

**UNIVERSITÉ DU QUÉBEC**

**MÉMOIRE DE RECHERCHE PRÉSENTÉ À  
L'UNIVERSITÉ DU QUÉBEC À TROIS-RIVIÈRES**

**COMME EXIGENCE PARTIELLE  
À L'OBTENTION DE LA MAÎTRISE EN BIOPHYSIQUE  
ET BIOLOGIE CELLULAIRES**

**PAR**

**VICKY BEAUDOIN, B. Sc. (Biologie médicale)**

**RECHERCHE DE NOUVELLES PHOSPHOLIPASES A<sub>2</sub> CYTOSOLIQUES  
DANS L'ÉPITHÉLIUM PIGMENTAIRE RÉTINIEN ET LA RÉTINE**

**SEPTEMBRE 2002**

2164

Université du Québec à Trois-Rivières

Service de la bibliothèque

Avertissement

L'auteur de ce mémoire ou de cette thèse a autorisé l'Université du Québec à Trois-Rivières à diffuser, à des fins non lucratives, une copie de son mémoire ou de sa thèse.

Cette diffusion n'entraîne pas une renonciation de la part de l'auteur à ses droits de propriété intellectuelle, incluant le droit d'auteur, sur ce mémoire ou cette thèse. Notamment, la reproduction ou la publication de la totalité ou d'une partie importante de ce mémoire ou de cette thèse requiert son autorisation.

## REMERCIEMENTS

Ce n'est pas facile d'être concis lorsque vient le temps des remerciements... tant de personnes finissent par contribuer d'une façon ou d'une autre à l'accomplissement d'un travail de recherche. Je m'efforcerai donc de mentionner les personnes qui ont contribué de plus près à ce travail de maîtrise, mais je n'oublie tout de même pas ceux et celles qui m'ont aidée à un moment ou à un autre lors de ces deux années de maîtrise.

Je remercie d'abord ceux et celles qui m'ont toujours soutenue moralement et qui étaient là lorsque rien ne semblait déboucher, ceux et celles qui me permettaient de décrocher... mes amis et ma famille.

Merci à ceux qui ont fait de l'ambiance du laboratoire un milieu de travail agréable, je parle ici de toute l'équipe du Dr Salesse, sans vouloir tous les nommer.

Merci aussi à ceux qui ont su m'éclairer par leurs conseils et leur expertise... Je parle ici du Dr Hamel et de ses étudiants de Montpellier.

Merci aussi au Dr Salesse qui m'a permis de travailler sur un projet de recherche très intéressant et qui m'a toujours fait confiance.

Enfin, merci au FCAR pour son soutien financier tout au long de mon travail de maîtrise.

C'est donc grâce à un peu de vous tous que ce travail fut mené à terme et que je suis présentement en train de remettre les résultats de mon travail... mon mémoire de maîtrise. Merci.

## RÉSUMÉ

Les phospholipases A<sub>2</sub> (PLA<sub>2</sub>) constituent une grande famille d'enzymes qui sont responsables de l'hydrolyse de l'acide gras situé en position sn-2 des phospholipides des membranes cellulaires. Ces enzymes pourraient être impliquées dans le recyclage des acides gras des photorécepteurs qui sont composés de phospholipides hautement insaturés. L'équipe du Dr Salesse a récemment démontré (Van Themsche *et al.*, 2001) que des phospholipases A<sub>2</sub> cytosoliques (cPLA<sub>2</sub>) de l'épithélium pigmentaire rétinien (EPR) présentaient des propriétés biochimiques différentes de toutes les phospholipases identifiées jusqu'à présent dans la littérature. Ceci suggère que ces cPLA<sub>2</sub> sont spécifiques à la rétine et, par conséquent, que leur mal fonctionnement pourrait être associé à diverses dysfonctions rétinienne.

L'objectif de ce projet de maîtrise consistait donc à isoler l'ADN complémentaire complet d'au moins une de ces cPLA<sub>2</sub> afin d'obtenir ultérieurement l'ADN du gène complet. Le présent mémoire détaille les différentes démarches entreprises afin d'atteindre cet objectif.

Ainsi, dans un premier temps, la vérification par RT-PCR de la présence des cPLA<sub>2</sub> déjà connues au niveau de l'EPR et de la rétine neurale a permis d'établir que la cPLA<sub>2</sub> gamma est présente au niveau de ces deux tissus, alors que la cPLA<sub>2</sub> alpha en est absente. Cette approche ne s'est toutefois pas avérée concluante pour la cPLA<sub>2</sub> bêta; d'autres amorces devront être essayées, puisqu'aucun fragment de son ADNc n'a pu être amplifié, que ce soit dans la rétine neurale, l'EPR ou les différents tissus testés.

De plus, lors de ce travail de maîtrise, deux nouvelles séquences présentant une identité avec les cPLA<sub>2</sub> connues ont été partiellement identifiées grâce aux banques de données de la NCBI (*National Center for Biotechnology Information*). Ces informations ont permis de cloner une partie de l'ADNc d'une de ces phospholipases par le criblage par RT-PCR de l'ADNc rétinien (rétine neurale). Par conséquent, plusieurs expériences ont ensuite été réalisées afin de cloner l'ADNc complet de cette nouvelle cPLA<sub>2</sub>, mais se

sont avérées vaines. Enfin, la présence de l'ADNc de cette nouvelle cPLA<sub>2</sub> dans plusieurs autres tissus permet d'exclure la possibilité d'un gène spécifique à la rétine.

## TABLE DES MATIÈRES

<b>REMERCIEMENTS</b>	<b>ii</b>
<b>RÉSUMÉ</b>	<b>iii</b>
<b>TABLE DES MATIÈRES</b>	<b>v</b>
<b>Liste des tableaux</b>	<b>x</b>
<b>Liste des figures</b>	<b>xi</b>
<b>Liste des symboles et abréviations</b>	<b>xiv</b>
<b>CHAPITRE I INTRODUCTION</b>	<b>I</b>
<b>1.1 Composantes de l’oeil</b>	<b>1</b>
1.1.1 La rétine	3
1.1.1.1 La rétine neurale	3
1.1.1.1.1 Les photorécepteurs	4
1.1.1.2 L’épithélium pigmentaire rétinien	9
1.1.1.2.1 Les fonctions de l'EPR	11
1.1.1.2.2 La phagocytose des photorécepteurs	12
1.1.1.2.3 Implication du récepteur mannose au niveau de la phagocytose	13
<b>1.2 Les phospholipases A<sub>2</sub></b>	<b>13</b>
1.2.1 La classification des phospholipases A <sub>2</sub>	14
1.2.2 Les phospholipases A <sub>2</sub> sécrétées	16
1.2.2.1 Inhibiteurs des sPLA <sub>2</sub>	18
1.2.3 Les phospholipases A <sub>2</sub> des groupes VI, VII, et VIII	19
1.2.3.1 Le groupe VIA	19
1.2.3.2 Le groupe VIB	20
1.2.3.3 Le groupe VIIA	21
1.2.3.4 Le groupeVIIB	21
1.2.3.5 Les groupes VIIIA et VIIIB	22



2.1.2.2.1 Criblage bioinformatique .....	47
2.1.2.2.2 Criblage par RT-PCR .....	47
2.1.2.2.3 Criblage de banques d'ADNc .....	47
2.1.2.2.4 Criblage d'une banque génomique.....	48
2.1.2.2.5 Criblage à l'aide d'EST marqués .....	48
2.1.2.2.6 Criblage par RT-PCR de différents tissus.....	48
<b>CHAPITRE 3 MATÉRIEL ET MÉTHODES .....</b>	<b>50</b>
<b>3.1 Préparation de l'ADNc de l'EPR et de la rétine neurale.....</b>	<b>50</b>
3.1.1 Dissection des globes oculaires humains .....	50
3.1.2 Extraction de l'ARN rétinien .....	51
3.1.3 Extraction de l'ARN de l'EPR.....	54
3.1.4 Transcription de l'ARN en ADNc .....	54
<b>3.2 Réaction de polymérase en chaîne (PCR) .....</b>	<b>55</b>
<b>3.3 Clonage des bandes PCR.....</b>	<b>60</b>
3.3.1 Purification des bandes .....	60
3.3.2 Clonage des bandes.....	60
3.3.2.1 Préparation des bactéries compétentes.....	62
<b>3.4 Purification de plasmides .....</b>	<b>62</b>
<b>3.5 Digestion des plasmides .....</b>	<b>64</b>
<b>3.6 Criblage de banques d'ADNc.....</b>	<b>64</b>
3.6.1 Préparation de la sonde .....	65
3.6.1.1 Marquage aléatoire.....	65
3.6.1.2 Incorporation de la digoxigénine-11-dUTP durant le PCR.....	67
3.6.1.3 Quantification de la sonde.....	69
3.6.2 Titrage des banques d'ADNc.....	70
3.6.2.1 Culture des cellules hôtes XL1-Blue.....	71
3.6.2.2 Préparation des pétris .....	71
3.6.3 Le traitement des membranes.....	72
3.6.3.1 Le traitement des membranes avec la protéinase K .....	74



3.6.4	Hybridation et détection.....	74
3.6.5	Purification des clones positifs.....	77
3.6.6	Conversion de l'ADN du phage en ADN plasmidique.....	78
<b>3.7</b>	<b>Criblage d'une banque génomique.....</b>	<b>80</b>
3.7.1	Titration de la banque génomique .....	80
3.7.1.1	Culture des cellules hôtes <i>E. coli</i> K802.....	80
3.7.2	Isolation de l'ADN à partir des lysats de phages .....	81
<b>3.8</b>	<b>Amplification rapide des extrémités de l'ADNc (RACE).....</b>	<b>82</b>
<b>3.9</b>	<b>PCR sur différents ADNc de tissus humains .....</b>	<b>85</b>
<b>3.10</b>	<b>Criblage bioinformatique .....</b>	<b>85</b>
3.10.1	BLAST à l'aide d'une séquence protéique connue.....	87
3.10.2	Analyse des clones trouvés .....	87
3.10.2.1	Analyse des HTGS par la combinaison de divers algorithmes de recherche .....	88
3.10.2.2	Analyse des HTGS par l'algorithme GENSCAN .....	88
3.10.3	Sélection de différents EST pour le criblage.....	89
<b>CHAPITRE 4</b>	<b><i>RÉSULTATS ET DISCUSSION.....</i></b>	<b>90</b>
<b>4.1</b>	<b>Recherche des cPLA<sub>2</sub> connues dans l'EPR et la rétine neurale.....</b>	<b>90</b>
4.1.1	RT-PCR à l'aide d'amorces spécifiques .....	90
<b>4.2</b>	<b>Recherche de nouvelles cPLA<sub>2</sub> dans l'EPR et la rétine neurale .....</b>	<b>92</b>
4.2.1	Approche expérimentale .....	93
4.2.1.1	RT-PCR à l'aide d'amorces dégénérées .....	93
4.2.1.2	Criblage de banques à l'aide d'une sonde d'ADNc .....	93
4.2.2	Approche bioinformatique .....	94
4.2.2.1	Criblage bioinformatique .....	94
4.2.2.2	Criblage par RT-PCR.....	100
4.2.2.2.1	Obtention d'un fragment pour le gène 2 .....	100
4.2.2.2.2	Analyse du fragment obtenu par RT-PCR .....	101
4.2.2.3	Criblage de banques d'ADNc .....	104

4.2.2.4 Criblage d'une banque génomique.....	105
4.2.2.5 Criblage par RT-PCR de différents tissus.....	109
4.2.2.6 Amplification rapide des extrémités de l'ADNc (RACE) .....	110
<b>CHAPITRE 5 CONCLUSIONS.....</b>	<b>112</b>
5.1 Récapitulation des résultats .....	112
5.2 Conclusions .....	114
5.2.1 Atteinte des objectifs fixés .....	115
5.3 Perspectives de recherche.....	116
<b>RÉFÉRENCES.....</b>	<b>118</b>
<b>ANNEXE A ARTICLE SOUMIS.....</b>	<b>133</b>
<b>ANNEXE B SÉQUENCES DES NOUVEAUX GÈNES 1 ET 2 .....</b>	<b>141</b>

## LISTE DES TABLEAUX

1.1	Compilation des différentes PLA <sub>2</sub> à histidine catalytique .....	17
1.2	Résumé des caractéristiques des PLA <sub>2</sub> des groupes VI, VII et VIII à sérine catalytique .....	23
1.3	Résumé des caractéristiques des cPLA <sub>2</sub> .....	29
3.1	Concentrations utilisées lors des réactions de PCR .....	56
3.2	Compilation des différentes amorces utilisées.....	56
3.3	Concentrations utilisées lors des réactions de PCR pour la préparation des sondes .....	68
3.5	Les différents programmes de l'algorithme BLAST .....	86
4.1	Compilation de différents EST.....	108

## LISTE DES FIGURES

1.1	Coupe transversale d'un œil humain.....	1
1.2	Organisation laminaire de la rétine . ....	3
1.3	Organisation laminaire de la rétine présentant différents types cellulaires. ....	4
1.4	Schéma d'un bâtonnet et de ses différentes composantes.....	5
1.5	Localisation de la molécule de rhodopsine .....	6
1.6	Segments externes des photorécepteurs .....	8
1.7	La cascade de la phototransduction.....	9
1.8	Cellule de l'EPR.....	10
1.9	Coupe transversale d'une monocouche d'EPR de singe.....	10
1.10	Structure d'une cellule différenciée de l'EPR.....	10
1.11	Schéma illustrant le mécanisme de phagocytose des segments externes des photorécepteurs par les cellules de l'épithélium pigmentaire rétinien .....	12
1.12	Hydrolyse d'un phospholipide par la phospholipase A2 .....	14
1.13	Comparaison structurale entre les trois isoformes de cPLA <sub>2</sub> connues dans la littérature jusqu'à présent.....	25
1.14	Schéma de la cPLA <sub>2</sub> alpha montrant les acides aminés les plus importants pour son activité .....	34
1.15	Modèle d'activation de la cPLA <sub>2</sub> alpha .....	37
1.16	La synthèse des différents dérivés de l'acide arachidonique et du lysophospholipide .....	40
3.1	Vérification de l'intégrité de l'ARN par migration sur gel d'agarose .....	53
3.2	Structure de la digoxigénine-11-dUTP .....	66

3.3	Deux des différents types de marquages possibles et la détection au CSPD par le système DIG de <i>Boehringer Mannheim</i> .....	66
3.4	Résultat d'une quantification. ....	70
3.5	Résultat typique d'un criblage de banque d'ADNc (banque de phage) où les clones positifs sont nombreux.....	77
3.6	Conversion du phage recombinant $\lambda$ TriplEx en pTriplEx .....	79
3.7	Schématisation des procédures pour le système du RACE.....	83
3.8	Schéma explicatif pour le clonage des différents produits obtenus lors des amplifications des extrémités de l'ADNc par le système <i>CloneAMP</i> de <i>GIBCO BRL</i> .....	84
4.1	Obtention de la cPLA <sub>2</sub> alpha dans le poumon .....	91
4.2	Obtention de la cPLA <sub>2</sub> gamma dans l'EPR.....	92
4.3	Schéma représentant les HTGS trouvés.....	95
4.4	Séquences codantes des nouvelles cPLA <sub>2</sub> obtenues à partir des nouveaux gènes 1 et 2.....	99
4.5	Obtention d'un fragment d'environ 1100 pb correspondant à l'ADNc du nouveau gène 2 trouvé par des analyses bioinformatiques .....	100
4.6	Séquence du fragment obtenu du nouveau gène 2 à l'aide des amorces « gène 2b » et « gène 2c ».....	102
4.7	Alignement du fragment obtenu par RT-PCR pour le nouveau gène 2 avec le segment correspondant de chacune des trois cPLA <sub>2</sub> connues (alpha, bêta et gamma).....	103
4.8	Représentation schématique des trois isoformes connues (cPLA <sub>2</sub> alpha, bêta et gamma) avec le fragment du nouveau gène 2 .....	104
4.9	Séquence détaillée de chacune des deux sondes internes au nouveau gène 2 utilisées pour le criblage. ....	106
4.10	Position des amorces utilisées pour la fabrication des sondes sur la séquence du fragment obtenu pour le gène 2 .....	108

4.11	Résultat obtenu pour l'amplification par PCR des deux EST utilisés pour la fabrication d'une sonde .....	108
4.12	Amplification du fragment de l'ADNc de la nouvelle cPLA <sub>2</sub> (gène 2) à partir des amorces <i>gène 2b</i> et <i>gène 2c</i> dans différents tissus.....	109
4.13	Résultats obtenus avec la technique du RACE .....	111

## LISTE DES SYMBOLES ET ABRÉVIATIONS

AA	acide arachidonique
ADN	acide désoxyribonucléique
ARNm	acide ribonucléique messenger
BLAST	<i>basic local alignment search tool</i>
pBPB	bromure de para-bromophénacyle
CaLB	domaine de liaison au calcium ( <i>Ca<sup>2+</sup>-dependent phospholipid-binding</i> )
DAG	diacylglycérol
DEDA	acide 7,7-diméthyl-5,8-eicosadiénoïque
DTT	dithiothréitol
EGF	facteur de croissance épidermique
EPR	épithélium pigmentaire rétinien
ERK	<i>extracellular signal-regulated kinase</i>
EST	<i>expressed sequence tags</i>
GFP	protéine fluorescente verte
GMP	guanosine 5'-monophosphate
GMPc	guanosine monophosphate cyclique
GRE	<i>glucocorticoid response element</i>
HDL	lipoprotéine de haute densité
HTGS	<i>high throughput genomic sequences</i>
IL-1	interleukine-1
INF- $\gamma$	interféron gamma
IP3	inositol 1,4,5-trisphosphate
LDL	lipoprotéine de faible densité
LUF	luffariellolide
MAPK	<i>mitogen-activated protein kinase</i>
MLD	monoalide
MNK1	<i>MAP kinase interacting kinase-1</i>
MP	membranes périnucléaires

NCBI	<i>national center for biotechnology information</i>
OA	acide oléique
OPC	oléyloxyethyl phosphorylcholine
PAF	facteur d'agrégation plaquettaire
PAF-AH	acétylhydrolase du facteur d'activation plaquettaire
PH	<i>pleckstrin homology</i>
PIP2	phosphatidylinositol 4,5-biphosphate
sPLA <sub>2</sub>	phospholipase A <sub>2</sub> sécrétée
cPLA <sub>2</sub>	phospholipase A <sub>2</sub> cytosolique
PLC	phospholipase C
PKC	protéine kinase C
PVP	polyvinylpyrrolidone
RE	réticulum endoplasmique
SCL	12-épi-scala-radial
SDS	dodécylsulfate de sodium
TNF- $\alpha$	facteur de nécrose tumorale alpha
TPC	thioéthéramide phosphatidylcholine

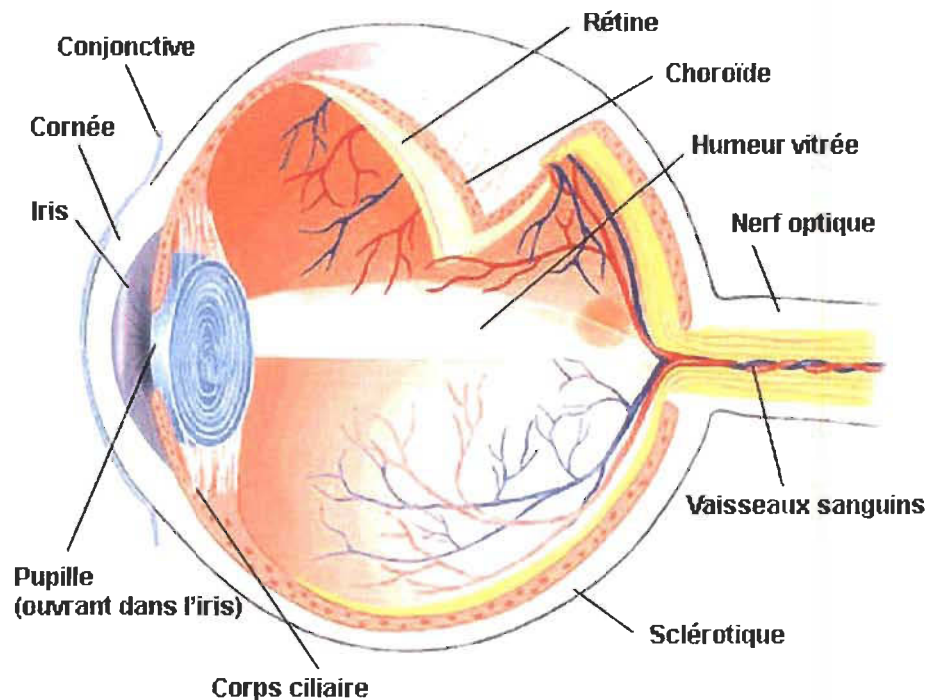


# CHAPITRE 1

## INTRODUCTION

### 1.1 Composantes de l'œil

L'œil est constitué de trois tuniques différentes : la tunique fibreuse, la tunique vasculaire et la tunique interne (Tortora, 1994). La tunique fibreuse est la membrane externe du globe oculaire; elle est formée d'une partie antérieure, la cornée, et d'une partie postérieure, la sclérotique. La cornée est une couche transparente non vascularisée et très innervée qui recouvre la partie colorée de l'œil ou iris. En raison de sa forme incurvée, la cornée contribue à focaliser les rayons lumineux. La face externe de la cornée est recouverte d'un épithélium qui se prolonge avec celui de la conjonctive bulbaire. La sclérotique est une couche blanche de tissu fibreux dense qui enveloppe tout le globe oculaire, à l'exception de la cornée. La sclérotique donne sa forme au globe oculaire; elle le rend plus rigide et protège ses composantes internes (figure 1.1).



**Figure 1.1** – Coupe transversale d'un œil humain. Image tirée du site <http://www.driesen.com>.

La tunique vasculaire est la couche moyenne du globe oculaire et est appelée l'uvée. Elle se compose de trois parties : la choroïde, le corps ciliaire et l'iris. La partie postérieure de la tunique vasculaire, la choroïde, est richement vascularisée et tapisse presque entièrement la face interne de la sclérotique. Ses vaisseaux sanguins nourrissent la surface postérieure de la rétine. Les mélanocytes, cellules produisant la mélanine, sont responsables de la couleur brun foncé caractéristique de la choroïde. La choroïde devient le corps ciliaire à la partie antérieure de la tunique vasculaire. Le corps ciliaire est la région la plus épaisse de la tunique vasculaire. Il couvre la région s'étendant de l'ora serrata jusqu'à un point situé derrière la jonction scléro-cornéenne. L'ora serrata est souvent décrite comme étant la région antérieure de la rétine en dents de scie. Le corps ciliaire est composé des procès ciliaires et du muscle ciliaire. Les procès ciliaires sont des saillies ou replis à la face interne du corps ciliaire, où le revêtement de cellules épithéliales sécrète l'humeur aqueuse, alors que le muscle ciliaire est une bande circulaire de muscle lisse qui modifie la courbure du cristallin pour l'adapter à la vision de loin ou de près. L'iris est la partie colorée du globe oculaire et il a la forme d'un anneau aplati. L'iris est suspendu entre le cristallin et la cornée; son extrémité externe est rattachée au procès ciliaire. Il est composé de fibres musculaires lisses radiaires et circulaires. La pupille est l'ouverture située au centre de l'iris. L'une des fonctions de l'iris consiste à régler l'entrée de la lumière dans la cavité postérieure du globe oculaire par la pupille (Tortora, 1994).

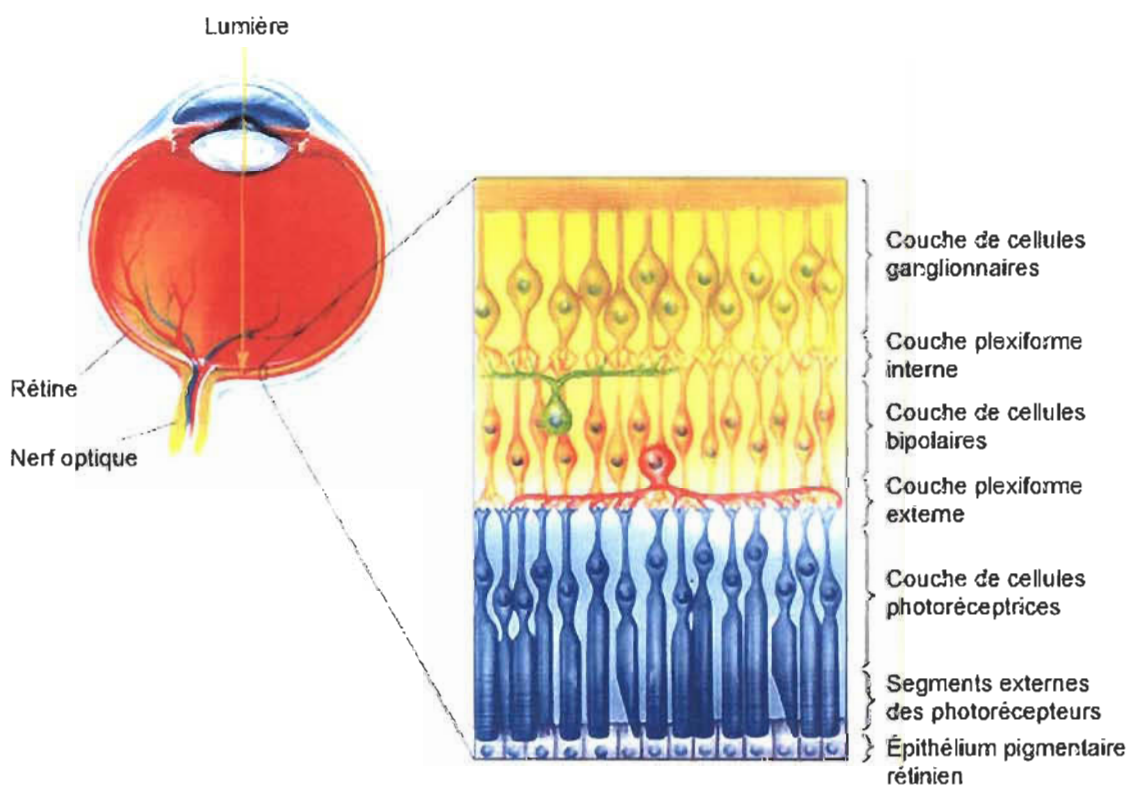
Enfin, la tunique nerveuse, la rétine, ne recouvre que les trois quarts postérieurs de l'œil et elle correspond au début de la voie nerveuse visuelle. À l'aide d'un ophtalmoscope, il est possible de faire l'examen du fond de l'œil et d'obtenir une image agrandie de la rétine et des vaisseaux sanguins qui traversent sa surface antérieure. La rétine est le seul endroit du corps où l'on peut voir directement les vaisseaux sanguins et les examiner pour détecter des modifications vasculaires pathologiques liées à différentes dystrophies rétinienne (Tortora, 1994).

### 1.1.1 La rétine

La rétine se compose de l'épithélium pigmentaire, qui est la partie non visuelle, et d'une partie neurale, visuelle.

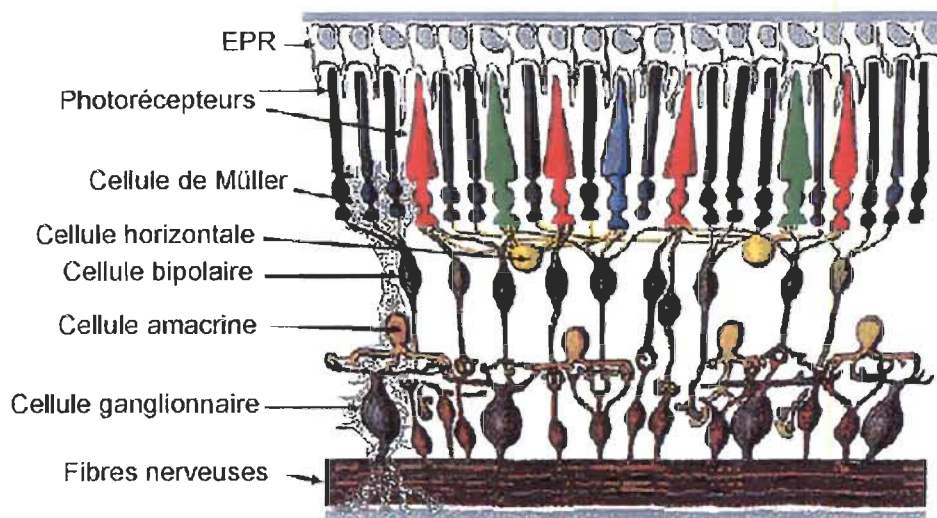
#### 1.1.1.1 La rétine neurale

La partie neurale de la rétine est une excroissance à couches multiples du cerveau. Elle traite largement les données visuelles avant de transmettre des influx nerveux au thalamus, qui les transmet ensuite au cortex visuel primaire. La rétine présente trois couches distinctes de neurones séparées par deux zones où se retrouvent les contacts synaptiques : les couches synaptiques (plexiformes) interne et externe. Les couches de neurones rétinien, nommées selon l'ordre de traitement de l'information visuelle, sont la couche des photorécepteurs, la couche des cellules bipolaires et la couche des cellules ganglionnaires (figure 1.2) (Tortora, 1994).



**Figure 1.2** – Organisation laminaire de la rétine. Image tirée du site Internet <http://www-old.physiol.usyd.edu.au>.

La lumière traverse donc les couches de cellules ganglionnaires avant d'atteindre la couche des photorécepteurs. Trois autres types de cellules sont présents dans la rétine : les cellules horizontales, les cellules amacrines et les cellules de Müller (figure 1.3).



**Figure 1.3** - Organisation laminaire de la rétine présentant différents types cellulaires. Image tirée du site Internet de WebVision.

Les deux premières forment des voies dirigées latéralement qui modifient les signaux transmis le long de la voie depuis les photorécepteurs jusqu'aux cellules bipolaires et ganglionnaires. Les cellules de Müller ne sont pas d'origine neurale et servent plutôt d'échafaudage dans les espaces interneuronaux (Lodish, 1997). Elles parcourent la rétine neurale d'une extrémité à l'autre. Ce sont des cellules gliales et elles exercent non seulement un rôle de soutien, mais sont aussi responsables du transport de métabolites du système sanguin vers les quatre types de cellules neuronales : les cellules bipolaires, horizontales, amacrines et ganglionnaires.

#### **1.1.1.1.1 Les photorécepteurs**

Les photorécepteurs sont constitués d'un segment externe, d'un segment interne et d'une terminaison synaptique (figure 1.4). La terminaison synaptique se trouve du côté du segment interne, là où elle entre en contact avec d'autres terminaisons nerveuses de la rétine neurale. Le segment interne comprend les différents organites cellulaires tels que

l'appareil de Golgi et les mitochondries. Le segment externe des bâtonnets est plus long que celui des cônes.

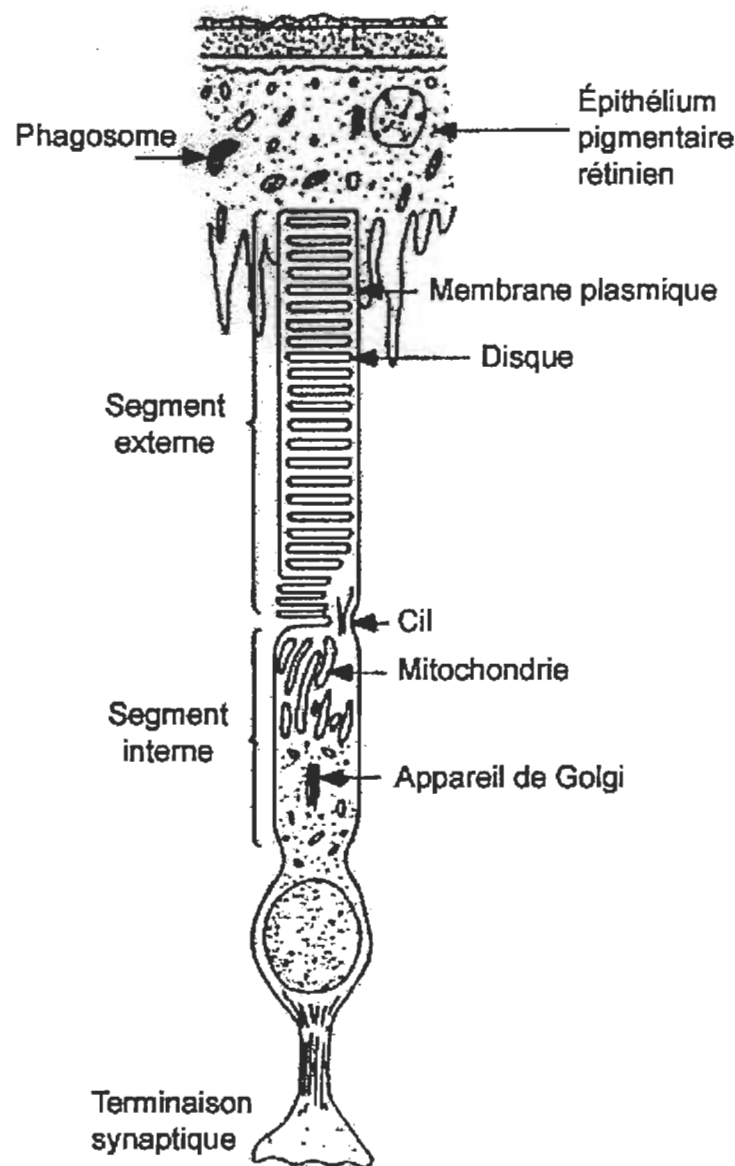
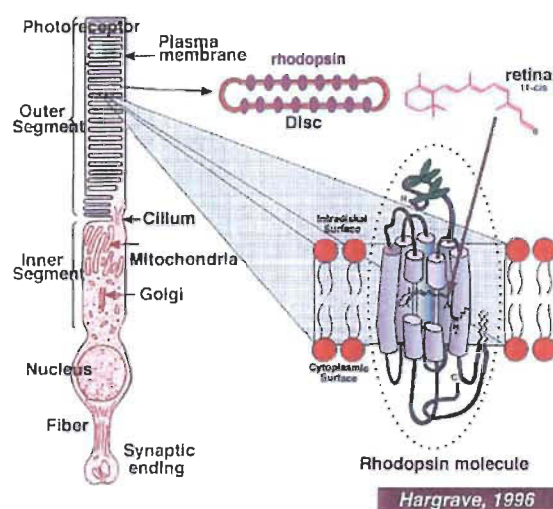


Figure 1.4 – Schéma d'un bâtonnet et de ses différentes composantes. Image tirée de Giusto (1990).

Le segment externe du bâtonnet consiste en un empilement de disques aplatis riches en rhodopsine et entourés par une membrane plasmique. Ces disques sont en fait le site d'initiation de la cascade d'événements de la transduction visuelle puisque la rhodopsine, le pigment visuel, est insérée dans la membrane du disque (figure 1.5).





**Figure 1.5** – Localisation de la molécule de rhodopsine. La rhodopsine se situe à l'intérieur du segment externe du bâtonnet, plus précisément dans la membrane du disque (Bernstein, 1999).

La membrane des disques du segment externe du bâtonnet a une composition phospholipidique intéressante. En effet, elle a une composition particulièrement élevée en acides gras polyinsaturés; plus de 60% des acides gras des phospholipides sont polyinsaturés. Parmi ces acides gras polyinsaturés, l'acide docosahexaénoïque (22:6 $\omega$ 3) compte pour 74% (Miljanich *et al.*, 1979; Salesse *et al.*, 1984). Il s'agit donc de la membrane naturelle la plus insaturée à être connue jusqu'à maintenant chez les mammifères vertébrés (Fliesler et Anderson, 1983; Giusto *et al.*, 2000). Puisque les acides gras polyinsaturés sont très susceptibles à la peroxydation (van Kuijk et Dratz, 1987), la proportion importante d'acides gras insaturés rend les disques membranaires très vulnérables aux dommages reliés à la peroxydation lipidique, qui sont reconnus pour altérer la structure et la fonction membranaire (Sevanian *et al.*, 1988; Wratten *et al.*, 1989).

La connexion entre les segments interne et externe est assurée par un cil modifié qui contient neuf paires de microtubules. Le photopigment est synthétisé dans le segment interne, puis incorporé dans les membranes du segment externe. Dans les bâtonnets, ce pigment s'incorpore dans des membranes nouvellement synthétisées (trois disques par heure) (Guénard, 1996). Ces disques migrent vers l'extrémité du segment externe, où ils sont progressivement éliminés et phagocytés par l'épithélium pigmentaire rétinien.

Dans les cônes, en revanche, le photopigment se distribue dans les replis de la membrane plasmique du segment externe puisqu'il n'y a pas de disques.

Les segments externes des photorécepteurs sont donc caractérisés par une synthèse constitutive de nouveaux disques membranaires à leur extrémité basale (Young et Bok, 1969). Ce renouvellement rapide (Feeney, 1973) est nécessaire pour compenser la dégradation de leur état fonctionnel, qui est engendrée par le photoblanchiment de la rhodopsine, de même que par la peroxydation probable des acides gras polyinsaturés qui les composent. Il a été démontré en 1969, par Young et Bok, que suite à leur synthèse, les nouveaux disques membranaires migrent vers l'extrémité distale des segments externes des photorécepteurs pour finalement être phagocytés par les cellules de l'épithélium pigmentaire rétinien. Il y a donc un équilibre entre la quantité de disques synthétisés et la quantité de disques phagocytés afin de maintenir constante la longueur des segments externes des photorécepteurs. Il a aussi été démontré que, suite à la digestion lysosomiale des disques phagocytés par les cellules de l'épithélium pigmentaire rétinien, certains de leurs éléments constitutifs sont retournés vers le segment interne des photorécepteurs (Gordon *et al.*, 1992; Chen et Anderson, 1993). Ainsi, l'acide docosahexanoïque provenant de disques membranaires phagocytés suit cette boucle de recyclage, c'est-à-dire qu'il est retourné au segment interne des photorécepteurs pour être recyclé. Cette boucle de recyclage pourrait impliquer les cPLA<sub>2</sub>, particulièrement au niveau d'un processus de déacylation-réacylation permettant la réparation des acides gras des phospholipides peroxydés avant qu'ils ne reprennent la boucle du recyclage. Les phospholipides ainsi recyclés migrent ensuite vers la portion basale des segments externes pour être de nouveau incorporés dans les disques membranaires en formation (Bazan *et al.*, 1992).

Les photorécepteurs sont des récepteurs visuels hautement spécialisés dans la transduction des rayons lumineux en potentiels générateurs. Les deux types de photorécepteurs sont appelés bâtonnets et cônes à cause de la forme de leur segment externe (figure 1.6). La rétine comprend environ six millions de cônes et 120 millions de bâtonnets. Les bâtonnets sont les récepteurs sensoriels responsables de la vision en

noir et blanc en semi-obscurité (vision scotopique). Ce sont donc eux qui permettent de distinguer les différentes nuances de lumière et d'ombre et de percevoir les formes et les mouvements. Cette caractéristique est attribuable à leur nombre important et à leur composition importante en photopigments, c'est-à-dire qu'ils ont un meilleur système d'amplification comparativement aux cônes. Les cônes sont des cellules spécialisées dans la vision des couleurs et l'acuité visuelle (netteté de la vision) sous une lumière vive (vision photopique). Ils ont un seuil d'activation par la lumière plus élevé et ne peuvent fonctionner en cas de faible éclairage (Tortora, 1994; Giusto *et al.*, 2000).

