GC substantiellement mieux que l’ADN polymérase *Thermus aquaticus* (*Taq*) et de façon similaire à *Pfu* *exo-*. Nous démontrons que le ratio molécule d’ADN / ADN polymérase est significativement plus haut pour *Vent* *exo-* que pour *Pfu* *exo-*, ce qui est reflété dans la sensibilité de *Vent* *exo-* à efficacement amplifier de l’ADN génomique. De plus, l’étendue de l’efficacité de *Vent* *exo-* souligne l’importance de conduire des tests évaluatifs pour identifier le dosage optimal de polymérase à utiliser pour réussir avec succès des amplifications PCR. Nous avons aussi déterminé la concentration optimale de MgSO<sub>4</sub> à utiliser avec *Vent* *exo-* ainsi que l’effet du nombre de cycles PCR sur sa thermostabilité. Nos résultats montrent que l’ADN polymérase *Vent* *exo-* produit des bandes d’intensité uniforme et peut être utilisée efficacement dans l’analyse *in vivo* de l’ADN par LMPCR.

## 2.2 Abstract

An optimized procedure for the ligation-mediated polymerase chain reaction (LMPCR) technique using *Thermococcus litoralis* *exo-* DNA polymerase (*Vent* *exo-*) was developed. The optimal dosage of *Vent* *exo-* at the primer extension and PCR amplification steps as well as the optimal DNA quantity to use were established. We showed that *Vent* *exo-* can efficiently create the blunt-ended termini required for subsequent linker ligation. *Vent* *exo-* proves to be more efficient than *Pyrococcus furiosus* *exo-* (*Pfu* *exo-*) for this task. *Vent* *exo-* resolves highly GC-rich sequence substantially better than *Thermus aquaticus* DNA polymerase (*Taq*) and with a similar efficiency as *Pfu* *exo-*. The DNA / DNA polymerase



activity ratio is significantly higher for *Vent* *exo*- than *Pfu* *exo*-, which is reflected by the sensibility of *Vent* *exo*- in efficiently amplifying genomic DNA. Furthermore, the range of efficiency of *Vent* *exo*- demonstrates the importance of conducting evaluative testing to identify the optimal dosage of use of this polymerase to obtain successful PCR amplification. Optimal  $\text{MgSO}_4$  concentrations to use with *Vent* *exo*- were established. Our results show that *Vent* *exo*- DNA polymerase produces bands of uniform and strong intensity and can efficiently be used for the analysis of DNA in living cells by LMPCR.

## 2.3 Introduction

Since its development in 1989 [1, 2], LMPCR (ligation-mediated polymerase chain reaction) has slowly gained recognition throughout the scientific community. Together with its derivative, TDPCR (terminal transferase-dependent PCR) [32, 33], LMPCR is one of the rare techniques available for *in vivo* DNA analysis. Compared to *in vitro* experiments that are very disruptive of the cellular environment, *in vivo* studies have the advantage of providing a representative picture of the chromatin structure in living cells. LMPCR has been used to map DNA-protein interactions in living cells that has furthered our understanding of gene regulation [3-8]. It has also proven to be useful in studying DNA repair rates [13-15], chromosomal methylation analysis [2, 9, 10] and nucleosome positioning [3, 11, 12]. LMPCR is the most sensitive technique available to map single-stranded DNA breaks at the nucleotide level of resolution using genomic DNA. The principle driving *in vivo* DNA analysis by LMPCR is to assess the local reactivity of DNA towards modifying agents like dimethylsulfate (DMS), ultraviolet light (UV) and DNase I inside living cells, compared to that of purified DNA [34, 35]. However, LMPCR is a complex technique that requires optimization of a multitude of variables, the most obvious being the combination of DNA polymerases for the different steps of the protocol. Previous studies have developed improved protocols in order to obtain efficient and reliable results [16, 18, 36]. The discovery and development of new DNA polymerases over the years has made it necessary to test their efficiency in LMPCR protocols.

The LMPCR procedure can be summarized in ten steps (Fig. 1): (I) conversion of modified bases to single-stranded DNA breaks characterized by a 5'-phosphate termini; (II) heat

denaturation of genomic DNA; (III) hybridization and extension of gene-specific oligonucleotide (primer 1) for the bottom or upper DNA strand to produce DNA molecules with an unknown double-stranded blunt-end; (IV) ligation of an asymmetrical double-stranded DNA linker to the blunt-ended DNA fragments providing a common sequence on the 5' end of all molecules; (V-VI) linear and exponential PCR amplifications of the DNA fragments using a gene-specific nested oligonucleotide (primer 2) and the linker-specific oligonucleotide (linker primer). Because there is no primer annealing to the complementary strand during the linear amplification, this strand will not be submitted to the subsequent PCR amplification; (VII) size-fractionation of the PCR products on a sequencing polyacrylamide gel followed by transfer of the DNA to a nylon membrane by electroblotting and UV crosslinking of the DNA to the membrane; (VIII) hybridization with a gene-specific labeled probe generated using primer 2 with a PCR product corresponding to the sequence to be analysed; and finally washing of the membrane and revealing of the sequence ladder by autoradiography. LMPCR preserves the quantitative representation of each fragment in the original population of cleaved molecules, allowing quantification on a phosphorimager.

Primer extension of DNA fragments and PCR amplification are the two major steps where the use of DNA polymerases needs to be optimized. For the primer extension step, the best DNA polymerase would be one that is thermostable, has no terminal transferase activity and efficiently produces blunt-ended fragments. Furthermore, it has to be able to polymerize through highly GC-rich sequences and DNA secondary structures. Similarly, for the amplification step, the ideal polymerase needs to be thermostable, very efficient, able to amplify indiscriminately a mixture of DNA fragments of different lengths (between 50 and 750 bp) and of varying GC-richness (from 5% to 95 %), as well as being able to efficiently resolve DNA secondary structures [17].

T7-modified DNA polymerase (*Sequenase*) and *Thermus aquaticus* DNA polymerase (*TAQ*) were initially used in LMPCR protocols [1, 2], but quickly *Thermococcus litoralis* DNA polymerase (*Vent*) was proved to be more efficient in primer extension and PCR amplification [16]. Our group has recently established the optimal conditions to use *Pyrococcus furiosus* exo- (*Pfu* exo-) [18], which was shown to be also more efficient than

*Sequenase* and *Taq* DNA polymerase. Meanwhile, these experiments have shown that each single polymerase possesses some peculiarities for their use. For example, optimal DNA-unit ratio needs to be identified to obtain highly efficient PCR results. Presently, some groups are using *Pfu* *exo-* in LMPCR protocols while others prefer to use *Vent* DNA polymerase; however, no comparison has been performed to determine which one of these two polymerases is the most efficient for LMPCR experiments.

Because of its 3'-5' exonuclease activity, we believed that *Vent* DNA polymerase would be less efficient than *Pfu* *exo-* in producing blunt-ended fragments at the primer extension step. Blunt-ends are necessary for efficient linker ligation preceding the PCR amplification step. Thus, we decided to compare *Pfu* *exo-* to *Vent* *exo-* DNA polymerase. *Vent* *exo-* has been genetically engineered to eliminate the 3'-5' proofreading exonuclease activity associated with *Vent* DNA polymerase [22]. In order to rigorously compare those two polymerases, we needed to establish the optimal conditions to use *Vent* *exo-* DNA polymerase in LMPCR protocols. Our first objective was to determine the ideal DNA to polymerase activity ratio during primer extension and PCR amplification steps. Because KCl stabilizes complex secondary DNA structures *in vitro* more efficiently than NaCl [37], we then tested if the substitution of KCl for NaCl in the DNA polymerase buffers would affect the efficiency of the polymerase. Thermostable DNA polymerases are prone to replication slippage when encountering DNA templates carrying hairpin structures, but *Vent* *exo-* demonstrated strand displacement capability, affected in part by temperature and magnesium concentration [28]. Therefore, we had to determine the optimal concentration of MgSO<sub>4</sub> to be used in LMPCR protocols with *Vent* *exo-*. Once those optimal conditions were established, our second objective was to compare the efficiency of *Vent* *exo-* to *Pfu* *exo-* when used with highly GC-rich sequences, with or without the addition of DMSO (dimethylsulfoxide) and 7-deaza-dGTP. The effect of the number of cycles in PCR amplification was evaluated as well as the identification of the optimal combination of polymerases to use in the primer extension and amplification steps.

## 2.4 Materials and methods

### 2.4.1 Chemicals, enzymes and equipment

*Vent* exo- and *Vent* DNA polymerase were purchased from New England BioLabs (Pickering, Canada). *Pfu* exo- and *Herculase* DNA polymerase were ordered from Stratagene (LaJolla, CA). *Taq* DNA polymerase and T4 DNA ligase were purchased from Roche Molecular Biochemicals (Laval, Canada). Dimethylsulfate (DMS), piperidine,  $K_2PdCl_4$  and hydrazine were purchased from Sigma-Aldrich Canada (Guelph, Canada). The desalted oligonucleotide primers were synthesized by Invitrogene (Burlington, Canada). All primer extensions and PCR amplifications were carried out on T gradient thermocycler from Biometra (Kirkland, Canada). Band intensities on autoradiograms were quantified using Fuji BAS 1000 phosphorimager from Fuji Medical Systems (Stamford, CT) and analyzed using software Image Gauge v3.0.

### 2.4.2 Synthetic Oligonucleotide Primers for LMPCR Analysis

The DNA sequence and  $T_m$  of the different primers used are shown in table 1.

### 2.4.3 Cell culture and DNA preparation

DNA was purified from peripheral blood lymphocytes of normal human males as described previously [17]. DNA concentration was determined using a spectrophotometer at 260 nm. Purified DNA was treated with DMS for guanine methylation or hydrazine with or without salt for pyrimidines in a standard Maxam-Gilbert cleavage reaction [38] or with  $K_2PdCl_4$  to reveal the adenine residues [39]. Hot piperidine (1 M, 82°C for 30 min) was then used to chemically convert the methylated guanines and damaged bases to single-strand DNA breaks. The single-strand break frequency was estimated using alkaline gel electrophoresis [40].

### 2.4.4 Ligation-mediated polymerase chain reaction (LMPCR)

The LMPCR protocol has been previously described in detail [17, 18]. Various regions of approximately 200 bp of the *P53*, *FMRI* and *C-JUN* genes were analyzed using the primer sets presented in table 1. Briefly, a gene specific primer was annealed to genomic fragments

of variable sizes and then extended using different DNA polymerases to produce double-stranded blunt-ends. Primer extension was performed through a thermal cycle of 5 min at 98°C, 4 min at the  $T_m$  of the primer minus 1°C, and 20 min at 75°C. The samples were immediately put on ice. An asymmetric double-stranded linker was then ligated to the phosphate terminal end of each fragment, providing a common sequence on the 5'-end of all fragments. With a selected DNA polymerase, a second nested gene specific primer was used for a single round of linear amplification, followed by PCR amplification using the second gene specific primer in conjunction with the linker primer. When using *Vent* exo-, PCR amplification was performed for 27 cycles of 2 min at 98°C, 2 min at annealing temperature, and 3 min at 75°C, with these modifications: (i) first-round of denaturation was 4 min at 98 °C; (ii) last-round extension was of 15 min at 75°C; (iii) annealing temperature started at  $T_m$  minus 1°C, and decreased with each cycle by 1°C until  $T_m -5^\circ\text{C}$  was reached.  $T_m -5^\circ\text{C}$  was maintained for 19 cycles and raised by 1°C per cycle until completion of the amplification; (iv) when required, modifications in cycle numbers were made at the  $T_m -5^\circ\text{C}$  annealing temperature; (v) Amplification using *Taq* DNA polymerase or *Pfu* exo- was performed as previously described [17, 18]. The resulting products were subjected to electrophoresis on 8% polyacrylamide gels alongside a Maxam and Gilbert sequencing ladder, followed by electrotransfer to nylon membranes, hybridization to a  $^{32}\text{P}$ -labeled gene specific probe, and visualization by autoradiography. Each experimental condition was assayed in duplicate and then run on a screening sequencing gel using a portion of the DNA to ensure that there was no significant variation between samples. The duplicates were then pooled on a combined gel which served for analysis.

#### **2.4.5 Phosphorimager quantification**

Phosphorimager analyses were performed to measure precisely the intensity of individual bands as described previously [18, 41]. Quantification of band intensities was used to determine relative intensity and signal background ratio. LMPCR protocols were repeated twice. For each protocol, a total of ten sets of bands were selected for each experimental condition of both autoradiograms. A set is defined as the bands representing DNA fragments of the same length for each experimental condition. Using the phosphorimager software package provided by the manufacturer (Image Gauge V3.0), a rectangle was

drawn around each selected band on the LMPCR autoradiogram. The band intensity value corresponding directly to the amount of radioactive material was read within each rectangle as counts per band surface. In each set, an area corresponding to the experimental condition that did not generate a band (either experimental condition without DNA polymerase or without DNA) was delimited and defined as “control value”. This control value was subtracted from all the measured bands of a set. The relative intensity value of each band within a set was then determined by dividing the intensity value of each band (minus the control value) over the cumulated intensity value of all bands composing that same set. From the ten sets of bands selected for both autoradiograms, the mean relative intensity values were calculated for each experimental condition and plotted against amounts of DNA, DNA polymerase, or MgSO<sub>4</sub> concentration

The signal to background ratio was determined as follow: An area between two sets of bands for each experimental condition, determined as the highest background noise found throughout the autoradiogram, was quantified using the phosphorimager software. The mean relative intensities of each experimental condition, calculated earlier, were then divided by the intensity value of this highest background noise found for each experimental condition. The ratio obtained, representing band intensity over background noise for each experimental condition, was plotted against amounts of DNA, DNA polymerase, or MgSO<sub>4</sub> concentration

## **2.5 Results**

### **2.5.1 Primer extension step**

The efficiency of *Vent* *exo-* for the primer extension step was compared with the optimal condition of *Pfu* *exo-* established previously by our group [18]. Three different activities of *Vent* *exo-* DNA polymerase (0.5 U, 1.0 U and 1.5 U) were tested for primer extension using purified genomic DNA treated with DMS to produce a global SSB frequency of 1 break per 400 bases and were compared to 1.5 U of *Pfu* *exo-*. DNA sequences of approximately 200 bp with relatively high GC content (varying from 60% to 80% of GC) located in exon 5 of *P53* gene, and the promoter region of *FMRI* and *C-Jun* genes were mapped by LMPCR using 3 U of *Taq* DNA polymerase at the amplification step. The results, shown in figure

2A, represent the selected area for each of those three genes. Globally, the intensity of the bands increases as the DNA quantity rises from 0.4  $\mu\text{g}$  to 2.4  $\mu\text{g}$ . The samples processed with *Vent* *exo*- showed a greater increase in signal intensity with rising DNA quantity than samples processed with *Pfu* *exo*-. Compared to *Pfu* *exo*-, we observed that the band intensity was less uniform with *Vent* *exo*-. In fact, some specific bands were absent at the 0.4  $\mu\text{g}$  DNA dose and only appeared at higher DNA quantities.