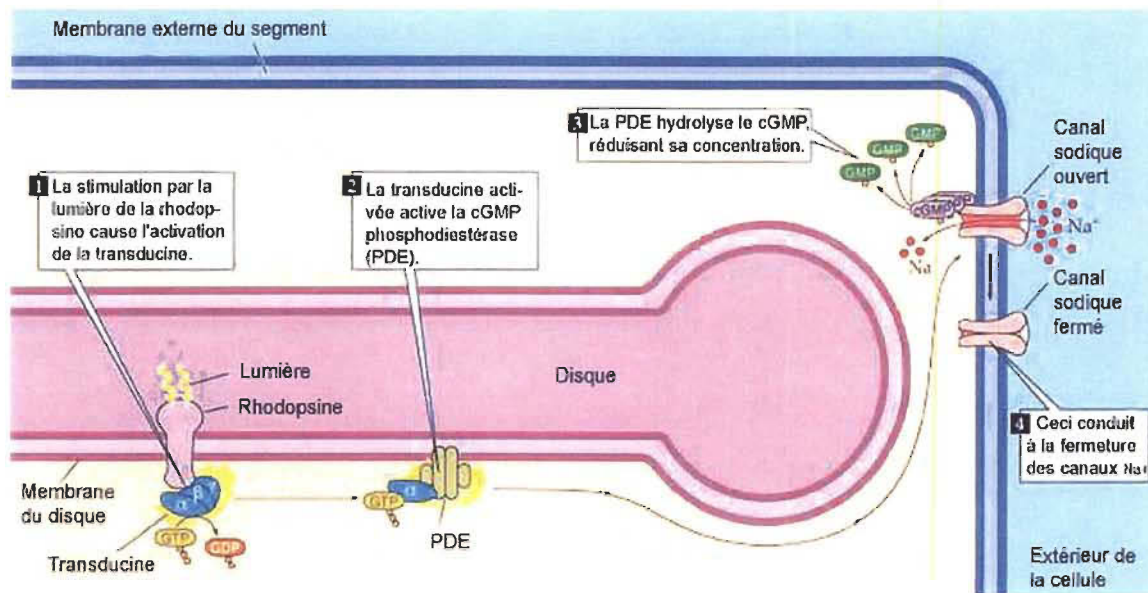


**Figure 1.6** – Segments externes des photorécepteurs. Segment externe du bâtonnet à gauche et du cône à droite (Young, 1970).

L'absorption de la lumière est nécessaire à la vision. Cette absorption est assurée par les pigments qui se trouvent dans le segment externe des bâtonnets et des cônes. Tel que mentionné plus tôt, le pigment qui est présent dans le segment externe des bâtonnets est la rhodopsine. La rhodopsine contient du rétinol, l'aldéhyde de la vitamine A. Un isomère du rétinol, le 11-cis rétinol, combiné avec l'opsine, une glycoprotéine, forme la rhodopsine. La lumière provoque l'isomérisation du 11-cis-rétinol en tout-trans-rétinol (métarhodopsine II), entraîne la dissociation de l'opsine et du rétinol et conduit à la conversion du rétinol en rétinol. Cette isomérisation du 11-cis-rétinol en tout-trans rétinol entraîne un changement conformationnel de l'opsine, et par conséquent l'activation de la transducine. Cette protéine G activera la *phosphodiesterase (PDE)* qui hydrolysera le GMPc. La diminution subséquente de la concentration en GMPc dans le cytoplasme va alors entraîner la dissociation du GMPc des canaux à  $\text{Na}^+/\text{Ca}^{++}$ , et la fermeture de ces canaux provoquera l'hyperpolarisation des photorécepteurs (figure 1.7).



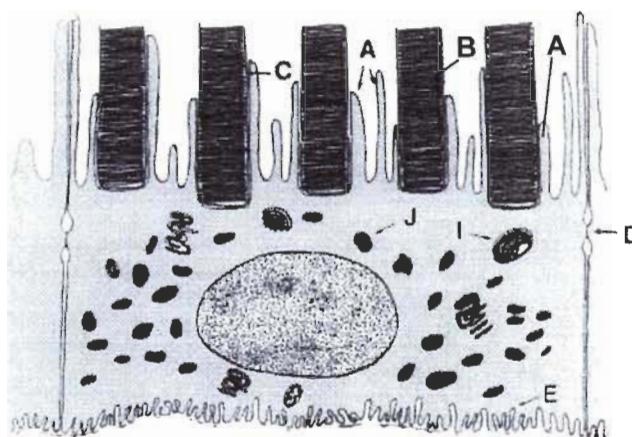
Le tout-trans-rétinal se dissocie de l'opsine et, une fois transformé par la *tout-trans rétinol déshydrogénase* en tout-trans rétinol, il est transporté vers l'épithélium pigmentaire, où il est ultimement transformé en 11-cis rétinol, qui est ensuite transporté vers la couche des photorécepteurs, incorporé dans les segments externes et combiné à nouveau à l'opsine pour régénérer la rhodopsine (Guénard, 1996; Saari, 2000).



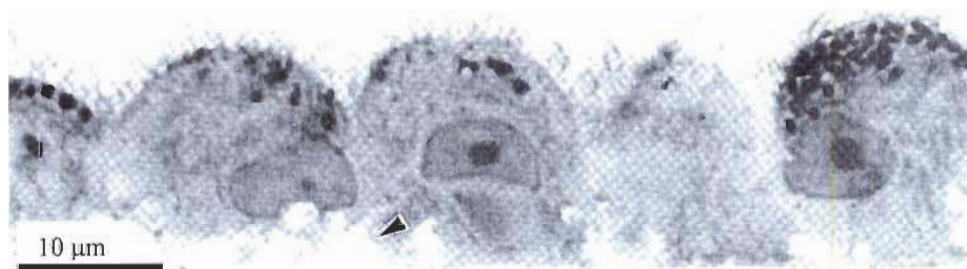
**Figure 1.7** - La cascade de la phototransduction. La lumière entraîne la photoisomérisation de la rhodopsine, activant la transducine, une protéine G hétérotrimérique (1). La sous-unité alpha active la *phosphodiésterase* (PDE) (2), laquelle dégrade le GMPc en 5'GMP (3). La diminution de la concentration en GMPc cause la fermeture des canaux sodiques (4). Cette diminution de l'entrée de sodium entraîne l'hyperpolarisation du photorécepteur. Image tirée du site Internet [www.utdallas.edu/~tres/integ/sen3/thum\\_7\\_09.jpg](http://www.utdallas.edu/~tres/integ/sen3/thum_7_09.jpg).

### 1.1.1.2 L'épithélium pigmentaire rétinien

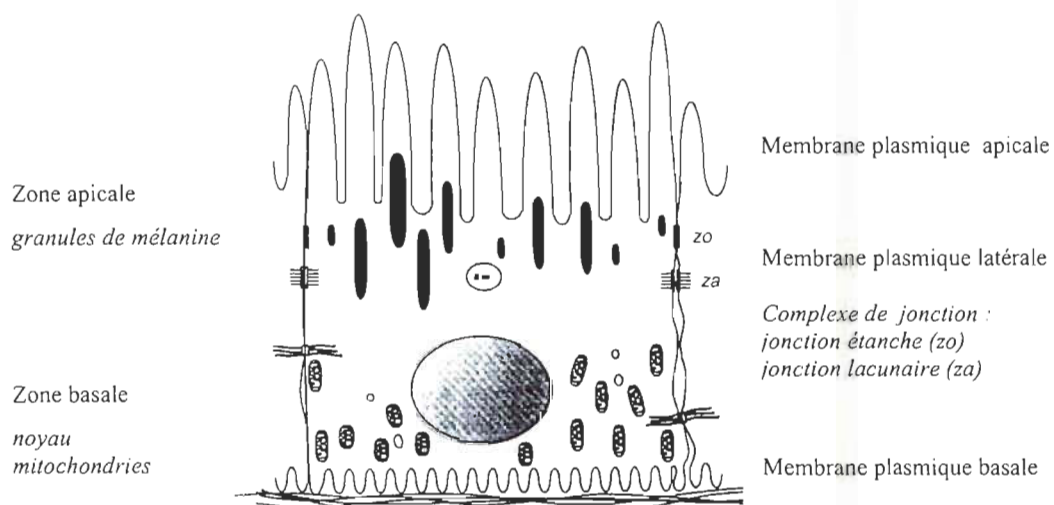
L'épithélium pigmentaire rétinien (EPR) est une monocouche de cellules épithéliales contenant de la mélanine (niveau apical), se trouvant entre la choroïde et la partie neurale de la rétine. Les cellules de l'épithélium pigmentaire rétinien ont une structure particulière (figures 1.8, 1.9 et 1.10). Elles possèdent une polarité morphologique et fonctionnelle; leurs différents organites et protéines membranaires sont distribués asymétriquement de part et d'autre de l'axe séparant l'apex de la région baso-latérale.



**Figure 1.8** – Cellule de l'EPR. Ce schéma montre les microvillis courts et longs (A), le segment externe des photorécepteurs (B), la matrice interphotoréceptrice (C) et les jonctions étanches intercellulaires (D). Image tirée de (Clark, 1986).



**Figure 1.9** – Coupe transversale d'une monocouche d'EPR de singe. Les processus apicaux sont évidents de même que les citernes basolatérales (pointe de flèche). Quelques-unes des cellules contiennent des mélanosomes ovoïdes dans la partie apicale (haut de la photo) (Pfeffer, 1991).



**Figure 1.10** - Structure d'une cellule différenciée de l'EPR (Marmor et Wolfensberger, 1998).

La membrane plasmique apicale (face à la rétine) présente de courts et longs microvillis (figure 1.8). Les microvillis courts phagocytent les segments externes des cônes et des bâtonnets, alors que les longs se projettent à la manière de doigts le long de ces segments externes (Hogan *et al.*, 1971). L'échange de nutriments et de déchets entre ces deux types de cellules est donc facilité par la proximité et la grande surface entre les membranes plasmiques des photorécepteurs et des cellules de l'EPR (Miller et Steinberg, 1979). Toutefois, la membrane plasmique des photorécepteurs et celle des cellules de l'EPR ne présentent aucun point de contact puisqu'elles sont séparées par une très fine couche de matrice extracellulaire nommée matrice interphotoréceptrice.

Sur leur surface latérale, chaque cellule de l'EPR est liée à ses cellules voisines par une ceinture continue de jonctions lacunaires et de jonctions étanches intercellulaires (voir figure 1.10) (Shakib *et al.*, 1972; Hudspeth et Yee, 1973). Les jonctions lacunaires permettent aux petites molécules de diffuser d'une cellule vers ses voisines, alors que les jonctions étanches se trouvent près de la région apicale de la membrane plasmique et préviennent la diffusion de macromolécules via les espaces intercellulaires au niveau de l'EPR. C'est principalement cette structure qui assure la barrière hémato-oculaire (Clark et Hall, 1986) et qui contrôle le flot de nutriments qui se déplace des choriocapillaires vers les photorécepteurs. C'est donc cette barrière sélective qui est responsable du maintien de la composition des fluides intracellulaires des photorécepteurs.

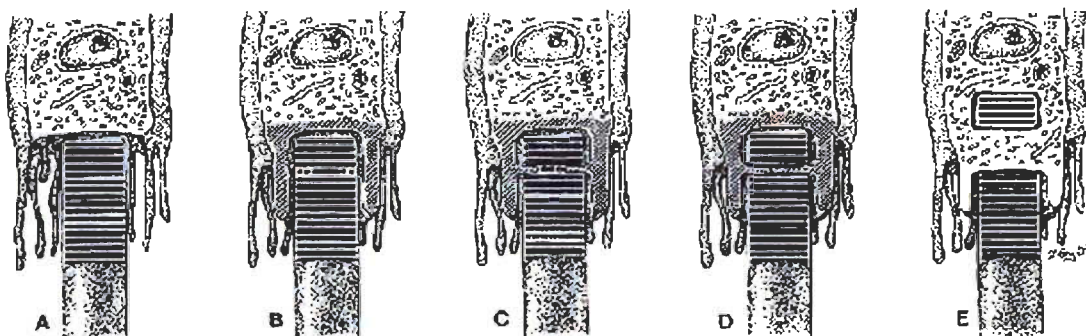
#### **1.1.1.2.1 Les fonctions de l'EPR**

L'EPR est l'une des plus importantes couches cellulaires dans le système visuel et est essentiel à l'intégrité structurale et fonctionnelle des cônes et des bâtonnets de la rétine. En plus d'être responsable de l'échange de métabolites et de déchets entre la rétine et les capillaires de la choroïde (Miller et Steinberg, 1979), l'EPR participe à plusieurs fonctions complexes telles que l'adhésion de la rétine (Miller *et al.*, 1987) et elle représente la principale structure responsable de la barrière hémato-oculaire (Clark et Hall, 1986). Aussi, la mélanine de l'EPR absorbe les rayons lumineux épars, ce qui prévient la réflexion et la dispersion intraoculaire de la lumière (Sarna, 1992). Cela

permet d'assurer que l'image produite sur la rétine par la cornée et le cristallin reste nette. De plus, les cellules de l'épithélium pigmentaire rétinien participent au métabolisme et à l'entreposage de la vitamine A (rétinol). En effet, une grande partie du rétinol qui a été photoblanchi lors du processus visuel est estérifié et emmagasiné dans les cellules de l'épithélium pigmentaire rétinien (Saari, 1987). Enfin, l'EPR exerce un rôle essentiel dans le maintien de l'intégrité structurale et fonctionnelle de la rétine par la phagocytose et la digestion des segments externes des photorécepteurs (Young et Bok, 1969), tel que mentionné plus haut.

#### 1.1.1.2.2 La phagocytose des photorécepteurs

Le mécanisme de la fonction phagocytaire de l'épithélium pigmentaire rétinien comporte cinq principales étapes (Matsumoto *et al.*, 1987) (figure 1.11). Dans un premier temps, il y a l'attachement des segments externes des photorécepteurs aux cellules de l'EPR via des récepteurs (figure 1.11A). Vient ensuite l'initiation de l'excision de l'extrémité distale des segments externes et l'allongement des pseudopodes des cellules de l'EPR (figure 1.11B). Dans un troisième temps, les pseudopodes entourent entièrement l'extrémité distale des segments externes à phagocyter (figure 1.11D). Enfin, le phagosome qui vient d'être créé migre à l'intérieur des cellules de l'EPR (figure 1.11E) pour ensuite fusionner avec des lysosomes et former le phagolysosome, là où les disques membranaires phagocytés subiront l'activité des enzymes hydrolytiques lysosomiales.



**Figure 1.11** – Schéma illustrant le mécanisme de phagocytose des segments externes des photorécepteurs par les cellules de l'épithélium pigmentaire rétinien. L'extrémité distale des segments externes des photorécepteurs est enveloppée par des pseudopodes de l'EPR et il y a formation progressive du phagosome qui migre alors vers les lysosomes cellulaires (A-E) (Matsumoto *et al.*, 1987).

### 1.1.1.2.3 Implication du récepteur mannose au niveau de la phagocytose

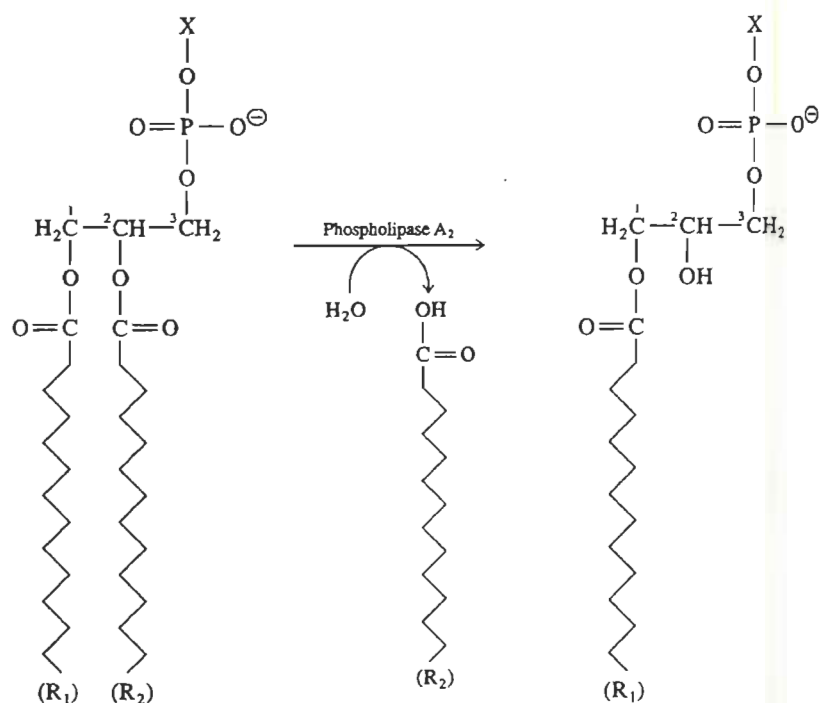
L'implication du récepteur mannose dans la phagocytose des segments externes des photorécepteurs par les cellules de l'épithélium pigmentaire rétinien a été démontrée en 1991 par l'équipe de Boyle (Boyle *et al.*, 1991). Au cours de leurs expériences, la préincubation des segments externes des photorécepteurs dans une solution d'anticorps dirigés contre les récepteurs mannose a permis d'observer une diminution de 80% du processus phagocytaire comparativement aux cellules de l'EPR incubées avec les segments externes des photorécepteurs seulement. De plus, la préincubation des segments externes des photorécepteurs avec le récepteur mannose purifié permet de noter une réduction de 93% de l'activité phagocytaire. Ces expériences permettent donc de croire en l'implication du récepteur mannose dans le processus phagocytaire.

Plus récemment, un récepteur de 180 kDa appelé *M-type receptor* a été étudié et les sPLA<sub>2</sub> apparaissent être les ligands endogènes normaux de ce récepteur. Ce dernier a une organisation moléculaire similaire au récepteur mannose des macrophages qui médie le recaptage des ligands mannose-glycosylés et la phagocytose des microorganismes parasitaires. Le récepteur *M-type* est internalisé par un processus d'endocytose, mais le rôle de cette internalisation demeure obscur (Zvaritch *et al.*, 1996).

## 1.2 Les phospholipases A2

Les phospholipases A2 constituent une superfamille d'enzymes définies par leur habileté à hydrolyser la liaison ester en position *sn*-2 des phospholipides membranaires (Dennis, 1983; Waite, 1987). Cette réaction libère un acide gras et un lysophospholipide (voir figure 1.12).





**Figure 1.12** - Hydrolyse d'un phospholipide par la phospholipase  $A_2$ . La chaîne en position *sn*-2 du phospholipide (à gauche) est hydrolysée par la phospholipase  $A_2$  et conduit à la production d'un lysophospholipide (à droite) et d'un acide gras (au centre). Image tirée de (Horton *et al.*, 1994).

### 1.2.1 La classification des phospholipases $A_2$

L'activité phospholipasique a été pour la première fois étudiée dans le venin de cobra à la fin du 19<sup>e</sup> siècle (Stephens *et al.*, 1898). Toutefois, la première  $PLA_2$  ne fut découverte qu'en 1945 dans le liquide pancréatique de mammifères (Fairbairn, 1945). Depuis, plusieurs autres phospholipases ont été isolées et caractérisées. Par conséquent, la classification des  $PLA_2$  a évolué au cours des années dans la littérature. Par la suite, plusieurs  $PLA_2$  ont été obtenues à partir de venins variés de serpents et d'abeilles et de pancréas de différents mammifères. Plus tard, ces petites enzymes sécrétées sont décrites comme étant calcium-dépendantes, présentant plusieurs ponts disulfure et possédant des acides aminés histidine et aspartate catalytiques (Davidson et Dennis, 1990). Au cours des années qui suivirent, d'autres  $PLA_2$  sécrétées ( $sPLA_2$ ) ont été découvertes et ces enzymes ont été divisées en deux groupes, selon la position des liaisons disulfure et les extensions et *loop* uniques (Heinrikson *et al.*, 1977; Dufton et Hider, 1983). Subséquemment, le groupe II des  $PLA_2$  s'est étendu aux  $PLA_2$  non

pancréatiques, parfois appelées sPLA<sub>2</sub>, en relation avec leur première isolation en tant qu'enzymes sécrétées dans le liquide synovial (Kramer *et al.*, 1989; Seilhamer *et al.*, 1989), et maintenant classifié dans le groupe IIA. Ensuite, l'unique PLA<sub>2</sub> du venin d'abeille fut classifiée dans le groupe III. Plusieurs formes additionnelles de PLA<sub>2</sub> sécrétées utilisant une histidine catalytique ont été découvertes ces dernières années. Ces PLA<sub>2</sub> sont clairement reliées aux groupes I, II et III des PLA<sub>2</sub>, mais ne se classent toutefois pas facilement dans ces derniers. Ceci a donc conduit à l'établissement des groupes V, IX, X, XI et XII.

La classification des PLA<sub>2</sub> en tant que petites enzymes sécrétées devint désuète en 1991 avec le séquençage et le clonage d'une PLA<sub>2</sub> (Clark *et al.*, 1991; Sharp *et al.*, 1991) dont l'activité phospholipasique A2 avait été détectée dans le cytosol de neutrophiles (Alonso *et al.*, 1986) et de plaquettes (Kramer *et al.*, 1986). Le séquençage de cette nouvelle PLA<sub>2</sub> cytosolique révéla une protéine de 85 kDa tout à fait différente de celles qui étaient connues à ce moment; l'absence de ponts disulfure de même que la présence d'une sérine catalytique furent observées (Clark *et al.*, 1991; Sharp *et al.*, 1991).

La présence au niveau cytosolique de ces PLA<sub>2</sub> leur a valu l'abréviation de cPLA<sub>2</sub>, mais elles sont actuellement classifiées dans le groupe IV (IVA puisque actuellement deux autres cPLA<sub>2</sub> ont été identifiées). Enfin, trois autres différents groupes de PLA<sub>2</sub> ont été découverts et classifiés, les groupes VI, VII et VIII, menant ainsi le nombre total de groupes à 12.

Actuellement, la classification des PLA<sub>2</sub> en douze groupes selon quatre critères différents prévaut dans la littérature (Six et Dennis, 2000). Le premier critère : l'enzyme doit catalyser l'hydrolyse de la liaison ester *sn*-2 d'un substrat phospholipidique naturel, tel que les phospholipides à longues chaînes d'acides gras (2 à 20 carbones ou plus), le facteur d'activation plaquettaire (PAF) ou des phospholipides à courtes chaînes d'acides gras oxydés. Deuxième critère : la séquence complète en acides aminés de la protéine doit être connue. Troisième critère : chaque groupe de PLA<sub>2</sub> doit inclure toutes les enzymes qui présentent une homologie de séquence importante. Si plus d'un gène de

PLA<sub>2</sub> homologue existe à l'intérieur d'une même espèce, alors on ajoute simplement une lettre pour le sous-groupe, comme c'est le cas pour les groupes IVA, IVB et IVC. Quatrièmement : les différentes formes actives d'un gène issues de l'épissage alternatif sont classées dans le même groupe et sous-groupe, mais sont distinguées par l'ajout d'un nombre arabe, comme c'est le cas par exemple pour les groupes VIA-1 et VIA-2. C'est donc sur la base de ces quatre critères que la classification actuelle comprend 12 groupes (I-XII) de PLA<sub>2</sub>.

### 1.2.2 Les phospholipases A2 sécrétées

Les phospholipases dites sécrétées (sPLA<sub>2</sub>) ou à histidine catalytique comprennent les groupes I, II, V, X, qui présentent un ensemble de caractéristiques très similaires, en plus des groupes III, IX, XII et XI, ce dernier groupe comprenant des sPLA<sub>2</sub> provenant de plantes et non de mammifères, contrairement aux précédentes.

Les groupes I, II, V et X sont intimement reliés et partagent un mécanisme commun pour le clivage du lien ester *sn*-2 des phospholipides. L'hydrolyse semble procéder via l'activation et l'orientation d'une molécule d'eau par une liaison hydrogène au site histidine actif, lequel est dépendant d'un pH variant de 7 à 9 (Scott *et al.*, 1990; White *et al.*, 1990; Scott *et al.*, 1991; Arni et Ward, 1996). L'aspartate conservée adjacente à l'histidine catalytique (His48) forme la dyade His/Asp. Ainsi, l'enzyme utilise l'histidine-48 dans son site actif et l'aspartate-49 polarisera une molécule d'eau, laquelle attaquera ensuite le groupement carbonyle de l'acide gras en position *sn*-2 du substrat phospholipidique (Scott *et al.*, 1991). L'aspartate est responsable de la liaison calcique qui permet de stabiliser l'état de transition négativement chargé de la réaction par la formation d'un trou oxy-anion, d'où la dépendance en calcium de l'ordre du millimolaire (mM). Enfin, d'autres acides aminés conservés participent à ce système d'activation. Notamment, les tyrosines et les glycines de la boucle liant le calcium et un deuxième aspartate qui active et oriente l'histidine catalytique (Six et Dennis, 2000).

Les groupes I, II, V et X des sPLA<sub>2</sub> présentent non seulement une bonne identité au niveau de leur séquence, mais une homologie plus grande encore au niveau structurel



(Six et Dennis, 2000). En plus des acides aminés hautement conservés du site catalytique, ces sPLA<sub>2</sub> présentent toutes six ponts disulfure liés à des acides aminés précis et, jusqu'à deux ponts disulfure supplémentaires uniques pour chacun des membres d'un groupe. De plus, toutes ces enzymes ont des séquences signal qui sont clivées dans le processus sécrétoire de la protéine active mature. La seule exception, la PLA<sub>2</sub> sécrétée du groupe IB est munie d'un propeptide et doit être clivée par la trypsine pour devenir une enzyme active mature. Les PLA<sub>2</sub> de ces quatre groupes possèdent toutes un poids moléculaire approximatif variant de 13 à 20 kDa.

Puisque les sPLA<sub>2</sub> se font de plus en plus nombreuses et qu'elles ne sont pas vraiment l'intérêt du travail de recherche présenté, une compilation résumant leurs principales caractéristiques et leur localisation permet d'avoir une vision plus globale de ces différents groupes dans le tableau 1.1.

**TABLEAU 1.1**  
**Compilation des différentes PLA<sub>2</sub> à histidine catalytique<sup>a</sup>**

<b>Phospholipases A2 utilisant une histidine catalytique (sPLA<sub>2</sub>)</b>					
<b>Groupe</b>	<b>Sous-groupe</b>	<b>Sources</b>	<b>Taille (kDa)</b>	<b>Ponts disulfure</b>	<b>Source de l'archétype</b>
I	A	Cobra, venin de "krait"	13-15	7	Cobra
	B <sup>b</sup>	Pancréas de mammifères	13-15	7	Humain
II	A	Liquide synovial, plaquettes de serpent à sonnette, venin de vipère	13-15	7	Humain
	B <sup>c</sup>	Venin de vipère gaboon	13-15	6	Vipère
	C <sup>d</sup>	Testicules de souris/rat	15	8	Souris
	D	Pancréas, rate de souris/humain	14-15	7	Humain
	E	Cerveau, cœur, utérus de souris/humain	14-15	7	Humain
	F <sup>e</sup>	Testicules, embryon de souris	16-17	7	Souris
V		Cœur, poumon, macrophage de mammifères	14	6	Humain
X		Rate, thymus, leucocyte d'humain	14	8	Humain
III <sup>f</sup>		Abeille, lézard, scorpion, humain	15-18	5	Abeille à miel

### Phospholipases A2 utilisant une histidine catalytique (sPLA<sub>2</sub>)

Groupe	Sous-groupe	Sources	Taille (kDa)	Ponts disulfure	Source de l'archétype
IX		Venin d'escargot (conodipine M)	14	6	Escargot de mer
XI	A	Pousse de riz vert (PLA <sub>2</sub> -I)	12.4	6	Riz
XI	B	Pousse de riz vert (PLA <sub>2</sub> -II)	12.9	6	Riz
XII		Cœur, muscle squelettique, rein et pancréas humains	20	nd	T2 helper murin

Inspiré de (Six et Dennis, 2000)

<sup>a</sup> Ce sont typiquement de petites PLA<sub>2</sub> extracellulaires requérant des concentrations calciques de l'ordre du millimolaire et un site actif comprenant l'histidine et la paire d'aspartates. Notez que les groupes V et X sont listés après les groupes I et II à cause de leur identité avec plusieurs acides aminés conservés incluant 6 ponts disulfure, une histidine et deux aspartates, de même qu'un peptide signal N-terminal qui est clivé pour obtenir une PLA<sub>2</sub> mature.

<sup>b</sup> Le groupe IB présente une insertion de 5 acides aminés connus sous le nom de boucle pancréatique.

<sup>c</sup> Le groupe IIB ne présente que 5 des 6 ponts disulfure hautement conservés.

<sup>d</sup> Le groupe IIC chez l'homme n'est qu'un pseudogène

<sup>e</sup> Le groupe IIF présente une cystéine additionnelle dans son extension C-terminale

<sup>f</sup> Le groupe III chez l'homme (55 kDa) semble posséder de nouveaux domaines C-terminal et N-terminal additionnels.

#### 1.2.2.1 Inhibiteurs des sPLA<sub>2</sub>

Les inhibiteurs présentés ici concernent uniquement les sPLA<sub>2</sub> les plus étudiées, c'est-à-dire les sPLA<sub>2</sub> des groupes I et II.

Le bromure de para-bromophénacyle (*p*BPB), un agent dénaturant modifiant spécifiquement les acides aminés histidine d'une séquence peptidique, est évidemment un inhibiteur des sPLA<sub>2</sub> puisqu'il se lie à l'acide aminé His48 au niveau du site actif et bloque ainsi l'accès aux molécules de substrat (Waite, 1987). Les sPLA<sub>2</sub> se caractérisent aussi par leur sensibilité à l'action dénaturante du dithiothréitol (DTT) (Bennett *et al.*, 1990) - à l'exception du groupe I - qui est un agent réducteur des ponts disulfure. De plus, les sPLA<sub>2</sub> peuvent être soumises à un traitement au H<sub>2</sub>SO<sub>4</sub> sans être inactivées (Apitz-Castro *et al.*, 1979) et résistent à la chaleur (Kozumplik *et al.*, 1989).

L'EGTA, puisqu'il est un chélateur des ions calcium, rend non disponible cet ion pour

son utilisation par les enzymes. Par conséquent, l'EGTA est un inhibiteur des sPLA<sub>2</sub> et cPLA<sub>2</sub> calcium-dépendantes (Tischfield, 1997).

Il existe plusieurs autres inhibiteurs de sPLA<sub>2</sub>. Les plus étudiés sont les produits marins naturels tels que le manoalide (MLD), le luffariellolide (LUF) et le 12-épi-scala-radial (SCL), les substrats analogues dont le thioéthéramide phosphatidylcholine (TPC) et l'oléxyéthyl phosphorylcholine (OPC) et un produit analogue à la PLA<sub>2</sub>, l'acide 7,7-diméthyl-5,8-eicosadiénoïque (DEDA). Ces composés stimulent, inhibent faiblement ou n'ont simplement pas d'effet sur l'activité de la cPLA<sub>2</sub> (Hope *et al.*, 1993).

### 1.2.3 Les phospholipases A2 des groupes VI, VII, et VIII

Bien que ces phospholipases soient classées dans des groupes à part, elles font partie des phospholipases à sérine catalytique, tout comme les PLA<sub>2</sub> du groupe IV qui seront détaillées à la sous-section 1.2.4.

#### 1.2.3.1 Le groupe VIA

La première PLA<sub>2</sub> clonée et caractérisée de ce groupe est souvent référée en tant que iPLA<sub>2</sub>, pour son indépendance au calcium. Elle est actuellement classifiée dans le groupe VIA-1. Cette PLA<sub>2</sub> a été pour la première fois isolée chez les macrophages P388D (Ackermann *et al.*, 1994). Il s'agit d'une protéine de 85-88 kDa contenant 750 acides aminés et comprenant huit répétitions ankyrine à l'extrémité N-terminale. Elle présente la séquence consensus lipasique, GTSTG, contenant la sérine (465) catalytique. La PLA<sub>2</sub> du groupe VIA présente une activité PLA<sub>2</sub> et lysophospholipasique et elle est capable d'hydrolyser une grande variété de substrats phospholipidiques (Tang *et al.*, 1997; Lio et Dennis, 1998).

Il existe de multiples isoformes de cette protéine, dû à l'épissage alternatif du gène. Ainsi, les PLA<sub>2</sub> du groupe VIA-2 présentent une forte identité avec les PLA<sub>2</sub> du groupe VIA-1; il n'y a qu'une insertion de 54 acides aminés au niveau de la huitième répétition

ankyrine en N-terminal. Cette insertion élève le poids moléculaire de cette protéine à 88 kDa (Winstead *et al.*, 2000).

Une troisième isoforme, connue sous l'appellation de PLA<sub>2</sub> du groupe VIA-3 a été identifiée. Sa séquence est tronquée en C-terminal et elle ne contient que 640 acides aminés. Il n'est pas encore déterminé si cette séquence encode une PLA<sub>2</sub> fonctionnelle (Winstead *et al.*, 2000).

Deux autres isoformes, toutefois inactives, ont aussi été identifiées. Les séquences de ces deux protéines contiennent 479 et 427 acides aminés. Il a été démontré que la première forme inactive entraînait une diminution de l'activité de la PLA<sub>2</sub> VIA-2 lorsqu'elles étaient cotransfectées. Par conséquent, il y aurait formation d'hétéro-oligomères et inhibition de l'isoforme active par l'inactive (Larsson *et al.*, 1998). Ces formes inactives, de par les règles de la nouvelle classification, ne figurent dans aucun des groupes de PLA<sub>2</sub>, bien que plusieurs les nomment respectivement groupes VIA-ankyrine-1 et VIA-ankyrine-2.

Plusieurs expériences ont démontré que les PLA<sub>2</sub> du groupe VIA jouent un rôle dans le remodelage phospholipidique et l'homéostasie par la production de lysophospholipides (Balsinde *et al.*, 1997; Winstead *et al.*, 2000). En plus de ce rôle, plusieurs rapports ont signalé l'implication des PLA<sub>2</sub> du groupe VIA dans les processus de signaux de transduction, de même que dans d'autres processus physiologiques (Winstead *et al.*, 2000).

### **1.2.3.2 Le groupe VIB**

Cette deuxième PLA<sub>2</sub> du groupe VI compte 437 acides aminés et présente 25% d'identité avec la PLA<sub>2</sub> VIA-1. L'ARNm de cette protéine a été détecté au niveau du cœur, du placenta, des reins, du foie, du cerveau et des muscles squelettiques (Six et Dennis, 2000). Son poids moléculaire calculé est de 88 kDa. Il s'agit aussi d'une PLA<sub>2</sub>

à sérine catalytique et sa spécificité pour l'hydrolyse des acides gras en position *sn*-2 a été vérifiée (Mancuso *et al.*, 2000).

Certaines autres protéines qui présentent une activité PLA<sub>2</sub> (*sn*-2) pourraient être classifiées dans l'un ou l'autre des sous-groupes du groupe VI. En effet, les patatines, des enzymes provenant, entre autres, de la pomme de terre et du concombre présentent plusieurs éléments, dont la séquence consensus lipasique (Gly-Thr-Ser-Thr-Gly), qui leur permettraient de faire partie de la classification actuelle.

### 1.2.3.3 Le groupe VIIA

La PLA<sub>2</sub> du groupe VIIA est en fait l'acétylhydrolase du facteur d'activation plaquettaire (PAF-AH). La PAF a été clonée pour la première fois en 1995 (Tjoelker *et al.*, 1995). Il s'agit d'une protéine de 45 kDa et de 441 acides aminés présentant le motif consensus des lipases : Gly-X-Ser273-X-Gly. Elle est retrouvée dans le flux sanguin de la plupart des animaux et, chez les humains, est associée à l'apolipoprotéine B100 (apoB100) des lipoprotéines de faible densité (LDL) et de haute densité (HDL) (Stafforini *et al.*, 1987; Stafforini *et al.*, 1999). La présence de phospholipides oxydés au niveau des LDL est associée à des conditions pathologiques telles que l'athérosclérose. Par conséquent, la localisation de cette PLA<sub>2</sub> permet à cette dernière de cliver la chaîne acyle oxydée et ainsi de s'avérer protectrice pour l'organisme (Parthasarathy *et al.*, 1985).

### 1.2.3.4 Le groupe VIIB

La PLA<sub>2</sub> du groupe VIIB est une enzyme intracellulaire qui a été purifiée et caractérisée pour la première fois à partir d'un cerveau bovin (Hattori *et al.*, 1995a). Cette enzyme présente une identité importante avec le groupe VIIA; elle contient le motif lipasique Gly-X-Ser-X-Gly et est aussi un monomère d'environ 40 kDa et compte 392 acides aminés (Hattori *et al.*, 1996). De la même façon que le groupe précédent, ce groupe est capable d'hydrolyser les chaînes acyles *sn*-2 contenant deux ou plusieurs carbones. Il n'est pas encore précisé si ce groupe d'enzymes, comme le précédent, ne peut seulement

qu'hydrolyser des substrats monomériques (Min *et al.*, 1999). Les PLA<sub>2</sub> de ce groupe s'avèrent fortement exprimées dans le foie et le rein et, de façon moins importante, dans les autres tissus. Enfin, des études suggèrent que les PLA<sub>2</sub> de ce groupe, alors que le groupe précédent était impliqué dans la protection contre l'oxydation au niveau plasmatique, joueraient un rôle protecteur contre les dommages oxydatifs au niveau du foie et des reins (Matsuzawa *et al.*, 1997).

#### 1.2.3.5 Les groupes VIIIA et VIIIB

Ces deux PLA<sub>2</sub> sont des enzymes qui ont initialement été clonées à partir de cerveaux bovins. Elles sont toutefois présentes chez différentes espèces et au niveau de divers tissus, à des stades variés du développement (Watanabe *et al.*, 1998). Ces deux protéines de 26 kDa présentent 62% d'identité entre elles et constituent en fait deux sous-unités présentes à l'intérieur de l'hétérotrimère PAF-AH1b (Hattori *et al.*, 1994a; Hattori *et al.*, 1995b; Adachi *et al.*, 1997). Ces trois protéines sont extrêmement bien conservées à travers les espèces, peut-être même les protéines les plus conservées connues jusqu'à présent (Watanabe *et al.*, 1998). La troisième unité de cet hétérotrimère est la protéine LIS-1, la protéine encodée par le gène impliqué dans la lissencéphalie de Miller-Decker, une maladie caractérisée par une migration neurale anormale résultant en une malformation dévastatrice du cortex cérébral (Hattori *et al.*, 1994b; Sapir *et al.*, 1999). LIS-1 est un homologue de la sous-unité bêta de la protéine G trimérique.

La première sous-unité catalytique clonée fut la PLA<sub>2</sub> du groupe VIIIA, ce qui lui avait valu à l'époque, le nom de sous-unité alpha-1. De même, le groupe VIIIB porta le nom de sous-unité alpha-2, suivant l'ordre logique des découvertes. Ces séquences d'approximativement 230 acides aminés contiennent toutes les deux une sérine à l'intérieur d'une pseudo-séquence consensus lipasique GXSXV, maintenant identifiée comme étant la sérine catalytique (Hattori *et al.*, 1995b; Ho *et al.*, 1997).

Enfin, le rôle exact de ce groupe d'enzymes et son activité n'est pas encore clair dans le système neuronal, mais des défauts génétiques associés aux protéines LIS-1/PAF

permettent de croire à l'implication de ces protéines dans divers troubles neuronaux génétiques.

### 1.2.3.6 Résumé comparatif des PLA<sub>2</sub> des groupes VI, VII et VIII

Le tableau 1.2 permet de reconnaître les principales différences entre les caractéristiques distinctes de chacune des PLA<sub>2</sub> des groupes VI, VII et VIII.

**TABLEAU 1.2**  
**Résumé des caractéristiques des PLA<sub>2</sub> des groupes VI, VII et VIII à sérine catalytique<sup>a</sup>**

Phospholipases A2 des groupes VI, VII et VIII						
Groupe	Sous-groupe	Sources	Nom commun	Taille (kDa)	Caractéristiques	Localisation chromosomique
VI	A-1	Macrophages P388D1, CHO	iPLA <sub>2</sub> , iPLA <sub>2</sub> -A	84-85	Forme courte épissée, 8 répétitions ankyrine	22q13.1 (Pickard <i>et al.</i> , 1999)
	A-2	Lymphocytes-B humains, testicules	iPLA <sub>2</sub> -B	88-90	Forme longue épissée, 7 répétitions ankyrine	22q13.1 (Pickard <i>et al.</i> , 1999)
	B	Cœur et muscle squelettique humains	iPLA <sub>2</sub> γ, iPLA <sub>2</sub> -2	88	Liée à la membrane	7q31 (Mancuso <i>et al.</i> , 2000)
VII	A	Plasma humain / bœuf / souris / porc	PAF-AH	45	Sécrétée, α/β hydrolase, triade ser/his/asp pour VIIA et B	ND
	B	Humain / bœuf / foie / rein	PAF-AH (II)	40	Intracellulaire, myristoylée	ND
VIII	A	Cerveau humain	PAF-AHb α <sub>1</sub>	26	Intracellulaire, arrangement protéine G, triade ser/his/asp, dimère	ND
	B	Cerveau humain	PAF-AHb α <sub>2</sub>	26	Intracellulaire, prot.G, triade ser/his/asp, active en hétéro/ homodimère	11q23 (Lecoite <i>et al.</i> , 1999)

Adapté de Six et Dennis (2000)

<sup>a</sup> Grosses enzymes, typiquement intracellulaires, qui utilisent une sérine nucléophile pour le clivage hydrolytique mais ne présentant pas de ponts disulfure et ne requérant pas de calcium pour la catalyse.



### 1.2.4 Les phospholipases A2 cytosoliques (groupe IV)

C'est en 1986 que les premières identifications et caractérisations de PLA<sub>2</sub> cytosoliques ont été rapportées chez les neutrophiles et les plaquettes humaines (Alonso *et al.*, 1986; Kramer *et al.*, 1986). Ce n'est qu'en 1991 que l'équipe de Sharp (Sharp *et al.*, 1991) séquença entièrement cette enzyme qui fut nommée PLA<sub>2</sub> cytosolique, selon la nomenclature de l'époque. Ces PLA<sub>2</sub> ne présentent aucune homologie de séquence avec les PLA<sub>2</sub> connues et constituent actuellement le groupe IVA de l'ensemble des PLA<sub>2</sub> connues dans la nomenclature. Ce groupe IVA s'explique par l'existence de deux autres PLA<sub>2</sub> cytosoliques (IVB et IVC) qui s'ajoutent dorénavant au groupe IV.

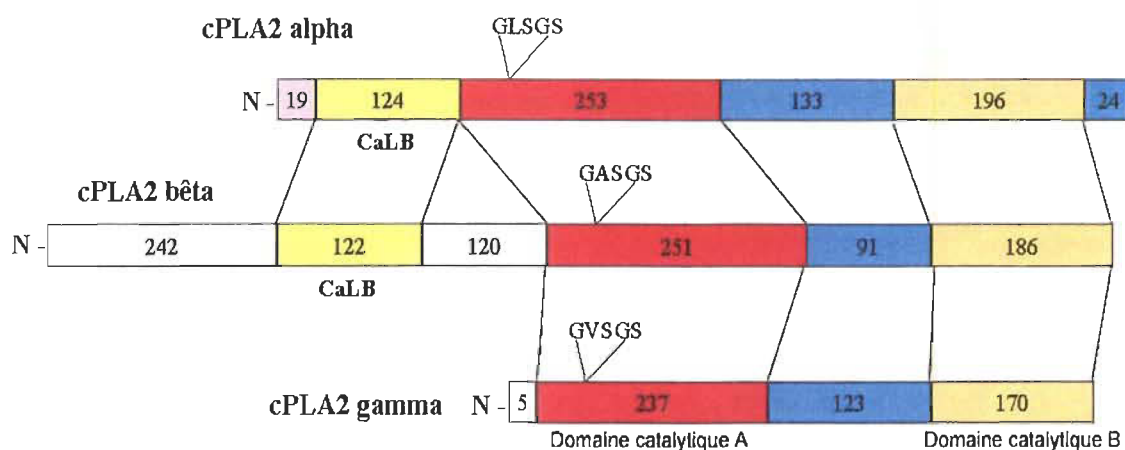
Les trois cPLA<sub>2</sub> sont souvent appelées cPLA<sub>2</sub> alpha, bêta et gamma, ce qui correspond en fait aux groupes IVA, IVB et IVC de la nomenclature utilisée actuellement.

#### 1.2.4.1 Isoformes des cPLA<sub>2</sub>

Les trois isoformes (alpha, bêta et gamma) contiennent toutes deux domaines catalytiques homologues, le domaine A et le domaine B, séparés par un fragment de séquence propre à chacune des isoformes (Hirabayashi et Shimizu, 2000). La séquence consensus lipasique, GXSGS, est localisée dans la région N-terminale du domaine catalytique A de chacune des isoformes. Le domaine catalytique A comprend 253, 251 et 237 acides aminés pour les cPLA<sub>2</sub> alpha, bêta et gamma, respectivement. De même, le domaine catalytique B contient 196, 186 et 170 acides aminés suivant le même ordre. La région comprise entre les deux domaines catalytiques comprend 91, 123 et 133 acides aminés, pour les cPLA<sub>2</sub> alpha, bêta et gamma respectivement, et ne présente que peu d'identité entre les trois isoformes. Les acides aminés impliqués dans la triade catalytique, qui sera abordée plus loin, se retrouvent tous dans les deux domaines catalytiques conservés, alors que les sérines connues pour être impliquées dans le processus de la phosphorylation se retrouvent dans les régions non conservées en C-terminal (Song *et al.*, 1999).



Les cPLA<sub>2</sub> alpha et bêta possèdent le domaine CaLB en N-terminal, alors que la cPLA<sub>2</sub> gamma en est dépourvue, expliquant ainsi son indépendance au calcium. La cPLA<sub>2</sub> bêta présente un domaine additionnel de chaque côté du domaine CaLB, comparativement aux deux autres isoformes (figure 1.13). Ainsi, un insert de 120 acides aminés se trouve entre le domaine catalytique A et le domaine C2 et 242 acides aminés constituent le domaine N-terminal de la cPLA<sub>2</sub> bêta, ce domaine qui ne compte que 19 et 5 acides aminés chez les cPLA<sub>2</sub> alpha et gamma, respectivement (Song *et al.*, 1999).



**Figure 1.13** – Comparaison structurale entre les trois isoformes de cPLA<sub>2</sub> connues dans la littérature jusqu'à présent. La cPLA<sub>2</sub> bêta est l'isoforme la plus longue, avec deux domaines supplémentaires, alors que la cPLA<sub>2</sub> gamma se reconnaît à sa plus petite taille, vu l'absence du domaine de liaison au calcium (CaLB) en N-terminal. Le domaine catalytique A apparaît en rouge, alors que le domaine B apparaît en orange.

La cPLA<sub>2</sub> alpha présente un motif PXSP pour son activation via les MAP kinases et la sérine-505 de ce motif est un site de phosphorylation. Les cPLA<sub>2</sub> bêta et gamma sont dépourvues de ce motif, suggérant que le mécanisme de phosphorylation connu chez la cPLA<sub>2</sub> alpha – qui sera abordé plus loin – ne s'observe pas chez ces dernières. Il a donc été conclu que s'il y avait des sites de phosphorylation pour les cPLA<sub>2</sub> bêta et gamma. Ils sont probablement spécifiques à chaque enzyme, et non pas une caractéristique conservée à l'intérieur du groupe (Pickard *et al.*, 1999).

Enfin, les cPLA<sub>2</sub> présentent une dernière région homologue qui est encore à l'étude actuellement. Il s'agit de la région se trouvant entre les acides aminés 350 et 400 chez la

cPLA<sub>2</sub> alpha et n'est pas présentée à la figure 1.13. Cette région présente une grande similarité avec la séquence de la phospholipase C delta-1 au niveau du site de liaison au PIP2 (phosphatidylinositol 4,5-biphosphate). Ce domaine est connu sous le nom de domaine PH (pleckstrin homology) chez la phospholipase C (Mosior *et al.*, 1998).

#### 1.2.4.2 Domaine CaLB

La cPLA<sub>2</sub> alpha présente une homologie de séquence avec d'autres protéines possédant un mécanisme de translocation calcium-dépendant. Ce domaine est abrégé CaLB pour *Ca<sup>2+</sup>-dependent phospholipid-binding*. On le nomme aussi le domaine C2, pour *constant region 2*, la région 2 faisant référence à la protéine kinase C qui contient aussi ce domaine particulier. Chez la cPLA<sub>2</sub> alpha, ce domaine particulier s'étend des acides aminés 20 à 143 (Song *et al.*, 1999).

L'augmentation intracellulaire des concentrations de l'ordre du micromolaire en calcium entraîne la translocation de la cPLA<sub>2</sub> alpha vers les membranes intracellulaires et les vésicules lipidiques (Channon et Leslie, 1990; Clark *et al.*, 1991; Nalefski *et al.*, 1994). La cPLA<sub>2</sub> alpha transloque vers les membranes périnucléaires et le réticulum endoplasmique, présentant alors son domaine catalytique calcium-indépendant au substrat séquestré au niveau de la bicouche (Hirabayashi *et al.*, 1999; Evans *et al.*, 2001).

Enfin, une étude assez récente explique la localisation périnucléaire de la cPLA<sub>2</sub> alpha par une interaction spécifique entre le domaine CaLB et la vimentine, un filament intermédiaire du cytosquelette qui est aussi localisé dans la région périnucléaire. L'interaction entre la cPLA<sub>2</sub> alpha et la vimentine dépend de trois éléments : du calcium, de la moitié N-terminale du domaine CaLB de la cPLA<sub>2</sub> alpha, contenant les quatre premières structures en feuillets bêta, et du domaine de la tête de la vimentine (Nakatani *et al.*, 2000). Par conséquent, il est spéculé que la cPLA<sub>2</sub> alpha pourrait lier simultanément la vimentine et les phospholipides membranaires, d'une façon calcium-dépendante.

#### 1.2.4.3 La cPLA<sub>2</sub> alpha (groupe IVA)

La cPLA<sub>2</sub> alpha est l'enzyme la plus étudiée de son groupe pour avoir été découverte la première. La cPLA<sub>2</sub> alpha est une protéine monomérique d'approximativement 85 kDa. L'ADNc de la cPLA<sub>2</sub> alpha encode une protéine de 749 acides aminés qui migre à 100-110 kDa sur un gel d'électrophorèse de SDS-polyacrylamide (poids apparent) (Clark *et al.*, 1990; Clark *et al.*, 1991).

Le gène de la cPLA<sub>2</sub> alpha a été localisé sur le chromosome 1q25 (Kramer et Sharp, 1997a). Il possède 18 exons et l'ADNc total comprend 2880 nucléotides, incluant environ 200 nucléotides pour la région 5' non traduite et environ 500 nucléotides pour la région 3' non traduite. La région promotrice est dépourvue de la boîte TATA mais se distingue toutefois des gènes « housekeeping » sans boîte TATA par le fait qu'elle ne contient pas de sites de liaison pour SP1 (Tay *et al.*, 1994b; Wu *et al.*, 1994a; Miyashita *et al.*, 1995).

#### 1.2.4.4 La cPLA<sub>2</sub> bêta (groupe IVB)

La cPLA<sub>2</sub> bêta a été découverte suite à l'examen de différentes bases de données d'ADN. Elle a été clonée pour la première fois en 1999 par l'équipe de Pickard. Il s'agit d'une protéine mature de 1012 acides aminés présentant une identité d'approximativement 30% avec la cPLA<sub>2</sub> alpha (Pickard *et al.*, 1999). Elle a un poids moléculaire d'environ 114 kDa. Elle possède une grande identité avec la cPLA<sub>2</sub> alpha dans la région des acides aminés catalytiques Ser228 et Asp549, mais ne présente toutefois pas les quatre sérines connues pour être impliquées dans la phosphorylation de la cPLA<sub>2</sub> alpha tel que mentionné précédemment (Pickard *et al.*, 1999; Song *et al.*, 1999). De plus, sa séquence protéique contient un motif histidine caractéristique du centre catalytique des protéases caspases de la cascade apoptotique, mais aucune région caractéristique de la cystéine catalytique propre à ces enzymes (Pickard *et al.*, 1999).

Le gène de la cPLA<sub>2</sub> bêta a été localisé sur le chromosome 15 (Pickard *et al.*, 1999), plus

précisément entre 15q11.2 et 15q21.3, selon les bases de données du *National Center for Biotechnology Information* (NCBI), dans une région près du gène de la phosphoinositol biposphate phosphatase. Enfin, l'ARNm de la cPLA<sub>2</sub> bêta est épissé de façon très faible; des *Northern blot* dans 14 tissus ont montré exclusivement la forme non épissée (8,5 kb) plutôt que la forme épissée (3 kb) (Pickard *et al.*, 1999).

#### 1.2.4.5 La cPLA<sub>2</sub> gamma (groupe IVC)

La cPLA<sub>2</sub> gamma a été clonée pour la première fois en 1998 dans une banque d'ADNc du muscle squelettique humain (Underwood *et al.*, 1998). Sa séquence encode une protéine de 541 acides aminés et son poids moléculaire calculé est d'approximativement 61 kDa. Elle présente une identité légèrement inférieure à 30% avec la cPLA<sub>2</sub> alpha. Les acides aminés conservés de la triade catalytique chez cette cPLA<sub>2</sub> sont la sérine 82, l'aspartate 385 et l'arginine 54. Le gène de la cPLA<sub>2</sub> gamma est localisé sur le chromosome 19 près du gène de la calmoduline. Le transcrit a une longueur d'environ 3 kb.

La cPLA<sub>2</sub> gamma présente un motif d'isoprénnylation (-CCLA) ou boîte -CAAX à l'extrémité carboxy-terminale et un site potentiel de myristoylation (MGSSEV-) en N-terminal (Hirabayashi et Shimizu, 2000). Le motif -CCLA a été identifié en tant que signal pour la prénylation, où le C est la cystéine modifiée, A, un acide aminé aliphatique et X, un acide aminé quelconque (Clarke, 1992). Le site potentiel de la myristoylation pourrait réguler la localisation de la cPLA<sub>2</sub> gamma à l'intérieur de la cellule (Underwood *et al.*, 1998).

#### 1.2.4.6 Résumé récapitulatif des trois cPLA<sub>2</sub>

Le tableau 1.3 résume les principales différentes caractéristiques des trois cPLA<sub>2</sub> connues jusqu'à présent.

Pour la sous-section suivante et celles qui suivront dans la sous-section 1.2.4, l'ajout de

la « cPLA<sub>2</sub> alpha » dans le titre indique simplement que les études n'ont été réalisées que sur cette cPLA<sub>2</sub> alpha et qu'il n'y a encore rien de publié dans la littérature concernant les cPLA<sub>2</sub> bêta et gamma.

**TABLEAU 1.3**  
**Résumé des caractéristiques des cPLA<sub>2</sub>**

Phospholipases A2 cytosoliques							
Groupe	Sous-groupe	Sources	Nom commun	Taille (kDa)	Effet du Ca <sup>++</sup>	Caractéristiques	Localisation chromosomique
IV	A	U937 <sup>a</sup> humain, plaquettes RAW 264.7 <sup>b</sup> , rein de rat	cPLA <sub>2</sub> α	85	<μM ; translocation membranaire	Domaine CaLB, α/β hydrolase, phosphorylation	1q25 (Tay <i>et al.</i> , 1995)
	B	Pancréas, cœur, foie et cerveau humains	cPLA <sub>2</sub> β	114	<μM ; translocation membranaire	Domaine CaLB, α/β hydrolase	15q14 (Pickard <i>et al.</i> , 1999)
	C	Cœur, muscle squelettique humains	cPLA <sub>2</sub> γ	61	Aucun	α/β hydrolase	19 (Pickard <i>et al.</i> , 1999)

Adapté de Six et Dennis (2000)

<sup>a</sup>U937 est une lignée de cellules leucémiques monocytaires humaines.

<sup>b</sup>RAW 264.7 est une lignée cellulaire de macrophages provenant de la souris.

#### 1.2.4.7 Régulation transcriptionnelle de la cPLA<sub>2</sub> alpha

Des stimuli extracellulaires peuvent altérer les niveaux d'ARN messager (ARNm) et de la protéine cPLA<sub>2</sub> alpha. Une variété de cytokines et des mitogènes tels que l'interleukine-1 (Lin *et al.*, 1992a), le facteur alpha de nécrose tumorale (TNF-α) (Wu *et al.*, 1996), le facteur de croissance épidermique (EGF) (Clark *et al.*, 1995), le ligand c-Kit (Murakami *et al.*, 1995) et l'interféron gamma (INF-γ) (Wu *et al.*, 1994b) sont maintenant connus pour induire l'activation et l'augmentation de la synthèse de cPLA<sub>2</sub> alpha dans divers modèles cellulaires. La stimulation de la synthèse de la cPLA<sub>2</sub> alpha survient en terme d'heures et résulte en une prolongation de la libération d'acide arachidonique et de la production d'eicosanoïdes. Enfin, l'interleukine-1β entraîne une augmentation des niveaux d'expression de la cPLA<sub>2</sub> alpha (Lin *et al.*, 1992a; Schalkwijk *et al.*, 1993).

La région promotrice du gène de la cPLA<sub>2</sub> alpha contient plusieurs sites de liaison potentiels pour AP-1, AP-2, NF-κB, NF-IL-6, PEA3, OCT, C/CEBP et le *glucocorticoid response element*. La région 5' flanquante possède une séquence polypyrimidine de 27 paires de bases responsable du faible niveau constitutif d'expression (Miyashita *et al.*, 1995) et une répétition de CA de 48 paires de bases qui semble conférer un effet inhibiteur sur la transcription (Wu *et al.*, 1994a). De plus, la régulation post-transcriptionnelle de la synthèse de la cPLA<sub>2</sub> alpha a été établie. Par exemple, de multiples séquences AUUUA dans la région 3' non codante de l'ARNm de la cPLA<sub>2</sub> alpha semblent être responsables de l'instabilité de l'ARNm, mais la stimulation à l'aide d'agents mutagènes stabilise l'ARNm endogène de la cPLA<sub>2</sub> alpha chez les cellules mésangiales glomérulaires du rat (Tay *et al.*, 1994a).

#### 1.2.4.8 Localisation cellulaire, intracellulaire et tissulaire

La cPLA<sub>2</sub> alpha est présente dans le cytosol de différentes cellules et tissus. Au niveau des cellules, notons les plaquettes, les macrophages, les neutrophiles, les cellules endothéliales, les cellules mésangiales rénales, les cellules vasculaires des muscles lisses, les kératinocytes, les monoblastes, les cellules épithéliales alvéolaires et la lignée des cellules macrophagiques. Le gène de la cPLA<sub>2</sub> alpha est largement exprimé, en particulier au niveau du cerveau, des poumons, du cœur, de la rate, du placenta et des reins (Kramer et Sharp, 1997a; Pickard *et al.*, 1999).

Des analyses par *Northern* montrent que la cPLA<sub>2</sub> bêta est exprimée ubiquitairement, mais plus fortement au niveau du pancréas, du cerveau, de cœur et du foie, alors que la cPLA<sub>2</sub> gamma est sélectivement exprimée au niveau du muscle squelettique et du cœur et, à un plus faible niveau, dans le cerveau, la rate, le placenta et le pancréas (Underwood *et al.*, 1998; Pickard *et al.*, 1999; Song *et al.*, 1999).

La localisation intracellulaire de la cPLA<sub>2</sub> alpha a été l'objet de nombreuses études afin de déterminer la membrane intracellulaire à laquelle l'enzyme se lie. Les études suggèrent que la cPLA<sub>2</sub> alpha transloque du cytosol vers la membrane nucléaire (Glover

*et al.*, 1995) ou vers la membrane nucléaire et la membrane du réticulum endoplasmique (Schievella *et al.*, 1995) ou encore, vers l'espace périnucléaire et la membrane nucléaire (Sierra-Honigmann *et al.*, 1996). De même, des études ont démontré que la cPLA<sub>2</sub> alpha pouvait aussi être transloquée au niveau de la membrane plasmique (Sierra-Honigmann *et al.*, 1996) ou être localisée ponctuellement dans des corps lipidiques cytoplasmiques dans certaines cellules (Yu *et al.*, 1998). Ces études étant immunohistochimiques, il est possible qu'il y ait eu une perte de signal causée soit par la fixation, la perméabilisation ou la coloration des cellules. C'est pourquoi l'équipe de Hirabayashi et ses collaborateurs ont procédé à un marquage plus direct à l'aide de la *green fluorescent protein* (GFP) sur des cellules vivantes. Les conclusions de leur étude indiquent que la cPLA<sub>2</sub> alpha transloque de façon restreinte au niveau de la région périnucléaire et de l'enveloppe nucléaire (Hirabayashi et Shimizu, 2000). Au cours de cette même étude, la localisation intracellulaire de la cPLA<sub>2</sub> gamma a aussi été déterminée. Ainsi, le marquage de la cPLA<sub>2</sub> gamma est apparu en fluorescence sous forme de structures ponctiformes cytoplasmiques, lesquelles apparaissent être des peroxysomes. En ce qui a trait à la localisation intracellulaire de la cPLA<sub>2</sub> bêta, il semble qu'aucune expérience n'a été tentée jusqu'à maintenant.

Une étude encore plus récente arrive à la conclusion que la translocation de la cPLA<sub>2</sub> alpha avait lieu non seulement au niveau du réticulum endoplasmique (RE) et des membranes périnucléaires (MP), mais aussi au niveau de l'appareil de Golgi (Evans *et al.*, 2001), tel que l'avaient déjà suggéré quelques travaux (Perisic *et al.*, 1999; Choukroun *et al.*, 2000). Ainsi, l'équipe d'Evans (Evans *et al.*, 2001) démontre que la cPLA<sub>2</sub> alpha est préférentiellement transloquée vers l'appareil de Golgi en réponse à des variations physiologiques des concentrations en calcium de courte durée. Aussi, leurs expériences les amènent à la conclusion que le ciblage des membranes est déterminé par le domaine C2 et que la translocation vers l'appareil de Golgi ou le RE et les MP est régulée par l'amplitude de la concentration calcique intracellulaire.

#### **1.2.4.8.1 Localisation stratégique pour la biosynthèse des eicosanoïdes**

L'acide arachidonique est métabolisé pour produire divers eicosanoïdes, incluant les



prostaglandines et les thromboxanes, les leucotriènes et les lipoxines, ou les acides époxyeicosatriénoïques via, respectivement, la cyclo-oxygénase, la lipoxygénase et le cytochrome P450, selon le type cellulaire (Shimizu et Wolfe, 1990). La majorité des protéines impliquées dans la production d'eicosanoïdes est localisée constitutivement, ou après stimulation, au niveau de l'enveloppe nucléaire et/ou du réticulum endoplasmique, suggérant donc que la cPLA<sub>2</sub> alpha transloquée au niveau de la région périnucléaire se trouve en position idéale pour procurer l'acide arachidonique libre aux enzymes du métabolisme des eicosanoïdes (Hirabayashi et Shimizu, 2000).

#### 1.2.4.9 Activité spécifique envers les phospholipides et leurs acides gras

Les expériences utilisant des vésicules lipidiques ont démontré que la cPLA<sub>2</sub> alpha présente une sélectivité marquée pour les phospholipides contenant l'acide arachidonique en position *sn*-2 (Clark *et al.*, 1990; Gronich *et al.*, 1990; Kramer *et al.*, 1991). Bien que les phospholipides contenant l'acide alpha-linolénique ou l'acide eicosapentanoïque soient aussi de bons substrats pour les cPLA<sub>2</sub> alpha, ils représentent des acides gras polyinsaturés présents en faible quantité au niveau cellulaire, rendant ainsi les phospholipides contenant l'acide arachidonique, les substrats majeurs des cPLA<sub>2</sub> en conditions physiologiques. La cPLA<sub>2</sub> alpha ne montre pas de préférence pour les têtes polaires de glycérophospholipides ou pour la position *sn*-1 (Diez *et al.*, 1992; Diez *et al.*, 1994), bien que le phosphate en position *sn*-3 soit nécessaire pour que la molécule puisse être un substrat potentiel. La cPLA<sub>2</sub> alpha a également une activité lysophospholipase *sn*-1 envers le 1-acyl-lyso-phosphatidylcholine et une faible activité transacylase *in vitro* (Reynolds *et al.*, 1993; de Carvalho *et al.*, 1995; Hanel et Gelb, 1995). Toutefois, l'activité biologique majeure de la cPLA<sub>2</sub> alpha demeure l'hydrolyse de la liaison *sn*-2 puisque les membranes cellulaires naturelles ne contiennent que peu de lysophospholipides (Hirabayashi et Shimizu, 2000).

La caractérisation biochimique des propriétés de la cPLA<sub>2</sub> bêta et la cPLA<sub>2</sub> gamma indique qu'elles ont une régiosélectivité différente de la cPLA<sub>2</sub> alpha. En effet, la cPLA<sub>2</sub> bêta est moins sélective pour le clivage de la liaison *sn*-2 que pour la liaison *sn*-1



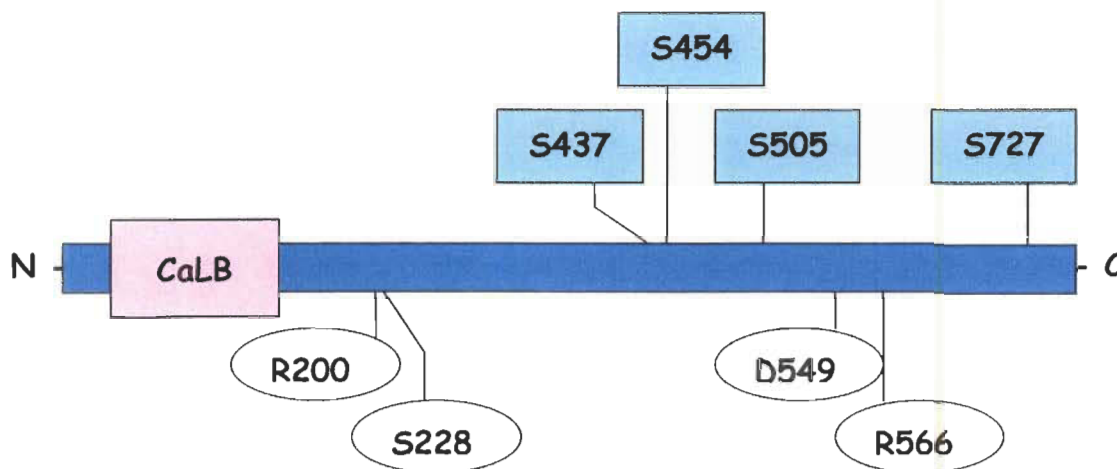
lors des expériences réalisées à partir de lysats cellulaires de cellules COS exprimant cette protéine (Song *et al.*, 1999). Elle semble posséder une activité PLA1 et lysoPLA<sub>2</sub> plus importante que son activité PLA<sub>2</sub> (Song *et al.*, 1999). La cPLA<sub>2</sub> gamma montre une activité PLA<sub>2</sub> et possiblement PLA1, et préfère l'acide arachidonique à la position *sn*-2 deux à trois fois et demi plus que les autres acides gras *in vitro* (Underwood *et al.*, 1998), ce qui est modeste comparativement à la préférence de la cPLA<sub>2</sub> alpha pour ce même acide gras qui est 24,5 fois plus importante que pour l'acide palmitique dans les mêmes conditions expérimentales (Underwood *et al.*, 1998).

#### 1.2.4.10 Mécanisme catalytique de la cPLA<sub>2</sub> alpha

Les lipases possèdent une triade catalytique *sérine-aspartate/glutamate-histidine* bien conservée chez les sérines protéases. Le mécanisme catalytique procède via un intermédiaire sérine-acyle (Hanel et Gelb, 1995). Le motif GLSGS présent chez toutes les PLA<sub>2</sub> du groupe IV (cytosoliques) rappelle le motif GX SXG classique des lipases et arbore la sérine-228 nucléophile (Sharp *et al.*, 1994). En plus de la sérine-228, l'aspartate-549 et l'arginine-200 de la cPLA<sub>2</sub> alpha ont été identifiés par mutagenèse dirigée comme étant les acides aminés de la triade catalytique (figure 1.14) (Pickard *et al.*, 1996). De même, Rhee et Bae, en 1997, ont démontré que l'arginine-566 était elle aussi essentielle à la catalyse (Rhee et Bae, 1997). Par contre, aucune des 19 histidines ne s'est avérée essentielle au processus catalytique.

Dans le mécanisme catalytique proposé, l'aspartate-549 agit en tant que base activant directement l'acide aminé Ser228. Ainsi, suite à une augmentation du calcium intracellulaire et de la translocation subséquente de la cPLA<sub>2</sub> alpha à la membrane, une molécule individuelle de substrat peut se lier au site actif, de telle sorte que son lien ester *sn*-2 est à proximité de l'acide aminé Ser228. Le phosphate de la tête phospholipidique est stabilisé par le guanidinium de l'arginine-200. Suivant la formation du complexe substrat-enzyme, l'aspartate-549 attaque le lien *sn*-2 ester et capture un proton, activant ainsi la sérine-228. L'aspartate-549 transfère un proton à la molécule phospholipidique, entraînant la transformation de cette molécule intermédiaire en une molécule sérine-acyl.

L'hydrolyse de cet intermédiaire par l'eau se produit alors via un mécanisme analogue. Par la suite, la cPLA<sub>2</sub> alpha peut soit se dissocier de l'interface de la bicouche ou lier un autre substrat phospholipidique, répétant ainsi le cycle (Dessen, 2000).



**Figure 1.14** - Schéma de la cPLA<sub>2</sub> alpha montrant les acides aminés les plus importants pour son activité. Les différents acides aminés impliqués dans la phosphorylation apparaissent dans les carrés verts et les acides aminés impliqués dans l'activité catalytique apparaissent dans les cercles oranges. Le domaine de liaison du calcium (CaLB) apparaît en rose, dans la région N-terminale. Image tirée de (Leslie, 1997).

En ce qui a trait au mécanisme catalytique des cPLA<sub>2</sub> bêta et gamma, les détails sont encore inexistant dans la littérature, ce qui est probablement attribuable au fait que les démarches expérimentales n'ont pas encore été entreprises à ce niveau.

Enfin, la cristallisation de la cPLA<sub>2</sub> alpha a permis de déterminer quels acides aminés étaient en contact avec les ions calcium (Gijon et Leslie, 1999). Ainsi, ce sont les acides aminés D40, T41, D43, N65, D93, A94 et N95 qui interagissent avec deux ions calcium.

#### 1.2.4.11 La phosphorylation de la cPLA<sub>2</sub> alpha

La stimulation des cellules par divers agonistes qui induisent une libération d'acide arachidonique promouvoit aussi la phosphorylation de la cPLA<sub>2</sub> alpha (Lin *et al.*, 1992b; Clark *et al.*, 1995). Il a été démontré que la cPLA<sub>2</sub> alpha est principalement phosphorylée sur les acides aminés Ser437, Ser454, Ser505 et Ser727 (figure 1.14) après stimulation à l'aide d'agonistes (de Carvalho *et al.*, 1996). Par contre, dans certaines

conditions, ce ne sont que les acides aminés Ser727 et Ser505 qui sont phosphorylés (Borsch-Haubold *et al.*, 1998). Le rôle de la phosphorylation de la Ser437 et de la Ser454 n'est pas encore clair puisque ces acides aminés ne sont pas conservés dans la cPLA<sub>2</sub> alpha chez différentes espèces (humain, souris, rat, poulet, poisson-zèbre). Par contre les acides aminés Ser505 et Ser727 sont conservés parmi des espèces d'évolution distincte (Hirabayashi et Shimizu, 2000).

La sérine-505 est phosphorylée par la *mitogen-activated protein kinase* (MAPK) (Lin *et al.*, 1993) et la sérine-727 est probablement phosphorylée par la *MAP kinase interacting kinase-1* (MNK1) (Hefner *et al.*, 2000).

La sérine-727 se trouve dans le motif consensus pour une kinase basotropique et peut être phosphorylée par la protéine kinase C et la protéine kinase A (Leslie, 1997). Toutefois, cette phosphorylation ne conduit pas à une augmentation significative de l'activité de la cPLA<sub>2</sub> (Nemenoff *et al.*, 1993). Puisque l'activation de la protéine kinase C peut déclencher la cascade des kinases et conduire à l'activation de la MAPK, l'activation de la cPLA<sub>2</sub> alpha semble être médiée indirectement par la famille de la MAPK (Qiu et Leslie, 1994). Une expérience où un inhibiteur de la p38 MAPK s'est avéré capable d'inhiber non seulement la phosphorylation de la Ser505, mais aussi celle de la Ser727, suggère que la Ser727 est phosphorylée par une kinase se trouvant plus loin dans la cascade des kinases que la MAPK (Borsch-Haubold *et al.*, 1998). Ce résultat concorde donc avec l'hypothèse de Hefner *et al.* (2000) qui suggère que la Ser727 soit phosphorylée par la MNK1.

La phosphorylation de la sérine-505 de la cPLA<sub>2</sub> alpha augmente l'activité intrinsèque de l'enzyme (Lin *et al.*, 1993) et la libération d'acide arachidonique (Lin *et al.*, 1992b). L'acide aminé Ser505 se trouve dans le motif PXSP qui constitue une séquence consensus pour le site de phosphorylation par la MAPK (Hirabayashi et Shimizu, 2000). Les études portant sur cette kinase via des cellules en culture et *in vitro* indiquent que la cPLA<sub>2</sub> alpha est un substrat pour la MAPK. L'activation maximale de la cPLA<sub>2</sub> alpha requiert une phosphorylation soutenue de la sérine-505 par un membre de la famille des

MAPK et joue un rôle important dans la libération d'acide arachidonique dans les cellules stimulées (Lin *et al.*, 1993; Qiu *et al.*, 1998). En outre, d'autres membres de la famille des MAPK, tels que la p38 MAPK, peuvent phosphoryler la sérine-505 de la cPLA<sub>2</sub> alpha chez les plaquettes stimulées par la thrombine (Kramer *et al.*, 1996).

Il est à noter que la phosphorylation de l'acide aminé Ser505 n'est pas suffisante pour l'activation de la cPLA<sub>2</sub> alpha et la libération d'acide arachidonique en absence d'une augmentation de la concentration calcique intracellulaire (Qiu *et al.*, 1998; Gijon *et al.*, 1999), puisque cet ion est nécessaire pour la liaison au substrat membranaire. L'activation optimale de la cPLA<sub>2</sub> nécessite la phosphorylation de l'enzyme et une augmentation de la concentration intracellulaire calcique, c'est-à-dire que la phosphorylation et l'augmentation calcique doivent agir de concert afin d'assurer l'activation optimale de la cPLA<sub>2</sub> menant à la libération d'acide arachidonique.

#### 1.2.4.12 Activation de la cPLA<sub>2</sub> alpha et son interaction avec certains éléments

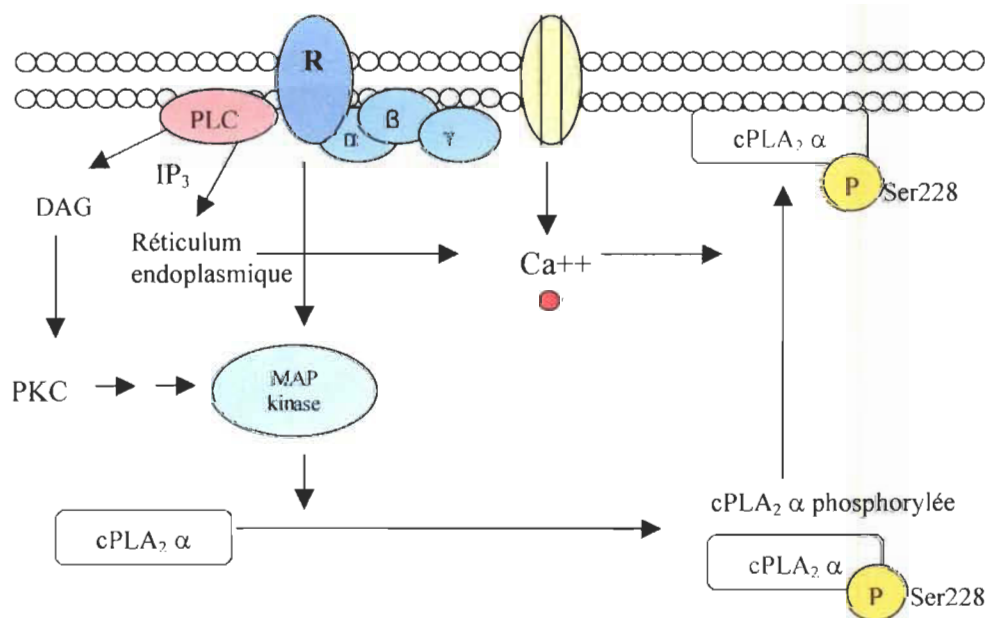
La cPLA<sub>2</sub> alpha est activée en réponse à divers stimuli tels que les cytokines, les hormones, les neurotransmetteurs, les agents mitogènes, les antigènes, les endotoxines et les constituants de la matrice extracellulaire (pour une revue, voir Hirabayashi, 2000). Certains stimuli physiques et générateurs de stress incluant l'oxydation, l'hyperglycémie et les rayons UV peuvent aussi induire l'activation de l'enzyme (Kramer et Sharp, 1997b). Enfin, l'activation est médiée par plusieurs agonistes et signaux intracellulaires distincts impliquant une protéine G (Winitz *et al.*, 1994; Burke *et al.*, 1997), une augmentation de la concentration calcique intracellulaire et l'activation de kinases telles que les *mitogen activated protein kinase* (MAPK) et la *protéine kinase C* (PKC), tel que mentionné plus tôt. L'intensité, la durée et la coïncidence coopérative des signaux semblent déterminer l'activation immédiate de la cPLA<sub>2</sub> alpha (Qiu *et al.*, 1998; Hirabayashi *et al.*, 1999).