When using low amounts of DNA (0.4 and 0.8  $\mu\text{g}$ ), 0.5 unit of *Vent* *exo*- DNA polymerase was the most efficient amount of enzyme for the autosomal (2 copies per genome) *P53* and *C-Jun* gene DNA sequences, whereas 1.0 unit was optimal for the gonosomal (1 copie per genome) *FMRI* gene sequences (Fig. 2A). When using higher amounts of DNA (1.6 and 2.4  $\mu\text{g}$ ), 1.0 unit of *Vent* *exo*- DNA polymerase was the most efficient amount of enzyme for all three genes tested (Fig. 2A). These results underline the importance of the ratio of DNA molecules per unit of DNA polymerase during primer extension. Compared to samples processed with *Pfu* *exo*-, the samples processed with *Vent* *exo*- produced bands with similar or greater intensity only when using 2.4  $\mu\text{g}$  of DNA.

Quantification of band intensity with the phosphorimager allowed us to analyze the relative band intensity produced by each polymerase at specific concentrations for each amount of DNA (Fig. 2B). This indicates if the efficiency of a DNA polymerase varies with the amount of DNA. Globally, the relative band intensity increased with the amount of DNA. It should be noted that the increase was less important for samples processed by *Pfu* *exo*-, mainly for the *FMRI* gene sequences. By careful measurement of the signal to background ratio, it was shown that the increase in the band intensity with the amount of DNA was the result of a higher efficiency of primer extension, rather than an increase in the background noise (data not shown). At higher amounts of DNA (1.6 and 2.4  $\mu\text{g}$ ), the samples processed by *Pfu* *exo*- showed a high background noise causing a decrease in the signal to background ratio (data not shown). This decrease in the signal to background ratio was not observed for *Vent* *exo*-.

In conclusion, for primer extension, when using 1.0  $\mu\text{g}$  of DNA (commonly used in LMPCR protocols), 1.5 U of *Pfu* *exo*- shows the best relative intensity and signal to background ratio and therefore would be the best choice. On the other hand, when using

more than 1.25  $\mu\text{g}$  of DNA, 1.0 U of *Vent* *exo*- shows a similar intensity and signal to background ratio as 1.5 U of *Pfu* *exo*- and therefore would be a very good alternative to *Pfu* *exo*-.

### 2.5.2 PCR Amplification step

Using *Pfu* *exo*- at the primer extension step, the efficiency of *Vent* *exo*- at the PCR amplification step was studied as a function of two parameters, i.e., varying the number of units of DNA polymerase (Fig. 3) or the amount of DNA (Fig. 4). For these two parameters, we used the same primer sets as in primer extension step. For each primer set, the experiment was repeated twice for a total of 6 autoradiograms (for *FMRI* gene, only primer set X was used). Purified DNA was treated with DMS and piperidine to produce a global SSB frequency of 1 break per 400 bases. Phosphorimager quantification was used to determine relative intensity and signal to background ratio as explained in the materials and methods section. By varying the number of units of *Vent* *exo*- used to amplify 1  $\mu\text{g}$  of DNA (Fig. 3), a characteristic normal curve of the DNA molecules per unit of DNA polymerase ratio was observed, with the optimal quantity of enzyme being around 0.5 to 2.0 U.

Concerning the *P53* gene (Fig. 3A), as little as 0.13 U of *Vent* *exo*- was sufficient to produce bands of fairly good and uniform intensity. From this low amount of enzyme, the intensity of the bands increased to a maximum reached at around 1 U (the optimum dose varying between 0.5 to 2.0 U). At higher doses of *Vent* *exo*-, the intensity of the bands decreased to almost nothing at around 8 U. The quantification performed using the phosphorimager confirmed that the maximum intensity of the bands was reached at 1 U of *Vent* *exo*- (Fig. 3B). The signal to background ratio showed an increase from 0.13 to 6 U (Fig. 3C), suggesting that the increase in band intensity was evidence of an improved amplification and not only an increase in the background noise. For the *FMRI* gene, the range of units giving very efficient amplification of DNA fragments was smaller than the one observed with the *P53* gene (Fig. 3A). Bands were seen when using 0.25 U, and their intensity increased to a maximum at around 1 U. There was a high background and the intensity of the bands decreased to no signal at around 6 U. The graph illustrating the



relative intensity (Fig. 3B) clearly showed the maximum intensity at 1 U, but this was not caused by a high background noise, as confirmed by signal to background ratio (Fig. 3C). Concerning the *C-Jun* gene, bands started to be seen when using 0.13 U of *Vent* *exo-* and were observable at up to 6 U. The maximum intensity of the bands was reached at 2.0 U (the optimum dose varying between 0.5 to 3.0 U) (Fig. 3A). The maximum intensity was found at 2 U of *Vent* *exo-* according to the calculation of the relative intensity (Fig. 3B). The signal to background ratio (Fig. 3C) was lower than for the other two genes. In fact, the intensity of the bands did not increase much for dose over 0.5 U compared to *P53* and *FMRI* genes.

Similarly, we tested the efficiency of *Vent* *exo-* on various amounts of DNA (Fig. 4). We used *Pfu* *exo-* for primer extension followed by amplification using 2.5 U of *Vent* *exo-*. The amount of DMS treated DNA to produce a global SSB frequency of 1 break per 400 bases varied from 0 to 5  $\mu$ g. The same three sets of primers as above were tested and quantification of bands intensity was conducted in the same way as described before.

For all three primer sets, *Vent* *exo-* was efficient with very low amounts of DNA (Fig. 4A). At very low amounts of DNA, such as 0.1 $\mu$ g, the intensity of the bands was sufficient to be analyzed, and there was uniform intensity between bands (Fig. 4A, lane 2). Concerning the *P53* primer set, the relative intensity of the bands increased quickly to reach a maximum at 2  $\mu$ g of DNA, then the intensity quickly decreased at a lower level than for the *FMRI* and *C-Jun* primer sets (Fig. 4B). In the case of *FMRI*, the intensity increased more slowly compared to other genes, but the maximum intensity was also reached at 2  $\mu$ g of DNA. When using 5  $\mu$ g of DNA, the intensity of the bands was still sufficient to be analyzed in LMPCR experiments. Concerning *C-Jun*, the intensity of the bands rose quickly with the amount of DNA, but reached a lower maximum intensity than for *P53* or *FMRI*, but the maximum was at around 2  $\mu$ g. The decrease occurring with higher amounts of DNA was similar to the one observed with *FMRI*, but happened more slowly compared to *P53*. Finally, the signal to background ratio of *P53* and *FMRI* was very similar throughout the range of DNA doses (Fig. 4C). In the case of *C-Jun*, we observed a very strong signal with low background noise giving us a very high ratio. The high relative intensity in the case of

*P53* noted above (Fig. 4B) was associated with a low signal to background ratio; this was not the case for *C-Jun*.

### 2.5.3 Optimal MgSO<sub>4</sub> concentration and effect of salt on PCR efficiency

The MgSO<sub>4</sub> concentration can greatly affect the efficiency of PCR amplification performed with *Vent* *exo-* [28]. Therefore, it was important to identify the optimal MgSO<sub>4</sub> concentration for the most efficient *Vent* *exo-* DNA amplification. For the primer extension step, 1.5 U of *Pfu* *exo-* and 1 µg of DNA were used, then followed by the amplification step using 1.5 U of *Vent* *exo-* to map a 200 bp sequence of the exon 5 of the *P53* gene. *Vent* polymerase ThermoPol reaction buffer was made without MgSO<sub>4</sub>, which was subsequently added to the amplification mix at various concentrations. Relative intensity was determined by phosphorimager quantification.

Between 0 and 0.5 mM of MgSO<sub>4</sub>, no amplification can be seen, but at 1 mM MgSO<sub>4</sub> bands can be observed. The intensity of the bands increased to a maximum at around 3 mM, and then decreased slowly to a minimum at around 25 mM (Fig. 5A). The relative intensity, calculated from phosphorimager quantification data, showed the same pattern of increasing and decreasing efficiency modulated by MgSO<sub>4</sub> concentrations (Fig. 5B). The analysis of signal to background ratio between 1 to 6 mM of MgSO<sub>4</sub> showed that the increase in relative intensity was modulated by a small increase in background noise (Fig. 5C). Finally, the overall variation in signal to background ratio was quite small throughout the concentrations of MgSO<sub>4</sub> tested.

It has been shown that K<sup>+</sup> cations promote the stabilization of unusual secondary structures in the DNA *in vitro* (ex: tetrahelix) better than Na<sup>+</sup> cations [37]. Because most commercial polymerase buffers are composed of KCl, we decided to test the efficiency of *Vent* *exo-* compared to *Pfu* *exo-* polymerase under NaCl conditions. We did not observe any significant difference in amplification efficiency when using KCl or NaCl, either at primer extension or amplification for any particular combination of polymerase in any of the three genes tested (data not shown).

#### 2.5.4 Number of cycles in the PCR amplification step

Because DNA fragments are exponentially amplified in PCR, the intensity of the signal is directly affected by the number of cycles. Therefore, we compared *Vent* *exo*- with *Pfu* *exo*- and *Taq* DNA polymerase in their ability to efficiently amplify DNA fragments when submitted to various numbers of cycles in LMPCR amplification step. Primer extension of 1 µg of genomic DNA was performed using *Pfu* *exo*- for all samples and amplified with either 1.5 U of *Vent* *exo*-, 3.5 U of *Pfu* *exo*- or 3 U of *Taq* DNA polymerase, to successfully mapped a 200 bp stretch of the *P53* gene (Fig. 6).

For *Pfu* *exo*- and *Vent* *exo*-, 12 and 17 cycles (22 for the *Taq* DNA polymerase) were insufficient to produce any visible bands for the exposure time used (Fig. 6, lanes 1-2, 7-8 and 13-15). Faint bands appeared after 22 cycles (27 for the *Taq* DNA polymerase) (Fig. 6, lanes 3, 9 and 16) and increased quickly at 27, 32 and 37 cycles (Fig. 6, lanes 4-6, 10-12 and 17-18). As the number of cycles increases, the background noise increases, mostly caused by overexposure of the autoradiogram. Compared to *Vent* *exo*-, when amplification was performed using *Pfu* *exo*- (lanes 7-12), the variation in band intensity between 32 and 37 cycles was greater (Fig. 6, compare lanes 11 and 12 vs. 5 and 6). With samples amplified using *Taq* DNA polymerase, the intensity of the bands was weaker than the ones produced using the other polymerases independently of the number of amplification cycles.

#### 2.5.5 Highly GC-rich sequences

GC-rich sequences often impair PCR amplification, especially in LMPCR protocols. We tested the efficiency of *Vent* *exo*- in resolving particular DNA structures created in highly GC-rich sequences like in exon 1 of the *FMR1* gene, known to contain a polymorphism for the CGG repeats that ranges from a few to several hundred repeats [5, 19]. We compared the efficiency of *Vent* *exo*- with *Pfu* *exo*- at the extension step with a combination of various polymerases at amplification, such as *Pfu* *exo*-, *Vent* *exo*-, *Taq* and *Herculase*. This experiment was conducted with and without the addition of 7-deaza-dGTP and/or DMSO to determine if these products would improve *Vent* *exo*- efficiency compared to the standard LMPCR protocol.

With the standard LMPCR protocol, *Taq* was very inefficient in amplifying DNA fragments regardless of the polymerase used at the primer extension step (Fig. 7A, lanes 5, 6, 13 and 14). Amplification using *Pfu* *exo-* allowed us to sequence 19 and 20 CGG repeats when extended with *Pfu* *exo-* (Fig. 7A, lanes 1 and 2) and *Vent* *exo-* (Fig. 7A, lanes 9 and 10) respectively. Amplification using *Vent* *exo-* allowed us to sequence 21 repeats (Fig. 7A, lanes 3, 4, 11 and 12), but artifacts derived from amplified fragments consisting of one extra nucleotide, that we refer to as shadow bands, were observed throughout the C ladders and at the beginning of the G ladders before the beginning of the repeats. These shadow bands created the impression of double CC overlapping one of the two G in the CGG repeats. This was noted independently of the polymerase used at the extension step. *Herculase* amplified DNA showed a short but intense G ladder (Fig. 7A, lane 7) but no C ladder (lane 8) when combined with *Pfu* *exo-* at primer extension. When the extension step was performed using *Vent* *exo-*, *Herculase* amplified samples showed the most intense G and C repeated sequences (Fig. 7A, lanes 15 and 16).

When using 7-deaza-dGTP, we observed that just a few repeats were added to the ones already mapped by standard protocols (Fig. 7B). The C ladder extended with *Pfu* *exo-* was this time successfully amplified with *Herculase* (Fig. 7B, lane 8). When *Vent* *exo-* was used at primer extension, the C ladder amplified with *Pfu* *exo-* was missing (Fig. 7B, lane 10), and a total of 24 repeats were successfully mapped with *Vent* *exo-* at amplification (Fig. 7B, lanes 11 and 12). *Taq* amplified samples were still unsuccessfully mapped with this protocol (Fig. 7B, lanes 5, 6, 13 and 14).

By adding 7-deaza-dGTP and DMSO (Fig. 7C), we noted that up to 33 repeats on the G ladder were mapped when using *Herculase* coupled with *Pfu* *exo-* at the primer extension step (Fig. 7C, lane 7), but for all our samples the C ladders were missing. When the extension step was carried out using *Vent* *exo-*, only samples amplified with *Herculase* were visible. Thus, we decided to test our protocol with DMSO in absence of 7-deaza-dGTP (Fig. 7D). When *Pfu* *exo-* was used at extension with *Vent* *exo-* at amplification, we successfully mapped up to 33 repeats including a strong C ladders containing shadow bands (Fig. 7D, lanes 9 and 10). Finally, we found that a very intense and clear 33 repeats were

also mapped using a combination of *Vent* *exo-* at primer extension and *Herculase* at the amplification step (Fig. 7D, lanes 5 and 6).

### 2.5.6 Polymerases efficiency at primer extension and amplification steps

Specific characteristics of a DNA polymerase, such as processivity and terminal transferase activity, greatly influence LMPCR results. By comparing various combinations of polymerases at the primer extension and amplification steps, we can evaluate the relative efficiency of a series of polymerases. Therefore, we tested the efficiency of *Pfu* *exo-*, *Vent* *exo-* and *Vent* DNA polymerase at the primer extension step for their ability to produce blunt-ended DNA fragments required for efficient linker ligation. We then subjected these samples to amplification with different polymerases such as *Pfu* *exo-*, *Vent* *exo-*, *Vent* and *Taq* DNA polymerase to successfully map a region of the *P53* gene (Fig. 8).

Using *Pfu* *exo-* at the primer extension step (Fig. 8, lanes 1-4), weak bands were observed for samples amplified with *Taq* (lane 4) compared to the ones obtained when using *Pfu* *exo-*, *Vent* *exo-* and *Vent* polymerase (Fig. 8, lanes 1, 2 and 3). In fact, these three polymerases yield very similar results, but slightly less background noise with *Vent* *exo-* for relatively similar band intensity. Surprisingly, this was not the case when using *Vent* *exo-* for primer extension (Fig. 8, lanes 5-8). We observed an increase in band intensity for all polymerases except for *Vent* *exo-* amplified samples (Fig. 8, lane 6), where the decrease in band intensity was quite significant. Interestingly, increasing the number of units of *Vent* *exo-* from 1.5 to 3 U at amplification greatly improved the signal (Fig. 8, compare lanes 6 vs. 14).