La cPLA<sub>2</sub> alpha est donc particulièrement intéressante en ce qui concerne son habileté à interagir directement avec différents éléments impliqués dans les signaux de

cellules activées, l'inhibition de la synthèse des PIP2s corrèle avec une diminution de la libération d'acide arachidonique. Enfin, une augmentation des niveaux de PIP2 dans les cellules au repos est suffisante pour activer la cPLA<sub>2</sub> alpha, résultant en une augmentation de l'acide arachidonique libéré (Balsinde *et al.*, 2000). Il a été démontré que la cPLA<sub>2</sub> alpha se lie avec une grande affinité et spécificité au PIP2 dans un rapport stoechiométrique de 1 : 1 dans les vésicules lipidiques (Mosior *et al.*, 1998). De plus,

37

transduction intracellulaire, notamment, les protéines G et les kinases. En effet, comme il est mentionné un peu plus haut, la cPLA<sub>2</sub> alpha est activée en partie par la MAPK. De même, puisque ces kinases sont régulées par les protéines G (Xing et Mattera, 1992), la cPLA<sub>2</sub> alpha est indirectement modulée par ces dernières. La figure 1.15 présente le modèle d'activation de la cPLA<sub>2</sub> alpha proposé par Lin *et al.* (1993). Dans ce modèle, l'activation de la MAPK peut être dépendante ou non de la PKC, tel que le suggèrent des études sur le sujet (Cobb *et al.*, 1991; Posada et Cooper, 1992).



**Figure 1.15** – Modèle d'activation de la cPLA<sub>2</sub> alpha. Dans ce modèle, l'activation rapide de la cPLA<sub>2</sub> alpha est causée par l'action synergique du calcium et de la phosphorylation de cette cPLA<sub>2</sub> par la MAP kinase. Lorsqu'un ligand se lie au récepteur approprié (R), ce dernier active la phospholipase C (PLC) via soit une voie dépendante ou indépendante d'une protéine G, conduisant à la formation de l'inositol triphosphate (IP<sub>3</sub>) et du diacylglycérol (DAG). L'IP<sub>3</sub> entraîne la mobilisation du calcium intracellulaire. Alternativement, l'augmentation intracellulaire la concentration calcique peut aussi résulter de l'activation de récepteurs couplés aux canaux calciques. L'augmentation de calcium entraîne la translocation de la cPLA<sub>2</sub> alpha du cytosol vers la membrane intracellulaire. Le DAG active la protéine kinase C (PKC), conduisant à l'activation de la MAP kinase. La MAP kinase activée phosphoryle alors la cPLA<sub>2</sub> alpha au niveau de l'acide aminé Ser228. De concert, la phosphorylation et l'augmentation calcique entraînent l'activation optimale de la cPLA<sub>2</sub> alpha. Image tirée de (Lin *et al.*, 1993).

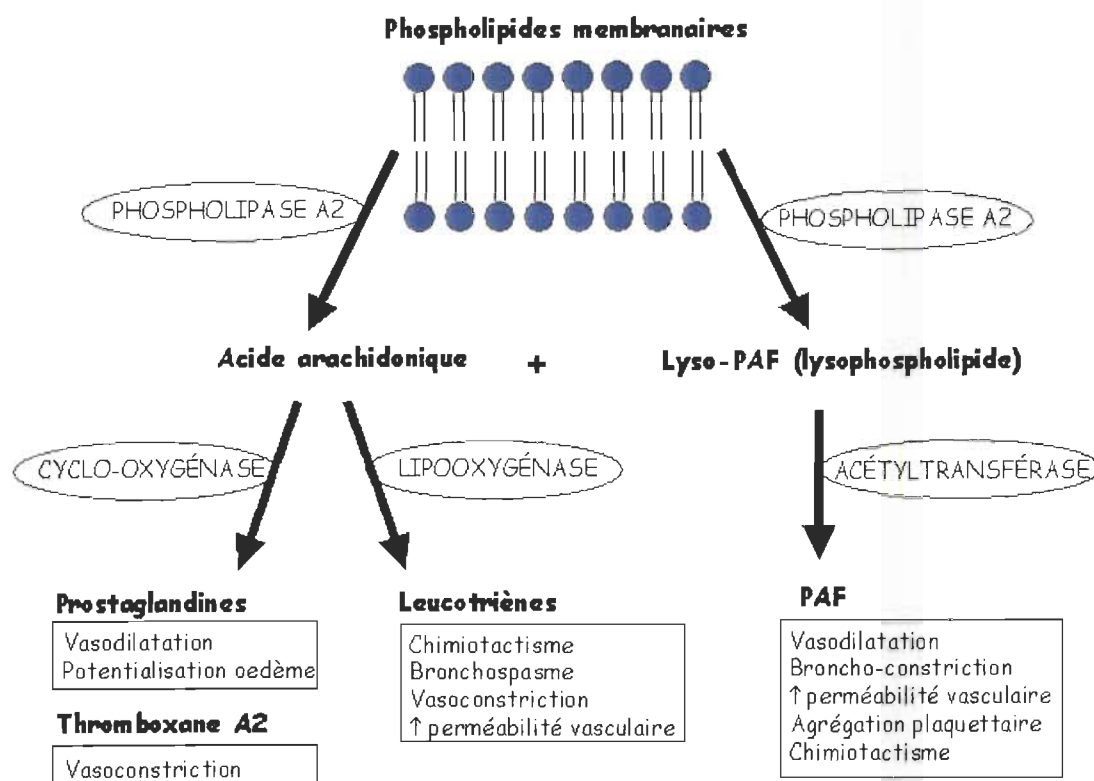
5, W-7, W-12 et W-13 s'avèrent être des inhibiteurs tant pour la cPLA<sub>2</sub> alpha que pour les sPLA<sub>2</sub> (Hope *et al.*, 1993).

#### 1.2.4.14 Implication des PLA<sub>2</sub> dans diverses pathologies

L'action hydrolytique des PLA<sub>2</sub> libère, tel que mentionné plus haut, un acide gras libre et un lysophospholipide. Ces acides gras libérés, plus particulièrement l'acide arachidonique (AA) et l'acide oléique (OA), peuvent constituer des stocks importants d'énergie, mais l'AA peut aussi agir en tant que second messager (Damron *et al.*, 1993; Kapus *et al.*, 1994; Trotti *et al.*, 1995; Sauvadet *et al.*, 1997) et en tant que précurseur des eicosanoïdes (Serhan *et al.*, 1996), lesquels sont des médiateurs potentiels de l'inflammation et impliqués dans les signaux de transduction (figure 1.16) (Shimizu et Wolfe, 1990). En effet, les eicosanoïdes sont impliqués dans différentes fonctions cellulaires telles que la régulation de la circulation sanguine, la migration, la sécrétion, l'apoptose et d'un point de vue plus médical, dans la stimulation de divers processus inflammatoires, allant de l'asthme à l'arthrite (Prescott, 1997).

Le deuxième produit de l'action de la PLA<sub>2</sub>, le lysophospholipide, est impliqué dans la signalisation cellulaire, le remodelage phospholipidique et les perturbations membranaires (Balsinde *et al.*, 1997; Moolenaar *et al.*, 1997). Le facteur d'agrégation plaquettaire (PAF), par exemple, fait partie de cette classe de lysophospholipides ou, plus précisément, d'un produit issu de l'hydrolyse du lysophospholipide (Hanahan, 1986; Venable *et al.*, 1993).

La surproduction de ces médiateurs lipidiques est responsable du phénomène de l'inflammation de même que de désordres tissulaires. Il n'est donc pas étonnant de voir dans la littérature l'implication des PLA<sub>2</sub> dans un grand nombre de situations pathologiques. De ce fait, les PLA<sub>2</sub> constituent une cible très intéressante pour la découverte de nouveaux médicaments; des inhibiteurs spécifiques de PLA<sub>2</sub> pourraient intervenir dans une grande variété de processus physiopathologiques.



**Figure 1.16** – La synthèse des différents dérivés de l'acide arachidonique et du lysophospholipide. Informations tirées du site Internet [http:// www.mbi.ufl.edu/~hnick/ cPLA<sub>2</sub>.htm](http://www.mbi.ufl.edu/~hnick/cPLA2.htm) de la *McKnight Brain Institute* de l'Université de Floride.

À l'aide de souris *knock-out* cPLA<sub>2</sub><sup>-/-</sup>, différents rôles ont été attribués à la cPLA<sub>2</sub> alpha (Sapirstein et Bonventre, 2000). Les résultats indiquent que ces souris *knock-out* sont partiellement protégées des dommages cérébraux causés par diverses conditions pathologiques. De même, l'étude des réponses induites par les allergènes aériens ont montré que la cPLA<sub>2</sub> est un important effecteur de l'hyperréactivité à ce niveau (Sapirstein et Bonventre, 2000).

Enfin, l'expression de la cPLA<sub>2</sub> alpha est élevée dans le poumon, le rein, la rate, le cerveau, la moelle épinière et l'ovaire en réponse à des stimuli pathologiques (Kol *et al.*, 1997; Fukuda *et al.*, 1999). L'ensemble des résultats sur ce sujet met donc en évidence l'implication de la cPLA<sub>2</sub> dans un nombre important de situations pathologiques.

#### 1.2.4.15 L'implication des cPLA<sub>2</sub> pour l'intégrité des photorécepteurs

Tel que décrit plus tôt dans la sous-section 1.1.1.1.1 portant sur les photorécepteurs, les dommages oxydatifs causés par la peroxydation des lipides membranaires sont reconnus pour altérer la structure et la fonction membranaire. Tel que vu précédemment, les acides gras polyinsaturés qui composent une fraction très importante des membranes des disques des photorécepteurs sont particulièrement susceptibles à la peroxydation (van Kuijk et Dratz, 1987; Van Themsche, 1999). De plus, la consommation d'oxygène nécessaire aux réactions métaboliques cellulaires génère la production et la libération de radicaux libres et d'espèces très réactives telles que les radicaux hydroxyle (OH<sup>•</sup>) et superoxyde (O<sub>2</sub><sup>•-</sup>) et le peroxyde d'hydrogène (H<sub>2</sub>O<sub>2</sub>) (Berman, 1991).

Le processus de déacylation-réacylation des phospholipides membranaires constitue un mécanisme cellulaire de protection des membranes contre les dommages causés par la peroxydation des acides gras. Ce mécanisme de déacylation-réacylation serait impliqué dans la réparation des membranes lipidiques peroxydées (van Kuijk et Dratz, 1987). Dans ce modèle, l'hydrolyse des lipides peroxydés par une PLA<sub>2</sub> résulte en l'excision des acides gras peroxydés qui seront ensuite réduits, réparés et réestérifiés. En effet, puisque les acides gras polyinsaturés sont majoritairement acylés à la position *sn*-2 des phospholipides membranaires, les PLA<sub>2</sub> constituent une voie enzymatique de choix pour les hydrolyser, et ainsi contribuer au maintien de l'intégrité membranaire (Van Themsche, 1999), d'autant plus qu'il a été démontré que certaines PLA<sub>2</sub> hydrolysent préférentiellement les lipides peroxydés (Sevanian et Kim, 1985; Sevanian et McLeod, 1987).

En conséquence, puisque les cPLA<sub>2</sub>, une fois activées, se trouvent au niveau des membranes, elles pourraient être impliquées dans un processus de déacylation-réacylation responsable de la protection membranaire contre la peroxydation lipidique au niveau des disques des segments externes des photorécepteurs, c'est-à-dire qu'elles seraient en partie responsables de l'intégrité membranaire des photorécepteurs.

Dans le même ordre d'idées, il est également possible que les cPLA<sub>2</sub> puissent jouer un



rôle dans le processus inflammatoire de l'œil par l'hydrolyse des acides gras peroxydés membranaires.

#### **1.2.4.16 Implication des cPLA<sub>2</sub> dans l'inflammation oculaire**

Tel que détaillé plus tôt, l'acide arachidonique est le précurseur de tous les membres de la famille des eicosanoïdes, d'importants médiateurs du processus de l'inflammation, et est issu de l'hydrolyse d'un phospholipide par la PLA<sub>2</sub>.

Au niveau de l'œil, lorsqu'il y a infection ou blessure, l'EPR n'exerce pas seulement une fonction passive de barrière hémato-oculaire, mais intervient aussi activement dans la mise en place de la réponse immunitaire locale (Marmor et Wolfensberger, 1998). Ainsi, en plus de son activité phagocytaire, l'EPR exprime plusieurs des récepteurs de surface et des médiateurs solubles requis pour la coordination de la réponse immunitaire (pour une revue, voir Marmor et Wolfensberger, 1998). Les médiateurs solubles que sont les eicosanoïdes sont alors produits, suivant le métabolisme de l'acide arachidonique. Par conséquent, de par leur localisation et leur activité intracellulaire, les cPLA<sub>2</sub> sont susceptibles d'être en partie responsables du processus inflammatoire oculaire.

#### **1.2.4.17 Implication des cPLA<sub>2</sub> de l'EPR dans la digestion lysosomiale**

Il a été mentionné précédemment que les disques membranaires de la portion apicale des segments externes des photorécepteurs phagocytés par les cellules de l'EPR sont par la suite soumis à l'action hydrolytique des enzymes lysosomiales de ces cellules. En effet, les lysosomes contiennent près d'une soixantaine d'enzymes hydrolytiques, dont des PLA<sub>2</sub> (Barrett et Heath, 1977). Aussi, il a été proposé que les enzymes lysosomiales soient spécialisées dans la digestion des disques membranaires des segments externes des photorécepteurs (Zimmerman *et al.*, 1983).

L'hydrolyse des acides gras polyinsaturés, tels que l'acide docosahexaénoïque, des phospholipides qui composent les disques membranaires constitue la première étape

d'une boucle de recyclage au cours de laquelle ces acides gras sont retournés vers les photorécepteurs pour être réacylés (Birkle et Bazan, 1986). Les phospholipides ainsi formés sont dirigés vers la portion basale des segments externes des photorécepteurs, où ils serviront à l'assemblage de nouveaux disques membranaires.

Par contre, puisque la rétine est richement vascularisée et abondamment oxygénée, les dérivés d'oxygène hyperréactifs sont susceptibles d'y être générés en grande quantité (Berman, 1991; Wu *et al.*, 1992). En conséquence, les disques membranaires des segments externes des photorécepteurs peuvent subir des dommages suite à la peroxydation des acides gras polyinsaturés qu'ils contiennent. Il devient donc important que ces acides gras polyinsaturés ne soient pas recyclés avant d'avoir été réparés.

Ainsi, si certaines cPLA<sub>2</sub> se trouvent à l'intérieur des lysosomes des cellules de l'EPR, elles pourraient jouer un rôle important dans le processus de renouvellement des disques membranaires des segments externes des photorécepteurs. Elles hydrolyseraient les acides gras peroxydés des membranes phospholipidiques phagocytées avant qu'ils ne soient réparés et renvoyés vers les photorécepteurs (Van Themsche, 1999).

Rappelons à ce sujet que la cPLA<sub>2</sub> alpha a déjà été localisée ponctuellement dans des corps lipidiques cytoplasmiques de certaines cellules (Yu *et al.*, 1998) et que la cPLA<sub>2</sub> gamma est apparue en fluorescence sous la forme de structures ponctiformes cytoplasmiques, lesquelles avaient été associées aux peroxysomes (Hirabayashi et Shimizu, 2000), mais qui pourraient tout aussi bien s'avérer être des lysosomes.

### 1.2.5 Nouvelles phospholipases A2

Des expériences assez récentes menées par l'équipe du Dr Salesse suggèrent la présence de deux nouvelles PLA<sub>2</sub> cytosoliques au niveau de l'épithélium pigmentaire rétinien (Jacob *et al.*, 1996; Jacob *et al.*, 1997; Van Themsche *et al.*, 2001). En effet, suite à une chromatographie sur échangeur de cations, différentes fractions présentant une activité PLA<sub>2</sub> ont été obtenues. La sensibilité de ces fractions envers le pBPB, l'EGTA, le DTT,

la chaleur et le  $\text{H}_2\text{SO}_4$  indique que les cellules de l'épithélium pigmentaire rétinien humain en culture contiendraient deux  $\text{PLA}_2$  intracellulaires différentes. De plus, des expériences d'immunobuvardage de type Western ne montrent aucune réactivité entre les  $\text{PLA}_2$  cytosoliques des cellules de l'épithélium pigmentaire rétinien humain et plusieurs antisérums dirigés contre les  $\text{PLA}_2$  sécrétées et la  $\text{cPLA}_2$  alpha (Van Themsche *et al.*, 2001), ce qui suggère que les  $\text{PLA}_2$  des fractions identifiées sont des  $\text{PLA}_2$  différentes des  $\text{sPLA}_2$  et  $\text{cPLA}_2$  connues à ce moment et probablement spécifiques à l'EPR (Jacob *et al.*, 1997; Jacob *et al.*, 1998; Van Themsche *et al.*, 2001).

Puisque l'EPR exerce d'importants rôles dans le maintien de la fonction normale de la rétine, de nouvelles  $\text{PLA}_2$  spécifiques à l'EPR pourraient être impliquées dans ces fonctions-clé de l'EPR. De même, l'implication possible des  $\text{cPLA}_2$  à divers niveaux de la fonction visuelle permet de croire qu'une ou plusieurs  $\text{PLA}_2$  spécifiques à l'EPR pourraient sans aucun doute être liées à différentes dégénérescences rétiniennes si leur fonctionnement n'était pas adéquat.

C'est donc cette avenue de recherche qui a été à l'origine du travail détaillé dans le présent mémoire et des différents objectifs de recherche fixés pour la réalisation du projet de maîtrise.

## **CHAPITRE 2**

### **OBJECTIFS DE RECHERCHE**

#### **2.1 Objectif global**

L'objectif global du présent projet de maîtrise était donc d'identifier les PLA<sub>2</sub> cytosoliques présentes au niveau de la rétine et/ou de l'EPR. Ces cPLA<sub>2</sub> comprennent non seulement les nouvelles cPLA<sub>2</sub>, mais aussi celles qui sont déjà détaillées dans la littérature et qui pourraient se trouver au niveau de la rétine et de l'EPR.

Par conséquent, différents objectifs spécifiques ont été élaborés afin de mener à bien le projet de maîtrise.

##### **2.1.1 Premier objectif spécifique**

Le premier objectif spécifique consiste à déterminer les différents isoformes de cPLA<sub>2</sub> présents au niveau de la rétine et/ou de l'EPR.

###### **2.1.1.1 RT-PCR à l'aide d'amorces spécifiques**

L'utilisation d'amorces spécifiques permet une plus grande stringence au niveau des conditions utilisées lors du PCR. Ainsi, la séquence des amorces a été déterminée en fonction de chacune des séquences des cPLA<sub>2</sub> connues.

##### **2.1.2 Deuxième objectif spécifique**

Le deuxième objectif spécifique consiste à identifier de nouvelles cPLA<sub>2</sub> présentes dans l'EPR et/ou la rétine. Deux approches différentes sont envisagées; l'approche expérimentale, une approche réalisée par la mise en œuvre de diverses manipulations au laboratoire, et l'approche bioinformatique, une approche plus théorique utilisant en première ligne les diverses ressources bioinformatiques.

### **2.1.2.1 Approche expérimentale**

Tel que décrit ci-haut, cette approche comprend les diverses manipulations pouvant être réalisées au laboratoire; il s'agit d'un travail expérimental. Pour cette approche, le RT-PCR à l'aide d'amorces dégénérées et le criblage de banques par l'utilisation de sondes d'ADNc marquées sont les deux avenues envisagées.

#### **2.1.2.1.1 RT-PCR à l'aide d'amorces dégénérées**

L'utilisation d'amorces dégénérées par RT-PCR permet de cribler en un PCR plusieurs isoformes de cPLA<sub>2</sub> contenues dans l'ADNc de la rétine et de l'EPR. La séquence de ces amorces est déterminée par l'alignement des séquences des trois isoformes connues. En effet, c'est suite à l'alignement que la séquence des amorces est déterminée pour les régions consensus trouvées.

Cette approche pourrait donc permettre d'amplifier l'ADNc de nouvelles cPLA<sub>2</sub> en utilisant des conditions peu stringentes lors de l'amplification par PCR.

#### **2.1.2.1.2 Criblage de banques à l'aide d'une sonde d'ADNc**

L'utilisation d'un bout de séquence d'ADNc d'une cPLA<sub>2</sub> déjà connue en tant que sonde marquée pourrait permettre d'obtenir l'une ou l'autre des cPLA<sub>2</sub> connues lors du criblage d'une banque d'ADNc de rétine ou d'EPR. Ainsi, en utilisant des conditions de faible stringence, les possibilités d'obtenir un clone différent des cPLA<sub>2</sub> connues est probable.

### **2.1.2.2 Approche bioinformatique**

Tel que mentionné plus tôt, cette approche consiste d'abord en l'utilisation des diverses ressources bioinformatiques. Dans un deuxième temps, les résultats obtenus par ces démarches peuvent mener à différentes manipulations au laboratoire.

#### **2.1.2.2.1 Criblage bioinformatique**

Le criblage bioinformatique consiste à utiliser les séquences déjà connues de cPLA<sub>2</sub> pour cribler différentes banques de données. La principale banque d'intérêt ici est la banque de données du NCBI. Ainsi, il devient possible de repérer, parmi des milliers de données, des séquences indéterminées présentant des identités importantes avec les cPLA<sub>2</sub> connues. Ce sont précisément ces séquences qui pourraient s'avérer être de nouvelles cPLA<sub>2</sub>.

Ces séquences sont en fait de l'ADN génomique avec lequel aucune protéine n'a encore été associée. Par conséquent, l'analyse bioinformatique de ces séquences par divers programmes permet de déterminer la séquence d'ADNc probable de ces « nouvelles » cPLA<sub>2</sub>.

#### **2.1.2.2.2 Criblage par RT-PCR**

Suite à la découverte d'une ou plusieurs nouvelles séquences présentant des identités importantes avec les cPLA<sub>2</sub> connues, le PCR est le premier outil utilisé afin d'obtenir l'ADNc de ces « nouvelles » cPLA<sub>2</sub>. Ainsi, en se basant sur les séquences codantes déduites par les différentes analyses bioinformatiques, la séquence des amorces spécifiques ont été déterminées et ensuite utilisées en PCR.

Lorsqu'un ou plusieurs fragments sont obtenus, il est possible de les utiliser pour le criblage de banques d'ADNc.

#### **2.1.2.2.3 Criblage de banques d'ADNc**

Le criblage de banques d'ADNc constitue l'approche la plus utilisée lorsqu'un fragment d'ADNc de la séquence d'intérêt est obtenu. Ainsi, le fragment obtenu est utilisé en tant que sonde marquée pour cribler la banque d'ADNc de rétine ou d'EPR dans le but d'obtenir un clone plus complet, c'est-à-dire un clone comprenant entièrement les extrémités 5' et 3' de la séquence codante d'intérêt.

#### **2.1.2.2.4 Criblage d'une banque génomique**

Le criblage d'une banque génomique s'avère un peu moins fréquent, mais demeure toutefois intéressant dans les cas où le criblage de banques d'ADNc ne permet pas d'obtenir des résultats intéressants.

En fait, si le criblage des banques d'ADNc de rétine et d'EPR ne permet pas d'obtenir des clones intéressants, il est possible que le criblage d'une banque génomique, en utilisant différentes conditions de stringence, puisse résoudre le problème. De plus, si le clone n'était pas assez fortement exprimé dans les banques d'ADNc, il est possible qu'il soit plus facile de l'obtenir dans une banque génomique.

#### **2.1.2.2.5 Criblage à l'aide d'EST marqués**

Dans les cas où le criblage à l'aide d'un fragment d'ADNc ne s'avère pas concluant, il est recommandé d'utiliser différents EST (*expressed sequence tags*) en tant que sondes marquées. Une fois de plus, c'est par le criblage bioinformatique – à l'aide de la ou des séquence(s) trouvée(s) - des banques de données de la NCBI, plus particulièrement la banque des EST, qu'il devient possible de déterminer quels EST pourraient être utilisés en tant que sonde. C'est la localisation des EST sur la ou les séquences d'intérêt qui détermine la pertinence de son utilisation en tant que sonde.

#### **2.1.2.2.6 Criblage par RT-PCR de différents tissus**

Il est possible que le clone désiré ne soit pas présent en quantité suffisante dans l'une ou l'autre des banques utilisées pour pouvoir l'obtenir par le criblage à l'aide de sondes. Par conséquent, l'utilisation d'ADNc d'une grande variété de tissus pourrait permettre de définir dans quel(s) tissu(s) l'ADNc des séquences d'intérêt est le plus abondant et, par le fait même, indiquer si le type de banque d'ADNc utilisé est justifié.

Ainsi, ce sont les amorces dont la séquence a été déterminée pour la ou les séquences trouvées initialement qui seront utilisées pour cribler les différents tissus par RT-PCR.

#### **2.1.2.2.7 Amplification rapide des extrémités de l'ADNc**

La dernière approche envisagée pour obtenir l'ADNc complet d'un ou des clones obtenus par l'une ou l'autre des techniques mentionnées précédemment est le RACE-PCR.

Les informations concernant cette technique, de même que les informations ayant trait aux techniques évoquées précédemment paraissent dans le chapitre qui suit.



## CHAPITRE 3

### MATÉRIEL ET MÉTHODES

#### 3.1 Préparation de l'ADNc de l'EPR et de la rétine neurale

La préparation de l'ADNc est à la base de l'ensemble des manipulations réalisées au cours du travail de maîtrise décrit dans le présent mémoire. Cette préparation comprend plusieurs étapes à partir du moment où l'œil humain est reçu. D'abord, la dissection des globes oculaires, puis le prélèvement de l'EPR et de la rétine neurale, l'extraction de l'ARN et, enfin, la transcription de l'ARN en ADNc.

##### 3.1.1 Dissection des globes oculaires humains

Les yeux humains utilisés sont obtenus de la Banque Nationale d'Yeux (Sainte-Foy, Québec, Canada) et proviennent de donneurs âgés entre 35 et 70 ans. Les yeux sont disséqués selon une modification de la méthode de Pfeffer (1991).

Dans un premier temps, la sclère est découpée à approximativement 0,5 cm autour de la cornée. La cornée, le cristallin et le vitré sont ensuite retirés successivement. Le nerf optique est ensuite enlevé. Le globe restant est aussitôt rempli de la solution de dissection (0,3 mg/ml BSA, 0,5 mg/ml PVP (polyvinylpyrrolidone) et 100 U/ml de pénicilline/streptomycine dans une solution MHBSS (*Modified Hanks Balanced Salt Solution*)) (MHBSS : 0,5 mM  $\text{CaCl}_2$ , 8,3 mM glucose, 15 mM HEPES, 3,5 mM KCl, 0,27 mM  $\text{MgCl}_2$ , 0,37 mM  $\text{MgSO}_4$ , 138,5 mM NaCl, 4,3 mM  $\text{NaHCO}_3$ , 1,0 mM  $\text{NaH}_2\text{PO}_4$ , pH 7,2) pour éviter tout assèchement de la rétine. Le globe restant est découpé en deux pièces grossièrement égales. La rétine neurale (sans l'EPR) est prélevée à l'aide de pinces stériles et aussitôt placée dans la solution de dissection. Ensuite, les feuillets EPR-choroïde sont détachés de la sclère et placés dans une solution de dispase (2,4% dispase II, 10 mg/ml PVP, 2% sérum de poulet, 88 mM  $\text{CaCl}_2$  dans le MHBSS) et incubés de 45 à 60 minutes à 37°C, après avoir été rincés brièvement dans

la solution de dissection. Toutes les étapes subséquentes sont réalisées en conditions stériles. Les feuillets d'EPR-choroïde sont retirés de la dispase à l'aide de pinces stériles et transférés dans la solution de dissection. Les feuillets d'EPR sont ensuite décollés de la choroïde avec des petits jets - réalisés à l'aide d'une pipette de plastique stérile remplie de la solution de dissection - dirigés entre la choroïde et l'EPR. Les feuillets ainsi délogés sont transférés dans un tube stérile contenant la solution de dissection et enfin centrifugés et resuspendus directement dans le *Tri Reagent*<sup>TM</sup> (pour extraction directe de l'ARN). Dans le cas où les cellules sont mises en culture, les feuillets d'EPR sont déposés dans des pétris stériles de plastique contenant 10 ml de milieu de culture (200 mg/ml albumax, 45 mg/ml acide ascorbique, 1% extraits de rétines bovines, 5% de sérum de veau, 1 mg/ml carnitine, 112 mg/ml fructose, 500 mg/ml glucose, 5 mg/ml glutathion, 6 mg/ml hypoxanthine, 67 mg/ml acide oxalacétique, 0,15 mg/ml acétate de rétinol, 5 mg/ml taurine, 0,025 mg/ml D- $\alpha$ -tocophérol, 50 mg/ml transferrine, et 0,3 mg/ml uridine dans du milieu de culture de base *Kératinocyte-SFM*, pH 7,2). Ces feuillets isolés d'EPR sont incubés à 37°C pendant 48 à 72 heures dans une atmosphère en CO<sub>2</sub> de 5%. Les cellules qui n'ont pas encore adhéré après ce délai sont collectées, centrifugées, resuspendues dans dix millilitres de milieu de culture et remises en culture dans le premier pétri pour l'incubation. Cette manipulation permet d'éliminer les cellules autres que celles de l'EPR comme les fibroblastes et les mélanocytes. Après 48 à 72 heures d'incubation, le milieu de culture est remplacé par du milieu de culture frais. Les cellules demeurent dans le même pétri jusqu'à une confluence de 90% environ, où elles sont alors passées ou trypsinisées et collectées pour l'extraction de l'ARN.

### 3.1.2 Extraction de l'ARN rétinien

Dans la mesure du possible, les manipulations pour l'extraction sont réalisées sur la glace afin de prévenir toute dégradation de l'ARN. La rétine neurale prélevée est aussitôt déposée dans un volume de 1 à 3 ml de *Tri Reagent*, dans un tube stérile de 10 ml (*Falcon*), et congelée à -80°C pour usage ultérieur, ou tout de suite utilisée, selon le protocole indiqué par la compagnie *Sigma* pour le *Tri Reagent*.

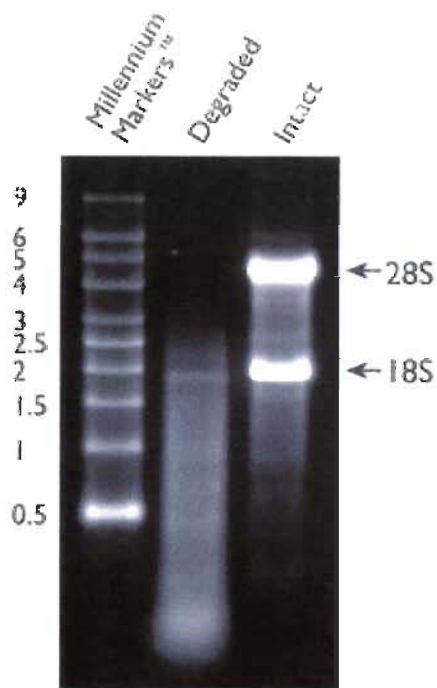
À l'aide d'une pipette appropriée dont l'embout est stérile, le tissu est homogénéisé par le va-et-vient du matériel dans l'embout. Le tube contenant l'homogénat est alors centrifugé à 12000g pendant dix minutes à 4°C. Le surnageant est transféré dans un microtube stérile et laissé à la température de la pièce pendant cinq minutes. Par la suite, 200 µl de chloroforme sont ajoutés au surnageant récupéré. Le microtube est alors vortexé vigoureusement pendant 15 secondes environ, puis laissé à la température de la pièce pendant 2 à 15 minutes pour être ensuite centrifugé de nouveau à 12000g pendant 14 minutes à 4°C. La phase aqueuse ainsi obtenue est transférée dans un nouveau microtube auquel sont ajoutés 500 µl isopropanol. Le tout est ensuite vortexé et laissé pendant cinq à dix minutes à la température de la pièce. Enfin, ce dernier microtube contenant l'isopropanol et la phase aqueuse est centrifugé à 12000g pendant 10 minutes à 4°C pour permettre de se débarrasser du surnageant et d'ensuite rincer le culot. Ainsi, le culot obtenu est rincé avec au moins 1 ml d'éthanol 75%, vortexé, centrifugé à 7500g pendant 5 minutes à 4°C puis, resuspendu dans de l'eau sans RNase après avoir été séché.

L'ARN ainsi obtenu est dosé par spectrophotométrie (biophotomètre de la compagnie *Eppendorf*) afin de déterminer approximativement la concentration d'ARN de chaque tube. En fait, puisque l'ADN et l'ARN présentent une absorption maximale à 260 nm et que pour une importante étendue de concentrations il existe une relation linéaire entre l'absorption de lumière et la concentration nucléique, il est possible de doser l'ARN et l'ADN par spectrophotométrie. À l'intérieur du spectrophotomètre, l'échantillon est exposé à une lumière ultraviolette de 260 nm et le photodétecteur mesure la quantité de lumière qui passe à travers l'échantillon. Plus la lumière est absorbée – moins le détecteur capte de lumière - plus la concentration en acides nucléiques est importante dans l'échantillon. La quantité de lumière absorbée par un acide nucléique est fonction du coefficient d'extinction molaire du type d'acides nucléiques présent en solution. La loi de Beer-Lambert établit la relation suivante entre le coefficient d'extinction molaire et la quantité de lumière absorbée :

$$A = \epsilon cl$$

où **A** est l'absorbance,  $\epsilon$ , le coefficient d'extinction molaire, **c**, la concentration du soluté et **l**, la distance parcourue par la lumière dans la solution (1 cm). Ainsi, pour un soluté particulier, l'ARN par exemple, un coefficient d'extinction précis  $(0,025 \text{ } (\mu\text{g/ml})^{-1} \text{ cm}^{-1})$  est attribué et il est possible de déterminer la concentration d'une solution homogène d'ARN. Par exemple, une absorbance de 1,0 à 260 nm pour l'ARN correspond à une concentration de 40  $\mu\text{g/ml}$ . De plus, puisque les protéines absorbent non seulement à 280 nm mais aussi à 260 nm, une seconde mesure de densité optique à 280 nm est effectuée par le spectrophotomètre. Un ARN pur doit avoir un rapport 260 / 280 compris entre 1,8 et 2, sans quoi on suppose la présence de protéines et/ou de phénol.

Enfin, la solution d'ARN sera déposée sur un gel d'agarose 1,5% TBE 1X pour vérifier la présence des deux bandes intactes d'ARN ribosomal (18S et 28S) et comparer leur intensité. En fait, la bande de 28S doit être approximativement deux fois plus intense que la bande de 18S pour confirmer l'intégrité de l'ARN obtenu (figure 3.1).



**Figure 3.1** – Vérification de l'intégrité de l'ARN par migration sur gel d'agarose. Dans la deuxième piste, l'ARN apparaît dégradé, alors que dans la dernière piste l'ARN intact apparaît sous forme de deux bandes bien distinctes. Image tirée du site Internet de la compagnie *Ambion*.

### 3.1.3 Extraction de l'ARN de l'EPR

À partir du moment où les cellules de l'EPR sont en suspension dans le *Tri Reagent*, le protocole suivi est le même que celui qui est décrit pour l'extraction de l'ARN rétinien, à l'exception de l'étape additionnelle consistant à prélever le matériel insoluble qui se trouve au-dessus de la phase aqueuse claire d'intérêt. Cette étape permet de se débarrasser d'un surplus de matériel lipidique avant de procéder à l'extraction proprement dite. Les cellules de l'EPR sont initialement suspendues dans un volume de 1 à 2 ml de *Tri Reagent*. Les manipulations précédant les centrifugations sont aussi réalisées sur la glace dans la mesure du possible.

### 3.1.4 Transcription de l'ARN en ADNc

À cette fin, le protocole suivi est celui qui est indiqué par *GibcoBRL*® pour l'utilisation de la *SuperScript<sup>TM</sup> II*.

Pour un volume réactionnel de 20 µl, 5 µg d'ARN total sont utilisés. Dans un microtube sans nucléase, 1 µl d'oligo (dT)<sub>12-18</sub> (500 µg/ml), 5 µg d'ARN total et 1 µl d'un mélange de dNTPs 10nM (10 mM de chacun : dATP, dGTP, dCTP et dTTP à pH neutre) sont déposés. Le volume total est complété à 12 µl avec de l'eau distillée stérile, puis chauffé à 65°C pendant cinq minutes, pour ensuite être refroidi rapidement sur la glace. Le microtube est ensuite brièvement centrifugé afin de récupérer tout le volume au fond du tube, puis 4 µl de *First-Strand Buffer* 5X, 2 µl de DTT 0,1 M et 1 µl de *RNaseOUT Recombinant Ribonuclease Inhibitor* (40 U/µl) sont ajoutés. Le contenu est ensuite mélangé légèrement et incubé à 42°C pendant deux minutes. Suite à cette incubation, 1 µl de *SuperScript II* est ajouté au volume du microtube et mélangé par un pipettage répété. Ce microtube est enfin incubé pendant 50 minutes à 42°C. La réaction est inactivée en plaçant le microtube à 70°C pendant 15 minutes.

L'ADNc ainsi obtenu est prêt à être utilisé pour les différentes amplifications par PCR. Il est conservé à 4°C s'il est utilisé fréquemment ou congelé à -20°C si son utilisation est

reportée à plus tard. La quantité d'ADNc obtenue dans chacun des tubes est déterminée par spectrophotométrie, selon le principe mentionné précédemment. Le rapport des densités optiques à 260 et 280 nm doit être supérieur à 1,9 pour pouvoir considérer que la solution d'ADN est dépourvue d'impuretés.

### 3.2 Réaction de polymérase en chaîne (PCR)

Les réactions de polymérase en chaîne (PCR) ont été réalisées en variant divers paramètres. En fait, puisqu'il est possible de varier plusieurs des conditions lors de cette réaction, il est inutile de préciser qu'une vaste gamme de conditions différentes ont été utilisées. Toutefois, puisqu'il serait superflu de présenter toutes ces conditions, seules les conditions utilisées le plus fréquemment sont détaillées. Par contre, il importe de mentionner que les réactions de PCR présentées dans cette section se rapportent aux réactions qui ont été réalisées à l'aide de matériel provenant uniquement de la rétine neurale ou de l'EPR.

Pour l'ensemble des réactions de PCR, 0,5 à 1 µg d'ADNc ont été utilisés dans le cas des RT-PCR, alors pour l'amplification de l'ADN plasmidique, 1 ng était suffisant. La quantité précise d'ADNc utilisée pour chacune des réactions a été déterminée en essayant différentes quantités d'ADNc. Ainsi, la quantité permettant d'obtenir une bande suffisamment intense lors de l'électrophorèse sur gel d'agarose était utilisée pour les réactions subséquentes.

La polymérase utilisée est celle de *Fermentas* (*Taq* DNA polymerase) et le tampon de PCR est celui contenant le  $(\text{NH}_4)_2\text{SO}_4$ , fourni avec l'enzyme. Les concentrations finales utilisées lors des manipulations sont présentées dans le tableau 3.1 qui suit.

Un nombre important d'amorces différentes a été utilisé lors des différentes réactions de PCR. Le tableau 3.2 présente les séquences des diverses amorces qui ont été déterminées lors de ce projet de maîtrise.

**TABLEAU 3.1**  
**Concentrations utilisées lors des réactions de PCR**

Réactif	Concentration finale
Eau stérile	-
Tampon 10X (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	1X
dNTP Mix 10 mM	250 µM
Amorce 1 (20 µM)	0,5 – 1 µM
Amorce 2 (20 µM)	0,5 – 1 µM
<i>Taq DNA polymerase</i>	1 – 2 U / 20 µl
MgCl <sub>2</sub> (25 mM)	2 – 2,5 mM
ADN plasmidique	1 ng / 20 µl
génomique/ADNc	0,5 – 1 µg / 20 µl

**TABLEAU 3.2**  
**Compilation des différentes amorces utilisées**

Nom de l'amorce	Séquence de l'amorce	T <sub>m</sub> (°C)	Produit attendu <sup>1</sup> (en pb)
<b>Amorces dégénérées pour les cPLA<sub>2</sub></b>			
cPLA <sub>2</sub> -1	5' GG(GT) (GC)(CT)(CT) TC(GT) GGC TCC ACC TGG 3'	70	1053
cPLA <sub>2</sub> -2	5' (AG)TA GTC (CA)AA (GT)GA SA(GT) (GT)AT GAG 3'	47	1053
<b>Amorces correspondant à la cPLA<sub>2</sub> alpha</b>			
Position 5-5'	5' CAT TTA TAG ATC CTT ACC AG 3'	60	5°
Position 51-5'	5' CCA CAA GTT (CT)AC GGT AGT GG 3'	66	51°
Position 204-5'	5' CCT GT(CG) TGG AA(CT) GAG ACC TT 3'	66	204°
Position 1798-3'	5' CTT (CT)GG AAA (GT)GG (CG)AG CTT GT 3'	65	1798°
Position finale 3'	5' CTA TGC TTT GGG TTT ACT TA 3'	60	2070e
cPLA <sub>2</sub> alpha courte (S)	5' ATT CTG CAC GTG ATG TGC CTG T 3'	67	1030



Nom de l'amorce	Séquence de l'amorce	T <sub>m</sub> (°C)	Produit attendu <sup>1</sup> (en pb)
cPLA <sub>2</sub> alpha courte (AS)	5' ACA GCT GCA TCC AGT TCA TCA T 3'	65	1030
cPLA <sub>2</sub> A1 (S)	5' GAG CTG AAA AAG GAT CCT GAC T 3'	65	440
cPLA <sub>2</sub> A2 (AS)	5' TGT CCC TAG AGT TTC ATC CA 3'	63	440
<b>Amorces correspondant à la cPLA<sub>2</sub> bêta</b>			
Bêta1	5' CAC TTC TAC CGG GAC TGG GTC T 3'	70	6761 / 2862
Bêta2	5' TCC TGG TTG TTG CAG ACA TTG T 3'	65	6761 / 2862
cPLA <sub>2</sub> bêta courte1 (S)	5' ATC AGC TTT GGA GGA TCT TGG A 3'	65	3565
cPLA <sub>2</sub> bêta courte2 (S)	5' ATC AGG AGC TCT AGT ATA GGC T 3'	65	2149
cPLA <sub>2</sub> bêta courte 3 (S)	5' TCT GAC TGC TAC GTG ACTTCTC T 3'	67	608
cPLA <sub>2</sub> bêta courte (AS)	5' CAA GCC TCA CCA CTT GAC CAG A 3'	68	-
Bêta 400 (S)	5' ATG TTG GCT ACC TCA TCA ATA 3'	62	381
Bêta 400 (AS)	5' TGT AGG TCA CCT TCG TGT AGT 3'	65	381
<b>Amorces correspondant à la cPLA<sub>2</sub> gamma</b>			
Gamma1	5' AGT TTC CAT AAT TCC TGG GCT C 3'	65	1564
Gamma2	5' GCC ACG TTC ATC AAC TCT CTA A 3'	65	1564
Gamma NYS	5' ATT ACT CTC TGA CCG ACT TC 3'	63	1036
Gamma FNY	5' CCA CAG GCA TCT ATG TTG AA 3'	63	1036
Gamma ext5	5' CAG TGC ACC ATG GGA AGC TCT G 3'	70	1649
Gamma ext3	5' AAG CTG AGG CTC ATC TAT GCC A 3'	67	1649
<b>Amorces correspondant au nouveau gène 1</b>			
Gène 1a	5' CTT GGA CAG TGA GCC AGA CAG A 3'	68	1000 <sup>c</sup>
Gène 1b	5' CAG CAG AGC TGG AGT TTC TGC A 3'	68	1813 <sup>c</sup>
Gène 1c	5' CAT CAG GTA GCA TTC CTT GAG A 3'	65	3441 <sup>c</sup>
Gène 1d	5' CTT CTT CTT CEC CAC TGC GAG C 3'	68	3795 <sup>c</sup>
1A	5' ATG GGT GAG CCT GGA AAA GGT A 3'	67	1 <sup>ère</sup>
1B	5' TTC TTA GTT GTG TGG CAG AGC T 3'	65	523 <sup>c</sup>
1C	5' ATG GGA TGG TTT GCT GAG GGT A 3'	67	658 <sup>c</sup>
1D	5' AAG CTG AGG ACA AGG ACC ATC T 3'	67	1060 <sup>c</sup>
1E	5' ATG GAA GAG CTG GAG GTG GAG T 3'	68	1279 <sup>c</sup>
1F	5' ACT GCC TTT CCG ATG GTC AGG T 3'	68	1790 <sup>c</sup>



Nom de l'amorce	Séquence de l'amorce	Tm (°C)	Produit attendu <sup>1</sup> (en pb)
1G (AS)	5' TGT AAA GGT GAC CCT GTA GCC T 3'	67	2298 <sup>e</sup>
1H (AS)	5' ACT TCT GCA GAC CCA CCT CGT A 3'	68	2656 <sup>e</sup>
1I (AS)	5' TTC CAC CTA GAG AAC TGG CCA T 3'	67	3062 <sup>e</sup>
1J (AS)	5' TGT CAA AGG TGG CCT CTG TGT A 3'	67	3709 <sup>e</sup>
<b>Amorces correspondant au nouveau gène 2</b>			
Gène 2a	5' TTT CAG CAA CCT CAG GTC TGT G 3'	67	1 <sup>ère</sup> 2
Gène 2b	5' GCT TCA ATC TCT GTG CAG AGG A 3'	67	1529 <sup>e</sup>
Gène 2c (AS)	5' TGG AGA AGT CTT TGT GGA TAC A 3'	65	2527 <sup>e</sup>
Gène 2d (AS)	5' TCT GAG CAA CCT CAG GTC TGT G 3'	68	3128 <sup>e</sup>
2END1	5' CAG GAC TAC TGT AGC CAC AAA G 3'	67	2518 <sup>e</sup>
2END2	5' TGT AGC CAC AAA GAC TTC TCC A 3'	65	2527 <sup>e</sup>
NEW2A	5' CTC ACG AAG ACT TCC TTA ACC T 3'	65	801 <sup>e</sup> 2
NEW2B	5' GAA TGT TCT GGA GCT TAG CAT C 3'	65	936 <sup>e</sup>
NEW2C	5' AGA TCA GGA CAA GCT GGA GCT G 3'	68	1215 <sup>e</sup>
NEW2D (AS)	5' TGA GAA GAG GTG GCA TTC CCT T 3'	67	2811 <sup>e</sup>
NEW2E (AS)	5' CTT CCT CCT TGT AGG TCA TGT T 3'	65	2979 <sup>e</sup>
2A	5' AAC AGG AAC TTC CAC TCA TTA A 3'	62	53 <sup>e</sup>
2B	5' AAC GAC TTA CAC AAT GTC TAA T 3'	60	261 <sup>e</sup>
2C	5' ATT AAA GCT ATA GAC CTG AGC T 3'	62	406 <sup>e</sup>
2D	5' TCC TGA AAC TCC AGG TCT GGA A 3'	67	501 <sup>e</sup>
2E	5' ATG AAG GAT GTG AAG CTG TTT T 3'	62	703 <sup>e</sup>
2F	5' TTC AAG GTT CTC TAT GAC ATC T 3'	62	994 <sup>e</sup>
2G	5' ATG TGG AGT TCC TGA TGG AAG A 3'	65	1085 <sup>e</sup>
2H	5' ACA TGG CAG CCC TAG AGA CAG A 3'	68	1313 <sup>e</sup>
2I (AS)	5' TCT CTT TGA CAT TGA GGA TCA A 3'	63	2068 <sup>e</sup>
SondeH	5' ACC TGG AGG GAC CTA TCA GAT A 3'	67	172
SondeH (AS)	5' TGC AGC ATG GAC TCC AGC ACT A 3'	68	172
SondeB	5' GAG AAG GAG CCC CTG ACC ACC T 3'	71	187
<b>Amorces spécifiques à différents vecteurs plasmidiques</b>			
T7 promoter primer	5' TAA TAC GAC TCA CTA TAG GGC G 3'	65	-
SP6 primer	5' ATT TAG GTG ACA CTA TAG AAT A 3'	59	-
M13 Forward (-21)	5' TGT AAA ACG ACG GCC AGT 3'	63	-

Nom de l'amorce	Séquence de l'amorce	T <sub>m</sub> (°C)	Produit attendu <sup>1</sup> (en pb)
M13 Reverse	5' AGC GGA TAA CAA TTT CAC AC 3'	61	-
T3 primer	5' CAA TTA ACCCTC ACTAAA GG 3'	61	-
Amorce SG5	5' TTG TGA AAT TTG TGA TGA TA 3'	58	-
<b>Amorces utilisées pour le RACE PCR</b>			
PolyT	5' TTT TTT TTT TTT TTT TTT TTT TTT TTT 3'	53	-
RACE S1	5' (CAU) <sub>4</sub> TTC CTG AGC AGG AGG AAC CA 3'	80	1558°
RACE S2	5' (CAU) <sub>4</sub> ACC TGC AGG AGG ATG AGG TA 3'	80	1622°
RACE S3	5' (CAU) <sub>4</sub> GAC CTA CTT CAT TGG CAT CT 3'	77	1743°
RACE AS1	5' (CAU) <sub>4</sub> TAA GCT CCT GGT CTT GTC CT 3'	78	2342°
RACE AS2	5' (CAU) <sub>4</sub> TGT CAG GAA GCC TTT AAA TG 3'	76	2444°
RACE AS3	5' (CAU) <sub>4</sub> CAG TAG TCC TGG TGC AGC TG 3'	81	2509°

<sup>1</sup> Puisque les amorces peuvent être pour la plupart utilisées pour différentes combinaisons, pour certaines, c'est la position du premier nucléotide de l'amorce (en 5') au niveau de l'ADNc qui est indiqué par un rang, considérant le A de l'ATG de départ comme la position du premier acide nucléique (ainsi A de l'ATG de départ est le 1<sup>er</sup> acide nucléique et T, le 2<sup>e</sup>).

<sup>2</sup> Indique que la séquence de l'amorce a été déterminée selon une séquence différente de la séquence présentée dans ce mémoire. L'amorce se trouve ainsi dans une région intronique de la séquence finale adoptée.

En ce qui a trait exclusivement à la réaction elle-même, voici les conditions utilisées pour la majorité des réactions réalisées :

- 1) Étape de dénaturation initiale de trois minutes à 95°C
- 2) Étape de dénaturation de 35 secondes à 95°C
- 3) Étape d'hybridation des amorces de 40 secondes à 50°C
- 4) Étape d'extension de trois minutes à 72°C
- 5) Étape finale d'extension de six minutes à 72°C

Les étapes 2-3-4 sont répétées 34 fois supplémentaires (35 cycles) avant d'en arriver à l'étape d'extension finale.

Les différents produits de PCR obtenus sont déposés sur un gel d'agarose de 1% TBE

1X (TBE 1X : 89 mM Tris base, 89 mM acide borique, 2 mM d'EDTA, pH 8,0) coloré au bromure d'éthidium (concentration finale de 0,01 mg/100 ml) et migrent à 100 Volts pendant environ une heure. Les marqueurs de poids moléculaire utilisés sont généralement le *GeneRuler 100 DNA Ladder Plus* (100 à 3000 pb) de la compagnie *Fermentas* et le  $\Phi$ X174 *Hae III* (118 à 1353 pb) de la compagnie *Promega*. Le résultat de la migration est observé et photographié à l'aide de l'appareil *Gel Doc* et du logiciel *Quantity One* de la compagnie *Biorad*.

### 3.3 Clonage des bandes PCR

Les différentes bandes obtenues par PCR qui se sont avérées intéressantes ont été découpées et purifiées, puis clonées dans un vecteur plasmidique.

#### 3.3.1 Purification des bandes

Les bandes d'intérêt observées grâce à un transilluminateur (*DyNA Light, Labnet*) sont découpées à l'aide d'une lame stérile et sont ensuite purifiées à l'aide du système *Ultrafree-DA* de *Millipore*. Lors de cette purification, la bande de gel d'agarose (à une concentration de moins de 1,25%) est déposée dans la partie « nébuliseur » du microtube. Le microtube est ensuite centrifugé pendant dix minutes à 5000 g. Par conséquent, la centrifugation force l'agarose à travers le nébuliseur de gel, le convertissant ainsi en fines gouttelettes qui sont capturées par le filtre inférieur. L'ADN extrait et la solution de TBE 1X traversent donc le filtre et sont collectés dans le fond du microtube sous forme de solution ADN/TBE 1X.

#### 3.3.2 Clonage des bandes

Le système de clonage utilisé est le *pGEM-T Easy Vector System* de *Promega*. Le protocole utilisé est conforme au protocole fourni par la compagnie.

Dans un microtube de 1,5 ml, sont déposés 3 µl du produit de PCR purifié (bande purifiée), 1 µl du vecteur *pGEM-T Easy*, 5 µl du *T4 DNA Ligase 2X Rapid Ligation Buffer* et 1 µl de *T4 DNA Ligase* (3 unités Weiss/µl). Cette solution est bien mélangée par pipettage répété, puis incubée pendant une heure à la température de la pièce ou une nuit à 4°C. Une fois la ligation terminée, la transformation est effectuée afin d'obtenir une quantité suffisante de plasmides pour le séquençage.

Pour la transformation, la souche *E.coli DH5α* est utilisée en tant que bactérie compétente. Ainsi, pour chacune des transformations, 50 µl d'une solution bactérienne d'*E.coli DH5α* et 2 µl du produit de ligation sont ajoutés à un microtube de 1,5 ml déjà placé sur la glace. Le microtube repose ainsi sur la glace pendant 20 minutes, puis les bactéries sont soumises à un choc thermique : 50 secondes à 42°C suivi immédiatement de deux minutes sur la glace. Ensuite, 950 µl de S.O.C (S.O.C. : 20 g/l bacto-tryptone, 5 g extrait de levure, 10 mM NaCl, 2,5 mM KCl, 20 mM glucose et 20 mM  $Mg^{2+}$  stock, pH 7,0) ( $Mg^{2+}$  stock : 20,33 g/100 ml  $MgCl_2 \cdot 6H_2O$  et 24,65 g/100 ml  $MgSO_4 \cdot 7H_2O$ ) sont alors ajoutés aux bactéries. Cette solution bactérienne est incubée à 37°C pendant 1h30 sous une agitation de 150 rotations par minute (rpm). Enfin, 100 µl de cette culture bactérienne sontensemencés sur chacun des pétris LB/X-Gal/IPTG/ampicilline préchauffés à la température de la pièce (pétris LB/X-Gal/IPTG/ampicilline : 10 g/l bacto-tryptone, 5 g/l extrait de levure, 5 g/l NaCl, 0,5 mM IPTG, 80 µg/ml X-Gal et 100 µg/ml d'ampicilline).

Ces pétris sont incubés pendant une nuit à 37°C, puis quelques colonies blanches bien isolées sont sélectionnées pour la purification plasmidique. Ce sont les colonies blanches qui sont sélectionnées puisqu'elles contiennent l'insert (le produit de PCR). En fait, cet insert vient s'insérer dans le gène de la *bêta-galactosidase* et empêche son expression et, par conséquent, empêche aussi l'hydrolyse de l'IPTG, responsable de la formation d'un produit bleu insoluble. Le produit insoluble n'étant pas présent dans ces colonies, elles demeurent blanches.

### 3.3.2.1 Préparation des bactéries compétentes

Les bactéries non compétentes *E. coli DH5α* sont d'abordensemencées sur un pétri LB/MgSO<sub>4</sub> (pétris LB/MgSO<sub>4</sub> : 10 g/l bacto-tryptone, 5 g/l extrait de levure, 5 g/l NaCl, 10 mM MgSO<sub>4</sub> et 15 g/l d'agar) sans antibiotique et incubées pendant une nuit à 37°C. Par la suite, une colonie isolée est repiquée dans 5 ml de milieu LB/MgSO<sub>4</sub> et incubé pendant deux à trois heures, jusqu'à ce que la densité optique à 600 nm atteigne 0,4. Les bactéries sont alors transférées dans deux tubes de 50 ml et centrifugées à 3000 rpm pendant dix minutes à 4°C. Par la suite, les solutions centrifugées sont décantées et chaque culot bactérien est resuspendu dans 20 ml de TfbI (TfbI : 30 mM KOAc, 50 mM MnCl<sub>2</sub>, 100 mM KCl, 10 mM CaCl<sub>2</sub> et 15% (p/v) glycérol) et gardé sur la glace pendant dix minutes.

Les tubes sont alors centrifugés à 3000 rpm pendant huit minutes à 4°C, décantés et les deux culots resuspendus doucement dans un seul tube contenant 4 ml de TfbII (TfbII : 10 mM Na-Mops, 75 mM CaCl<sub>2</sub>, 10 mM KCl, 15% (w/v) glycérol). Cette solution bactérienne est enfin aliquotée dans des microtubes préalablement placés dans la glace sèche. Ces aliquots sont conservés à -80°C.

Avant d'utiliser ces bactéries lors des différents clonages, leur efficacité de transformation est vérifiée. Ainsi, à l'aide d'une concentration connue de vecteur non coupé, les bactéries sont transformées, puisensemencées tel que décrit plus haut. Les colonies obtenues sont comptées afin de déterminer la quantité de *colony forming units* (cfu) obtenue par µg d'ADN plasmidique. Le rapport cfu/µg doit être d'au moins 1x10<sup>8</sup>/µg, sans quoi les bactéries dites compétentes ne sont pas utilisées.

### 3.4 Purification de plasmides

Une fois les bactéries blanches sélectionnées, elles sont mises en culture séparément, dans des tubes de 15 ml contenant 5 ml de milieu LB/ampicilline (10 g/l bacto-tryptone,

5 g/l extrait de levure, 5 g/l NaCl, 100 µg/ml ampicilline). Ces tubes sont ensuite incubés pendant 16 heures à 37°C sous agitation (220 rpm). Suite à cette incubation, les cultures bactériennes sont prêtes à être soumises à la purification plasmidique.

En fait, 4 ml servent à la purification plasmidique et deux volumes de 0,5 ml serviront à préparer deux stocks glycérol de la colonie prélevée (stock glycérol : 0,5 ml de culture bactérienne, 0,5 ml de milieu LB/glycérol) (LB/glycérol : 50% glycérol, 50% LB). Ces stocks glycérol sont conservés à -80°C.

La purification de plasmides est entièrement réalisée selon le protocole de *QIAGEN*, le *QIAprep Miniprep Handbook*. Il s'agit donc de centrifuger à vitesse maximale (14000 rpm) la solution bactérienne restante afin d'obtenir un culot bactérien. Le culot est ensuite resuspendu dans 250 µl de tampon contenant la RNase (*Buffer P1*), puis 250 µl du *Buffer P2* sont ajoutés à ce mélange. Le microtube est alors inversé quelques fois délicatement. Une fois que le tampon *P2* est ajouté, il ne faut pas dépasser cinq minutes d'incubation, au risque de fragmenter l'ADN génomique. À cette étape, la solution est claire et visqueuse. Par la suite, 350 µl de *Buffer N3* sont ajoutés au microtube et ce dernier est rapidement inversé plusieurs fois pour éviter une précipitation localisée des débris cellulaires. Le microtube est alors centrifugé pendant dix minutes à vitesse maximale. Le surnageant est ensuite prélevé à l'aide d'une pipette et déposé sur une colonne *QIAprep*. Cette colonne est alors centrifugée pendant une minute à vitesse maximale et l'éluat est jeté. L'ADN plasmidique alors fixé à la colonne est lavé à l'aide de 750 µl du *Buffer PE* qui sont ajoutés à la colonne qui est ensuite centrifugée pendant une minute à vitesse maximale. L'éluat est jeté et la colonne est recentrifugée à vitesse maximale pendant une minute. L'ADN plasmidique est finalement élué : 1,5 ml du *Buffer EB* (10 mM Tris-HCl, pH 8,5) est ajouté à la colonne qui repose ensuite à la température ambiante pendant 1 minute. La colonne est finalement centrifugée pendant une minute à vitesse maximale.

Les plasmides sont donc purifiés et récupérés dans une solution tampon Tris-HCl 10 mM (pH 8,5). Les concentrations obtenues sont ensuite mesurées par spectrophotométrie et

les plasmides sont digérés afin de vérifier si la longueur des inserts correspond à celle des produits de PCR de départ.

### 3.5 Digestion des plasmides

Les plasmides purifiés sont soumis à une digestion enzymatique afin de déterminer la longueur de l'insert de chaque solution plasmidique obtenue. Ainsi, 1 µl de l'enzyme *EcoRI* (10 U/µl, *Fermentas*) est ajouté aux 5 µl de la solution plasmidique, 2 µl de la solution *tampon O+ 10X* (*Fermentas*) (*tampon O+* : 50 mM Tris-HCl (pH 7,5), 10 mM MgCl<sub>2</sub>, 100 mM NaCl et 0,1 mg/ml BSA) et 12 µl d'eau ultrapure dans un microtube de 1,5 ml. Cette solution est incubée à 37°C pendant 3 heures, puis la réaction de digestion est inactivée par une incubation de 20 minutes à 65°C.

Chacune de ces solutions de plasmides digérés est déposée sur un gel d'agarose 1% TBE 1X. La migration se poursuit pendant environ 40 minutes à 100 Volts et les bandes sont visualisées et photographiées, tel que décrit dans la section 3.2.

Enfin, si la concentration des solutions plasmidiques est adéquate (de l'ordre du 100 µg/µl) et que les inserts présentent la longueur attendue sur gel, les plasmides peuvent alors être envoyés pour le séquençage automatique (tous les séquençages sont réalisés à l'Université Laval par le Service d'analyse et de synthèse d'acides nucléiques) afin de déterminer la séquence nucléotidique précise des inserts.

### 3.6 Criblage de banques d'ADNc

Le criblage des banques d'ADNc de rétine neurale et d'EPR comprend de nombreuses étapes. En premier lieu, la sonde doit être préparée et le titre de la banque à cribler doit être vérifié afin de déterminer la quantité de phages présents par millilitre de banque.



Viennent ensuite les étapes de préparation des pétris, le traitement des membranes et finalement, l'hybridation et la détection constituent les dernières étapes du criblage.

### 3.6.1 Préparation de la sonde

Les sondes pour le criblage ont toutes été préparées à l'aide d'ADN plasmidique. Il s'agissait soit de plasmides dont l'insert était un segment d'ADNc d'une cPLA<sub>2</sub> ou d'un plasmide dont l'insert était un EST (voir la section 3.10). Ainsi, des criblages ont été réalisés à l'aide d'un plasmide SG5, dont l'insert était la cPLA<sub>2</sub> alpha, qui provenait de la compagnie *Merck Frosst*. D'autres criblages ont été effectués à l'aide de plasmide *pGEM-T Easy (Promega)* contenant un fragment d'une « nouvelle » cPLA<sub>2</sub>. Inutile de préciser que différents fragments (3) de cette dernière cPLA<sub>2</sub> ont été essayés. Enfin, des criblages ont été réalisés à l'aide de différents plasmides dont les inserts étaient des EST. Ces différents plasmides (EST) ont été obtenus via la compagnie *Research Genetics*.

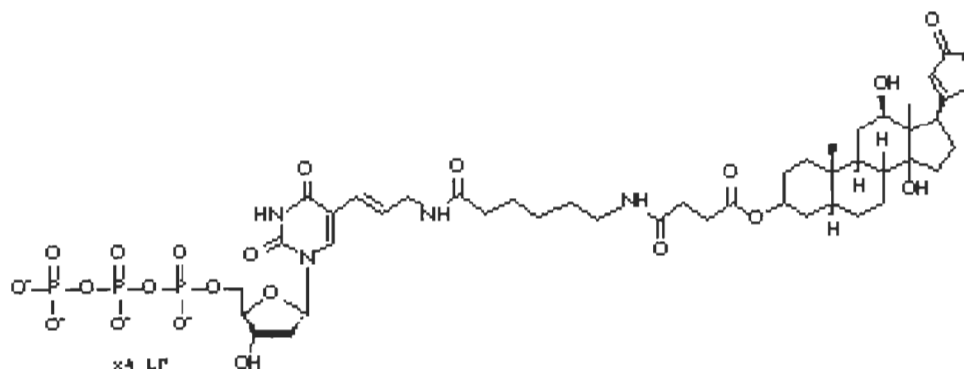
Dans le cadre du travail de maîtrise présenté ici, deux différentes techniques de préparation de la sonde ont été utilisées. Les protocoles suivis pour ces deux approches figurent dans *The DIG System User's Guide for Filter Hybridization* de *Boehringer Mannheim (Eisel, 1995)*. La première technique est le marquage aléatoire et la deuxième, l'incorporation de la digoxigénine-11-dUTP durant le PCR. La première technique fonctionnait relativement bien, mais ne permettait pas d'obtenir une quantité aussi appréciable de sondes marquées que la deuxième. C'est donc la raison pour laquelle les sondes fabriquées à partir de l'ADNc de la cPLA<sub>2</sub> alpha sont les seules qui ont été réalisées selon la première technique. Dans les deux cas, le matériel de départ était de l'ADN plasmidique où l'insert était de l'ADNc de cPLA<sub>2</sub> ou d'EST.

#### 3.6.1.1 Marquage aléatoire

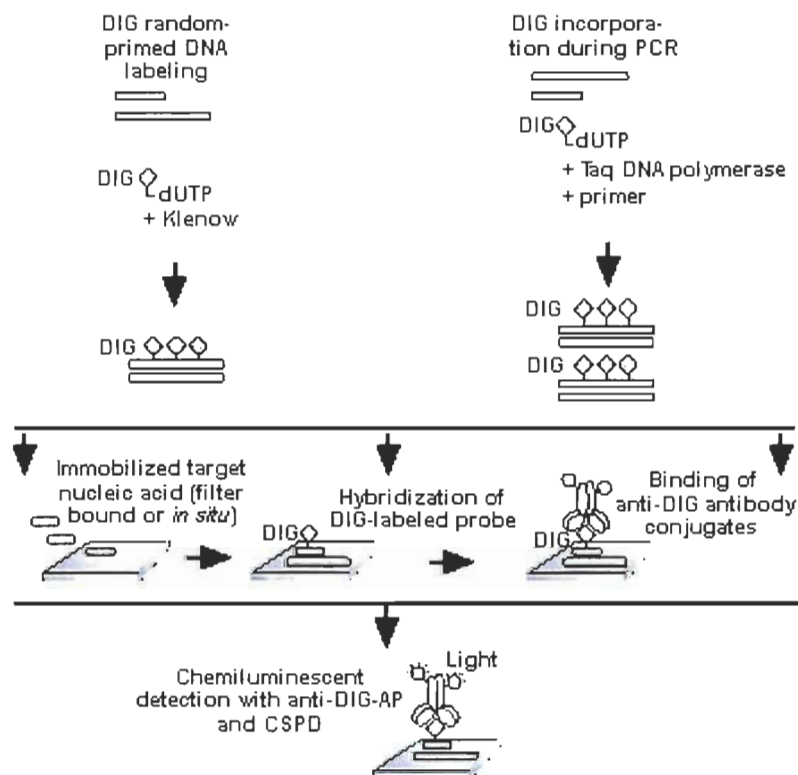
Au cours de cette technique, une molécule de digoxigénine-11-dUTP (DIG-dUTP) (figure 3.2) est incorporée à tous les 20 à 25 nucléotides lors de la synthèse du brin



complémentaire de la sonde. C'est l'enzyme *Klenow* qui synthétise ce brin complémentaire en incorporant de temps à autre la molécule DIG-dUTP (figure 3.3).



**Figure 3.2** - Structure de la digoxigénine-11-dUTP. Cette molécule est utilisée pour la plupart des applications de marquage non radioactif d'ADN. Image tirée du site Internet de Boehringer Mannheim.



**Figure 3.3** - Deux des différents types de marquages possibles et la détection au CSPD par le système DIG de *Boehringer Mannheim*. La première technique présentée (en haut à gauche : *DIG random-primed DNA labeling*) est celle qui est appelée marquage aléatoire dans le présent mémoire. La deuxième technique (en haut à droite : *DIG incorporation during PCR*) est la deuxième technique présentée dans ce mémoire et correspond à l'incorporation de la DIG durant le PCR. Schéma adapté d'une image tirée du site Internet de Boehringer Mannheim.

Il s'agit donc de diluer de 1 à 3 µg de l'ADN qui servira de sonde dans un volume total de 15 µl complété avec de l'eau ultrapure. Cet ADN est ensuite dénaturé dans l'eau bouillante pendant 10 minutes et rapidement plongé dans la glace. Ensuite, 2 µl d'*hexanucleotide mixture* et 2 µl de *dNTP labeling mixture* (fournis dans la trousse *DIG High Prime DNA Labeling and Detection Starter Kit II*) sont ajoutés au tube sur glace. Subséquemment, 1 µl de l'enzyme *Klenow* (concentration finale de 100 U/ml) est ajouté. Le tout est bien mélangé et incubé à 37°C pendant 20 heures. Cette réaction est stoppée par l'ajout de 2 µl d'une solution d'EDTA (200 mM EDTA, pH 8,0).

### 3.6.1.2 Incorporation de la digoxigénine-11-dUTP durant le PCR

Cette technique nécessite une quantité d'ADN de départ beaucoup moins importante que pour la technique précédente. De plus, elle permet d'obtenir une quantité importante de sonde marquée. Elle consiste en l'incorporation de la molécule DIG-dUTP par la *Taq DNA Polymerase (Expand High Fidelity)* durant le PCR (figure 3.3).

Dans un premier temps, la réaction qui servira à produire la sonde doit être optimisée. Par conséquent, plusieurs PCR doivent être réalisés afin de déterminer les conditions optimales. Ainsi, il faut varier les paramètres suivants : température d'hybridation, temps d'élongation, concentration en MgCl<sub>2</sub>, quantité d'ADN, concentration d'amorces et de dNTP, la concentration de la polymérase, température de dénaturation et temps de dénaturation et d'extension. Une fois que cette optimisation est effectuée, la sonde est fabriquée selon le protocole indiqué dans le *DIG System User's Guide for Filter Hybridization* pour le *DIG PCR Probe Synthesis Kit* de *Boehringer Mannheim*. Ainsi, le tableau 3.3 présente les différentes concentrations utilisées.

Les conditions d'amplification doivent aussi être optimisées. Par conséquent, ce ne sont pas exactement les conditions suggérées dans le protocole qui sont utilisées. Les conditions mises au point pour toutes les sondes préparées sont les suivantes :

- 1) Étape de dénaturation initiale de trois à quatre minutes à 95°C

- 2) Étape de dénaturation de 40 secondes à 95°C
- 3) Étape d'hybridation des amorces de 40 secondes à 57°C
- 4) Étape d'extension de trois minutes à 72°C
- 5) Étape finale d'extension de cinq à six minutes à 72°C

Les étapes 2-3-4 sont répétées 34 fois supplémentaires (35 cycles) avant d'en arriver à l'étape d'extension finale.

**TABLEAU 3.3**

**Concentrations utilisées lors des réactions de PCR pour la préparation des sondes**

Réactif	Concentration finale
Eau stérile	-
<i>Tampon PCR avec MgCl<sub>2</sub> 10X</i>	1X
<i>dNTP stock solution 10X</i>	250 µM
Amorce 1 (20 µM)	0,5 – 1 µM
Amorce 2 (20 µM)	0,5 – 1 µM
Polymérase <i>Taq</i> ( <i>Expand High Fidelity</i> )	2,5 U/50 µl
ADN plasmidique	2,5 ng/50 µl

Par la suite, 10 µl du produit de PCR sont utilisés afin de vérifier le produit obtenu par migration sur gel d'agarose. Généralement, le produit de PCR de l'amplification avec la digoxigénine migrera en parallèle avec la même amplification, mais sans la digoxigénine. Normalement, les deux bandes obtenues ne migrent pas tout à fait à la même vitesse, la digoxigénine ralentissant la vitesse de migration dans le cas de l'ADNc marqué. Ce décalage entre les deux bandes sur le gel d'agarose peut varier, mais il indique que la réaction de marquage a bien fonctionné. Ces étapes comprenant la migration et la photographie du gel se font exactement tel que décrit dans la section 3.2. Cependant, dans le cas où la bande désirée n'est pas la seule à apparaître sur le gel, elle est découpée et purifiée, selon les informations présentées dans la section 3.3.2.

### 3.6.1.3 Quantification de la sonde

Afin de déterminer la quantité réelle de sonde obtenue, il faut procéder à sa quantification. Cette étape est aussi décrite dans le manuel fourni avec la trousse de la digoxigénine (Eisel, 1995).

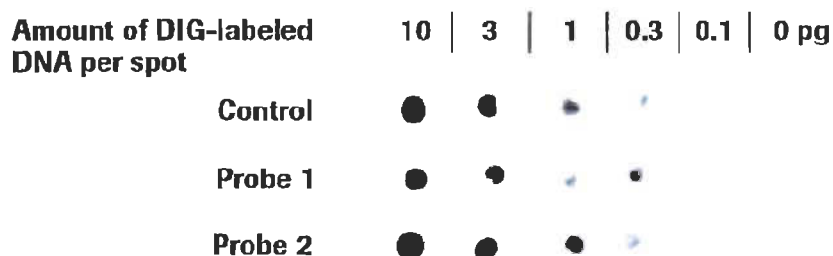
Dans un premier temps, la solution de sonde marquée doit être diluée à une concentration approximative de 1 µg/ml. Cette concentration peut être estimée selon les standards établis par la compagnie *Boehringer Mannheim* dans leur guide. Par la suite, cette dilution servira à préparer cinq autres dilutions (300 pg/µl, 100 pg/µl, 30 pg/µl, 10 pg/µl et 3 pg/µl). Parallèlement, des dilutions de l'ADN contrôle fourni seront aussi préparées. Un seul microlitre de chacune de ces dix dilutions sera déposé sur une pièce de membrane de nylon chargée positivement, laquelle sert aussi à l'hybridation (voir le traitement des membranes dans la section 3.6.3). Ainsi, cinq points correspondant aux différentes dilutions de la sonde marquée seront alignés parallèlement à cinq autres points correspondant aux dilutions équivalentes de l'ADN contrôle. Une fois les gouttes séchées, l'ADN est fixé à la membrane par exposition aux rayons ultraviolets (UV).

Ces membranes sont ensuite incubées successivement dans des solutions précises pendant un certain temps. Les détails concernant les différentes solutions sont présentés plus loin (voir la section 3.6.3 sur le traitement des membranes). Voici donc les différentes solutions et le temps d'incubation des membranes pour chacune des solutions correspondantes :

- |                         |           |
|-------------------------|-----------|
| 1) Solution de blocage  | 2 minutes |
| 2) Solution d'anticorps | 3 minutes |
| 3) Solution de blocage  | 1 minute  |
| 4) Tampon de lavage     | 1 minute  |
| 5) Tampon de détection  | 1 minute  |

Suite à la dernière incubation, la membrane sera traitée tel qu'indiqué dans la section

3.6.3 concernant le traitement des membranes. Il y aura donc incubation de la membrane avec le substrat, puis exposition au film pour finalement le développer et comparer les signaux (figure 3.4).



**Figure 3.4** - Résultat d'une quantification. Exemple du résultat obtenu dans le cas d'une quantification parallèle de deux sondes différentes (deux rangées du bas). La première rangée (le contrôle) est utilisée en tant que référence pour estimer la concentration des sondes fabriquées. Image tirée du site Internet de *Boehringer Mannheim*.

C'est donc la comparaison des signaux obtenus pour l'ADN contrôle avec les signaux de la sonde marquée qui permettra de déterminer la concentration initiale de la sonde qui devra être utilisée.

### 3.6.2 Titrage des banques d'ADNc

Les banques d'ADNc de rétine neurale et d'EPR ont toutes deux été construites dans le vecteur  $\lambda$ TriplEx2<sup>TM</sup> provenant de *Clontech*. Ces banques ont été préparées par le Dr Salesse et le Dr Hamel à l'INSERM U254 à Montpellier, en France. La préparation des pétris a été réalisée selon les indications du protocole fourni avec ce vecteur ( $\lambda$ TriplEx<sup>TM</sup> &  $\lambda$ TriplEx2<sup>TM</sup> Libraries User Manual).

Tout d'abord, les cellules hôtes adaptées pour ces banques sont les cellules dites XL1-Blue. C'est la souche XL1-Blue MRF' de *Clontech* qui a été utilisée. Cette souche présente un gène de résistance à la tétracycline.

### 3.6.2.1 Culture des cellules hôtes XL1-Blue

Dans un premier temps, il s'agit de préparer un pétri-mère sur lequel on prélèvera au besoin les colonies désirées pour préparer le criblage. Ce pétri-mère qui contient l'antibiotique approprié devra être refait à tous les deux semaines afin de toujours disposer de colonies fraîches.

Les cellules souches XL1-Blue MRF fournies par la compagnie sont conservées dans un microtube à -80°C. Il faut donc prélever un fragment de la glace du microtube contenant les bactéries et ensemer un pétri LB/tétracycline, qui sera le pétri-mère (pétris LB/tétracycline : 10 g/l bacto-tryptone, 5 g/l extrait de levure, 5 g/l NaCl, 15 g/l agar et 15 µg/ml tétracycline). À partir d'une des colonies isolées obtenues, un pétri LB/tétracycline/MgSO<sub>4</sub> (pétris LB/tétracycline/MgSO<sub>4</sub> : 10 g/l bacto-tryptone, 5 g/l extrait de levure, 5 g/l NaCl, 15 g/l agar, 15 µg/ml tétracycline et 10 mM MgSO<sub>4</sub>) est ensemené. En utilisant une de ces dernières colonies, un bouillon LB/MgSO<sub>4</sub>/maltose (bouillon LB/MgSO<sub>4</sub> : 10 g/l bacto-tryptone, 5 g/l extrait de levure, 5 g/l NaCl, 0,2% maltose et 10 mM MgSO<sub>4</sub>) de 15 ml est ensemené et incubé à 37°C, pendant une nuit avec agitation (140 rpm) ou jusqu'à ce que la densité optique à 600 nm atteigne 2,0. La solution bactérienne est alors centrifugée pendant 5 minutes à 5000 rpm, le surnageant est jeté et le culot bactérien est resuspendu dans 7,5 ml de MgSO<sub>4</sub> 10 mM.