Finally, we observed a very weak signal when using *Vent* DNA polymerase at the primer extension step for all four polymerases (Fig. 8, lanes 9-12). This was accompanied by the appearance of many shadow bands visible when combined with *Pfu* *exo-* at the amplification step (Fig. 8, lane 9). These shadow bands were much less pronounced for the other amplified samples (Fig. 8, lanes 10, 11 and 12). Finally, the characteristic migration shift of all samples amplified with *Taq* DNA polymerase due to its terminal transferase activity was observed (Fig. 8, lanes 4, 8 and 13).

## 2.6 Discussion

### 2.6.1 Primer extension step

The efficiency of DNA polymerases for the extension of genomic fragments is clearly affected by the amount of DNA. Our results showed compelling evidence of the importance of the DNA / polymerase ratio. Band intensities for *Vent* *exo*- samples were faint for low amounts of DNA compared to the ones with *Pfu* *exo*-, but they increased with DNA levels suggesting that the DNA / polymerase ratio of *Vent* *exo*- is higher than for *Pfu* *exo*-. Therefore, with low amounts of DNA, *Vent* *exo*- might not be the best choice for primer extension, but with higher amount of DNA, it becomes very efficient.

Since each guanine residue of a DNA sequence has the same probability of being methylated *in vitro* by DMS, all band intensities should be equal on the autoradiogram. Variable band intensity in the *P53* samples extended with *Vent* *exo*- could not have resulted from premature polymerization arrest or annealing problems, since bands corresponding to longer DNA molecules were much more intense than bands corresponding to smaller DNA molecules. Inefficient ligation and PCR amplification were not determining factors because these steps were identical in both *Vent* *exo*- and *Pfu* *exo*- LMPCR protocols. We suggest that *Vent* *exo*- was unable to produce ligatable blunt extremities with the same efficiency for all cleaved DNA molecules. This seems to only happen on G stretches and this was more obvious with *P53*, but still happened with *FMRI* and *C-JUN*. We could not find any sequence specificity but we should mention that this was not observed at optimal polymerase quantity.

One of the key characteristics we are looking for when selecting a polymerase for LMPCR studies, is its ability to create blunt-ended fragments at the primer extension step. Suppression of ligatable ends by DNA polymerases results on either from their inability to extend to the end of the template or their terminal transferase activity. In LMPCR, the addition of an extra non-templated nucleotide at the extremity of blunt double-stranded DNA molecules prevents ligation of the asymmetric linker, and subsequently inhibits the amplification of these molecules, thereby modifying the quantitative representation of the corresponding bands on the autoradiogram. Therefore, it becomes obvious that a

polymerase possessing a terminal transferase activity, like *Taq* DNA polymerase, is of no use at the primer extension step, since most of the DNA fragments extended with *Taq* will possess a non-templated 3' extra nucleotide [25, 26]. It was reported that *Vent* exo-polymerase produces an additional base corresponding to extension beyond the template by one nucleotide in approximately 30% of the total extension products and would not be expected to efficiently create blunt-ended termini [22]. This might explain why under sub-optimal conditions, we obtained non-uniform signals when using *Vent* exo- at primer extension. One would expect that if there was a sequence-dependent terminal transferase activity, certain bands would be missing, and if there were not such a sequence dependency, all the samples would have the same probability of being over extended. Therefore, all bands should appear 30% weaker when extended with *Vent* exo- compared to samples extended with *Pfu* exo- which is known to produce close to 100% blunt ended DNA molecules.

Nevertheless, linker ligation can also be inhibited if the polymerase possesses a 3'-5' exonuclease activity, because a fully extended DNA fragment could become 5' protuberant. Because *Vent* exo- is exonuclease deficient, we suggest that suppression of ligatable ends by this polymerase is most probably caused by non-extended products, due to its low level of processivity and its distributive behavior [22]. Because dissociation and reinitiation is a frequent event for a distributive enzyme, it is believed that *Vent* exo- would leave DNA fragments non-fully extended before switching to another template. Under optimal conditions, the remaining unfinished template would be extended by reinitiation events of other polymerases. Some templates could therefore remained incompletely extended when sub-optimal polymerase quantities are used. Furthermore, too much polymerase could compete to occupy DNA templates and therefore pause at specific sites until switching occurs.

Finally, we suggest that *Vent* exo- is more efficient with high amounts of DNA, but can be used in primer extension by reducing to a minimum its quantity. Because of the variability of the results for the three genes tested, we suggest that for 1  $\mu$ g of DNA, one would use 0.75 U of *Vent* exo-. Taken altogether, this demonstrates the importance of conducting evaluative testing on more than one gene to obtain the complete picture.

### 2.6.2 PCR amplification step

During amplification, the efficiency of a polymerase can be directly related to band intensity in LMPCR experiments. One could question if it is better to obtain maximum band intensity or simply to overexpose faint bands. The answer to this question is given by the relative background noise. It is of no use to reach maximum band intensity if there is a significant increase in background noise. Therefore, we should identify conditions which result in appropriate signal intensity, uniformity in band intensity and finally, in a high signal to background ratio.

Our results showed that under our conditions, the effective range of amplification efficiency using *Vent* *exo*- varies between 0.25 to 5 units of polymerase with an optimal quantity around 1.5 U, considering the relative intensity and signal to background ratio for three different genes. This becomes interesting when comparing these results to the ones previously published on *Pfu* *exo*- and *Taq* DNA polymerase [18]. *Pfu* *exo*- optimal quantity was established at 3.5 U in an effective range varying between 2 and 25 U, tested under similar experimental conditions. *Taq* optimal quantity was established at 3 U varying from 0.13 to 15 U. The range of *Vent* *exo*- amplification efficiency is therefore quite small compared to *Pfu* *exo*- and *Taq* DNA polymerase. With successful amplification of DNA fragments with only 0.13 U of *Vent* *exo*- (Fig. 3A, P53, lane 2), our results showed the very high DNA / polymerase ratio for this enzyme compared to *Pfu* *exo*-. This ratio being critical, it also explains that at more than 6 U, there is a decrease in band intensity whereas the low DNA / polymerases ratio of *Pfu* *exo*- permits successful amplification up to 25 U.

We also noticed that the supplier recommended using 4 U / 100  $\mu$ l reaction volume of *Vent* *exo*- when performing PCR amplification, which is obviously more than the optimal quantity as determined in our experimental conditions. In fact, with this very small range of efficiency, it becomes critical to use the appropriate amount of *Vent* *exo*- when carrying out PCR experiments.

Interestingly, low amounts of DNA, such as 0.1  $\mu$ g were sufficient to obtain successful amplifications using *Vent* *exo*-. It was reported that the  $K_m$  of *Vent* *exo*- for DNA (0.09) was one of the lowest found in the literature, and that it should give the polymerase an



enhanced ability to efficiently replicate extremely low amounts of DNA [22]. This is not an intrinsic property of thermostable polymerases since *Taq* and *Pfu* DNA polymerase have higher  $K_m$ 's (1.4) [22, 42]. Our results showed compelling evidence of the sensibility of *Vent* *exo-* to low amounts of DNA, particularly when comparing it to previous published results on *Pfu* *exo-* and *Taq* DNA polymerase, which did not present this kind of sensibility [18].

Furthermore, with increasing amounts of DNA, we still obtained an excellent amplification of DNA fragments. When using specific primers for *FMRI* and *C-JUN* genes, we successfully amplified 5.0  $\mu$ g of DNA with excellent intensity, suggesting that we could have succeeded at higher quantities if tested. With *P53* gene, relative intensity calculation clearly showed that higher ( $> 5.0 \mu$ g) amount of DNA, no further amplification was efficient. This indicates once again the importance of conducting optimization experiments using several genes because of intrinsic variations, potentially caused by secondary structures and/or primer characteristics.

### 2.6.3 Salt conditions

In addition to its thermal stability, what makes *Vent* *exo-* DNA polymerase an interesting candidate for LMPCR is its ability to progress through special secondary DNA structures, via its strand displacement activity [22]. *Pfu* *exo-* does not possess strand displacement activity and therefore undergoes replication slippage, which generates hairpin deletions during PCR amplification. *Vent* *exo-* is also able to slip; however, slippage can be inhibited when its strand-displacement activity is induced [28]. It was also demonstrated that high magnesium concentration (over 4 mM) could stabilize secondary structures and restrict the ability of *Vent* *exo-* to enter and polymerize through the hairpin structure, which would result in strand slippage. Deletion of a DNA segment could result in missing bands and therefore would interfere with proper LMPCR analysis.

Because most known DNA polymerases require magnesium ions to polymerize, it is critical to identify the optimal  $MgSO_4$  concentration for the most efficient *Vent* *exo-* activity. Our results showed that little variation in the amplification efficiency occurs between 1.5 to 12.5 mM of  $MgSO_4$ . Taking into account the relative intensity and background noise, we

determined that 3.0 mM should be used for efficient and optimal PCR amplification. In the same time, this concentration should allow efficient strand displacement activity and permit polymerization through particular secondary DNA structures.

#### **2.6.4 Thermostability of *Vent* exo-**

It is expected from any enzyme isolated from hyperthermophile to be thermostable. It was reported that *Vent* exo- half-life at 100°C is of 1.8 hours and of 6.7 hours at 95 °C. *Taq* DNA polymerase does not retain any activity after this 1.8-hour incubation time [22]. This potential stability over a wide range of extreme temperatures becomes a necessity for PCR amplification in LMPCR protocols. Frequently, highly GC-rich sequences need to be analyzed, which require very high denaturation and annealing temperature. This is the reason why we compared the capability of *Vent* exo-, *Pfu* exo- and *Taq* DNA polymerase to sustain numerous PCR cycles of high denaturing temperature at the amplification step.

Our results demonstrated the higher efficiency of *Pfu* exo- and *Vent* exo- compared to *Taq* DNA polymerase in successfully amplifying DNA fragments at high cycling temperatures. We believe that this is a result of thermal stability because *Taq* extension rates and the  $K_m$  for dNTPs are similar to those of *Vent* exo-, and the *Taq*  $K_m$  for DNA is similar to that of *Pfu* [22, 42]. Our standard protocol is based on 27 cycles, but we increased it up to 37 to test if a plateau in DNA amplification would be reached. We observed that at 37 cycles, the signal of samples amplified with *Pfu* exo- still increased, whereas the intensity of *Vent* exo-treated samples seems to top off. This particularity cannot result in reaching the PCR plateau phase faster in one sample compared to the other, which would still be in the exponential amplification phase, because we carefully made sure that all constituent (DNA, dNTPs, primers and buffers), except the polymerase, were from the same source and at the same concentration. To explain these results, we suggest that the thermal stability of *Pfu* exo- is superior to the one of *Vent* exo- and therefore can sustain many more cycles before the enzyme loses its polymerization capability.

### 2.6.5 Highly GC-rich sequence

The *FMRI* gene is characterized by a particular and complex DNA secondary structure composed of a succession of CGG triplet repeats varying between 6 and 250 [19]. This kind of structure, composed of highly GC-rich sequence, can often impair PCR analysis because of the difficulty of the DNA polymerase in resolving them. This particular sequence of the *FMRI* gene has previously been mapped with a modified LMPCR protocol using DMSO and 7-deaza-dGTP [18].

It was demonstrated that intramolecular base pairing between guanine and cytosine residues could cause anomalies in the migration behavior of certain DNA fragments in the polyacrylamide gel, and therefore results in band compression and difficulty in analyzing sequencing reaction [43]. The addition of 7-deaza-dGTP to the polymerization mixture was shown to resolve this particular problem by preventing base compressions, caused by stacking of the sequenced fragments, which lead to abnormal migration patterns during gel electrophoresis [44]. Our hypothesis was that the 7-deaza-dGTP could not only act on gel migration of the amplified DNA fragments but also in helping the polymerase resolving the complex structure associated with the CGG repeats found in the exon 1 of *FMRI* gene. The 7-deaza dGTP precludes Hoogsteen bond formation by virtue of replacement of the N-7 of the guanine ring with a methane moiety [45]. Furthermore, many studies have reported the use of additives like betaine, DMSO and formamide to enhance PCR amplification yield and specificity of GC-rich sequences [46, 47]. More specifically, DMSO has been used to rescue a failed PCR catalyzed by *Vent* DNA polymerase [48]. In our experiments, we wanted to compare the efficiency of *Vent* exo- with *Pfu* exo- in resolving this particular sequence when combined to various polymerases, without adding any DMSO or 7-deaza-dGTP. We also wanted to examine the individual and combined effects of these substances on these same polymerase combinations.

Ours results demonstrated the inefficacy of *Taq* DNA polymerase to efficiently resolve this particular CGG repeat. We believe that the high denaturing and annealing temperatures required for amplification of a 100% GC-rich sequence are beyond the thermal stability of this polymerase. This might also explain why no further improvements were obtained with the addition of 7-deaza-dGTP and/or DMSO.

Similar results were obtained using *Pfu* *exo-* and *Vent* *exo-* at the primer extension step with and without 7-deaza-dGTP. The only striking difference is the missing C ladder of the sample amplified with *Herculase* when primer extension was performed with *Pfu* *exo-*. The strand displacement activity of *Vent* *exo-* at the primer extension step might be involved in easing the secondary structure of this region, which then becomes more easily accessible by the polymerase used at the amplification step. No reports have appear concerning the efficiency of *Herculase* in resolving secondary structures, but it consists predominantly of *Pfu* DNA polymerase combined with *Taq* 2000 [49, 50], we believe it would react as these two polymerases do separately. Therefore *Herculase* could cause replication slippage when encountering particular DNA secondary structures.

Another explanation for the increased intensity of the signal using *Herculase* consists in the presence of the ArchaeMaxx polymerase-enhancing factor who would react with some *Vent* *exo-* DNA polymerase leftover from the primer extension step. The ArchaeMaxx factor functions as a dUTPase to overcome dUTP poisoning, which is caused by dUTP accumulation during PCR through dCTP deamination [29, 49, 50]. It was reported that archaeal DNA polymerases, such as *Pfu* and *Vent* DNA polymerase, were inhibited by dU-containing DNA [29, 30]. Therefore, some *Vent* *exo-* could be combined with the *Herculase* mix and improved the amplification efficiency of this region due to its strand displacement activity and in return the ArchaeMaxx factor would favor efficient polymerization with *Vent* *exo-* by inhibiting dUTP poisoning.

When amplification was performed with *Vent* *exo-* DNA polymerase, we noted the presence of intense shadow bands in the C ladder. These artifacts were always positioned one nucleotide above the expected ones and therefore, were interpreted as to be the results of the addition of an extra base to a high proportion of the DNA fragments amplified by *Vent* *exo-* polymerase. Nevertheless, this was surprizing because all other amplification experiments carried out in this project involving the use of *Vent* *exo-* DNA polymerase did not show any of this activity and always gave us clear and precise bands. There might be a relation with the particular structure and sequence of the *FMRI* gene, or maybe an undetected sequence dependency of the *Vent* *exo-* terminal transferase activity.

Clearly, the addition of 7-deaza-dGTP did not significantly improve the amplification efficiency for any of the polymerases tested. It actually reduced the efficiency when combine to DMSO. As for the addition of DMSO, it seems to either radically improve or inhibit the amplification in the various conditions examined. Therefore, further testing will be required to identify the optimal concentration of DMSO to use and the effect it has on polymerase efficiency.