De grands pétris (150 mm) LB/MgSO<sub>4</sub> préchauffés à 37°C dans un incubateur sont séchés pour les prochaines étapes en laissant le couvercle légèrement ouvert lors du préchauffage. Les pétris sont renversés (couvercle vers le bas) lorsqu'ils sont incubés.

### 3.6.2.2 Préparation des pétris

Diverses dilutions de la banque de phages sont préparées afin d'obtenir un nombre de plages de lyse pour les différents pétris variant de 0 à 500 pfus (*plaque forming units*) par pétri, zéro étant le contrôle négatif. Ces dilutions sont réalisées dans le tampon de

dilution lambda 1X (tampon de dilution lambda 1X : 0,1 M NaCl, 10 mM  $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ , 35 mM Tris-HCl (pH 7,5) et 0,01% de gélatine).

Un tube de 50 ml stérile est utilisé pour chacun des pétris qui servira au titrage. Différents volumes des dilutions préparées seront déposés dans chacun des tubes puis, pour chacun, 100  $\mu\text{l}$  de tampon de dilution lambda 1X et 200  $\mu\text{l}$  de la solution bactérienne préparée précédemment sont ajoutés. Ces tubes sont ensuite incubés à 37°C pendant 15 minutes, sous faible agitation (110 rpm). Huit millilitres de LB/ $\text{MgSO}_4$  top agarose (45°C) sont alors ajoutés à chacun des tubes (LB/ $\text{MgSO}_4$  top agarose : 10 g/l bacto-tryptone, 5 g/l extrait de levure, 5 g/l NaCl, 10 mM  $\text{MgSO}_4$  et 7,2 g/l d'agarose), puis rapidement inversé quelques fois et aussitôt versés sur les grands pétris préchauffés et séchés. Le top agarose doit être versé uniformément sur le pétri et sans bulle. Les pétris sont laissés à la température de la pièce le temps que le mélange d'agarose se durcisse, puis incubés inversés à 37°C pendant environ neuf heures, ou jusqu'à que les plages de lyse apparaissent bien définies, sans se toucher. À ce stade, les pétris sont conservés à 4°C. Les plages de lyse sont alors comptées et le titre déterminé. Le titre est généralement donné en terme de pfus/ml (*plaque forming units* par millilitre) et doit s'élever au-dessus de  $1 \times 10^6$  pour être intéressant; un titre de  $1 \times 10^9$  est excellent.

Cette étape est préalable à toute utilisation de banques de phages afin de déterminer si le titre est adéquat pour le criblage et préciser les volumes qui devront être utilisés. Il s'agit donc d'une vérification. Par conséquent, les résultats des différents titrages ne seront pas présentés dans la partie des résultats.

### 3.6.3 Le traitement des membranes

Une fois les pétris préparés, comme décrit ci-dessus, et refroidis quelques heures, il est possible de les utiliser pour préparer les différentes membranes qui serviront à l'hybridation. Les membranes utilisées sont des membranes de nylon chargées positivement (*Hybond-N+*, *Amersham Pharmacia Biotech*).



Toutes les étapes relatives au traitement des membranes nécessitent le port de gants sans poudre et des instruments propres. Le protocole suivi pour ces étapes est celui qui est indiqué par la compagnie *Boehringer Mannheim (The DIG System User's Guide for Filter Hybridization)* puisque c'est le *DIG High Prime DNA Labeling and Detection Starter Kit II* de *Roche* qui est utilisé pour l'hybridation.

Les membranes de nylon sont numérotées à l'aide d'un stylo à bille. La membrane, tenue par deux extrémités, est ensuite déposée - face numérotée vers le pètri - sur le pètri en déposant initialement la partie centrale, puis en laissant progressivement les côtés adhérer au milieu. La membrane est ensuite marquée à quatre endroits différents à l'aide d'une aiguille stérile, afin de conserver des points de repère sur la gélose et la membrane. Après deux minutes, la membrane est retirée et une deuxième membrane y est déposée. Cette deuxième membrane sera marquée exactement aux mêmes endroits que la membrane précédente, puisqu'il s'agit d'un duplicata. Elle est retirée du pètri après trois minutes.

Les deux membranes reposent ensuite brièvement sur du papier *Whatman 3MM*. Elles sont alors déposées, face vers le haut, sur une nouvelle feuille de papier *Whatman 3MM* imbibé de la solution de dénaturation pendant cinq minutes (solution de dénaturation : 0,5 N NaOH et 1,5 M NaCl). Elles sont déposées brièvement sur du papier *Whatman* sec entre tous les traitements. Par la suite, elles sont placées sur du papier *Whatman* imbibé de la solution de neutralisation pendant cinq à 15 minutes (solution de neutralisation : 1,0 M Tris-HCl (pH 7,5) et 1,5 M NaCl), puis 10 minutes sur du papier *Whatman* imbibé de tampon SSC 2X (SSC 2X : 0,3 mM NaCl et 30 mM citrate de sodium (pH 7,0)). Les membranes sont ensuite laissées sur du papier *Whatman* sec jusqu'à ce qu'elles soient sèches. Elles sont alors soumises aux rayons UV (120000  $\mu\text{J}/\text{cm}^2$  pendant 30 secondes avec l'appareil *UVC 500 Crosslinker* de *Hoefer*) afin de fixer l'ADN à la membrane.



### 3.6.3.1 Le traitement des membranes avec la protéinase K

Dans certains cas, les membranes ont été traitées à la protéinase K avant l'hybridation. Cette étape permet de se débarrasser des protéines chargées possiblement présentes sur la membrane, sans quoi leur charge pourrait interagir avec les sondes marquées. Lors de ce traitement, les membranes sont déposées sur une feuille de papier d'aluminium et environ un millilitre de la solution de protéinase K (2 mg/ml) est versé uniformément sur chacune des membranes. Une seconde feuille d'aluminium est utilisée pour recouvrir les membranes hermétiquement. Cet emballage est ensuite placé à 37°C pendant une heure. Lorsque cette incubation est achevée, une feuille de papier *Whatman* 3MM entièrement imbibée d'eau ultrapure est placée directement sur la face supérieure des membranes, puis une bouteille, avec un peu de pression, est roulée sur ces dernières. Les débris collent alors au papier et peuvent ainsi être enlevés au moment où le papier est délicatement retiré.

À partir de ce moment, les membranes sont prêtes à être utilisées pour l'hybridation. Dans le cas où ce traitement n'est pas effectué, les membranes sont utilisées tout de suite après l'étape de la fixation aux UV.

### 3.6.4 Hybridation et détection

Les membranes sont placées dans les bouteilles d'hybridation, et environ 20 ml (par membrane) de la solution de préhybridation sont ajoutés à chacune des bouteilles (solution de préhybridation : SSC 5X, 0,1% N-lauroylsarcosine, 1% *Blocking Reagent* - fourni dans la trousse de *Boehringer Mannheim* concentré à 10% - et 0,02% SDS). La préhybridation est réalisée à 68°C dans un four à hybridation rotatif (*Boekel : VWR Scientific Products*) pendant une heure. La solution de préhybridation est ensuite jetée et remplacée par la solution d'hybridation (solution de préhybridation contenant la sonde).

Environ dix millilitres par membrane sont ajoutés dans les bouteilles d'hybridation.

Cette solution d'hybridation doit être préchauffée à la température d'hybridation (68°C) avant d'être utilisée. Selon la stringence voulue, différentes températures d'hybridation peuvent être essayées. Dans les manipulations présentées ici, les plus faibles températures d'hybridation utilisées étaient de 50°C, mais la majorité des hybridations se sont déroulées à 68°C.

Suite à l'hybridation, les membranes sont soumises à deux lavages différents. Tout d'abord, elles sont lavées deux fois pendant cinq minutes à la température de la pièce dans la solution de lavage 2X (solution de lavage 2X : SSC 2X et 0,1% SDS) avec une faible agitation. Elles sont ensuite égouttées puis lavées deux fois pendant 15 minutes à 68°C dans la solution de lavage 0,5X (solution de lavage 0,5X : SSC 0,5X et 0,1% SDS). La concentration de cette dernière solution en SSC pourrait être diminuée à 0,1X pour augmenter la stringence. De même, il suffit d'augmenter la concentration en sels pour diminuer la stringence.

Suite aux lavages post-hybridation, les membranes sont équilibrées dans un tampon de lavage (tampon de lavage : tampon d'acide maléique (0,1 M acide maléique et 0,15 M NaCl dont le pH est ajusté à 7,5 à l'aide de NaOH solide) et 0,3% Tween 20) avec une faible agitation pendant une minute. Le substrat pour la chimiluminescence est alors placé à la température de la pièce afin qu'il puisse réchauffer avant son utilisation.

Les membranes sont ensuite incubées pendant 30 à 60 minutes avec une faible agitation dans la solution de blocage (solution de blocage : *Blocking Solution* 10X dilué à 1X à l'aide du tampon d'acide maléique). L'incubation suivante est effectuée dans la solution d'anticorps (solution d'anticorps : *Anti-Digoxigenin-AP* dilué 1 : 10000 dans la solution de blocage) pendant 30 minutes, sous agitation à la température de la pièce. Les étapes de lavage sont réalisées dans des plats en plastique propres, alors que les étapes de blocage et d'incubation avec l'anticorps sont réalisées séparément dans des pétris appropriés. L'utilisation des pétris permet de restreindre la quantité de solution et d'anticorps nécessaires. Par conséquent, pour chacune des membranes, 1 µl d'anticorps

dans 10 ml de solution de blocage est utilisé pour l'incubation dans la solution d'anticorps.

Une fois les membranes retirées de la solution d'anticorps, elles sont replacées dans un plat de plastique contenant le tampon de lavage. Les membranes sont alors lavées deux fois pendant 15 minutes, agitées délicatement, à la température de la pièce. Elles sont ensuite équilibrées dans le tampon de détection (tampon de détection : 100 mM Tris-HCl (pH 9,5) et 100 mM NaCl) pendant deux à cinq minutes. Il est très important qu'en aucun moment suivant l'hybridation les membranes ne s'assèchent, car il pourrait en résulter un bruit de fond très élevé.

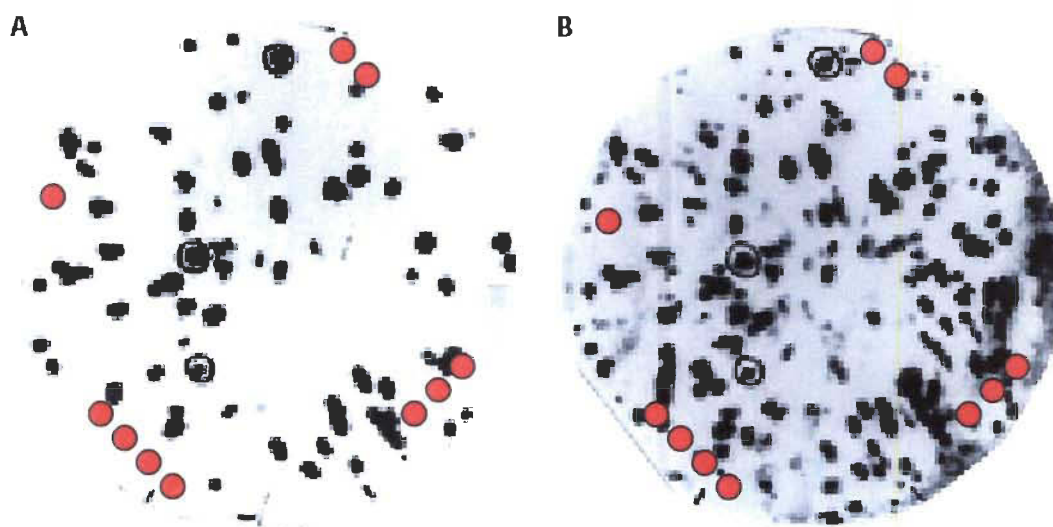
Suivant cette dernière équilibration, les membranes sont brièvement égouttées, puis placées sur une pellicule plastique (*Saran*) se trouvant dans une cassette utilisée pour le développement de films. Des gouttes du substrat utilisé, le *CSPD ready-to-use* fourni par la compagnie sont alors déposées uniformément sur les membranes, qui sont ensuite recouvertes d'une autre pellicule plastique, sans créer de bulles d'air. Les membranes reposent ainsi pendant cinq minutes à la température de la pièce dans la cassette, puis le surplus de substrat est enlevé à l'aide d'un essuie-tout propre passé doucement sur la pellicule plastique. Les membranes emprisonnées entre les deux pellicules de plastique sont ensuite incubées pendant 15 minutes à 37°C dans un four à incubation dans le but de diminuer de temps d'exposition au film.

Enfin, les membranes sont replacées dans la cassette et exposées à un film *Kodak X-OMAT AR* pendant environ 15 minutes. Ce film est ensuite développé selon les techniques traditionnelles.

Ainsi, les différentes étapes relatives à la détection consistent globalement en l'utilisation d'un anticorps qui reconnaît la molécule de digoxigénine et qui s'y fixe au niveau de l'ADN lié à la membrane, puis en la transformation du CSPD, par la phosphatase alcaline couplée à l'anticorps, en un composé instable (anion phénolate) qui

émet à une longueur d'onde de 477 nm. Cette émission est enfin détectée à l'aide d'un film *Kodak* approprié.

Le résultat obtenu pour chaque paire de membranes (*duplicata*) sont comparés afin de discriminer les signaux positifs et les faux positifs, qui sont en fait des signaux ne se trouvant que sur une seule des deux membranes plutôt que sur les deux. Pour ce faire, les films des *duplicata* sont placés de telle sorte qu'ils ont la même orientation, selon la position des points de repère faits initialement (figure 3.5).



**Figure 3.5** - Résultat typique d'un criblage de banque d'ADNc (banque de phage) où les clones positifs sont nombreux. La membrane de gauche serait le résultat du *duplicata* (A) alors que la membrane de droite serait la membrane initiale déposée sur le pètri (B). Remarquez les points rouges qui ont été ajoutés volontairement afin de donner une idée de la disposition des points de repère sur la membrane (la grosseur des points est évidemment très exagérée puisqu'il ne s'agit que de petits trous d'aiguille en réalité). Ce sont ces points qui permettent d'orienter les deux membranes exactement de la même façon pour déterminer les signaux positifs (encadrés). Adapté d'une image tirée du guide de *Boehringer Mannheim*.

### 3.6.5 Purification des clones positifs

Une fois les clones positifs repérés sur les pètris initiaux, ils sont prélevés et réutilisés pour un criblage secondaire. Le résultat de ce criblage devrait mener à des membranes contenant beaucoup plus de signaux positifs et, ultimement, seulement des signaux

positifs pour l'ensemble des plages de lyse si le clone était bien isolé lorsqu'il a été prélevé du pétri.

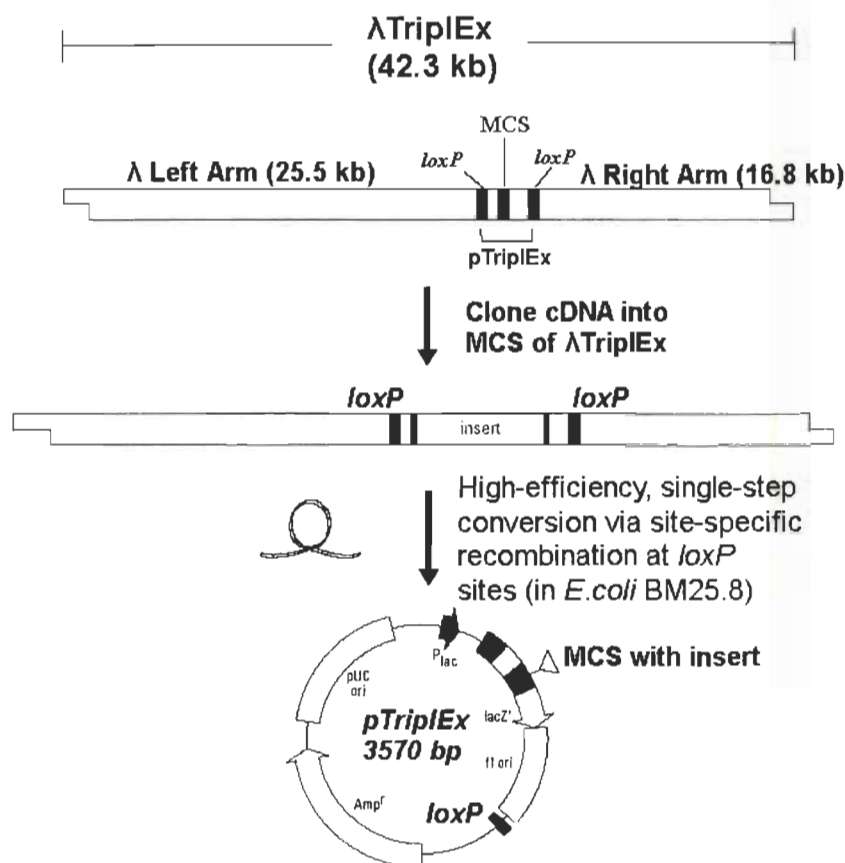
La technique consistant à prélever les clones positifs est réalisée à l'aide d'une pipette pasteur stérile. L'ouverture la plus large est utilisée pour prélever le ou les clones positifs présents dans la phase top agarose du pétri. Chaque clone prélevé est déposé séparément dans un microtube de 1,5 ml contenant 500  $\mu$ l de tampon de dilution lambda 1X et une goutte de chloroforme dans le but d'éliminer les contaminations bactériennes. Le microtube est ensuite vortexé vigoureusement pendant quelques minutes, pour permettre au phage d'être relâché dans la solution, puis placé à 4°C pour la nuit afin de permettre au processus d'élution de se poursuivre. Chacune des solutions de phages obtenues doit alors être titrée et le criblage doit être répété afin de vérifier si ces clones sont bel et bien positifs. Si lors de ce second criblage la majorité des clones s'avèrent positifs et très bien isolés, il est possible de prélever seulement deux ou trois clones positifs et de procéder à leur élution dans le tampon de dilution lambda. Ces phages sont considérés comme étant purs et chacune des solutions de ces phages élués peut alors être utilisée pour la conversion du phage en plasmide.

### 3.6.6 Conversion de l'ADN du phage en ADN plasmidique

Pour cette conversion, le protocole et le matériel bactérien fourni par la compagnie *Clontech* est utilisé. La conversion de  $\lambda$ TriplEx2 en pTriplEx2 nécessite l'excision et la circularisation du plasmide complet à partir du phage recombinant. Le plasmide est issu d'une recombinaison médiée par la *cre recombinase* aux sites *loxP* (figure 3.6). Cette recombinaison a lieu automatiquement lorsqu'un phage est transduit dans une bactérie hôte exprimant la *cre recombinase*. Dans le système utilisé, la souche bactérienne *E. coli* BM25.8 cultivée à 31°C présente l'activité de cette enzyme.

Il s'agit donc de prélever une colonie isolée d'un pétri préalablement ensemencé avec la souche bactérienne *E. coli* BM25.8, pour ensuite inoculer 10 ml de bouillon LB/MgSO<sub>4</sub> dans un tube de 50 ml. Cette solution est incubée à 31°C pendant une nuit, sous

agitation (150 rpm) jusqu'à ce que la densité optique à 600 nm atteigne 1,1-1,4. À ce moment, 100 µl de MgCl<sub>2</sub> 1M sont ajoutés à la culture bactérienne.



**Figure 3.6** - Conversion du phage recombinant  $\lambda$ TriplEx en pTriplEx. Le site de clonage multiple (MCS) est situé à l'intérieur de la région plasmidique, laquelle est flanquée par des sites *loxP* aux jonctions  $\lambda$ . La transduction de  $\lambda$ TriplEx dans la souche *E. coli* BM25.8 promouvait la libération de la *cre recombinase* et ainsi la circularisation de pTriplEx aux sites *loxP*.

Par la suite, dans un tube de 20 ml, 200 µl de cette solution bactérienne et 150 µl de la plage de lyse positive sont ajoutés. Cette solution est incubée à 31°C pendant 30 minutes, sans agitation, puis 400 µl de bouillon LB (bouillon LB : 10 g/l bacto-tryptone, 5 g/l extrait de levure et 5 g/l de NaCl) sont ajoutés. Ce milieu est alors incubé à 31°C pendant une heure additionnelle, sous agitation (225 rpm).

Enfin, 1 à 10 µl de cette suspension de cellules infectées sontensemencés sur des pétris

LB/carbénicilline (10 g/l bacto-tryptone, 5 g/l extrait de levure, 5 g/l de NaCl, 15 g/l agar et 50 µg/ml carbénicilline) afin d'obtenir des colonies isolées. Ces pétris sont incubés à 37°C pendant une nuit. Parmi les colonies obtenues, plusieurs sont prélevées et l'ADN plasmidique de chacune est isolé afin de procéder ensuite au séquençage automatique.

### **3.7 Criblage d'une banque génomique**

Le criblage d'une banque génomique requiert essentiellement les mêmes étapes que le criblage de banques d'ADNc. Ainsi, le titre est établi de la même façon, à l'exception des bactéries hôtes utilisées qui sont normalement spécifiques au phage utilisé. En ce qui a trait à la préparation de la sonde, aux quantifications, au traitement des membranes, à l'hybridation et à la détection, les étapes sont les mêmes que celles qui ont été décrites dans la section précédente. Ainsi, seules les étapes différant du criblage des banques d'ADNc sont présentées dans la présente section.

#### **3.7.1 Titrage de la banque génomique**

La banque génomique utilisée est la banque génomique humaine de *Clontech EMBL3 SP6/T7* (H106j : placenta). Le protocole suivi pour la préparation des bactéries hôtes et des pétris est donc le *Lambda Library User Manual* fourni par la compagnie avec la banque. Par conséquent, la cellule hôte utilisée ici est la souche K802 de *E. coli*.

##### **3.7.1.1 Culture des cellules hôtes *E. coli* K802**

Tel que mentionné précédemment, il s'agit dans un premier temps de préparer un pétri LB sur lequel on prélèvera, au besoin, les colonies qui seront utilisées pour le criblage. Ce pétri ne contient pas d'antibiotique et devra être refait à toutes les deux semaines afin de toujours disposer de colonies fraîches.



Les cellules souches *E. coli* K802 sont conservées à -80°C. Il faut donc prélever un fragment de la glace contenant les bactéries et ensemencer un pétri LB (pétri-mère) (10 g/l bacto-tryptone, 5 g/l extrait de levure, 5 g/l de NaCl et 15 g/l agar). À partir d'une des colonies isolées obtenues, un deuxième pétri LB est ensemencé. Enfin, à partir d'une des colonies de ce dernier pétri, un bouillon LB/MgSO<sub>4</sub>/maltose de 15 ml est ensemencé et incubé à 37°C pendant une nuit avec agitation (140 rpm), jusqu'à ce que la densité optique à 600 nm atteigne 2,0.

### 3.7.2 Isolation de l'ADN à partir des lysats de phages

Tel qu'expliqué plus tôt, les plages de lyse positives bien isolées sont prélevées à l'aide de l'embout d'une pipette pasteur et déposées dans des microtubes de 1,5 ml contenant 500 µl de tampon de dilution lambda et une goutte de chloroforme. Après quelques rondes de criblage afin d'obtenir uniquement des clones positifs, on procède à l'isolation de l'ADN à partir des lysats de phages. Le protocole utilisé est celui qui est indiqué par le *QIAGEN Lambda Handbook* pour le *QIAGEN Lambda Mini Kit*.

Tout d'abord, lorsque le lysat est pur et que son titre est déterminé, il faut préparer des pétris où les plages de lyse seront confluentes. À partir de ces pétris, les manipulations indiquées dans le manuel de *QIAGEN* sont suivies. Ainsi, 10 ml de tampon SM (*Suspension Medium*) (Tampon SM : 50 mM Tris-HCl (pH 8,0), 0,1 M NaCl, 8 mM MgSO<sub>4</sub> et 0,01% gélatine) sont ajoutés à chacun des pétris et placés à 4°C pendant une nuit. Ce tampon est récupéré dans un tube stérile, puis 2 ml additionnels sont ajoutés au pétri pour nettoyer la gélose. Par la suite, 245 µl de chloroforme (concentration finale dans le lysat de 2% v/v) sont ajoutés aux 12 µl de lysat récupérés. Cette solution est vortexée vigoureusement et centrifugée à 10000g pendant 10 minutes pour se débarrasser de l'agarose résiduel et récupérer le surnageant.

Les manipulations suivantes sont exactement celles qui sont décrites dans le protocole fourni par la compagnie *QIAGEN*. Le produit final est une solution d'ADN de phages, laquelle peut être directement envoyée au séquençage avec les amorces appropriées afin



de déterminer l'insert présent entre les bras du phage. De même, il est possible d'utiliser des enzymes de digestion spécifiques (EcoRI) au phage utilisé pour exciser l'insert et en déterminer la longueur, mais cette manipulation n'a pas été faite au cours de ce travail de maîtrise.

### 3.8 Amplification rapide des extrémités de l'ADNc (RACE)

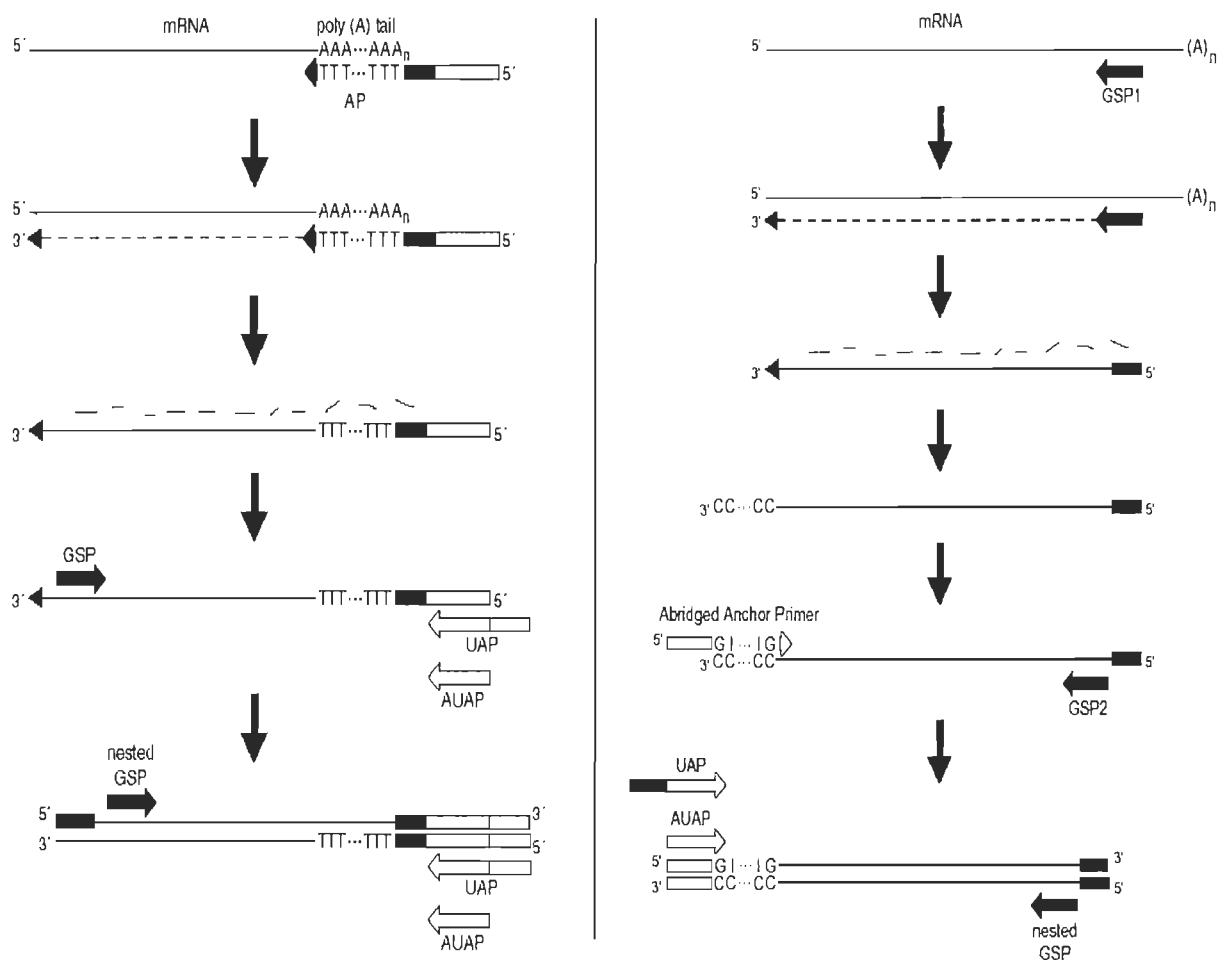
Lorsqu'un fragment d'ADNc pour un gène donné est obtenu et que les extrémités 5' et 3' sont manquantes, la technique généralement utilisée est le système RACE (*Rapid Amplification of cDNA Ends*). Cette technique permet, à l'aide de deux approches distinctes, d'obtenir les extrémités manquantes de l'ADNc. Les deux approches suivies ici sont exactement celles qui sont décrites dans les manuels d'instruction de *GIBCO BRL : 5' RACE System for Rapid Amplification of cDNA Ends Version 2.0* et *3' RACE System for Rapid Amplification of cDNA Ends*. La figure 3.7 est un schéma résumant les différentes étapes pour la procédure du RACE 5' et RACE 3'. Il importe de mentionner que pour chacune des étapes réalisées lors de cette technique, des contrôles positifs sont effectués. Ainsi, il est possible de vérifier que chacune des étapes a bien fonctionné.

Ainsi, à partir de l'ARN obtenu de la rétine neurale, ces deux approches ont été réalisées. Dans les deux cas, c'est la procédure permettant le clonage UDG (méthode de clonage utilisant l'*Uracyl DNA Glycosylase*), c'est-à-dire l'utilisation de l'amorce UAP (*Universal Amplification Primer*) qui a été choisie. Ainsi, les produits obtenus peuvent directement être clonés grâce au système *CloneAmp* de *GIBCO BRL* (figure 3.8).

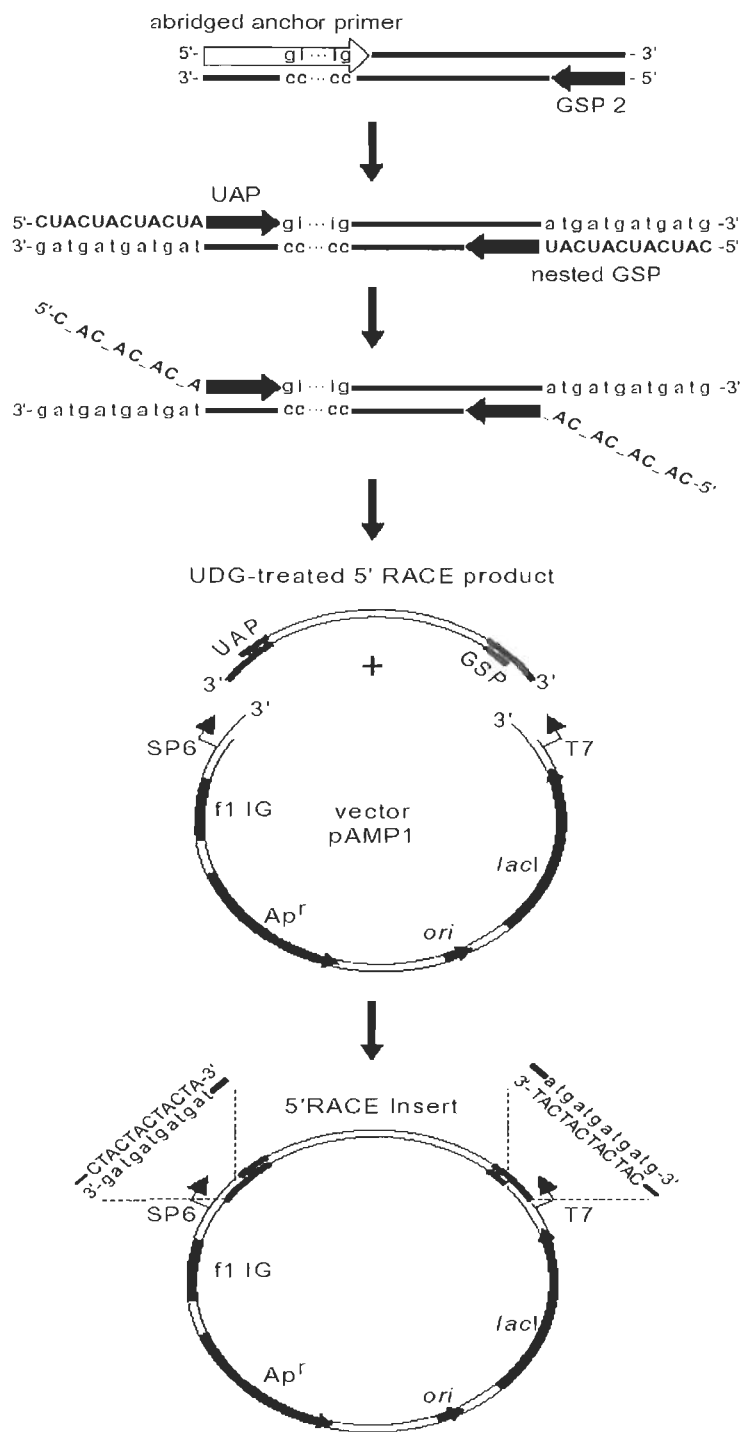
Les séquences des amorces utilisées lors de cette technique sont présentées dans le tableau 3.2.

Lorsque les divers produits obtenus par les différentes amplifications sont clonés dans le vecteur pAMP1, la transformation peut être réalisée selon les étapes décrites dans la

section 3.3.2. Cette fois, 1 à 5  $\mu$ l seront utilisés pour la transformation des bactéries compétentes *E. coli DH5 $\alpha$* .



**Figure 3.7** - Schématisation des procédures pour le système du RACE. Dans chacune des procédures, des amorces spécifiques sont utilisées (GSP : *gene specific primer*). La procédure pour obtenir l'extrémité 3' est présentée à gauche alors que la procédure pour l'extrémité 5' est présentée à droite. Dans le RACE 3', un oligo-dT appelé AP contenant une région adaptatrice permet de créer le premier brin d'ADNc via la SuperScriptII RT. L'ARN est ensuite dégradé, puis une seconde amplification est réalisée grâce à une amorce spécifique (GSP) et une amorce s'hybridant à la région adaptatrice (amorce UAP : *Universal Amplification Primer*). Le produit obtenu peut ensuite être amplifié à nouveau à l'aide d'amorces spécifiques plus internes (nested GSP). Dans le RACE 5', une amorce spécifique s'hybride d'abord pour créer le premier brin d'ADNc à l'aide de la SuperScriptII RT. L'ARN est ensuite dégradé par un mélange de RNases. Le produit obtenu est purifié et une queue poly-C y est ajouté en 3' par la *terminal déoxynucléotidyl transférase (TdT)*. Une amplification est ensuite faite à l'aide d'une amorce comportant une région poly-C jumelée à une région adaptatrice et une amorce spécifique plus interne (GSP2). Cette région adaptatrice permet une amplification subséquente à l'aide de l'amorce UAP et d'une amorce spécifique plus interne (nested GSP).



**Figure 3.8** - Schéma explicatif pour le clonage des différents produits obtenus lors des amplifications des extrémités de l'ADNc par le système *CloneAMP* de *GIBCO BRL*. D'abord un premier PCR utilisant l'amorce AAP (*Abridged Anchor Primer*) et une amorce spécifique (GSP2), puis une amplification subséquente avec une amorce spécifique à laquelle est ajoutée la séquence de clonage (CAU)<sub>4</sub>. L'enzyme UDG (*uracil DNA glycosylase*) enlève les bases déoxyuracyles. Les sites ne possédant plus de bases se détachent de l'autre brin ce qui crée une extrémité simple-brin. Le produit ainsi traité à l'UDG s'hybride au vecteur pAMP1 qui contient deux extrémités 3' à bouts cohésifs.

Ultimement, les divers plasmides seront séquencés, puis les séquences analysées afin de déterminer si les produits obtenus correspondent à la séquence cherchée. Cette analyse est réalisée à l'aide du système BLAST disponible sur le site Internet de la *NCBI*.

### 3.9 PCR sur différents ADNc de tissus humains

Des PCR ont été réalisés sur des tissus humains autres que la rétine neurale et l'EPR. Pour ce faire, ce sont les différents ADNc fournis dans la trousse *Human MTC<sup>TM</sup> Panel I* de *Clontech* qui a été utilisée. Cette trousse comprend en fait l'ADNc de différents tissus tels que le cerveau, le cœur, le placenta, le poumon, le foie, le muscle squelettique, le rein et le pancréas. De plus, l'ADNc de chacun de ces tissus est constitué d'un mélange d'ADNc provenant des tissus de 2 à 16 Caucasiens.

Lors de ces PCR, les conditions mentionnées précédemment (voir la section 3.2) ont été appliquées. Par contre, pour un volume réactionnel de 20 µl, seulement 3 ng d'ADNc du tissu ont été utilisés pour chacune des réactions. Les produits de PCR obtenus ont été analysés tel que décrit dans la section 3.2.

### 3.10 Criblage bioinformatique

Le criblage bioinformatique comporte plusieurs étapes différentes. La plupart des approches présentées ont été effectuées grâce à divers algorithmes de recherche disponibles sur Internet. Le premier algorithme utilisé est le BLAST (*Basic Local Alignment Search Tool*) qui permet, comme son nom l'indique, de repérer des identités de séquence entre la séquence soumise et les diverses séquences des banques de données de la *NCBI*. Ensuite, les algorithmes DIALIGN et TRANSLATE TOOL ont été utilisés respectivement pour l'alignement de différentes séquences et la traduction de séquences nucléotidiques; ils se trouvent sur le site d'Expasy (<http://www.expasy.ch>). Ces algorithmes sont les principaux outils utilisés lors des analyses bioinformatiques, bien que les algorithmes de GENSCAN (<http://genes.mit.edu/GENSCAN.html>) et

CLUSTALW (<http://www.ebi.ac.uk/clustalw/>), aussi disponibles sur Internet, aient aussi été utilisés.

L'algorithme BLAST comprend différents types d'analyses. Par exemple, en utilisant le programme blastx, il est possible de trouver des protéines similaires à la séquence nucléotidique soumise. Le tableau 3.4 résume les différents programmes pour l'algorithme BLAST.

**TABLEAU 3.4**  
**Les différents programmes de l'algorithme BLAST**

Nom du programme	Utilité
BLAST standard (blastn)	À partir des banques de données de séquences nucléotidiques, trouver des séquences nucléotidiques similaires à celle qui est soumise.
BLAST traduit (tblastx)	À partir des banques de données de séquences nucléotidiques traduites, trouver des protéines similaires à la protéine obtenue par la traduction de la séquence nucléotidique soumise.
BLAST traduit (blastx)	À partir des banques de données de séquences protéiques, trouver des protéines similaires à la protéine obtenue par traduction de la séquence nucléotidique soumise.
BLAST standard protéine (blastp)	À partir des banques de données de séquences protéiques, trouver des protéines similaires à la séquence protéique soumise.
BLAST traduit (tblastn)	À partir des banques de données de séquences nucléotidiques traduites, trouver des protéines similaires à la séquence protéique soumise.

Le but du criblage informatique est de repérer, parmi les milliers de séquences des banques de données de la NCBI, des séquences présentant des similitudes avec la ou les séquences soumises. Les séquences ainsi obtenues sont minutieusement analysées afin d'identifier un nouveau gène de la même famille, le groupe des cPLA<sub>2</sub>.

### 3.10.1 BLAST à l'aide d'une séquence protéique connue

Dans un premier temps, c'est la séquence protéique de chacune des cPLA<sub>2</sub> connues qui est utilisée afin de procéder à une recherche via le programme *tblastn*. Les paramètres utilisés sont de faible stringence, c'est-à-dire que même les séquences de plus faible identité seront sélectionnées. Cette analyse est réalisée avec la banque de données d'HTGS (*High Throughput Genomic Sequence*), puis sur la banque de données *nr* (*non redundant*). Les HTGS sont des séquences génomiques d'environ 200 000 paires de bases chacune qui ne sont pas complètement séquencées. Ils sont classés selon l'avancement du séquençage, en phase 0, 1, 2 ou 3, la phase 3 indiquant que le séquençage est complété. La banque de données *nr* comprend toutes les séquences des banques de la NCBI à l'exception des EST, des HTGS non complétés (phases 0, 1 et 2) et de certaines autres séquences sans intérêt pour ce travail de recherche.

Les séquences obtenues doivent être des séquences pour lesquelles aucune protéine n'a encore été attribuée. Cette vérification est effectuée grâce au programme *blastp*. Il s'agit en fait de vérifier que la ou les séquences trouvées présentent une identité seulement partielle aux cPLA<sub>2</sub> déjà connues et non pas une identité parfaite, ce qui indiquerait qu'il s'agit simplement de la séquence d'une PLA<sub>2</sub> déjà connue. Dans le cas où la séquence présente une identité importante, mais imparfaite, elle sera utilisée pour de plus amples analyses puisqu'il pourrait s'agir de la séquence d'une nouvelle cPLA<sub>2</sub>.

### 3.10.2 Analyse des clones trouvés

Ces analyses consistent à déterminer avec précision la séquence codante probable de chacun des HTGS trouvés. Les séquences trouvées sont en fait des HTGS non complétés (faisant partie de la base de données des HTGS) ou des HTGS achevés (faisant partie de la base de données *nr*). Ces séquences ont généralement une longueur variant autour de 200000 nucléotides, ce qui est énorme à analyser.

Au cours de ce travail de maîtrise, deux approches ont été utilisées. La première

consiste en une analyse des HTGS grâce à l'utilisation de plusieurs algorithmes en parallèle, alors que la deuxième consiste en à l'utilisation de l'algorithme GENSCAN, tous ces algorithmes étant accessibles via Internet.

### **3.10.2.1 Analyse des HTGS par la combinaison de divers algorithmes de recherche**

Dans cette approche, chacune des séquences nucléotidiques des clones trouvés (HTGS) est traduite en acides aminés. Pour ce faire, chacune des séquences doit être soumise au programme TRANSLATE TOOL (<http://www.expasy.ch/tools/dna.html>) par tranche de 20000 paires de bases (nombre maximal permis par traduction). Chaque traduction est analysée afin de déterminer quelle pourrait être la séquence codante du clone HTGS trouvé, c'est-à-dire la séquence codante pour le nouveau gène potentiel de la cPLA<sub>2</sub>. La détermination de ces régions codantes est basée sur le résultat obtenu lors des premiers BLASTs. Ainsi, il devient possible de supposer une séquence en acides aminés et une séquence nucléotidique qui pourraient correspondre à la nouvelle cPLA<sub>2</sub>.

Cette séquence en acides aminés, qui est en fait une protéine hypothétique, est ensuite soumise au programme *blastp* afin de déterminer son identité avec les cPLA<sub>2</sub> déjà connues. Cette analyse permet de déterminer quelles régions codantes sont possiblement manquantes pour la séquence protéique hypothétique. Par conséquent, lorsqu'une région semble manquante, le clone est repris et traduit à nouveau afin de repérer une région codante additionnelle qui pourrait permettre de compléter la séquence protéique hypothétique.

### **3.10.2.2 Analyse des HTGS par l'algorithme GENSCAN**

Une façon beaucoup plus rapide de déterminer les séquences protéique et nucléique des séquences codantes des HTGS, mais qui permet beaucoup moins l'analyse, est l'utilisation de l'algorithme GENSCAN. Cet algorithme, en utilisant la transformée de Fourier, permet de repérer l'emplacement de gènes potentiels sur chacun des HTGS, situant ainsi les exons sur la séquence de chaque clone HTGS.

Ainsi, lorsque les deux approches sont utilisées en parallèle, les différentes possibilités sont confrontées et la séquence codante la plus probable est déterminée. Enfin, lorsque la ou les séquences protéiques sont mises au point grâce aux diverses comparaisons et analyses bioinformatiques, elles servent à déterminer la séquence de diverses amorces spécifiques, afin de vérifier la présence de cette « nouvelle » cPLA<sub>2</sub> au niveau de la rétine neurale et de l'EPR par PCR.

### **3.10.3 Sélection de différents EST pour le criblage**

Tel que mentionné dans la section 3.6, des criblages ont été réalisés à l'aide d'EST. Ces EST ont été choisis à l'aide de l'algorithme de recherche BLAST, mais en sélectionnant spécifiquement la base de données des EST pour la recherche. La séquence utilisée pour la recherche était l'un ou l'autre des HTGS obtenus par le criblage bioinformatique.

L'emplacement des EST sur la séquence génomique est déterminant pour le choix des EST lors du criblage. Ainsi, les EST choisis devaient tous se trouver à proximité d'un gène d'intérêt, c'est-à-dire dans les régions flanquantes ou dans la zone de l'HTGS où se trouvent les différentes régions codantes de la « nouvelle » cPLA<sub>2</sub>.



## CHAPITRE 4

### RÉSULTATS ET DISCUSSION

#### 4.1 Recherche des cPLA<sub>2</sub> connues dans l'EPR et la rétine neurale

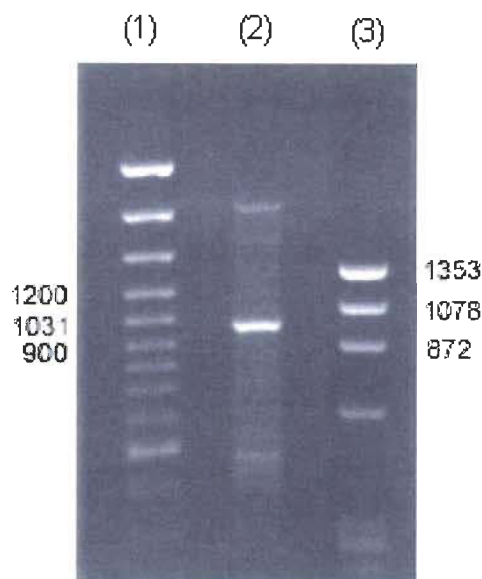
La détermination des cPLA<sub>2</sub> connues présentes dans l'EPR et la rétine neurale a été réalisée par deux approches. D'abord, par l'utilisation d'amorces spécifiques aux isoformes connues dans la littérature, puis par des amorces dégénérées, en se basant sur les différentes régions conservées présentes dans le groupe des cPLA<sub>2</sub>.

##### 4.1.1 RT-PCR à l'aide d'amorces spécifiques

Afin de trouver quelles cPLA<sub>2</sub> connues étaient présentes dans l'EPR et la rétine neurale, des amorces spécifiques aux trois différentes isoformes de cPLA<sub>2</sub> connues dans la littérature jusqu'à présent ont été utilisés en RT-PCR. Ainsi, différentes combinaisons d'amorces pour les cPLA<sub>2</sub> alpha, bêta et gamma ont été essayées. Les détails concernant ces différentes amorces sont présentés à la section 3.2.

En ce qui a trait à la cPLA<sub>2</sub> alpha, les amorces utilisées (voir tableau 3.2) n'ont pas permis de démontrer sa présence au niveau de l'EPR ou de la rétine neurale. Toutefois, par les amorces désignées « alpha courte », une amplification avec l'ADNc du poumon a permis de montrer que cette paire d'amorces était bel et bien fonctionnelle (figure 4.1). En effet, une bande d'environ 1000 pb a été obtenue et le clonage a permis de confirmer qu'il s'agissait bien de la cPLA<sub>2</sub> alpha.

La cPLA<sub>2</sub> bêta n'a pas été obtenue dans l'EPR ou la rétine neurale, mais les nombreuses combinaisons d'amorces utilisées (voir tableau 3.2) ne permettent pas non plus d'obtenir cette cPLA<sub>2</sub> dans d'autres tissus. D'autres amorces devront donc être utilisées, plus courtes encore, afin d'obtenir cette cPLA<sub>2</sub> dans les tissus où elle est exprimée en grande quantité, comme c'est le cas dans le pancréas (Pickard *et al.*, 1999).



**Figure 4.1** - Obtention de la cPLA<sub>2</sub> alpha dans le poumon. La piste (1) contient le marqueur de poids moléculaire GeneRuler 100bp DNA Ladder Plus. La piste (2) présente le fragment de la cPLA<sub>2</sub> alpha obtenu et la piste (3) contient le marqueur de poids moléculaire ΦX174 Hae III.

La cPLA<sub>2</sub> gamma a été obtenue à la fois dans la rétine neurale et dans les cellules de l'EPR. Toutes les combinaisons d'amorces utilisées (voir tableau 3.2) se sont avérées efficaces pour cloner cette cPLA<sub>2</sub>. La figure 4.2 présente la bande obtenue d'environ 1500 pb avec les amorces *gamma1* et *gamma2* (voir tableau 3.2) dans l'EPR. La bande obtenue dans la rétine neurale avec ces mêmes amorces est identique à celle qui a été obtenue dans l'EPR et le clonage a permis de confirmer dans les deux cas qu'il s'agissait bien de la cPLA<sub>2</sub> gamma.

Ainsi, une seule des cPLA<sub>2</sub> connues a été clonée au niveau de l'EPR et de la rétine neurale, la cPLA<sub>2</sub> gamma. La cPLA<sub>2</sub> alpha n'est définitivement pas présente dans ces tissus, mais d'autres amorces doivent être essayées en ce qui concerne la cPLA<sub>2</sub> bêta. Un des problèmes possibles avec l'amplification de cette cPLA<sub>2</sub> repose sur le fait qu'elle est majoritairement exprimée sous sa forme non épissée. Par conséquent, les différentes amorces utilisées pourraient amplifier des fragments trop longs pour les paramètres d'amplification utilisés. Toutefois, des amorces situées dans une partie exonique pour amplifier un fragment d'environ 600 pb n'ont pas permis de remédier à ce problème. D'autres amorces pour amplifier un fragment encore plus court pourraient être essayées

### 4.2.1 Approche expérimentale

L'approche expérimentale consiste en l'utilisation de deux techniques différentes : le RT-PCR via des amorces dégénérées et le criblage de banques à l'aide d'une sonde d'ADNc.

#### 4.2.1.1 RT-PCR à l'aide d'amorces dégénérées

Les séquences des amorces dégénérées ont été déterminées de façon à pouvoir cribler l'ensemble des cPLA<sub>2</sub>, c'est-à-dire qu'elles sont dessinées de telle sorte qu'elles peuvent s'hybrider sur n'importe laquelle des cPLA<sub>2</sub> connues, et par conséquent, de toute autre nouvelle cPLA<sub>2</sub> présentant les identités de séquence ciblées.

Cette approche n'a malheureusement donné aucun des résultats escomptés. Bien que de nombreux paramètres aient été essayés, aucune des bandes obtenues et clonées ne s'est avérée être une cPLA<sub>2</sub>.

Puisque même la cPLA<sub>2</sub> gamma n'a pas pu être obtenue par cette approche, il est possible que les amorces ne soient pas adéquates pour permettre une bonne hybridation lors de l'amplification par PCR, peu importe les conditions utilisées.

#### 4.2.1.2 Criblage de banques à l'aide d'une sonde d'ADNc

Suite à l'obtention du plasmide SG5 de la compagnie *Merck Frosst* contenant la séquence codante de la cPLA<sub>2</sub> alpha, un criblage des banques d'ADNc de rétine neurale et d'EPR a été réalisé. Pour chacune des banques et pour chaque condition particulière, environ 500000 clones ont été criblés. En fait, une première ronde de criblage a été réalisée dans des conditions de stringence moyenne afin de maximiser les chances d'obtenir des clones positifs. Toutefois, cette première ronde n'a pas permis d'obtenir de clones spécifiques. Tous les clones intéressants obtenus ont été isolés, purifiés et séquencés, mais aucun ne s'est avéré être l'ADNc d'une cPLA<sub>2</sub>. Par conséquent, trois autres rondes de criblage ont été effectuées dans des conditions de plus en plus

stringentes. Malheureusement, aucun de ces criblages n'a permis d'obtenir un clone de cPLA<sub>2</sub>.

Il est possible que le criblage d'un plus grand nombre de clones permette l'obtention d'un clone positif. Toutefois, le nombre de clones criblés permet normalement d'obtenir un clone spécifique, à moins qu'il ne soit que très faiblement représenté dans la banque.

Un dernier point pourrait aussi permettre d'expliquer l'inefficacité de cette approche. En fait, seulement une année après les premiers criblages réalisés sur les banques d'ADNc de rétine neurale et d'EPR humains, le titre de la banque est passé de l'ordre de  $10^9$  à  $10^6$ . Cette chute du titre pourrait donc être en partie responsable du faible nombre de copies de l'ADNc recherché dans ces banques, bien qu'elle soit difficile à expliquer, puisque le titre de la banque d'EPR bovin dont nous disposons aussi n'a presque pas changé en dix ans.

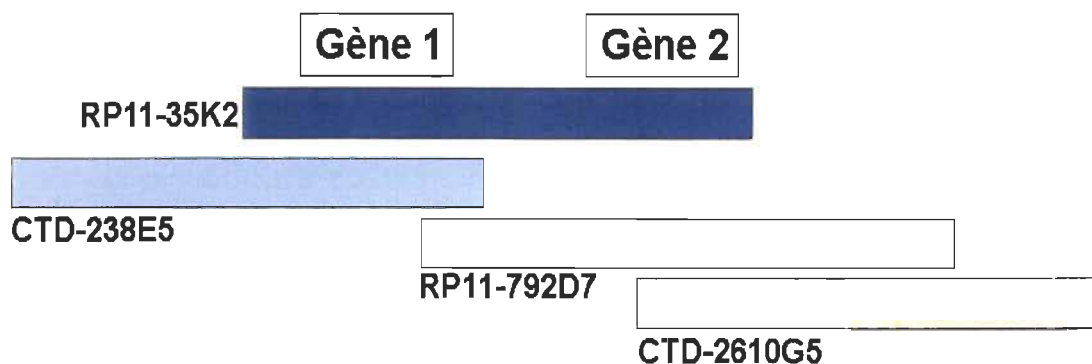
#### **4.2.2 Approche bioinformatique**

L'approche bioinformatique consistait uniquement à repérer de nouvelles séquences de cPLA<sub>2</sub> à partir des séquences génomiques connues, suite au séquençage du génome humain. Une fois les séquences potentielles identifiées, diverses techniques pouvaient être utilisées afin d'obtenir l'ADNc de ces « nouvelles » cPLA<sub>2</sub>. Les résultats obtenus pour le criblage bioinformatique, puis pour chacune des techniques subséquentes utilisées, sont présentés ci-après.

##### **4.2.2.1 Criblage bioinformatique**

Le criblage bioinformatique a permis de trouver un HTGS (ou clone), sur lequel se trouvaient deux séquences potentielles pour de nouvelles cPLA<sub>2</sub>. Ce clone est appelé RP11-35K2 et se trouve sur le chromosome 15; il s'étend sur 173660 nucléotides et est plus précisément situé sur le bras long du chromosome, dans la région entre 15q11 et 15q15.

L'une des deux séquences d'intérêt semblait incomplète; le clone se termine avant que l'on puisse déterminer s'il y a une autre partie codante en 5'. Par conséquent, des recherches sur le site de la NCBI au niveau du génome ont permis de trouver des HTGS se superposant à l'HTGS initial. Ainsi, ce sont les clones identifiés sous les appellations CTD-238E5, RP11-792D6 et CTD-2610G5 qui ont été repérés (voir figure 4.3). En fait, c'est le clone CTD-2610G5, se trouvant légèrement en aval du clone initial RP11-35K2 et possédant une extrémité 5' plus importante que les autres clones pour le nouveau gène 2, qui a permis de vérifier que toutes les parties codantes appartenant au nouveau gène 2 se trouvaient sur le clone RP11-35K2. Les deux séquences génomiques des nouveaux gènes trouvés à l'aide de ces HTGS sont présentées en annexe (annexe B).



**Figure 4.3** - Schéma représentant les HTGS trouvés. Il montre la position des différents HTGS trouvés en superposition avec le clone RP11-35K2 au niveau de chromosome 15, plus précisément dans la région s'étendant de 15q11 à 15q15. Les clones sont présentés dans le sens 5' vers 3', de la gauche vers la droite, alors que les gènes sont présentés dans le sens contraire, l'ATG de l'extrémité 5' se trouvant à droite.

L'analyse du clone RP11-35K2 par l'utilisation combinée de divers programmes bioinformatiques a permis de déterminer deux séquences codantes initiales pour des cPLA<sub>2</sub> présentant une identité d'environ 40% avec la cPLA<sub>2</sub> bêta. Par la suite, l'analyse du clone RP11-35K2 par le programme GENSCAN a permis de trouver que ce clone contenait trois gènes potentiels. Deux de ces gènes correspondaient sensiblement aux séquences codantes initiales, mais présentaient une extrémité plus longue en 5'. De plus, les séquences initiales présentaient, à un ou deux endroits, un exon de plus que les séquences obtenues via GENSCAN. Par conséquent, la combinaison des résultats obtenus de part et d'autre, c'est-à-dire par GENSCAN et l'utilisation combinée de programmes bioinformatiques, a mené à l'établissement de deux nouvelles séquences codantes potentielles pour de nouvelles cPLA<sub>2</sub>. La séquence se trouvant le plus en

amont sur le chromosome 15 a été appelée « nouveau gène 1 », alors que la séquence plus en aval a été nommée « nouveau gène 2 ». Les deux séquences sont présentées ci-après, à la figure 4.4.

#### cPLA<sub>2</sub> nouveau gène 1

```

atgggtgagcctggaaaaggtaccacaagttccattctgatctacaggactggcagcct
M G E P G K G T T S S H S D L Q D W Q P      20
ggcctgaagggcggtccagggtggtgtggtgagcgcctggccagcctttgggagtg
G L K G R S R V V W M A G P G P A F G S
gtcaggctaggcggtcaccagatgcagggtggacctggggacggcgccctgagcagcgc
V R L G G H P D A G W T W G R G P E Q R
cgagcagagcctccaccggaggcacaggaacccggcatcctcgcgatggcagcaatggc
R A E P P P E A Q E P G I L G D G S N G
tgccccgtggctgggtcccaaaagcggcaccgctcccactttcagaccgggccccagaa
C P A G W S P K R H R S H F Q T R A P E      100
gcgcatccccagttccccctcctggggctcctgctcccagtgcaagagcagcaccgctct
A H P Q F P L L G L L L P V Q E Q H R S
gggcccagttccacttcaggggcctgctctcccagtcgcgctgctccccagctgcgctg
G P S S T S G P C S P Q S R C S P A A L
ctcctctgccagcgggagacctggcccagccccatcgccgggccccagggcgagaca
L L C Q R A T W P S P I A A A P R A Q T
gttctgaggggtacgagggcagatactacagggggtcagccagttcttagttgtgtggca
V A E G Y E A D T T G G Q P V L S C V A
gagctctgtaaaccaggggcccccaacccctggccacagaccagtagcagctccgtggcc
E L C K P G A P N P L A T D Q Y Q S V A      200
tgtttggaaccagggccacagcaggtggtgaacagtgggcgagctcccagacttcatg
C L E P G H T A G G E Q W A S S P D F M
ggatggtttgctgaggggtatgtctggcccagcgatgggcatcgatccaccacctctgaa
G W F A E G M S G P A M G I G S T T S E
atgtggtgtgacacccctccccagacccactccccccaggagcctgacccctgggagc
M W L S D P S P D P T P T Q E P D P G S
ccttttgaggggctgtctccatgccacctgttgacagtgaggggtcatccggatgaaaaat
P L E G L S P C H L L T V R V I R M K N
gtccggcagggtgatatgcaaccagtaggtatagagctggcaccctgcctgcagggtccc
V R Q A D M Q P V G I E L A P C L Q A P      300
agcgtaccggagacagacctgaaggggtggtggtccaggccccgggtggggggccaggtt
S V P E T D L K G V V Q A R G G G A S V
ctggaaaagccaaggggaaggggtcaagaggggtgagcaggttcctgtgagccagacagac
L E K P R E G F K R A E Q V P V S Q T D
tgttttgtgagcctctgggtgccaccgcctctcagaagaagctgaggacaagaccatc
C F V S L W L P T A S Q K K L R T R T I
tccaactgccccaaatccagagtgaatgaaagcttcaacttccagatccagagccgagtg
S N C P N P E W N E S F N F Q I Q S R V
aagaacgtgctagagttgagtgctgtgtagaagacacagtgacaccagatgacctctc
K N V L E L S V C D E D T V T P D D H L      400
ctgacagttctctatgacctcaccaagctctgtttccgaaagaaaaccacgtgaagttt
L T V L Y D L T K L C F R K K T H V K F
ccactcaaccgcagggcatggaagagctggaggtggagttcctgctggaggagagtgccc
P L N P Q G M E E L E V E F L L E E S P
tctccacctgagaccctcgtcaccaatggcgtgctgggtgtaattatcttctgggttcc
S P P E T L V T N G V L V V I I F L G S
tgtagctccagagggccaggtggctgctgctcaggggaacaggaccaagggagaaaaa
C S S R G H G W L L L S G E Q D Q G R K
cagtgggcccagcttggtctctgtcctatcctgacctctgcaggagtttagactaaacgag
Q W A Q L G L C P I L T S A G V R L N E      500
gccagccaaatggggcacaggcagcactggggcacgagctggggcttctgtacagagga
A S Q M G H R Q H W G T S W G F C T E G
ggagtgaaggacctcctggtgatggtgaacgaatcctttgagaacacccagcgtgtccgg
G V K D L L V M V N E S F E N T Q R V R
ccctgcttggaaacctgtgccccacctctgcctgcttccaaaccgtgctgcttccac
P C L E P C C P T S A C F Q T A A C F H
taccccaagtacttccaggtccaggtgcacgtggaagtgcccaagagtcactggagctgt
Y P K Y F Q S Q V H V E V P K S H W S C
gggctttgctgccgctctcgcaagaagggccccatcagccagccctcgactgcctttcc
G L C C R S R K K G P I S Q P L D C L S      600

```

gatggtcaggatgatgaccctgctgtgggtgagagttatgaattacacatgaagtctaca  
 D G Q V M T L P V G E S Y E L H M K S T  
 ccctgccctgagacactggacgtgcggtgggttcagcctgtgcccagcagagctggag  
 P C P E T L D V R L G F S L C P A E L E  
 tttctgcagaagcggaaggtcgtggtggccaaggccctgaagcaggtgctgcagctggag  
 F L Q K R K R K V V V A K A L K Q V L Q L E  
 gaagacctgcaggaggacgaggtgccgtgatagccatcatggccactgggggtgaaca  
 E D L Q E D E V P L I A I M A T G G G T  
 agatccatgacctccatgtatggccacctgctggggctgcagaagctgaacctcctggac  
 R S M T S M Y G H L L G L Q K L N L L D 700  
 tgtgccagctacatcactggtctatcaggggccacctggacctgactacctgtaccgt  
 C A S Y I T G L S G A T W T M A T L Y R  
 gaccctgactggctcctccaaaacttgagcctgctatctttgaggtcggagacatgtg  
 D P D W S S K N L E P A I F E A R R H V  
 gtaaggacaagctaccctccctgttccagaccagctccgcaaattccaggaggagctc  
 V K D K L P S L F P D Q L R K F Q E E L  
 cggcagcgagccaggaaggtacaggggtcacctttacagacttctggggcctgctgata  
 R Q R S Q E G Y R V T F T D F W G L L I  
 gagacctgctggggagcaggttagaccaagacctgggactctgagagactgctctaaa  
 E T C L G D E V D Q D P G T L R D C P K 800  
 ctgaggatccagctcccgccctgtctggtggtgggcagcccgagggggtggtgctatg  
 L R I Q S R P L S G G G Q P E G V V P M  
 ccattggtgcctttcagagtgatttggctggtgctcagctggggcgatacagcagcccg  
 P L V P F Q S D L A G V S W G R Y S S P  
 agaaatgaatgcaactgtcagatcagcgtgctgctttgagctgcggccagaacccctg  
 R N E C K L S D Q R A A L S C G Q N P L  
 cccatctacctcaccatcaatgtcaaggatgatgtaagcaaccaggacttcagagagtg  
 P I Y L T I N V K D D V S N Q D F R E W  
 ttcgagttctccccctacgaggtgggtctgcagaagtatggggccttcatccctccgag  
 F E F S P Y E V G L Q K Y G A F I P S E 900  
 ctcttcggctccgagttcttcatggggcggtggtgaagaggatcccgaggtctcgaatc  
 L F G S E F F M G R L V K R I P E S R I  
 tgctacatgctaggcctgtggagcagcatcttctccctgaacctgctggatgcttgaac  
 C Y M L G L W S S I F S L N L L D A W N  
 ctgtcacacacctcgaggaggtttttccacaggtggacaaggagaaagtcagagacatc  
 L S H T S E E F F H R W T R E K V Q D I  
 gaagacgagccgatcctgcctgaaatcccaaatgtgatgctaacctcctggagaccag  
 E D E P I L P E I P K C D A N I L E T T  
 gtatgatcccagggtcatggtgtccaattctttccgagaaatccttaccatcggtcc  
 V V I P G S W L S N S F R E I L T H R S 1000  
 ttctgtctgagtttcacaaacttctgtctgggtgcagctgcacaccaactacctccag  
 F V S E F H N F L S G L Q L H T N Y L Q  
 aatggccagttctctaggtgaaagtgcctctgtccattgggtgtcagggcaccaggtgt  
 N G Q F S R W K V P L S I G C Q G T R C  
 gttggtaacgggaaggcttggctttcagacacagtgctagatggtttccaaacagctg  
 V G N G K A W L S D T V L D G F P N Q L  
 accgagtcgcgcaaccacctgtgcctgctggacactgcgttctttgtcaactccagctac  
 T E S A N H L C L L D T A F F V N S S Y  
 ccgcccctcctcaggccagagcgaagccgacctcatcatccacctcaactactgtgct  
 P P L L R P E R K A D L I I H L N Y C A 1100  
 ggtcccagacaaaggcaacctggttccagtcgccccttcagggtaccaccacatccac  
 G S Q T K A T W F Q S P F Q A T H H I H  
 ccaagtctgtgttggatgcgaacaccaacgaaagggttttgtccattttacttggga  
 P S L C L D A N T N E R A F V H F T W G  
 agcagcttgcaactggagcccctgaaacaaacctgtgagtactgactgtgcagaacatc  
 S S L Q L E P L K Q T C E Y C T V Q N I  
 cccttcccaaatcagagctgccagatgagaatgaaatctcaagggaatgctacctgatg  
 P F P K Y E L P D E N E N L K E C Y L M  
 gagaaccccggaacccgatgccccatcgtagcttttctccactcatcaatgacact  
 E N P Q E P D A P I V T F F P L I N D T 1200  
 ttccgaaaatacaaggcaccaggtgtagagcgaagccctgaggagctggagcagggccag  
 F R K Y K A P G V E R S P E E L E Q G Q  
 gtggacatttatggtcccaaaactccctatgccaccaaggagctgacatacacagaggcc  
 V D I Y G P K T P Y A T K E L T Y T E A  
 accttgacaagctggtgaaactctcagagtataacatcctgaataataaggacactctc  
 T F D K L V K L S E Y N I L N N K D T L  
 ctccaggtctgcggctgcagtgagagaagaagcgctgaaggccaggtgtccctcc  
 L Q A L R L A V E K K K R L K G Q C P S 1280  
 tag  
 -

cPLA<sub>2</sub> nouveau gène 2

```

atgactagatccagaacactaacgacaccaaatgctggcgaagacatggagcaacaggaa
M T R S R T L T T P N A G E D M E Q Q E 20
cttccactcattaatgctgcaaagactgacaccgagctgcaataaggcaccaggagct
L P L I N A A K T D T E S A I R H P G A
caactgaccgctgatggcaagctgattatacaggaccccttctgtcctgaaaaggacaat
Q L T A D G K L I I Q D P F C P E K D N
tcatacttactgaaatggacacacattctgtatgtggattacctttccggcttcctgcac
S Y L L K W T H I L Y V D Y L S G F L H
ctcagcaagcagcacaatccaacgacttacacaatgtctaattacatatgactgtgcat
L S K H D N P T T Y T M S N L H M T V H 100
aacagcaacttagaaaccctgcggctctcttcaagtgggggtatccacgtgtgacccat
N S N L E T L R L S S S G G Y P R V T H
cgagcaatgggtttccatctttttcaactctttttgagaatctgattaaagctatagac
R A M V F H L F Q T L F E N L I K A I D
ctgagctctctgcctcaggggaaccagcttccctgccaccactcctgggactcgtgcct
L S S L P Q G N Q L P A T T P G T R L P
ccccctggctgcatcttctgtcctgaaactccaggctctggaagaggtgctccgctgcag
P P G C I F C P E T P G L E E V L R L Q
gtcctggagccctgccagaggtttctggagatgctggttgaccagcctcttggggaaga
V L E P C Q R F L E M L V D P A S W G R 200
ggggcagggcgccagtgccagcagacgacaggggcaccggtgctagcagggaaacagggcg
G A G R Q C Q Q T T G A P V L A G N R A
ggctctgctcatttccctcttctgtcaaagaggactttggcatgaaggatgtgaagctg
G L L I S L F L S K R D F G M K D V K L
ttttctggggaggcctctacctgctggcagctcacagtggggtcctggaggcgcggaac
F S G E A S T C W Q L T V R V L E A R N
ctgcgctgggctgacctgttgagtgggaccccttacgtgatcctacagctgtcgacc
L R W A D L L S E A D P Y V I L Q L S T
gcacctggaatgaagttaagaccaagacgctcaccgacaccagtcacctgtgtggaat
A P G M K F K T K T L T D T S H P V W N 300
gaggccttccgtttccttatccaaagtccaggtcaagaatgttctggagcttagcatctat
E A F R F L I Q S Q V K N V L E L S Y
gatgaggactcagtcacggaggatgacatctgcttcaaggttctctatgacatctcagaa
D E D S V T E D D I C F K V L Y D I S E
gtcctccctggcaagctgctccgaaaaccttctccagagtcctccaggagaggaggag
V L P G K L L R K T F S Q S P Q G E E
ctggatgtggagtctctgatggaagaaacgtcagatcgccagaaaaacctcatcaccaac
L D V E F L M E E T S D R P E N L I T N
aaagtcatgtggcccgagagctgtcatgcctggatgtgcatctggacagcagggagc
K V I V A R E L S C L D V H L D S T G S 400
accgctgtggttgagatcaggacaagctggagctggagctggtgctgaaggggtcctat
T A V V A D Q D K L E L E L V L K G S Y
gaggacacacagacatccttctgggcacagcctctgccttccgcttccactacatggca
E D T Q T S F L G T A S A F R F H Y M A
gccctagagacagagctgagcgggccctgaggagctccagaagcaatggctggatggg
A L E T E L S G R L R S S R S N G W N G
gacaactcagctgggtacctcactgtgccctgagggccttgaccattgggaaggaggtg
D N S A G Y L T V P L R P L T I G K E V
actatggtgttctgtccaaatgccccaggagtgaggctgcagctcaaggcagagggc
T M D V P A P N A P G V R L Q L K A E G 500
tgccctgaggagctggccgtgcacctgggcttcaatctctgtgcagaggagcaggccttc
C P E E L A V H L G F N L C A E E Q A F
ctgagcaggaggaagcaggtggtggccaagccctgaagcaggccctgcagctggacaga
L S R R K Q V V A K A L K Q A L Q L D R
gacctgcaggaggatgaggtacctgtgtggcatcatggccacaggaggaggtgccgg
D L Q E D E V P V V G I M A T G G G A R
gccatgacctcactctacggccacctattggccttgagaagctgggcctcctagactgt
A M T S L Y G H L L A L Q K L G L L D C
gtgacctacttcagtggcatctctggtctacgtggacaatggccacctgtacggggac
V T Y F S G I S G S T W T M A H L Y G D 600
cctgagtggtgcagagggacctggagggacctatcagatacgcccgaggacacctggcc
P E W S Q R D L E G P I R Y A R E H L A
aagagcaagctggaggtcttttcccagagcgctggcgagctaccgcccggagctggag
K S K L E V F S P E R L A S Y R R E L E
ctgcgggctgagcagggccacccacgacctttgtggacctgtggcgctagtgtggag
L R A E Q G H P T T F V D L W A L V L E
tccatgctgcacggccaggtgatggatcagaagctgtcaggacagagagccgacctggaa
S M L H G Q V M D Q K L S G Q R A A L E
cggggtcagaacctctgcccctctacttgagcctcaatgtcaaagagaacaatctggag

```



```

R G Q N P L P L Y L S L N V K E N N L E 700
aactggacttcaaggagtggttgagttctccccctatgaggtcgggtttcctgaagtac
T L D F K E W V E F S P Y E V G F L K Y
ggggccttcgtccctcctgagctcttcggctccgagttcttcatgggacggctgatgagg
G A F V P P E L F G S E F F M G R L M R
aggatccccggagccccggatctgctttctggaagccatctggagcaacattttctccctg
R I P E P R I C F L E A I W S N I F S L
aacctgctggatgctggtatgacctcaccagttctggggagtcctggaacagcacatc
N L L D A W Y D L T S S G E S W K Q H I
aaggacaagaccagagcttagagaaggagccccctgaccacctcggggacctcctcgagg
K D K T R S L E K E P L T T S G T S S R 800
ctggaggcctcgtggtgcagccaggcacggcgctggccaggcatttaaaggcttctctg
L E A S W L Q P G T A L A Q A F K G F L
acaggcaggccccctccaccagcgcagcccaacttccctccaggcctccagctgcaccag
T G R P L H Q R S P N F L Q G L Q L Q L Q
gactactgtagccacaaagactttctccacctgggcagactaccagcttgactccatgcc
D Y C S H K D F S T W A D Y Q L D S M P
agccagctgaccccccaaggagccccggctctgctggtggagcgccgctacttcatcaac
S Q L T P K E P R L C L V D A A Y F I N
accagctctccctccatgttcgggccaggccgagggctggacctcatcctctccttcgac
T S S P S M F R P G R R L D L I L S F D 900
tactccctatctgcgcccttcgaggcactgcagcagacggagctgtactgcccggggccgg
Y S L S A P F E A L Q Q T E L Y C R A R
gggctgcccttccccgggtggaacccagccctcaggaccagcaccagccaagggaatgc
G L P F P R V E P S P Q D Q H Q P R E C
cacctcttctcagacccccgctgccccgaggccccgatcctgctgcaactccccgctggtc
H L F S D P A C P E A P I L L H F P L V
aatgctccttcaaggaccactcagccccgggtgtccagcgcagccccgcagagctccag
N A S F K D H S A P G V Q R S P A E L Q
ggtggccaagtggatctcaccggggccacctgcccctacacctgtccaacatgacctac
G G Q V D L T G A T C P Y T L S N M T Y 1000
aaggaggaagacttcgagcgcctgctgcggtcagtgactacaacgtgcagaccagccag
K E E D F E R L L R L S D Y N V Q T S Q
ggtgccatcctgcaggccctgaggaccgcgtgaagaccggactctagaggcgaggcct
G A I L Q A L R T A L K H R T L E A R P 1040
ccaagggcacagacctga
P R A Q T -

```

**Figure 4.4** - Séquences codantes des nouvelles cPLA<sub>2</sub> obtenues à partir des nouveaux gènes 1 et 2. Le gène 1 a une séquence codante de 3843 pb alors que le gène 2 en compte 3138.

Lorsque chacune de ces séquences est soumise au programme BLAST, plus précisément au *blastp*, pour déterminer le pourcentage d'identité avec d'autres protéines dans les bases de données de la NCBI, des pourcentages d'identité de 48 et 51% avec la cPLA<sub>2</sub> bêta sont obtenus, respectivement, pour les protéines associées au nouveau gène 1 et au nouveau gène 2. Ces pourcentages appuient donc l'hypothèse de la découverte de nouveaux gènes pour la famille des cPLA<sub>2</sub>.

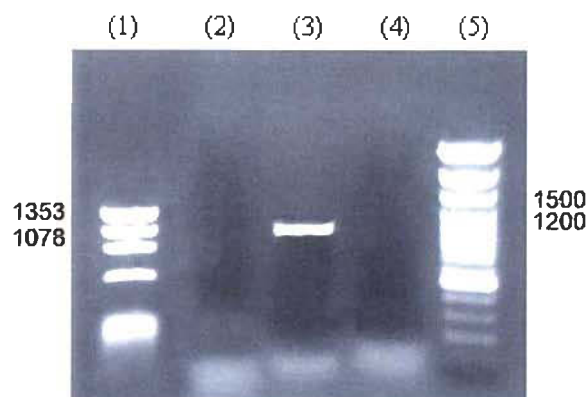
Une fois les séquences déterminées, il devient possible d'entreprendre différentes manipulations au laboratoire afin d'obtenir l'ADNc de ces clones. Comme l'objectif du projet consistait à trouver de nouvelles cPLA<sub>2</sub> spécifiques à la rétine neurale ou l'EPR, ce sont ces tissus qui ont été ciblés en premier lieu pour les manipulations suivantes.

#### 4.2.2.2 Criblage par RT-PCR

En se basant sur les deux séquences élaborées, de nombreuses combinaisons d'amorces spécifiques à l'une ou l'autre des deux séquences ont été utilisées. Les séquences de ces différentes amorces sont présentées dans le tableau 3.2.

##### 4.2.2.2.1 Obtention d'un fragment pour le gène 2

Suite à l'utilisation d'un nombre très important de combinaisons d'amorces et de paramètres, plusieurs bandes obtenues par RT-PCR ont été clonées et séquencées. Toutefois, une seule s'est avérée être un fragment de l'ADNc attendu. Ce sont les amorces « gène 2b » et « gène 2c » du gène 2 (voir tableau 3.2) qui ont permis l'obtention de cette bande d'environ 1100 pb (figure 4.5) au niveau de la rétine neurale. En effet, cette bande n'est pas obtenue au niveau de l'EPR.



**Figure 4.5** - Obtention d'un fragment d'environ 1100 pb correspondant à l'ADNc du nouveau gène 2 trouvé par des analyses bioinformatiques. Les pistes (1) et (5) contiennent respectivement les marqueurs  $\Phi$ X174 Hae III et le GeneRuler 100bp DNA Ladder Plus. La piste (3) présente le fragment obtenu par RT-PCR à l'aide des amorces *gène 2b* et *gène 2c* avec l'ADNc de rétine neurale. La piste (2) présente l'amplification avec ces mêmes amorces avec l'ADNc de l'EPR, alors que la piste (4) présente le contrôle négatif.