### **2.6.6 Combining polymerases at primer extension and amplification steps**

Previous research has demonstrated the efficiency of *Pfu* *exo-* at the primer extension in producing blunt-ended fragments [18]. Our results showed an evident increased in band intensity when substituting *Pfu* *exo-* for *Vent* *exo-* DNA polymerase at the primer extension step, except when the amplification was conducted with 1.5 U of *Vent* *exo-*, where we observed a significant decrease in band intensity. This was resolved by adjusting the amount of *Vent* *exo-* to 3 units at the amplification. Our results clearly showed that *Vent* *exo-* could efficiently create blunt-ended fragments in our experimental conditions, which is contradictory to what was published originally [22]. We therefore suggest that terminal transferase activity of *Vent* *exo-* is dependent of specific conditions yet to be identified.

Garrity and Wold have previously developed an LMPCR procedure using *Vent* DNA polymerase and confirmed its efficiency at the primer extension step when compared with *Sequenase* and *Taq* DNA polymerase [16]. It was also reported that 90 % of all extended products of *Vent* DNA polymerase were blunt ended [22]. Our results demonstrated that *Vent* DNA polymerase could effectively be used for primer extension but with far less efficiency than *Pfu* *exo-* or *Vent* *exo-*. Interestingly, many shadow bands were observed when *Pfu* *exo-* was used to amplify *Vent* extended samples. We suggest that the exonuclease activity of the polymerase might play a role in our experimental conditions, producing DNA fragments of variable lengths, but somehow there must be a link with the used of *Pfu* *exo-* because this was not observed when combined to other polymerase.

The efficiency of a specific polymerase can also be examined at the amplification step. The signal intensity can be correlated to the efficiency of a polymerase to amplify particular DNA fragments. Thus, using one polymerase at the primer extension, we can then compare

the efficiency of various polymerases at the amplification step. From there, we can assume that a polymerase characterized by a high terminal transferase activity will cause an upward migration shift in all fragments, like observed when using *Taq* DNA polymerase at the amplification. In the case of polymerase with exonuclease property, a migration shift could also be observed but will result in the appearance of artifactual bands located under the expected ones. We can presume that if this exonuclease activity is high, more shifted bands will be observed, but if it is mild, those nonspecific bands will be faint compared to the specific ones.

Our results showed compelling evidence of the lack of terminal transferase activity of *Vent* *exo-* DNA polymerase in this experiments (Fig. 8) when combined with various polymerases. No shadow bands or artifacts can be seen in any *Vent* *exo-* amplified samples, except with *FMRI* GCC repeats, when compared to the migration shifted observed with *Taq* amplified samples. *Vent* *exo-* seems very efficient at the amplification step compared to the other polymerases, except when primer extension was performed with this same polymerase. As mentioned earlier, the signal intensity was brought back to optimal efficiency by raising the amount of *Vent* *exo-* to 3 U. Taking all these combinations of polymerase into account, this suggests that there is a link between the polymerase used at the primer extension with the polymerase chosen for amplification. Further evidence is shown by the appearance of shadow bands in the *Pfu* *exo-* amplified samples previously submitted to extension by *Vent* DNA polymerase. If it was related to the specific properties of the *Vent* DNA polymerase during primer extension, it should have reacted in a similar way with all the other polymerases used for the amplification. In the same way, this effect was not observed with other amplifications performed with *Pfu* *exo-*. Therefore, something must have remained from the extension product that affected the following amplification.

Because all single steps of the LMPCR protocol were conducted exactly the same with the same constituents for all samples, the only difference being the polymerase used, we suggest that some active units of polymerase remained throughout ligation and DNA precipitation prior to the amplification phase. These remaining units of polymerase could then cooperate or compete with the other selected polymerase during PCR amplification. Because optimal ratio of *Vent* *exo-* DNA polymerase, amount of DNA and  $MgSO_4$  were all

determined using *Pfu* *exo-* at the extension phase, we believe that this optimal ratio would not have been the same if primer extension was conducted with another polymerase. This might explain why 3 U of *Vent* *exo-* was necessary to obtain efficient amplification when extension was performed with *Vent* *exo-*. Further testing will be required to identify if any polymerase do remain actively present, and how they interfere with added polymerases.

If the efficiency at the amplification is influenced by the presence of active remaining DNA polymerase from the primer extension step, it does not represent a defect of LMPCR procedure in itself. This is explained by the fact that all cells analyzed by LMPCR protocols would endure the same influence in amplification efficiency, permitting intrinsic results comparison. Therefore, understanding the effect of polymerase combination is not a premise for efficiently applying the procedure but does remain an appealing challenge at the technical level.

## 2.7 Conclusion

Throughout this project, we have demonstrated that *Vent* *exo-* can be an excellent DNA polymerase for LMPCR. In all conditions assayed, *Vent* *exo-* DNA polymerase was able to achieve efficient results comparable to the ones obtained with *Pfu* *exo-*. *Vent* *exo-* was shown to be superior to *Taq* DNA polymerase in all conditions examined. Its efficiency at creating blunt-ended fragments, required for linker ligation, was demonstrated. GC-rich sequences of *P53*, *FMRI* and *C-JUN* genes were easily resolved by *Vent* *exo-*, except for the particular CGG repeated sequence of *FMRI*.

We believe that *Vent* *exo-* is an appropriate choice for LMPCR experiments due to its potential strand displacement activity in resolving particular secondary DNA structure. Furthermore, this polymerase was shown to be less affected by variation in DNA quantities compared to *Pfu* *exo-*. Finally, because comparable results can be achieved with *Pfu* *exo-* and *Vent* *exo-*, and for reason of cost efficiency, we strongly recommend the use of *Vent* *exo-* DNA polymerase when conducting LMPCR analysis.

Using the PCR programs presented therein, and the supplied commercial amplification buffers, we recommend conducting primer extension using 1.5 U of *Pfu* *exo-* followed by

amplification using 1.5 U of *Vent* exo- DNA polymerase. According to our results, we also recommend adding 1 mM of MgSO<sub>4</sub> to the amplification mixture to reach a total concentration of 3 mM (the *Vent* amplification buffer already contain 2 mM of MgSO<sub>4</sub>). If optimal results are not achieved this way, the combination *Pfu* exo- / *Pfu* exo- at the primer extension and amplification step is not likely to improved in any way the experiments. Therefore, as an alternative, we suggest using 0.75 U of *Vent* exo- at the primer extension step, followed by 2.5 U of *Herculase* at the PCR amplification. Addition of DMSO should also be considered.

## 2.8 Acknowledgement

The authors are grateful to Martin Angers for inspiring this project, Patrick Rochette for his technical assistance with the phosphorimager and Dr. Darel Hunting for carefully editing the manuscript. This study was supported by grant from the Canadian Genetic diseases Network (MRC/NSERC NCE Program). Régen Drouin is a research scholar (Senior level) of the “Fonds de la Recherche en Santé du Québec” and holds the Canada Research Chair in « Genetics, Mutagenesis and Cancer ».

## 2.9 Legend of figures

**Figure 1.** Overview of the different steps in the LMPCR protocol.

**Figure 2.** Comparison of the efficiency of *Vent* exo- and *Pfu* exo- with different amounts of DNA at the primer extension step of LMPCR. (A) Representative autoradiograms of the *p53* exon 5 sequence, and *FMR1* and *C-Jun* gene promoter sequences were analyzed using primer set 5.1 and 5.2, primer set X1 and X2 and primer set JD1 and JD2 respectively. All PCR amplifications were done using 3 U of *Taq* DAN polymerase. LMPCR was performed on increasing quantities of purified genomic DNA treated with standard Maxam-Gilbert guanine cleavage reaction to produce a global SSB frequency of 1 break per 400 bases; 0.4, 0.8, 1.6 and 2.4 µg of DNA were used (lanes 1-4, 5-8, 9-12 and 13-16, respectively). Lanes 1, 5, 9 and 13 show LMPCR protocols carried out using 1.5 U of *Pfu* exo- at the primer extension step; lanes 2-4, 6-8, 10-12 and 14-16 show LMPCR protocols performed using 0.5 (lanes 2, 6, 10 and 14), 1.0 (lanes 3, 7, 11 and 15), 1.5 U (lanes 4, 8, 12



and 16) of *Vent* *exo*- at the primer extension step. Corresponding sequence ladder of the area mapped consists of a standard Maxam-Gilbert for the G, T+C and C lanes, and modified  $K_2PdCl_4$  reaction for the A lane, with primer extension performed with 1.5 U of *Pfu* *exo*- (lanes 17-20). **(B)** Graphical representation of the relative intensity of specific units of either *Vent* *exo*- or *Pfu* *exo*- at the primer extension step as the DNA dose increases. Each point represents the intensity for one DNA polymerase at a specific unit for a selected DNA amount relatively to the summation of the intensity of this same polymerase at the same specific unit for all DNA amounts, calculated from phosphorimager quantification. *p53*, *FMRI* and *C-Jun* genes results are shown for 1.5 U of *Pfu* *exo*- (open circles), 0.5 U of *Vent* *exo*- (filled circles), 1 U of *Vent* *exo*- (filled triangles) and 1.5 U of *Vent* *exo*- (filled squares).

**Figure 3.** Comparison of PCR amplification efficiency using different amounts of *Vent* *exo*- at the PCR amplification step. **(A)** Each primer extension step was performed using 1.5 U of *Pfu* *exo*- and 1  $\mu$ g of purified genomic DNA treated with standard Maxam-Gilbert guanine cleavage reaction to produce a global SSB frequency of 1 break per 400 bases. Each of the three protocols was repeated twice. Representative autoradiogram of the *p53*, *FMRI* and *C-Jun* gene sequence was analyzed using primer set 5.1 and 5.2, primer set X1 and X2 and primer set JD1 and JD2 respectively. The amount of *Vent* *exo*- used during the amplification step varied from 0 - 20 U (lanes 1-15). **(B)** Graphical representation of the relative intensity of using various units of *Vent* *exo*- at the PCR amplification step. Each point represents the intensity of one condition relatively to the summation of the intensity of all the conditions calculated from phosphorimager quantification. *p53* (filled circles), *FMRI* (open squares) and *C-Jun* primer set results (filled triangles) are shown. Global scale representation is also shown in the upper right corner. **(C)** Graphical representation of the signal to background ratio of using various units of *Vent* *exo*- at the PCR amplification step. Each point represents the signal intensity of one condition relatively to the background intensity of this same condition calculated from phosphorimager quantification. *p53* (filled circles), *FMRI* (open squares) and *C-Jun* gene results (filled triangles) are shown.

**Figure 4.** Comparison of PCR amplification efficiency of 2.5 U of *Vent* *exo*- using different amounts of purified genomic DNA. (A) Each primer extension step was performed using 1.5 U of *Pfu* *exo*- and 0 to 5  $\mu$ g (lanes 1–15) of purified genomic DNA treated with standard Maxam–Gilbert guanine cleavage reaction to produce a global SSB frequency of 1 break per 400 bases. Each of the three protocols was repeated twice. Representative autoradiogram of the *p53*, *FMRI* and *C-Jun* gene sequence were analyzed using primer set 5.1 and 5.2, primer set X1 and X2 and primer set JD1 and JD2 respectively. (B) Graphical representation of the relative intensity of using various amounts of purified genomic DNA amplified with *Vent* *exo*-. Each point represents the intensity of one condition relatively to the summation of the intensity of all the conditions calculated from phosphorimager quantification. *p53* (filled circles), *FMRI* (open squares) and *C-Jun* gene results (filled triangles) are shown. (C) Graphical representation of the signal to background ratio of variable amounts of purified genomic DNA amplified with *Vent* *exo*-. Each point represents the signal intensity of one condition relatively to the background intensity of this same condition calculated from phosphorimager quantification. *p53* (filled circles), *FMRI* (open squares) and *C-Jun* gene results (filled triangles) are shown.

**Figure 5.** Comparison of PCR amplification efficiency of *Vent* *exo*- using different concentrations of  $\text{MgSO}_4$ . (A) For each primer extension step, 1.5 U of *Pfu* *exo*- and 1  $\mu$ g of purified genomic DNA treated with standard Maxam–Gilbert guanine cleavage reaction to produce a global SSB frequency of 1 break per 400 bases were used. The concentration of  $\text{MgSO}_4$  at the amplification step varied from 0 to 25 mM (lanes 1–20). For all of the PCR amplifications, 1.5 U of *Vent* *exo*- was used. Representative autoradiograms of the exon 5 of the *p53* gene was analyzed using primer set 5.1 and 5.2. (B) Graphical representation of the relative intensity of samples amplified with *Vent* *exo*- at different concentrations in  $\text{MgSO}_4$ . Each point represents the intensity of one condition relatively to the summation of the intensity of all conditions calculated from phosphorimager quantification. *p53* gene results (filled circles) are shown. (C) Graphical representation of the signal to background ratio of samples amplified with *Vent* *exo*- in different concentrations of  $\text{MgSO}_4$ . Each point represents the signal intensity of one condition

relatively to the background intensity of this same condition calculated from phosphorimager quantification. *p53* gene results (filled circles) are shown.

**Figure 6.** Comparison of the effects of various numbers of cycles during PCR amplification. For each primer extension step, 1.5 U of *Pfu* *exo*- and 1 µg of purified genomic DNA treated with standard Maxam–Gilbert guanine cleavage reaction to produce a global SSB frequency of 1 break per 400 bases were used. Representative autoradiogram of the exon 5 of the *p53* gene was analyzed using primer set 5.1 and 5.2. The PCR amplification was carried out using 1.5 U of *Vent* *exo*- (lanes 1-6), 3.5 U of *Pfu* *exo*- (lanes 7-12) or 3 U of *Taq* (lanes 13-18). The number of cycles of the PCR amplification varied from 12 to 37 cycles (lanes 1-6, 7-12 and 13-18 respectively).

**Figure 7.** Comparison of the efficiency of various polymerases in resolving extremely GC-rich DNA in LMPCR. (A) LMPCR of purified genomic DNA with a high global SSB frequency (1 break per 400 bp) was performed on the CGG triplet repeat of the *FMR1* gene. The sequence shown was analyzed using primer set U1 and U2 selected from the *FMR1* gene. Primer extensions were performed using 1 µg of purified genomic DNA from a normal male who has a *FMR1* gene with a haplotype of 33 CGG triplet repeats (~100% GC-rich) was treated with standard Maxam–Gilbert guanine or cytosine cleavage reaction (G and C respectively). Primer extensions were carried out using 1.5 U of *Pfu* *exo*- (lanes 1-8) or 0.75 U of *Vent* *exo*- (lanes 9-16). Subsequent PCR amplifications were carried out using 3.5 U of *Pfu* *exo*- (lanes 1-2, 9-10), 1.5 U of *Vent* *exo*- (lanes 3-4, 11-12), 3 U of *Taq* (lanes 5-6, 13-14) or 2.5 U of *Herculase* (lanes 7-8, 15-16). (B) LMPCR was performed as described in “A”, but in the presence of 7-deaza-dGTP in the primer extension and PCR amplification buffers. (C) LMPCR was performed as described in “A”, but in the presence of 7-deaza-dGTP in the primer extension and PCR amplification buffers, with the addition of 10% dimethylsulfoxide (DMSO) at the PCR amplification step. (D) LMPCR was performed as described in “A”, but in the presence 10% DMSO at the PCR amplification step.

**Figure 8.** Comparison of the efficiency of *Pfu* *exo*-, *Vent* *exo*- and *Vent* at the primer extension step combined to *Pfu* *exo*-, *Vent* *exo*-, *Vent* and *Taq* at the PCR amplification step. Each primer extension step was performed using 1 µg of purified genomic DNA

treated with standard Maxam–Gilbert guanine cleavage reaction to produce a global SSB frequency of 1 break per 400 bases. Representative autoradiogram of the exon 5 of the *p53* gene was analyzed using primer set 5.1 and 5.2. Primer extensions were carried out using 1.5 U of *Pfu* *exo-* (lanes 1-4 and 13), 0.75 U of *Vent* *exo-* (lanes 5-8 and 14) or 0.75 U of *Vent* (lanes 9-12 and 15). Subsequent PCR amplifications were carried out using 3.5 U of *Pfu* *exo-* (lanes 1, 5, 9), 1.5 U of *Vent* *exo-* (lanes 2, 6, 10), 1.5 U of *Vent* (lanes 3, 7, 11), 3 U of *Taq* (lanes 4, 8, 12) or 3 U of *Vent* *exo-* (lanes 13-15).

## 2.10 Figures 1 à 8 et tables 1 de l'article

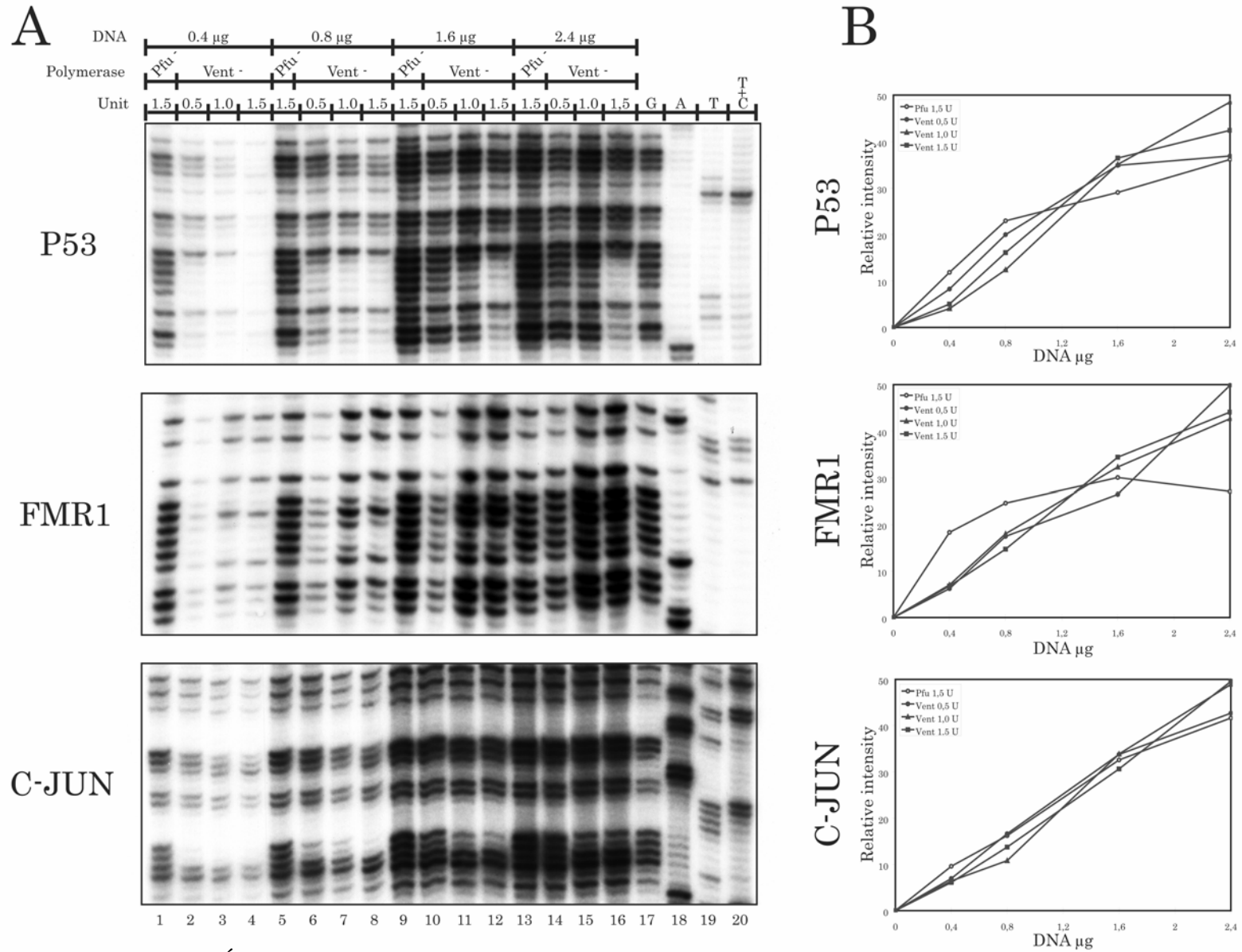
Gene	primer	Strand	Sequence (5'-3')	T <sub>m</sub> (°C) <sup>A</sup>
<i>p53</i>	5.1	upper	CACTTGTGCCCTGACTTTCAAC	54
	5.2		TGCCCTGACTTTCAACTCTGTCTCC	60
<i>C-JUN</i>	JD1	lower	CCGCGCACCTCCACTC	53
	JD2		ACCTCCACTCCCGCCTCGCTGC	67
<i>FMRI</i>	X1	upper	AAGTGAAACCGAAACGGAGC	57.7
	X2		GAGCTGAGCGCCTGACTGAGGCCG	70.5
	U1	upper	AAGTACCTTGTAGAAAGCGC	55.7
	U2		CACTTCCACCACCAGCTCCTCCAT	66.9

A. T<sub>m</sub> determined by the Gene Jockey software program

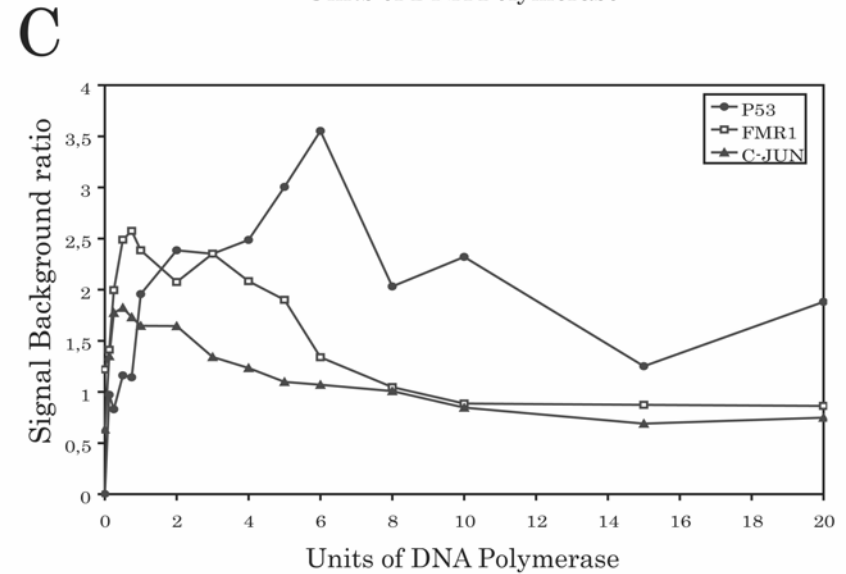
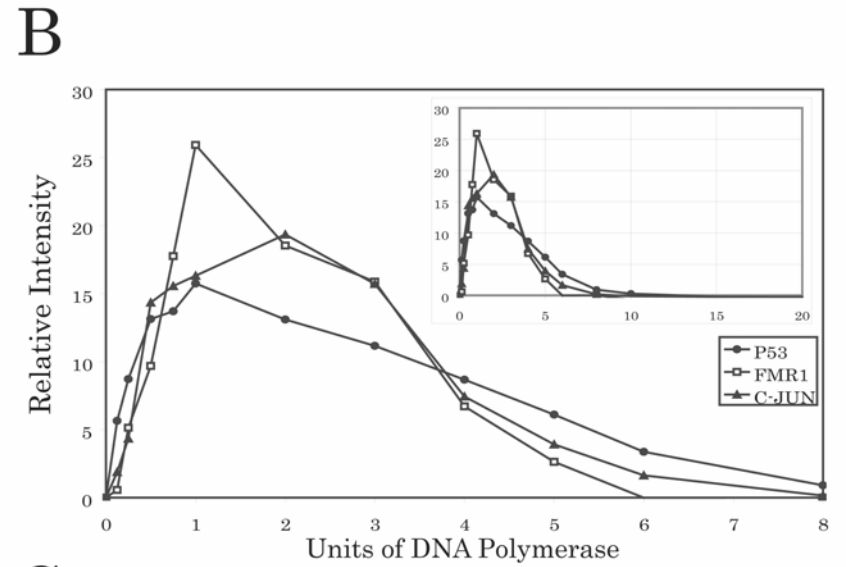
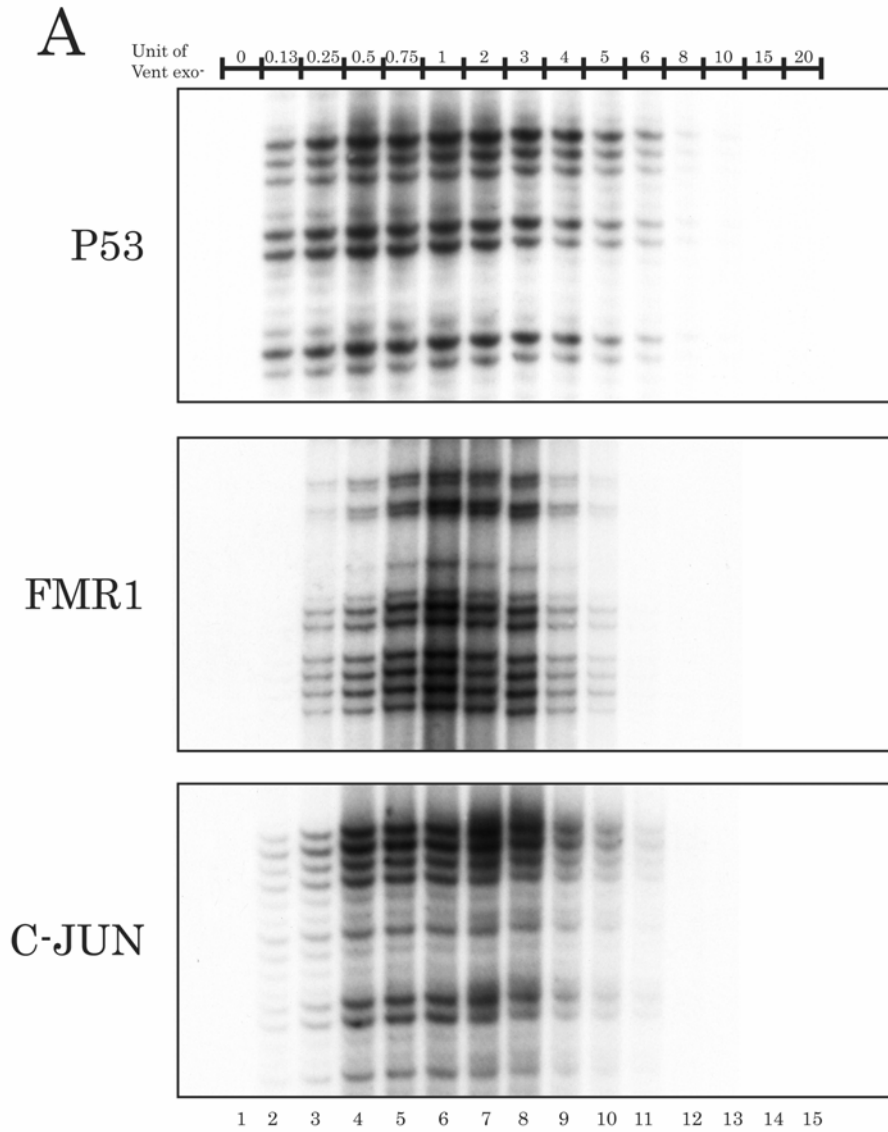
**Table 1** : Oligonucléotides sélectionnés pour l'étude

**Figure 1** : Illustration des différentes étapes de la technique LMPCR (voir Figure 2 de l'introduction, page 4)



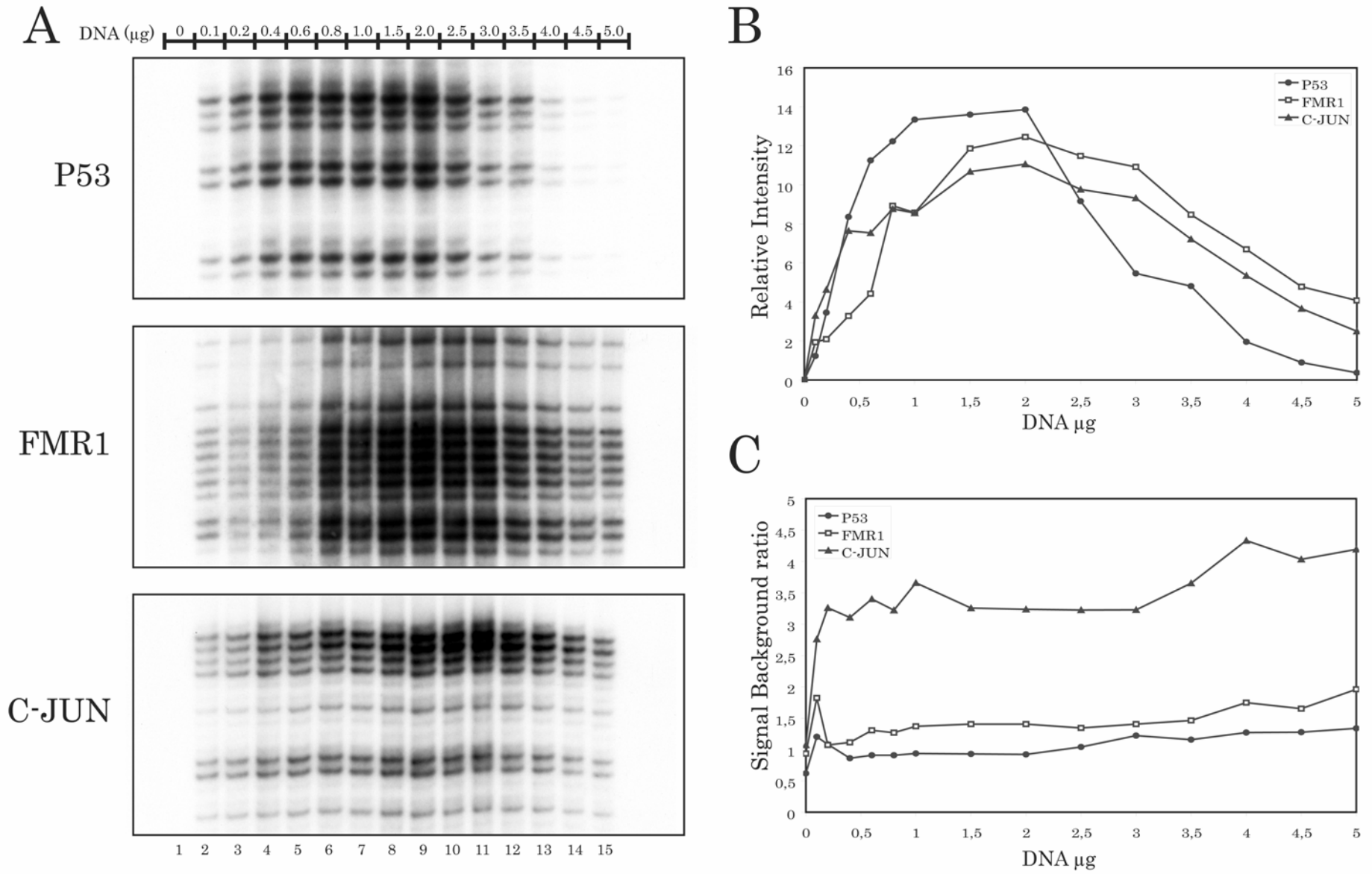


**Figure 2 :** Étude de dosage de *Vent* exo- en fonction de la concentration en ADN à l'extension d'amorces

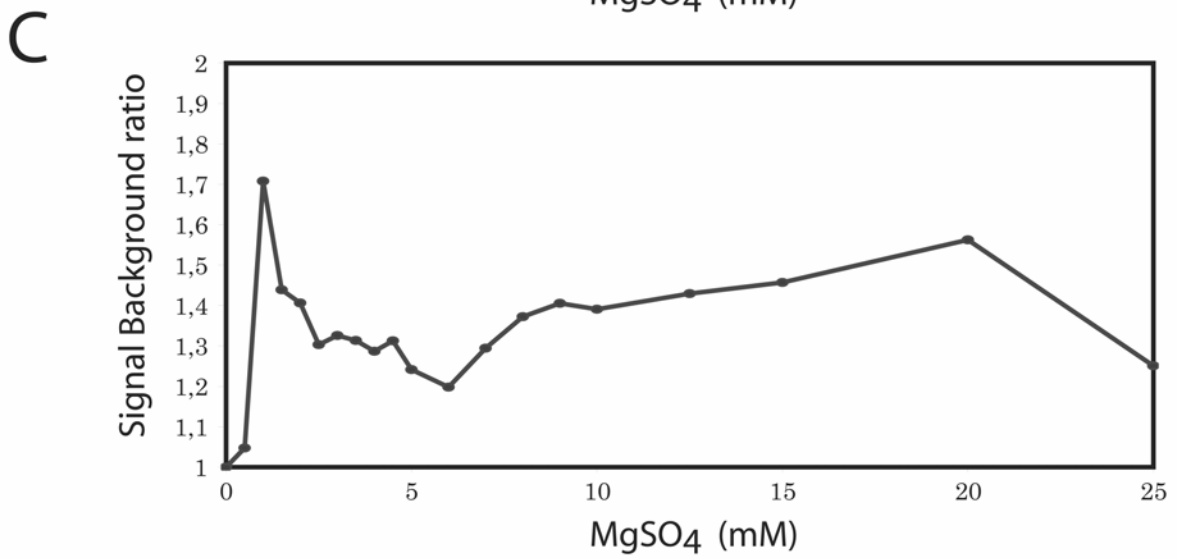
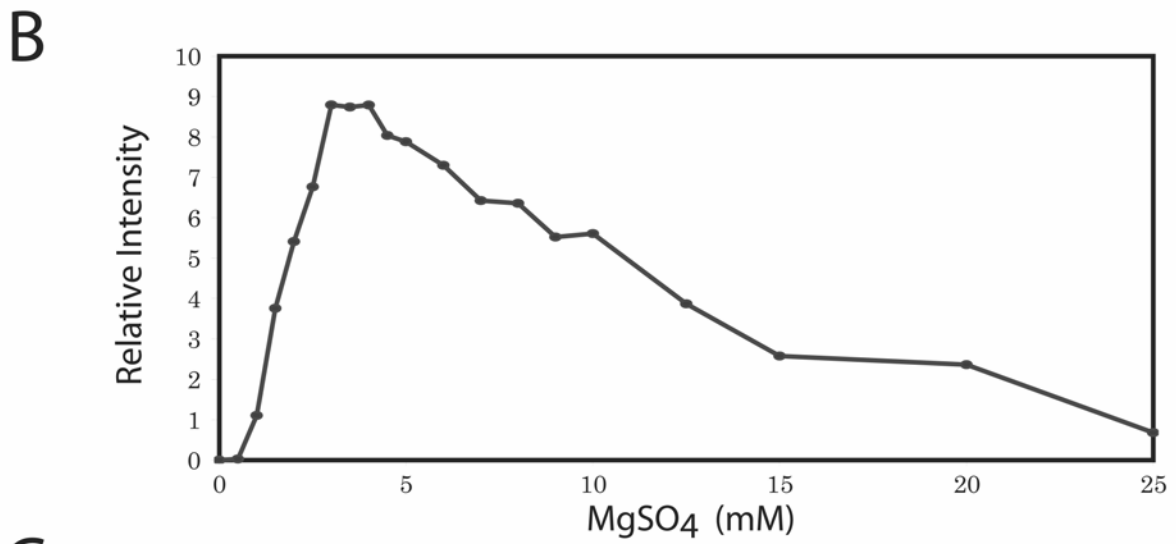
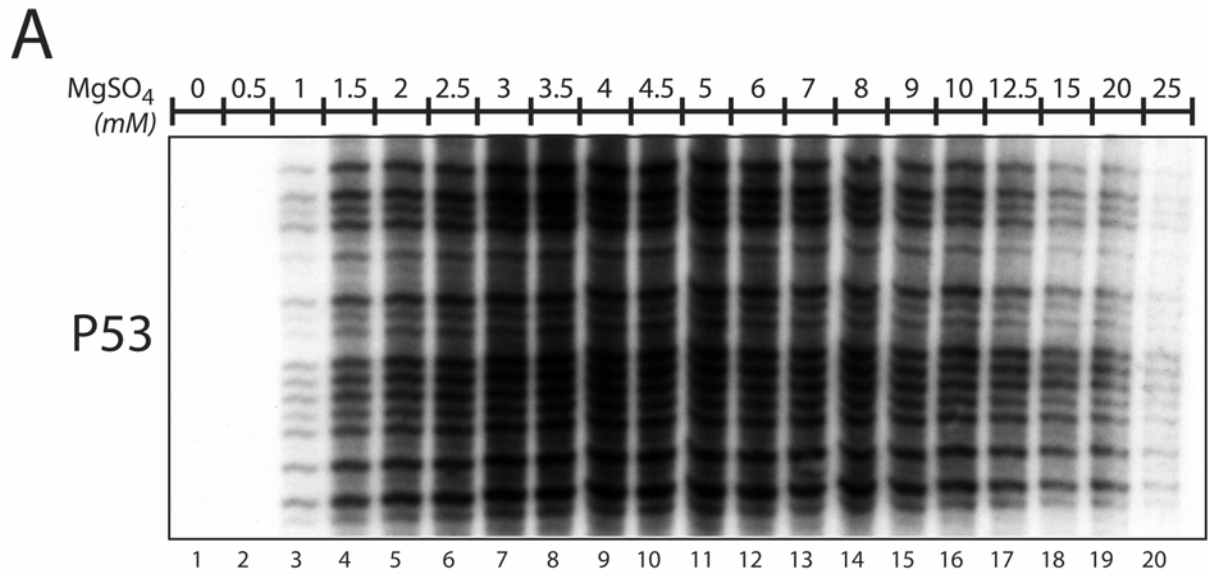


**Figure 3** : Étude de dosage de *Vent exo-* à l'amplification

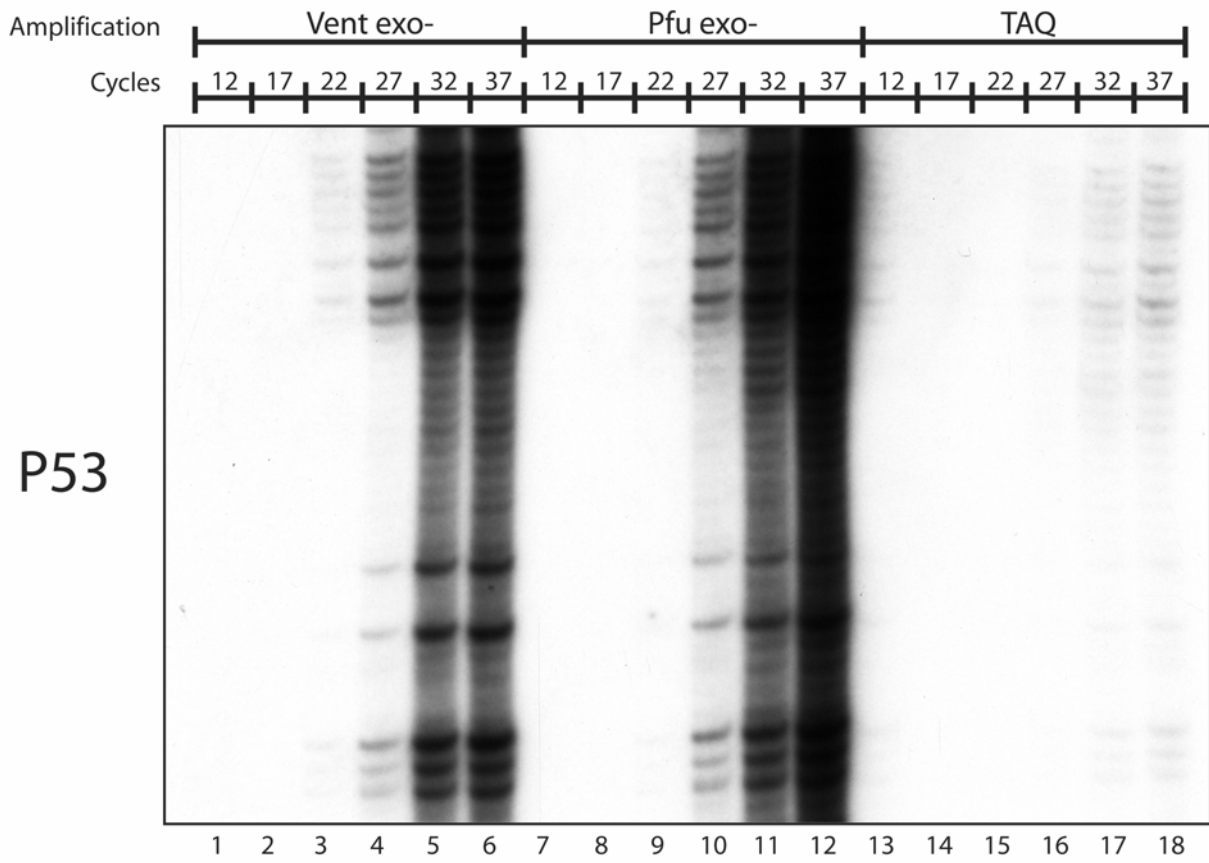




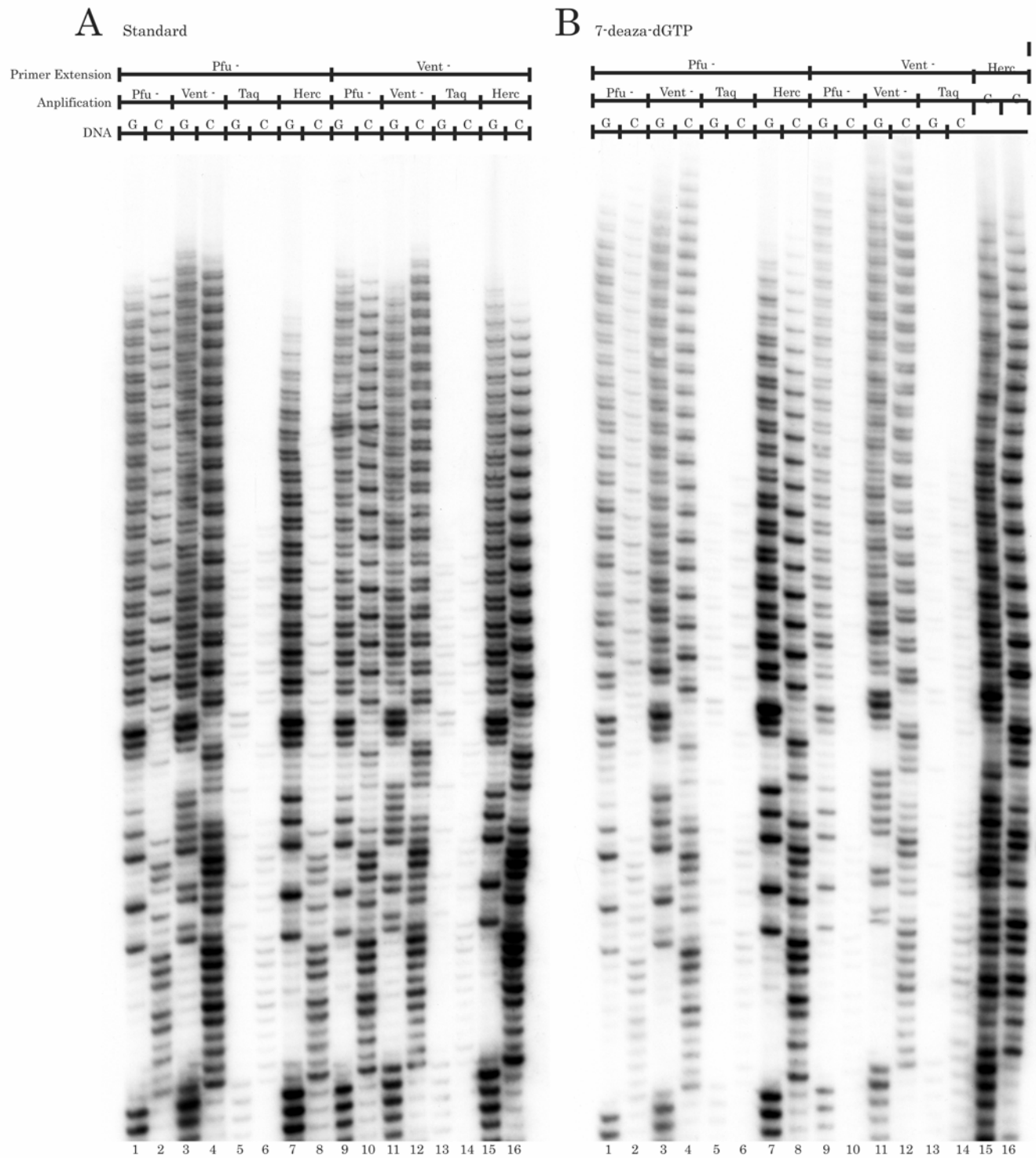
**Figure 4 :** Étude de concentration en ADN à l'amplification



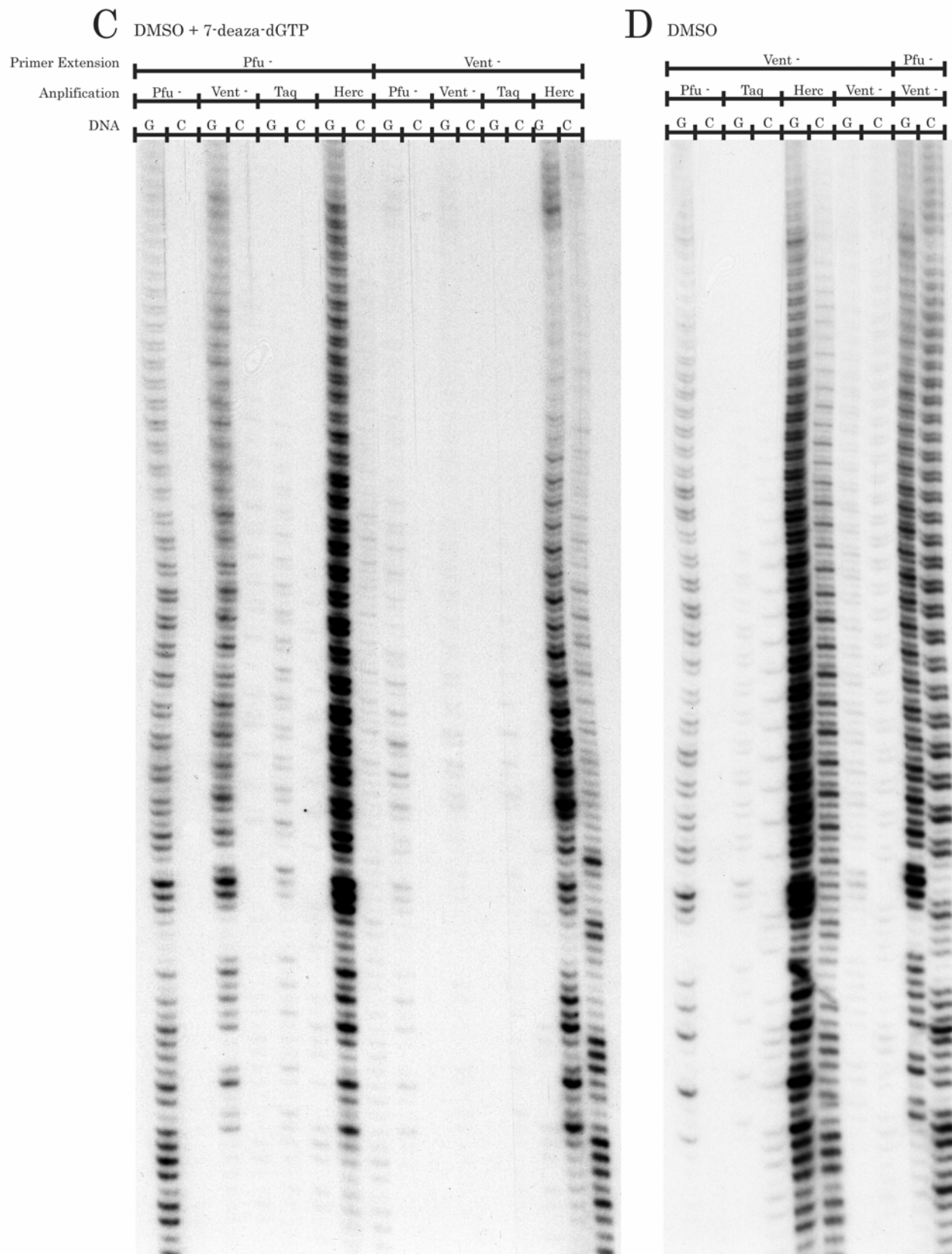
**Figure 5** : Étude de concentration en MgSO<sub>4</sub>



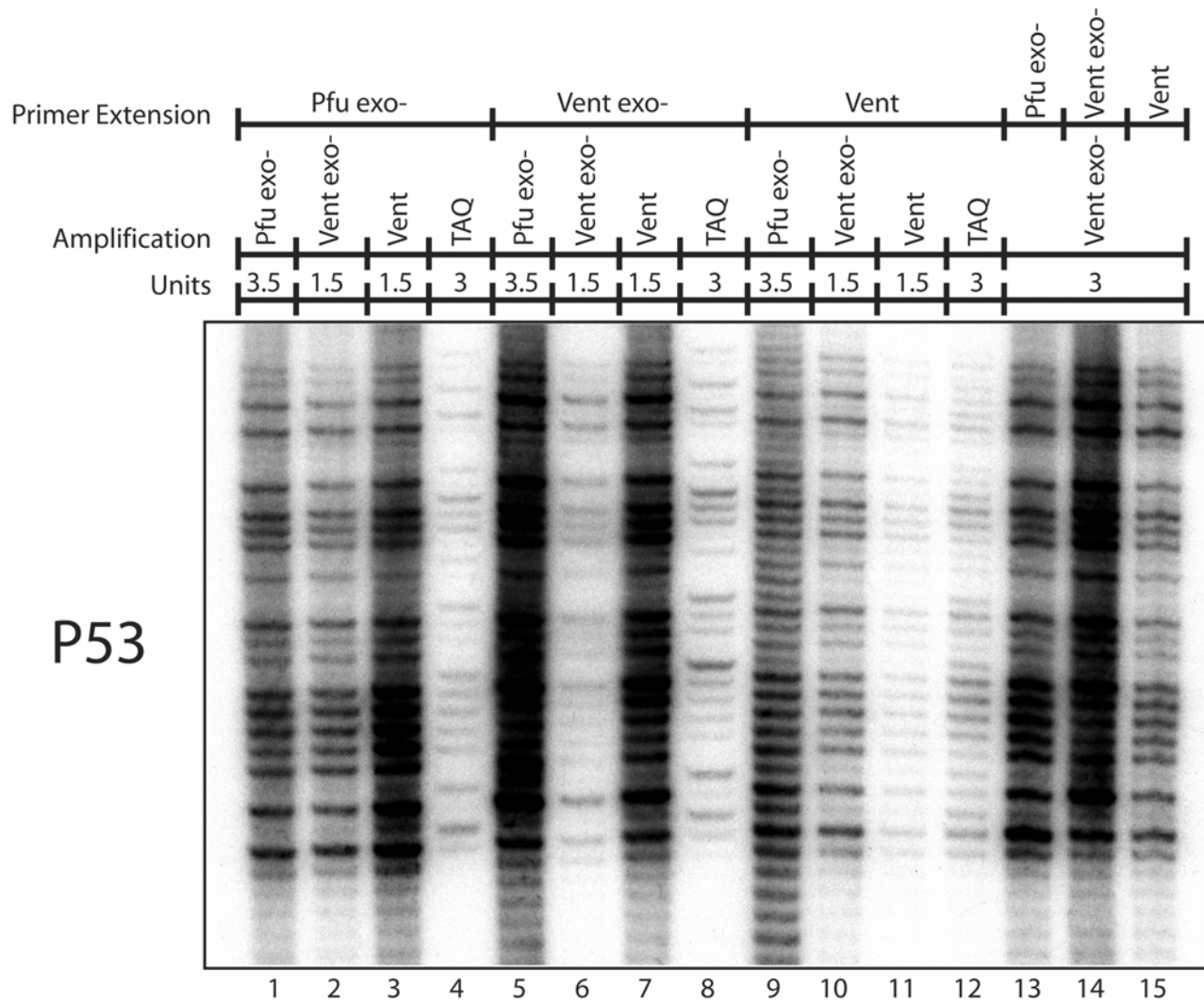
**Figure 6** : Effet du nombre de cycles sur la thermostabilité des polymérase



**Figure 7 A et B** : Efficacité de *Vent* exo- à résoudre la répétition CGG du gène



**Figure 7 C et D :** Efficacité de *Vent* exo- à résoudre la répétition CGG du gène



**Figure 8** : Combinaison des polymérases au cours du processus LMPCR

## Conclusion

Au cours de cette étude, nous avons tenté d'établir les conditions optimales d'utilisation de l'ADN polymérase *Vent* *exo*- dans la technologie LMPCR. Notre premier objectif consistait à déterminer la quantité optimale de *Vent* *exo*- à utiliser au cours de l'étape d'extension d'amorces, ainsi que la relation entre la quantité initiale d'ADN et la quantité de polymérase. Compte tenu de la variation d'efficacité obtenue dans l'étude des gènes *p53*, *FMRI* et *C-Jun*, nous avons évalué que 0,75 U de *Vent* *exo*- devrait être utilisé à l'étape d'extension d'amorces. Avec cette quantité, on s'assure d'obtenir des résultats adéquats et représentatifs de tous les fragments d'ADN présents dans l'échantillon de départ. L'absence de certaines bandes lors de l'utilisation de faibles quantités d'ADN, montre que *Vent* *exo*- se caractérise par un rapport molécules d'ADN / molécules de polymérase élevé, comparé au rapport de *Pfu* *exo*-. De ce fait, nos résultats indiquent que pour des quantités de l'ordre de 1 µg d'ADN et moins, l'utilisation de *Pfu* *exo*- serait préconisée à celle de *Vent* *exo*- pour l'étape d'extension d'amorces. Malgré la faible variation observée entre les deux polymérases, l'utilisation de *Vent* *exo*-, avec exposition adéquate de l'autoradiogramme, nous donnera des intensités satisfaisantes pour effectuer l'évaluation de résultats dans les protocoles LMPCR.

Kong et Coll. ont rapporté que 30 % des fragments polymérisés à l'aide de *Vent* *exo*- montraient la présence d'un nucléotide surnuméraire ajouté en 3', définissant ainsi la présence d'une activité Tdt pour cette polymérase [22]. Si l'on suppose que l'activité Tdt de *Vent* *exo*- affecte de façon uniforme et aléatoire tous les fragments présents, on s'attendra à observer une diminution globale de l'intensité des bandes de tous les fragments correspondants. Par contre, si l'on suppose que cette activité dépende de la séquence, on s'attendra alors à la disparition de certaines bandes correspondants à certains fragments spécifiques. Cette deuxième hypothèse pourrait, entre autres, expliquer la disparition de certaines bandes lorsque l'enzyme est utilisé avec de faibles quantités d'ADN. On notera qu'en conditions optimales d'utilisation, ce phénomène n'a pas été observé, et qu'étonnamment *Vent* *exo*- peut être utilisée efficacement pour la production d'extrémités franches, nécessaires à la liaison de l'adaptateur.

Notre second objectif portait principalement sur l'efficacité de *Vent* *exo*- lors de l'étape d'amplification PCR. Nos études de quantités en *Vent* *exo*- effectuées sur les gènes *p53*, *FMRI* et *C-Jun*, ont montré que 1,5 U était la quantité optimale de *Vent* *exo*- à utiliser pour l'étape d'amplification PCR, lorsque *Pfu* *exo*- était utilisée préalablement pour l'extension d'amorces. La perte d'efficacité qui survient à une quantité supérieure à 6 U de *Vent* *exo*-, dénote une fois de plus, le haut rapport molécules d'ADN / molécules de polymérase caractérisant *Vent* *exo*-, comparativement à *Pfu* *exo*- qui a été rapporté comme étant fonctionnel jusqu'à 25 U [18]. De plus, *Vent* *exo*- démontre une plus grande sensibilité que *Pfu* *exo*- en raison de la présence de bandes à aussi peu que 0.13 U comparativement à 2 U pour *Pfu* *exo*-. Les limites étroites de la quantité efficace d'utilisation de *Vent* *exo*- montrent l'importance de bien identifier la quantité optimale de polymérase à utiliser. Par contre, même si le haut rapport molécules d'ADN / molécules de polymérase de *Vent* *exo*- implique des limites étroites dans la quantité de polymérase à utiliser, il n'en va pas de même pour la quantité en ADN. Nos résultats montrent bien ce fait, où l'on observe que *Vent* *exo*- est très sensible à de faibles quantités d'ADN (0,1 µg), mais reste tout aussi efficace à de plus hautes quantités (5,0 µg).

Nous avons mentionné plus tôt que *Vent* *exo*- possédait une activité de déplacement de brin qui varie en fonction de la concentration de MgSO<sub>4</sub>. Plus précisément, il a été suggéré que les ions magnésium stabilisaient les structures secondaires particulières de l'ADN et qu'à des concentrations supérieures à 4 mM de MgSO<sub>4</sub>, *Vent* *exo*- ne pouvait plus pénétrer à l'intérieur des épingles causant ainsi le glissement de réplication provoquant la délétion d'une partie du fragment d'ADN à amplifier [27, 28]. Par contre, les ions magnésium étant essentiels au processus de polymérisation, on ne peut envisager d'éliminer leur présence. De ce fait, nous avons déterminé la concentration optimale en MgSO<sub>4</sub> à utiliser lors de l'étape d'amplification PCR avec *Vent* *exo*-. Nos résultats montrent qu'entre 1,5 et 10,0 mM de MgSO<sub>4</sub>, une amplification efficace des fragments d'ADN a été obtenue. Nous avons donc déterminé que 3,0 mM était la concentration optimale de MgSO<sub>4</sub> à utiliser, ce qui favorise en même temps l'activité de déplacement de brin de l'ADN polymérase *Vent* *exo*-.



Finalement, nous avons étudié l'effet du nombre de cycles à l'amplification PCR sur la thermostabilité des ADN polymérase *Vent* *exo-*, *Pfu* *exo-* et *Taq*. Nos résultats ont clairement montré que *Pfu* *exo-* et *Vent* *exo-* étaient plus thermostables que l'ADN polymérase *Taq*, favorisant ainsi en une plus grande efficacité d'amplification. De plus, *Pfu* *exo-* s'est avéré être un peu plus résistante à la chaleur que *Vent* *exo-*, d'où la saturation du signal après 37 cycles PCR pour *Vent* *exo-*.

Après avoir déterminé les conditions optimales d'utilisation de *Vent* *exo-*, nous avons pu comparer l'efficacité de cette dernière avec différentes polymérase, dans différentes conditions expérimentales. Notre troisième objectif était donc de vérifier, au niveau des étapes d'extension d'amorces et d'amplification, l'effet des cations, plus précisément le sodium et le potassium, présents dans les tampons de réaction sur l'efficacité de polymérisation des ADN polymérase *Pfu* *exo-*, *Vent* *exo-* et *Taq*. L'objectif de cette expérience reposait principalement sur l'hypothèse que le KCl stabilise les structures particulières de l'ADN plus efficacement que le NaCl [37]. Étant donné que la plupart des tampons réactionnels commerciaux sont composés de KCl, nous voulions aussi examiner si la substitution par du NaCl affecterait l'efficacité de la polymérase *Vent* *exo-* au cours des diverses étapes où elle est requise. Nos résultats n'ont montré aucune différence notable quant à la substitution de KCl par du NaCl ni à l'extension d'amorces, pas plus qu'à l'étape d'amplification. De plus, le retard de migration des échantillons amplifiés à l'aide de la *Taq*, dénotant l'activité Tdt de cette polymérase, nous permet de conclure que *Pfu* *exo-* et *Vent* *exo-* ne possèdent pas d'activité Tdt. En effet, si 30 % des fragments amplifiés par *Vent* *exo-* était soumis à l'activité Tdt de l'enzyme, on s'attendrait à voir cette même proportion de fragments subir un décalage total, comme dans le cas de *Taq* ou partiel, par la présence d'artéfacts de bande. Ces artéfacts de bandes, communément appelés « doubles bandes », reflètent la présence simultanée sur le gel d'un même fragments à deux positions différentes, principalement causé par une différence dans la longueur de ces deux fragments, influençant ainsi leur migration. C'est ce qui se produit lors de l'ajout d'un nucléotide par une polymérase sur une proportion d'un même fragment.

Dans un même ordre d'idée, nous avons évalué l'efficacité de différentes polymérase combinées lors de l'extension et lors de l'amplification, dans la résolution de la structure

particulière du gène *FMRI*. L'exon 1 de ce gène est caractérisé par la présence d'une succession de répétitions de triplets CGG dont le nombre varie entre 6 et 250 triplets [19]. Ce type de structure est composé d'une séquence 100 % riche en GC et s'avère être limitative pour l'activité de certaines polymérases. Nous voulions donc examiner si l'ADN polymérase *Vent* *exo*- était capable de résoudre ce genre de structure comparativement à ce qui avait déjà été obtenu avec *Pfu* *exo*- [18]. Nos résultats nous ont indiqué que *Vent* *exo*- montrait une efficacité similaire à celle de *Pfu* *exo*-. Par contre, beaucoup d'artéfacts de bande ont été observés lors de l'utilisation de *Vent* *exo*-. Comme mentionné plus tôt, nous suggérons qu'il s'agisse de l'addition d'un nucléotide en 3' des fragments amplifiés, causée par l'activité Tdt de la polymérase. En tenant compte des résultats obtenus plus tôt, il s'agit en fait du seul cas où ce genre de bandes a été observé avec *Vent* *exo*- en conditions optimales. De ce fait nous suggérons que l'activité Tdt de *Vent* *exo*-, rapportée par Kong et coll., est dépendante de certaines conditions particulières d'utilisation ou dépendante de la séquence à polymériser. Il se peut en effet, que la structure même du gène conduise à l'activation de l'activité Tdt de *Vent* *exo*-, ou bien que la proportion élevée en GC favorise l'accumulation de dUTP. Comme mentionné dans l'introduction, les hautes températures de PCR favorisent la désamination des dCTP en dUTP ce qui inhibe l'activité de polymérisation de *Vent* *exo*- [29, 30]. L'efficacité des enzymes testées et les résultats obtenues à l'aide de l'*Herculase* peuvent d'ailleurs être expliqués par la présence de dUTPase dans le mélange réactionnel, sous forme du facteur ArchaeMaxx [49, 50]. Pour ce qui est de l'ajout de 7-deaza-dGTP ou de DMSO dans les mélanges réactionnels, on constate que seul le DMSO semble améliorer l'efficacité des polymérases testées. Par contre, des expériences de concentration devront être effectuées afin d'approfondir le sujet davantage.

Finalement, notre dernier objectif consistait à évaluer quelle combinaison de polymérases était le plus efficace lors de l'extension d'amorces et lors de l'amplification. Les résultats obtenus montrent une fois de plus l'activité Tdt de *Taq* causant le retard de migration des fragments d'ADN amplifié à l'aide de cette polymérase. L'utilisation de *Vent* *exo*- à l'extension d'amorces semble être plus efficace que *Pfu* *exo*-, de part l'intensité des amplifications conduites à l'aide de *Pfu* *exo*-, *Vent* et *Taq*. Par contre, la chute d'intensité des fragments polymérisés par *Vent* *exo*- montre qu'il existe un lien direct entre la

polymérase utilisée à l'extension d'amorces et à l'amplification PCR. La présence d'artéfacts de bandes dans l'échantillon combinant *Vent* et *Pfu* *exo-* vient aussi supporter cette hypothèse.

Sachant que les conditions d'utilisation de *Vent* *exo-* à l'amplification ont été déterminées à partir de fragments polymérisés à l'aide de *Pfu* *exo-*, et sachant que l'efficacité d'amplification de *Vent* *exo-* est influencée par la polymérase utilisée à l'extension d'amorces, nous aurons comme perspectives futures d'identifier la cause de cette variation. D'ici là, nous recommandons l'utilisation de *Pfu* *exo-* à l'extension d'amorces suivit de l'amplification à l'aide de *Vent* *exo-*. Ceci s'avérera une décision judicieuse, particulièrement en raison des coûts associés à l'ADN polymérase *Vent* *exo-* qui ne représentent qu'une fraction de ceux de la *Pfu* *exo-*. Finalement, nos résultats laissent entrevoir que *Vent* *exo-* ne possède pas d'activité Tdt constitutive, mais nous permet d'envisager la possibilité d'activation de cette activité dans certaines conditions particulières. D'autres expériences seront requises pour mieux approfondir ce sujet.

## Références

1. Mueller, P.R. and B. Wold, *In vivo footprinting of a muscle specific enhancer by ligation mediated PCR*. Science, 1989. **246**(4931): p. 780-6.
2. Pfeifer, G.P., et al., *Genomic sequencing and methylation analysis by ligation mediated PCR*. Science, 1989. **246**(4931): p. 810-3.
3. Pfeifer, G.P. and A.D. Riggs, *Chromatin differences between active and inactive X chromosomes revealed by genomic footprinting of permeabilized cells using DNase I and ligation-mediated PCR*. Genes Dev, 1991. **5**(6): p. 1102-13.
4. Litt, M.D., I.K. Hornstra, and T.P. Yang, *In vivo footprinting and high-resolution methylation analysis of the mouse hypoxanthine phosphoribosyltransferase gene 5' region on the active and inactive X chromosomes*. Mol Cell Biol, 1996. **16**(11): p. 6190-9.
5. Drouin, R., et al., *Structural and functional characterization of the human FMR1 promoter reveals similarities with the hnRNP-A2 promoter region*. Hum Mol Genet, 1997. **6**(12): p. 2051-60.
6. Chin, P.L., J. Momand, and G.P. Pfeifer, *In vivo evidence for binding of p53 to consensus binding sites in the p21 and GADD45 genes in response to ionizing radiation*. Oncogene, 1997. **15**(1): p. 87-99.
7. Litt, M.D., et al., *5-Azadeoxycytidine-induced chromatin remodeling of the inactive X-linked HPRT gene promoter occurs prior to transcription factor binding and gene reactivation*. J Biol Chem, 1997. **272**(23): p. 14921-6.
8. Angers, M., et al., *In vivo protein-DNA interactions at the kinin B(1) receptor gene promoter: no modification on interleukin-1 beta or lipopolysaccharide induction*. J Cell Biochem, 2000. **78**(2): p. 278-96.
9. Spugnardi, M., et al., *Epigenetic inactivation of RAS association domain family protein 1 (RASSF1A) in malignant cutaneous melanoma*. Cancer Res, 2003. **63**(7): p. 1639-43.
10. Szabo, P.E., J.R. Mann, and G.P. Pfeifer, *Methylation analysis by chemical DNA sequencing*. Methods Mol Biol, 2002. **200**: p. 29-41.
11. Xu, J., J. Wu, and P.C. Dedon, *DNA damage produced by enediynes in the human phosphoglycerate kinase gene in vivo: esperamicin A1 as a nucleosome footprinting agent*. Biochemistry, 1998. **37**(7): p. 1890-7.
12. Chen, C. and T.P. Yang, *Nucleosomes are translationally positioned on the active allele and rotationally positioned on the inactive allele of the HPRT promoter*. Mol Cell Biol, 2001. **21**(22): p. 7682-95.
13. Gao, S., R. Drouin, and G.P. Holmquist, *DNA repair rates mapped along the human PGK1 gene at nucleotide resolution*. Science, 1994. **263**(5152): p. 1438-40.
14. Tornaletti, S. and G.P. Pfeifer, *Slow repair of pyrimidine dimers at p53 mutation hotspots in skin cancer*. Science, 1994. **263**(5152): p. 1436-8.
15. Rochette, P.J., et al., *Human cells bearing homozygous mutations in the DNA mismatch repair genes hMLH1 or hMSH2 are fully proficient in transcription-coupled nucleotide excision repair*. Oncogene, 2002. **21**(37): p. 5743-52.

16. Garrity, P.A. and B.J. Wold, *Effects of different DNA polymerases in ligation-mediated PCR: enhanced genomic sequencing and in vivo footprinting*. Proc Natl Acad Sci U S A, 1992. **89**(3): p. 1021-5.
17. Drouin, R., et al., *In vivo DNA analysis*. Methods Mol Biol, 2001. **148**: p. 175-219.
18. Angers, M., et al., *Optimal conditions to use Pfu exo(-) DNA polymerase for highly efficient ligation-mediated polymerase chain reaction protocols*. Nucleic Acids Res, 2001. **29**(16): p. E83.
19. Fu, Y.H., et al., *Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox*. Cell, 1991. **67**(6): p. 1047-58.
20. Perler, F.B., et al., *Intervening sequences in an Archaea DNA polymerase gene*. Proc Natl Acad Sci U S A, 1992. **89**(12): p. 5577-81.
21. Belkins, S. and H. Jannasch, *A new extremely thermophilic, sulfur-reducing heterotrophic marine bacterium*. Arch. Microbiol., 1985. **141**: p. 181-186.
22. Kong, H., R.B. Kucera, and W.E. Jack, *Characterization of a DNA polymerase from the hyperthermophile archaea Thermococcus litoralis. Vent DNA polymerase, steady state kinetics, thermal stability, processivity, strand displacement, and exonuclease activities*. J Biol Chem, 1993. **268**(3): p. 1965-75.
23. Saiki, R.K., et al., *Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase*. Science, 1988. **239**(4839): p. 487-91.
24. Mattila, P., et al., *Fidelity of DNA synthesis by the Thermococcus litoralis DNA polymerase--an extremely heat stable enzyme with proofreading activity*. Nucleic Acids Res, 1991. **19**(18): p. 4967-73.
25. Clark, J.M., *Novel non-templated nucleotide addition reactions catalyzed by procaryotic and eucaryotic DNA polymerases*. Nucleic Acids Res, 1988. **16**(20): p. 9677-86.
26. Hu, G., *DNA polymerase-catalyzed addition of nontemplated extra nucleotides to the 3' end of a DNA fragment*. DNA Cell Biol, 1993. **12**(8): p. 763-70.
27. Viguera, E., D. Canceill, and S.D. Ehrlich, *Replication slippage involves DNA polymerase pausing and dissociation*. Embo J, 2001. **20**(10): p. 2587-95.
28. Viguera, E., D. Canceill, and S.D. Ehrlich, *In vitro replication slippage by DNA polymerases from thermophilic organisms*. J Mol Biol, 2001. **312**(2): p. 323-33.
29. Hogrefe, H.H., et al., *Archaeal dUTPase enhances PCR amplifications with archaeal DNA polymerases by preventing dUTP incorporation*. Proc Natl Acad Sci U S A, 2002. **99**(2): p. 596-601.
30. Lasken, R.S., D.M. Schuster, and A. Rashtchian, *Archaeobacterial DNA polymerases tightly bind uracil-containing DNA*. J Biol Chem, 1996. **271**(30): p. 17692-6.
31. Greagg, M.A., et al., *A read-ahead function in archaeal DNA polymerases detects promutagenic template-strand uracil*. Proc Natl Acad Sci U S A, 1999. **96**(16): p. 9045-50.
32. Chen, H.H., et al., *Terminal transferase-dependent PCR (TDPCR) for in vivo UV photofootprinting of vertebrate cells*. Sci STKE, 2001. **2001**(77): p. PL1.
33. Komura, J. and A.D. Riggs, *Terminal transferase-dependent PCR: a versatile and sensitive method for in vivo footprinting and detection of DNA adducts*. Nucleic Acids Res, 1998. **26**(7): p. 1807-11.
34. Pfeifer, G.P., J. Singer-Sam, and A.D. Riggs, *Analysis of methylation and chromatin structure*. Methods Enzymol, 1993. **225**: p. 567-83.

35. Cartwright, I.L. and S.E. Kelly, *Probing the nature of chromosomal DNA-protein contacts by in vivo footprinting*. Biotechniques, 1991. **11**(2): p. 188-90, 192-4, 196 passim.
36. Hornstra, I.K. and T.P. Yang, *In vivo footprinting and genomic sequencing by ligation-mediated PCR*. Anal Biochem, 1993. **213**(2): p. 179-93.
37. Fry, M. and L.A. Loeb, *The fragile X syndrome d(CGG)<sub>n</sub> nucleotide repeats form a stable tetrahelical structure*. Proc Natl Acad Sci U S A, 1994. **91**(11): p. 4950-4.
38. Maxam, A.M. and W. Gilbert, *Sequencing end-labeled DNA with base-specific chemical cleavages*. Methods Enzymol, 1980. **65**(1): p. 499-560.
39. Iverson, B.L. and P.B. Dervan, *Adenine specific DNA chemical sequencing reaction*. Nucleic Acids Res, 1987. **15**(19): p. 7823-30.
40. Drouin, R., S. Gao, and G.P. Holmquist, *Agarose gel electrophoresis for DNA damage analysis.*, in *Technologies for Detection of DNA Damage and Mutations.*, G.P. Pfeifer, Editor. 1996, Plenum Press: New York, NY. p. 37-43.
41. Rochette, P.J., et al., *UVA-induced cyclobutane pyrimidine dimers form predominantly at thymine-thymine dipyrimidines and correlate with the mutation spectrum in rodent cells*. Nucleic Acids Res, 2003. **31**(11): p. 2786-94.
42. Lundberg, K.S., et al., *High-fidelity amplification using a thermostable DNA polymerase isolated from Pyrococcus furiosus*. Gene, 1991. **108**(1): p. 1-6.
43. Motz, M., S. Paabo, and C. Kilger, *Improved cycle sequencing of GC-rich templates by a combination of nucleotide analogs*. Biotechniques, 2000. **29**(2): p. 268-70.
44. Dierick, H., et al., *Incorporation of dITP or 7-deaza dGTP during PCR improves sequencing of the product*. Nucleic Acids Res, 1993. **21**(18): p. 4427-8.
45. McConlogue, L., M.A. Brow, and M.A. Innis, *Structure-independent DNA amplification by PCR using 7-deaza-2'-deoxyguanosine*. Nucleic Acids Res, 1988. **16**(20): p. 9869.
46. Chakrabarti, R. and C.E. Schutt, *Novel sulfoxides facilitate GC-rich template amplification*. Biotechniques, 2002. **32**(4): p. 866, 868, 870-2, 874.
47. Winship, P.R., *An improved method for directly sequencing PCR amplified material using dimethyl sulphoxide*. Nucleic Acids Res, 1989. **17**(3): p. 1266.
48. Berger, S.L., *Rescue of a failed polymerase chain reaction catalyzed by vent or deep vent DNA polymerase with 10% dimethyl sulfoxide*. Anal Biochem, 1994. **222**(1): p. 290-3.
49. Borns, S. and H.H. Hogrefe, *Herculase™ enhanced DNA polymerase delivers high fidelity and great performance*. Strategies, 2000. **13**(1): p. 1-3.
50. Borns, S., B. Scott, and H.H. Hogrefe, *Novel Herculase™ Hotstart DNA polymerase results in greater specificity and sensitivity*. Strategies, 2001. **14**(2): p. 41-42.