Puisque l'amplification par PCR ne permet pas toujours d'obtenir l'ADNc recherché, d'autres techniques ont été mises au point afin de contourner le problème. En fait, le problème évident qui se présente lors de l'amplification par PCR à l'aide des différentes combinaisons d'amorces a peut-être une explication. En effet, il était facile de

remarquer, lorsque les séquences des amorces ont été déterminées, que les séquences des nouveaux gènes 1 et 2 sont riches en bases G et C; leur composition en bases GC dépasse 60%. Ainsi, malgré la panoplie de conditions et d'amorces utilisées, il est probable que plusieurs paramètres nécessitent un ajustement adapté spécifiquement à ce type de séquences.

Les techniques de criblage de banques constituent donc une voie alternative pour obtenir l'ADNc recherché. Ainsi, les manipulations subséquentes ont davantage été axées sur le criblage de différentes banques (voir les sous-sections 4.2.2.3 et 4.2.2.4).

#### **4.2.2.2.2 Analyse du fragment obtenu par RT-PCR**

Le fragment obtenu a été cloné, séquencé et analysé afin de déterminer son identité avec les cPLA<sub>2</sub> connues. De même, il fallait vérifier l'exactitude de la séquence prédite comparativement à ce qui a été obtenu par le simple alignement de ces deux séquences, dans le but de s'assurer de la validité des approches bioinformatiques utilisées pour établir la séquence de ces nouveaux gènes de cPLA<sub>2</sub>. La figure 4.6 présente la séquence de fragment obtenu pour ce nouveau gène 2.

En comparant la séquence prédite à la figure 4.4 (acides aminés 511 à 849) avec celle obtenue en figure 4.6, il est évident que la séquence obtenue correspond exactement à la séquence prédite par les analyses bioinformatiques. En effet, l'alignement de la séquence prédite versus la séquence obtenue est parfait.

Ainsi, les exons supplémentaires présents dans la séquence codante initiale lors des premières analyses (par la combinaison de programmes bioinformatiques plutôt que par l'utilisation unique de GENSCAN) font sans doute partie des séquences introniques du gène. En effet, puisque la séquence obtenue (figure 4.4) est codante (ADNc), il est possible de déterminer, dans la séquence codante prédite, quels exons sont en fait des introns. Par conséquent, la déduction finale – qui n'incluait pas certains exons des analyses initiales – pour la séquence codante qui avait été établie s'est avérée juste.

```

gcttcaatctctgtgcagagga|gcaggccttcctgagcaggaggaagcagggtggtggccaag
F N L C A E E Q A F L S R R K Q V V A K
gccctgaagcaggccctgcagctggacagagacctgcaggaggatgaggtaccctgtgtg
A L K Q A L Q L D R D L Q E D E V P V V
ggcatcatggccacaggaggagggtgccgggcatgacctactctacggccacctattg
G I M A T G G G A R A M T S L Y G H L L
gccttcagaaagctgggctcctagactgtgtgacctacttcattggcatctctggctct
A L Q K L G L L D C V T Y F I G I S G S
acgtggacaatggccacacctgtacggggaccctgagtggtcgagagggacctggaggga
T W T M A H L Y G D P E W S Q R D L E G
cctatcagatacgccgggagcacctggccaagagcaagctggagggtcttttcccagag
P I R Y A R E H L A K S K L E V F S P E
cgctggcgagctaccgcccggagctggagctgcccgtgagcaggggccaccccacgacc
R L A S Y R R E L E L R A E Q G H P T T
tttgtggacctgtggcgctagtgtgctggagtccatgctgcacggccagggtgatggatcag
F V D L W A L V L E S M L H G Q V M D Q
aagctgtcaggacagagagccgcccctggaacggggtcagaacctctgcccctctacttg
K L S G Q R A A L E R G Q N P L P L Y L
agcctcaatgtcaaagagaacaatctggagacactggacttcaaggagtgggttgagttc
S L N V K E N N L E T L D F K E W V E F
tccccctatgaagtcggtttcctgaagtacggggccttcgtccctcctgagctcttcggc
S P Y E V G F L K Y G A F V P P E L F G
tccgagttcttcatgggacggctgatgaggaggatcccggagccccggatctgctttctg
S E F F M G R L M R R I P E P R I C F L
gaagccatctggagcaacattttctccctgaacctgctggatgcctggatgacctcacc
E A I W S N I F S L N L L D A W Y D L T
agttctggggagtcctggaacagcacatcaaggacaagaccaggagcttagagaaggag
S S G E S W K Q H I K D K T R S L E K E
cccctgaccacctcggggacctcctcgcggtggaggcctcggtggctgcagccaggcacg
P L T T S G T S S R L E A S W L Q P G T
gcgctggcccaggcatttaaggcttcctgacaggcaggccccctccaccagcgagcccc
A L A Q A F K G F L T G R P L H Q R S P
aacttcctccagggcctccagctgcaccaggactact|gtagccacaagacttctcca
N F L Q G L Q L H Q D Y C S H K D F S

```

**Figure 4.6** - Séquence du fragment obtenu du nouveau gène 2 à l'aide des amorces « gène 2b » et « gène 2c ». Ces deux amorces sont respectivement encadrées au début et à la fin de la séquence présentée. Ce fragment compte 1020 nucléotides.

La partie séquencée a été alignée, à l'aide du programme CLUSTALW, avec les trois autres cPLA<sub>2</sub> connues dans la littérature. Cet alignement est présenté dans la figure 4.7. Il démontre bien la similarité du fragment du nouveau gène 2 avec les cPLA<sub>2</sub> connues dans la littérature. Ce fragment du nouveau gène 2 présente une identité d'environ 53% avec le segment correspondant de la cPLA<sub>2</sub> bêta.

Des analyses plus poussées de ce fragment ont permis de vérifier la présence de la séquence consensus lipasique GXSGS propre aux cPLA<sub>2</sub>. En effet, le fragment présente une région lipasique GISGS entre les acides aminés 76 et 80 inclusivement, de la séquence obtenue. De même, par l'alignement de ce segment du nouveau gène 2 avec les trois autres cPLA<sub>2</sub> connues, les domaines caractéristiques aux cPLA<sub>2</sub> ont pu être repérés et les domaines manquants visualisés (figure 4.8).

Ces différentes analyses permettent de croire que le fragment obtenu est probablement un nouveau membre de la famille des cPLA<sub>2</sub> puisqu'il présente, à première vue, toutes les caractéristiques de cette famille.

```

nouveau gène 2  FNLCAEEQAFLSRRKQVVAKALKQALQLDRD---LQEDEVVVGIMATGGGARAMTSLYG  57
beta            FGPCAEEQAFLSRRKQVVAALRQALQLDGD---LQEDEIPVVAIMATGGGIRAMTSLYG  57
alpha           MALCDQEKTFRQQRKEHIRESMKKLLGPKNSEGLHSARDVPVVAIIGSGGGFRAMVGFSG  60
gamma           PGLQKEEKAVERRRRLHVLKALKKLR-----IEADEAPVVAVLGSGGGLRAHIACLG  52

nouveau gène 2  HLLALQKLGLLDCVITYFIGISGSTWTMAHLYGDPEWSQRDLEGPIRYAREHLAKSKLEVF  117
beta            QLAGLKELGLLDCVSYITGASGSTWALANLYEDPEWSQKDLAGPTELLKTQVTKNKLGLV  117
alpha           VMKALYESGILDCATYVAGLSGSTWYMSTLYSHPDFPEKGPEEINEELMKNVSHNPQLLL  120
gamma           VLSEMKEQGLLDAVITYLAGVSGSTWAISSLYTN----DGDMEALEADLKRHFTRQEWDLA  108

nouveau gène 2  SPERLASYRRELELRAEQGHPTTFVDLWALVLESMLHGQVMDQKLSGQRAALERGQNPLP  177
beta            APSQLQRYRQELAERARLGYPSCFTNLWALINEALLHDEPHDHKLSQDREALSHGQNPLP  177
alpha           TPQKVCRYVESLWKKKSSGQPVTFDTDFGMLIGETLIHNRMNTTLSSLKEKVNTAQCP  180
gamma           KSLQKT-----IQAARSENYSLTDFWAYMVISKQTRPELHSLNMKKPVEEGTLPYP  161

nouveau gène 2  LYLSLN--VKENNLETLDfKEWVEFSpyEVGFLKYGAfVpPELFGSEFFMGRlMRRIPEP  235
beta            IYCALN--TKGQSLTTFEFGEWCEFSpyEVGFpKYGAfIPSELFGSEFFMGLMKRLPES  235
alpha           LFTCLH--VKP-DVSELMFADWVEFSpyEIGMAKYGTfMAPDLFGSKFFMGTVVKKYEEN  237
gamma           IFAAIDNDLQPSWQEARAPETWFEFTPHHAGFSALGAFVSITHFGSKFKKGRlVRTHPER  221

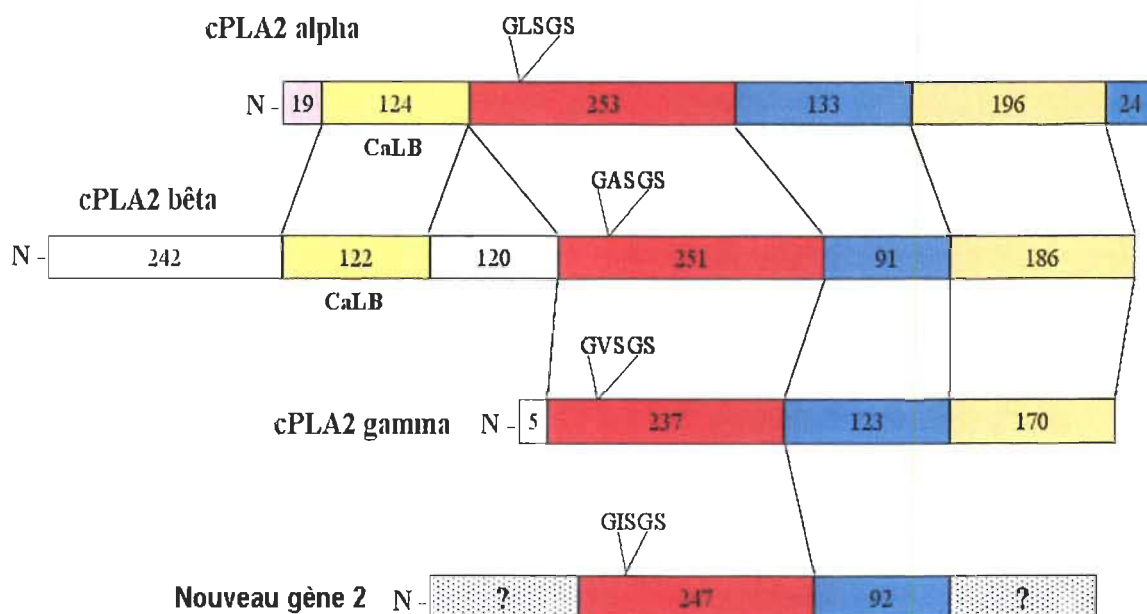
nouveau gène 2  RICFLEASSGESWK--QHIDKT-RSLEK-----EIWSNIFSLNLLDAW-----YDL  279
beta            RICFLEGIWSNLYA--ANLQDSLYWASEP-----SQFWRWRVNRQANLD-----KEQ  280
alpha           PLHFLMGVWGSafs--ILFNRLGVSGSQSRGSTMEEELENITTKHIVSNDSSDSDESH  295
gamma           DLTFLRGLWGSALGNTEVIREYIFDQLRN-----LTLKGLWRRAVANAKSIGHLIFAR  274

nouveau gène 2  TPLTTSSTSSRLASWLQPGTALAQAfKGfLTGRPLHQR-----SPNfLQG-LQLHQD  331
beta            VPLLKIEEPPSTAG-----RIAeffTDLLTWRPLAQa-----THNfLRG-LHFHKD  325
alpha           EPKGTEENEDAGSDYQSDNQASWIHRMIMALVSDSALFNTREGRAGKVHNFMLG-LNLNTS  354
gamma           LLRLQESSQGEHPPPEDEGGEPEHTWLTEMLENWTRTSLEKQEQPHEDPERKGSLSNlMD  334

nouveau gène 2  YCSHKDFS  339
beta            YFQHPhFS  333
alpha           YPLSPLSD  362
gamma           FVKKTGIC  342

```

**Figure 4.7** - Alignement du fragment obtenu par RT-PCR pour le nouveau gène 2 avec le segment correspondant de chacune des trois cPLA<sub>2</sub> connues (alpha, bêta et gamma). La protéine associée au fragment du nouveau gène 2 présente une identité de 53% avec la cPLA<sub>2</sub> bêta.



**Figure 4.8** - Représentation schématique des trois isoformes connues (cPLA<sub>2</sub> alpha, bêta et gamma) avec le fragment du nouveau gène 2. La nouvelle cPLA<sub>2</sub> (nouveau gène 2) semble être organisée de la même façon que les trois autres isoformes de la famille dont elle fait probablement partie.

#### 4.2.2.3 Criblage de banques d'ADNc

Le criblage de banques a été réalisé avec une banque d'ADNc de rétine neurale humaine et une banque d'ADNc d'EPR humain. Plusieurs hybridations ont été effectuées, en utilisant des conditions de stringence variées afin de maximiser les chances d'obtenir les clones positifs recherchés. Pour chacune des conditions utilisées, un minimum de 150000 clones ont été criblés. La quantité moins importante de clones criblés s'explique par le fait que les sondes utilisées pour cribler les banques de rétine neurale et d'EPR afin de « pêcher » l'ADNc de la nouvelle cPLA<sub>2</sub> sont beaucoup plus spécifiques à l'ADNc recherché que ne l'est la sonde fabriquée à partir de l'ADNc de la cPLA<sub>2</sub> alpha (voir sous-section 4.2.1.2).

Le criblage a été réalisé à l'aide d'une sonde fabriquée à partir du fragment du nouveau gène 2 (figure 4.6) inséré dans un vecteur plasmidique (*pGEM-T Easy*). Puisque cette sonde devrait présenter 100% d'identité avec le clone recherché et qu'elle a une longueur considérable (1100 pb), il était fort probable que le clone serait « pêché » par le

criblage. Cette sonde n'a toutefois pas permis d'obtenir les clones recherchés. En fait, bien que certains clones se soient avérés positifs lors de la détection, aucun des clones séquencés n'était une cPLA<sub>2</sub>.

Ainsi, les deux banques ne semblaient pas présenter le clone cherché en proportion suffisante pour permettre de le retrouver par cette technique de criblage. Cette hypothèse est fort plausible puisqu'il s'agissait dans les deux cas d'une banque amplifiée. Ainsi, si le gène est faiblement exprimé au niveau de ces tissus, le nombre de copies correspondant au gène sera faible et possiblement trop dilué par l'amplification pour permettre de le « pêcher » par le criblage.

L'alternative était donc de tenter le criblage d'une banque provenant d'un tissu différent et présentant possiblement un nombre plus représentatif de la copie de l'ADNc d'intérêt ou de tenter le criblage avec une banque génomique. C'est donc une banque génomique placentaire qui fut utilisée pour les criblages subséquents.

#### **4.2.2.4 Criblage d'une banque génomique**

Ici encore, comme dans tous les criblages réalisés lors de ce travail de maîtrise, de nombreuses hybridations ont été effectuées, sous différentes conditions de stringence.

Ce criblage a d'abord été fait à l'aide de la sonde fabriquée à partir du fragment du nouveau gène 2. Malheureusement, peu importe les conditions utilisées, aucune des hybridations n'a permis d'obtenir les clones recherchés. En fait, aucun clone n'a pu être obtenu par cette approche.

L'explication de cet échec repose peut-être sur le fait que la sonde est faite d'ADNc, alors que la banque est génomique. Par conséquent, il est possible que là où se trouvent les introns dans la séquence génomique, des boucles se forment au niveau de l'ADN et ne permettent pas une bonne hybridation de la sonde. Afin de remédier à ce problème, des sondes ont été synthétisées par PCR (à l'aide des paires d'amorces *sondeH-* *sondeH*

(AS) et gène 2c (AS) - sonde B (voir tableau 3.2)) dans des régions du fragment du nouveau gène 2 qui sont entièrement exoniques (voir figure 4.10). Ce sont donc des plasmides dont les inserts comprenaient ces fragments exoniques, de 151 pb ou de 187 pb (selon les amorces utilisées pour l'amplification) qui ont servi en tant que sondes (figures 4.9).

Insert de la sonde H (5' vers 3') :

ACCTGGAGGGACCTATCAGATACGCCCGGGAGCACCTGGCCAAGAGCAAGCTGGAGGTC  
TTTTCCCAGAGCGCCTGGCGAGCTACCGCCGGGAGCTGGAGCTGCGGGCTGAGCAGGG  
CCACCCACGACCTTTGTGGACCTGTGGGCGCTAGTGCTGGAGTCCATGCTGCA

Insert de la sonde B (5' vers 3') :

GAGAAGGAGCCCCTGACCACCTCGGGGACCTCCTCGCGGCTGGAGGCCTCGTGGCTGCA  
GCCAGGCACGGCGCTGGCCAGGCATTTAAAGGCTTCCTGACAGGCAGGCCCTCCACC  
AGCGCAGCCCCAACTTCCTCCAGGGCCTCCAGCTGCACCAGGACTACTGTAGCCACAAAG  
ACTTCTCCA

**Figure 4.9** - Séquence détaillée de chacune des deux sondes internes au nouveau gène 2 utilisées pour le criblage. La sonde H compte 172 nucléotides alors que la sonde B en compte 187.

Le criblage a donc été répété à l'aide de ces deux nouvelles sondes. Malheureusement, aucun clone positif n'a pu être obtenu, peu importe les paramètres utilisés.

La dernière avenue à explorer pour le criblage était l'utilisation des EST pour fabriquer des sondes. Ainsi, les hybridations ont été réalisées à l'aide de deux sondes différentes qui étaient en fait des plasmides contenant l'un ou l'autre des deux EST choisis. Plusieurs EST auraient pu être utilisés, mais deux ont été sélectionnés (apparaissent ombrés dans le tableau 4.1). Ce sont donc deux plasmides différents dont les inserts avaient une longueur approximative de 500 pb et 800 pb qui ont été utilisés pour le criblage. Le tableau 4.1 présente quelques exemples des EST qui auraient pu être utilisés pour la fabrication d'une sonde. Plusieurs autres EST sont aussi disponibles mais ne figurent pas dans le tableau, ce tableau se voulant avant tout un aperçu des EST trouvés à l'aide du programme *blastn*.

Les résultats de ce dernier criblage n'ont pas non plus mené à l'obtention de clones positifs. Faut-il en conclure ici encore que la banque génomique utilisée n'était pas



représentative du clone recherché? C'est effectivement possible puisqu'il s'agissait aussi d'une banque amplifiée.

```

gcttcaatctctgtgcagaggagcaggccttcctgagcaggaggaagcaggtggtggccaag
  F N L C A E E Q A F L S R R K Q V V A K
gccttgaagcaggccctgcagctggacagagacctgcaggaggatgaggtacccgttgtg
  A L K Q A L Q L D R D L Q E D E ♦ V P V V
ggcatcatggccacaggaggaggtgccgggccatgacctcactctacggccacctattg
  G I M A T G G G A R A M T S L Y G H L L
gccttgcagaagctgggcctcctagactgtgtgacctacttcattggcatctctggctct
  A L Q K L G L L D C V T Y F I G I S G S
Acgtggacaatggccacctgtacggggacctgagtggtcgagagggacctggaggga
  T W ♦ T M A H L Y G D P E W S Q R D L E G
Cctatcagatacgccccgggagcacctggccaagagcaagctggaggtcttttccccagag
  P I R Y A R E H L A K S K L E V F S P E
cgcttggcagctaccgccgggagctggagctgcgggctgagcagggccacccacgacc
  R L A S Y R R E L E L R A E Q G H P T T
tttgtggacctgtgggcgtagtgctggagtccatgctgcacggccaggtgatggatcag
  F V D L W A L V L E S M L H G ♦ Q V M D Q
aagctgtcaggacagagagccgcccttgaacggggtcagaacctctgccccctcacttg
  K L S G Q R A A L E R G Q N P L P L Y L
agcctcaatgtcaaagagaacaatctggagacactggacttcaaggagtgggttgagttc
  S L N V K E N N L E T L D F K ♦ E W V E F
tccccctatgaagtcggtttcctgaagtacggggccttcgtccctcctgagctcttcggc
  S P Y E V G F L K Y G A F V P P E L F G
tccgagttcttcatgggacggctgatgaggaggatcccgagccccggatctgctttctg
  S E F F M G R L M R R I P E P R I C F L
Gagccatctggagcaacattttctccctgaacctgctggatgcctggatgacctcacc
  E ♦ A I W S N I F S L N L L D A W Y D L T
agttctggggagtcctggaaacagcacatcaaggacaagaccaggagcttagagaaggag
  S S G E S W K Q H I K D K T R S L ♦ E K E
cccctgaccacctcggggacctcctcgcggtggaggcctcgctggctgcagccaggcacg
  P L T T S G T S S R L E A S W L Q P G T
gcgctggcccaggcatTTAAaggcttctgacaggcaggcccctccaccagcgagcccc
  A L A Q A F K G F L T G R P L H Q R S P
Aacttctccagggcctccagctgcaccaggactatgtagccacaagacttctcca
  N F L Q G L Q L H Q D Y C S H K D F S

```

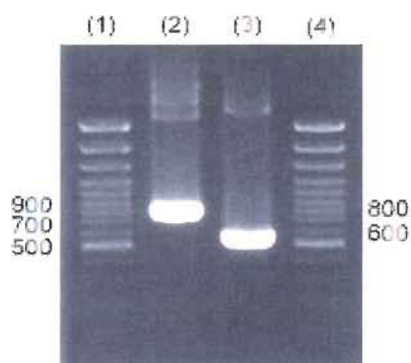
Amorce SondeH

Amorce SondeH (AS)

Amorce SondeB

Amorce gène 2c (AS)

**Figure 4.10** – Position des amorces utilisées pour la fabrication des sondes sur la séquence du fragment obtenu pour le gène 2. Les amorces sont encadrées et le symbole ♦ indique la position d'une limite intron-exon. Les séquences des sondes H et B apparaissent en italique. Pour plus de détails quant à la position des introns, voir l'annexe B.



**Figure 4.11** - Résultat obtenu pour l'amplification par PCR des deux EST utilisés pour la fabrication d'une sonde. Les pistes (1) et (3) contiennent le marqueur GeneRuler 100bp DNA Ladder Plus, alors que les pistes (2) et (3) contiennent respectivement l'insert amplifié des EST 2068898 et 2739188.

**TABLEAU 4.1**  
**Compilation de différents EST**

Numéro de l'EST (IMAGE) <sup>1</sup>	Provenance	Longueur lisible approximative <sup>2</sup>	Localisation probable <sup>3</sup>
119109	Poumon humain	693	RP11-35K2
2068898	Mélanocyte, cœur fœtal, utérus humain	443	RP11-35K2
2739188	Carcinome de la bouche chez l'humain	500	RP11-35K2
1184010	Cellule B germinale	428	RP11-35K2
5167964	Cerveau humain	667	nouveau gène 2

<sup>1</sup> Les EST sont couramment désignés par leur numéro IMAGE qui figure généralement suite à leur numéro d'accèsion sur le site de la NCBI.

<sup>2</sup> Certaines des séquences présentées en tant qu'EST sur les bases de données ne sont lues que partiellement lors du séquençage. Les auteurs précisent alors seulement la longueur des inserts selon la séquence de qualité qui a été lue. Ainsi, les inserts sont généralement plus longs que ce qui est précisé dans la description. C'est donc ce qui explique la différence entre les longueurs prévues pour les EST ombrés et les longueurs observées (figure 4.11).

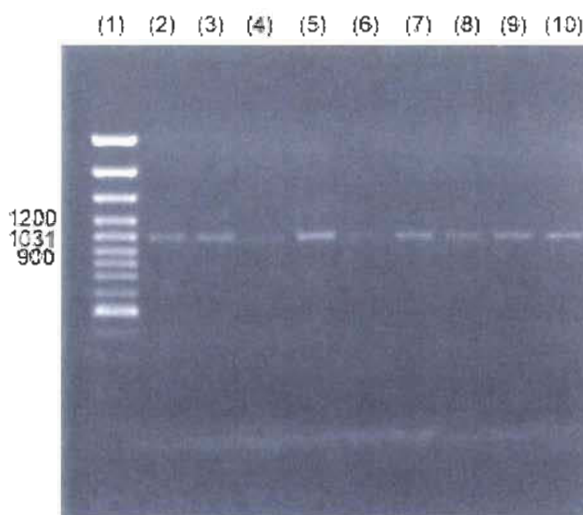
<sup>3</sup> La localisation correspond à la séquence soumise pour le *blastn*.

À l'heure actuelle, de nouveaux EST sont déjà disponibles et pourraient être essayés sur différentes banques d'ADNc. C'est donc dire que le nombre d'EST pouvant être utilisés pour la fabrication d'une sonde est très important.

Il s'agit donc de déterminer dans quel(s) tissu(s) le clone recherché pourrait être retrouvé en quantité suffisante pour préparer une banque d'ADNc. Le criblage par RT-PCR de différents tissus, en parallèle avec la rétine neurale pourrait apporter les réponses à ce sujet et permettre de déterminer si la nouvelle cPLA<sub>2</sub> (gène 2) est exprimée uniquement dans la rétine et/ou l'EPR.

#### 4.2.2.5 Criblage par RT-PCR de différents tissus

Le criblage de différents tissus a été réalisé à l'aide de l'ADNc de tissus fourni dans la trousse *Human MTC Panel I* de *Clontech*, tel que mentionné à la section 3.9. Les différents ADNc utilisés proviennent du cœur, du cerveau, du placenta, du poumon, du foie, du muscle squelettique, du rein et du pancréas. La figure 4.12 présente les résultats obtenus pour l'amplification du fragment du nouveau gène 2 à partir de ces différents ADNc et de l'ADNc de rétine neurale, à l'aide des amorces *gène 2b* et *gène 2c*.



**Figure 4.12** - Amplification du fragment de l'ADNc de la nouvelle cPLA<sub>2</sub> (gène 2) à partir des amorces *gène 2b* et *gène 2c* dans différents tissus. La première piste (1) contient le marqueur GeneRuler 100bp DNA Ladder Plus. Les pistes (2) à (9) contiennent le fragment du nouveau gène 2 amplifié respectivement dans le cœur (2), le muscle squelettique (3), le cerveau (4), le placenta (5), le foie (6), le poumon (7), le rein (8), le pancréas (9) et la rétine neurale (10).

Ce résultat ne permet pas de préciser dans quel tissu en particulier le nouveau gène 2 est préférentiellement exprimé, puisque les tissus présentent tous à peu près une bande de la

même intensité pour la même amplification. En fait, ce qui était cherché, c'est une bande très très intense, ou « patate » dans le jargon scientifique, et non pas une bande un peu plus foncée, comme c'est le cas pour le placenta (voir figure 4.12). Par contre, les bandes moins intenses, comme celles qui apparaissent pour le cerveau et le foie (voir figure 4.12), indiquent que ces tissus ne sont pas du tout adéquats pour les utiliser en tant que banque d'ADNc dans le but de « pêcher » l'ADNc complet du nouveau gène 2. En effet, puisque la même quantité initiale d'ADNc est utilisée pour toutes les amplifications, il est possible de déterminer dans quel(s) tissu(s) le nouveau gène est le plus fortement exprimé, en comparant l'intensité des bandes obtenues après amplification et migration. Ainsi, parce qu'aucune bande ne s'est pas encore vraiment démarquée des autres jusqu'à présent, d'autres tissus devront être testés pour tenter d'obtenir cette bande.

Puisque le criblage ne s'est pas avéré très concluant, une autre approche a été envisagée. Cette approche consiste à amplifier les extrémités d'un fragment d'ADNc à l'aide de la technique du RACE. Comme le nouveau gène 2 est définitivement présent au niveau de la rétine neurale, il est probable que l'approche du RACE avec l'ARN rétinien puisse permettre d'obtenir les extrémités 5' et 3' manquantes du clone cherché.

#### **4.2.2.6 Amplification rapide des extrémités de l'ADNc (RACE)**

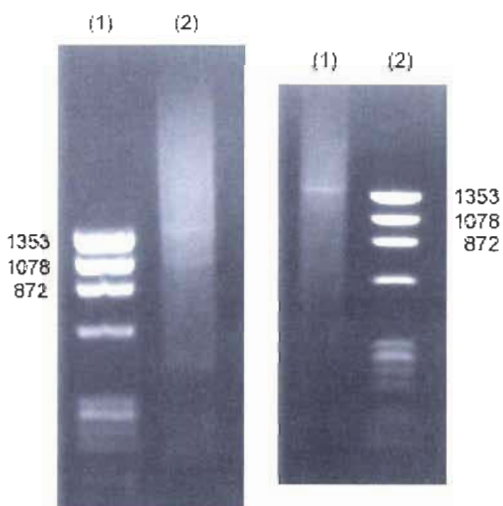
Le RACE a donc été utilisé dans le but d'obtenir les extrémités manquantes du nouveau gène 2. Cette technique consiste en deux approches différentes, tel que décrit dans la section 3.8. La première approche est la recherche de l'extrémité 5' en tirant profit de la queue poly-A commune aux ARNm. Diverses températures d'hybridation pour le PCR ont été utilisées afin de maximiser les chances d'obtenir un fragment, mais aucune des conditions essayées n'a permis d'obtenir un fragment, en dépit des essais contrôles qui confirment l'efficacité des paramètres utilisés.

En ce qui a trait à l'approche pour déterminer l'extrémité 3' (RACE 3'), la première ronde d'amplification a permis d'obtenir quelques bandes très pâles, en plus d'une

longue traînée. Ce résultat apparaît sur la figure suivante (figure 4.13).

La deuxième ronde d'amplification a permis d'obtenir une bande légèrement mieux définie, mais elle est identique à ce qui a été obtenu en troisième ronde (figure 4.13). Par conséquent, c'est ce fragment qui a été cloné, mais il ne s'agissait pas de l'ADNc d'intérêt.

Ainsi, il y a définitivement un problème dans l'obtention des extrémités de l'ADNc du nouveau gène 2 par cette approche. Le problème vient sans doute de la séquence elle-même puisque tous les contrôles positifs fonctionnent bien. C'est donc dire que la composition de la séquence recherchée n'est peut-être pas une séquence qui permet d'utiliser les conditions optimales suggérées normalement. Ainsi, la composition riche en bases G et C peut ici encore intervenir et empêcher l'optimisation des protocoles de façon conventionnelle.



**Figure 4.13** - Résultats obtenus avec la technique du RACE. Résultats obtenus pour la première ronde d'amplification pour le RACE 3' (à gauche) et pour la troisième ronde d'amplification (à droite). Le gel de gauche présente le marqueur  $\Phi$ X174 Hae III dans la piste (1) et le produit de PCR dans la piste (2). Le gel de droite présente le produit de PCR obtenu suite à la troisième ronde d'amplification pour le RACE 3' dans la piste (1) et le même marqueur que pour le gel de gauche dans la piste (2).

avec l'ADNc de différents tissus, en parallèle avec l'ADNc de la rétine neurale et de l'EPR.



**Figure 4.2** - Obtention de la cPLA<sub>2</sub> gamma dans l'EPR. Les marqueurs habituels sont encore utilisés ici (ΦX174 Hae III et GeneRuler 100bp DNA Plus) dans les pistes (1) et (3). La bande d'environ 1500 pb en piste (2) est l'ADNc de la cPLA<sub>2</sub> gamma.

Les différents résultats détaillés dans la présente section font aussi partie d'un article intitulé « *Cloning of the phospholipases A2 expressed by the human retinal pigment epithelium* » qui a été soumis le 5 octobre 2001. Cet article figure en annexe (annexe A) à la fin du présent mémoire.

## 4.2 Recherche de nouvelles cPLA<sub>2</sub> dans l'EPR et la rétine neurale

Suite à la vérification de la présence de cPLA<sub>2</sub> connues dans la rétine neurale et l'EPR, la recherche de nouvelles cPLA<sub>2</sub> a été entreprise. Cette recherche a été abordée de deux façons différentes. D'abord, par une approche dite expérimentale puisqu'elle consiste en l'utilisation de diverses techniques de laboratoire. Ensuite, par une approche dite bioinformatique puisque, comme le terme l'indique, elle reposait dans un premier temps sur la recherche bioinformatique via l'utilisation de divers algorithmes de recherche disponibles sur Internet (pour plus de détails, voir la section 3.10).

## CHAPITRE 5

### CONCLUSIONS

#### 5.1 Récapitulation des résultats

Dans un premier temps, divers RT-PCR ont été réalisés afin de déterminer quelles cPLA<sub>2</sub> connues sont présentes au niveau de l'EPR et de la rétine neurale. La cPLA<sub>2</sub> gamma s'est avérée présente dans ces deux tissus, alors que la cPLA<sub>2</sub> alpha en est absente. En ce qui a trait à la cPLA<sub>2</sub> bêta, aucune des amorces utilisées n'est arrivée à amplifier un de ses fragments, peu importe le tissu ou les conditions utilisées. Ce problème est probablement attribuable au fait que sa forme épissée – forme à partir de laquelle les séquences des amorces ont initialement été déterminées - soit peu présente au niveau des tissus comparativement à sa forme non épissée. De plus, la forte composition de l'ADNc de cette cPLA<sub>2</sub> en GC (61%) pourrait aussi être responsable de la difficulté à amplifier son ADNc et ne doit pas être négligée. Par conséquent, il n'est pas encore possible d'affirmer que la cPLA<sub>2</sub> bêta est absente ou présente au niveau de l'EPR et/ou de la rétine neurale.

Par la suite, la recherche de nouvelles cPLA<sub>2</sub> a été entreprise par diverses analyses bioinformatiques réalisées sur les banques de données de la NCBI. Dans un premier temps, les séquences de cPLA<sub>2</sub> connues ont été utilisées pour repérer des séquences présentant une identité avec ces cPLA<sub>2</sub>, mais pour lesquelles aucune protéine n'était encore attribuée. Ensuite, ces séquences homologues, appelées HTGS ou clones, ont été soumises à divers algorithmes de recherche, afin de déterminer les séquences codantes potentielles pour de nouvelles cPLA<sub>2</sub>. De ces analyses, deux séquences, appelées nouveau gène 1 et nouveau gène 2, ont été déterminées. À partir de ces deux séquences, des séquences d'amorces spécifiques ont été déterminées et celles-ci ont été utilisées en RT-PCR. Une seule combinaison – parmi une soixantaine de combinaisons - de ces 29 amorces a permis d'obtenir une séquence d'intérêt, soit un fragment de la séquence du nouveau gène 2.

L'analyse de ce fragment, via les divers algorithmes où l'alignement est possible, suggère fortement qu'il s'agirait d'une nouvelle cPLA<sub>2</sub>. En effet, non seulement l'alignement avec les autres cPLA<sub>2</sub> connues est presque parfait, mais ce fragment présente les domaines reconnus chez les autres cPLA<sub>2</sub>, en plus de la séquence consensus lipasique propre à cette classe d'enzymes. Ainsi, l'alignement de la protéine associée à la séquence clonée grâce aux divers algorithmes de recherche a permis d'estimer une identité de 52% avec la cPLA<sub>2</sub> bêta. De même, les domaines A et B sont pratiquement complets chez le fragment cloné, toujours selon l'alignement. Ces arguments suggèrent l'appartenance de ce nouveau gène 2 au groupe des cPLA<sub>2</sub>. Ainsi, l'ADNc complet de ce nouveau gène 2 devait être cloné afin de permettre des analyses plus poussées de cette probable nouvelle cPLA<sub>2</sub>.

Par conséquent, différentes banques ont été criblées en se basant sur la séquence obtenue du fragment cloné (nouveau gène 2). Diverses sondes ont été utilisées et différentes conditions essayées. Des sondes fabriquées à partir de l'ADNc de la cPLA<sub>2</sub> alpha et des sondes réalisées à l'aide d'EST ou de différents segments de l'ADNc cloné du nouveau gène 2 ont été utilisées pour le criblage. Toutefois, aucun de ces criblages n'a permis d'obtenir de clones positifs. Il est donc probable qu'aucune de ces banques ne présente le clone recherché en quantité suffisante pour qu'il puisse être « pêché » dans les conditions utilisées. L'utilisation d'une banque où ce clone est suffisamment représenté (un titre avant amplification de l'ordre de 10<sup>9</sup>, par exemple) pourrait peut-être remédier au problème.

Les RT-PCR réalisés sur différents tissus, dont la rétine neurale, n'ont pas permis de déterminer dans quel(s) tissu(s) le gène d'intérêt était beaucoup plus abondamment exprimé, puisque parmi tous les tissus examinés, aucun ne présentait une bande de l'intensité cherchée. Par conséquent, d'autres tissus devront être testés afin de tenter d'obtenir une bande vraiment plus intense que celles qui ont été obtenues jusqu'à présent. Lorsque cette bande sera obtenue, il sera possible d'utiliser le tissu en question pour fabriquer une banque d'ADNc dans le but de la cribler pour obtenir l'ADNc complet du nouveau gène 2, tel qu'expliqué à la sous-section 4.2.2.5.



Enfin, la technique du RACE a été utilisée afin de tenter d'obtenir les extrémités 5' et 3' de l'ADNc cloné pour le nouveau gène 2. Une bande a pu être obtenue dans la recherche de l'extrémité 3', mais s'est avérée être tout autre chose que ce qui était recherché. Ainsi, cette approche n'a pas permis d'obtenir les résultats attendus.

## 5.2 Conclusions

Les diverses démarches bioinformatiques ont mené à la découverte de deux nouveaux gènes potentiels pour le groupe IV des PLA<sub>2</sub>, ou cPLA<sub>2</sub>. Les différentes approches entreprises lors du travail de maîtrise détaillé dans le présent mémoire ont permis d'isoler une partie d'un de ces deux nouveaux gènes. Ainsi, un fragment d'environ 1100 nucléotides du gène appelé nouveau gène 2 a été obtenu. Des analyses bioinformatiques plus poussées ont pu permettre d'apprécier l'homologie de cette séquence avec les cPLA<sub>2</sub> déjà connues. En effet, tous les algorithmes utilisés indiquent que l'homologie de ce fragment cloné avec la cPLA<sub>2</sub> bêta se situe autour de 50%, ce qui est très considérable puisque l'identité présente entre les différentes cPLA<sub>2</sub> connues est d'approximativement 30%. Par conséquent, il est fort probable que cette nouvelle séquence soit une partie de la séquence d'une nouvelle cPLA<sub>2</sub>.

La recherche des extrémités du clone obtenu permettra non seulement de vérifier l'alignement de cette potentielle cPLA<sub>2</sub> en entier avec les autres cPLA<sub>2</sub> connues, mais aussi de déterminer l'activité phospholipasique de cette nouvelle protéine pour ainsi confirmer sa classification au sein de la famille des PLA<sub>2</sub>.

Le fait que cette séquence nucléotidique soit riche en bases G et C semble être une cause importante des problèmes rencontrés au cours de ce travail de maîtrise. Il n'est donc pas étonnant de voir que cette nouvelle séquence présente une forte identité avec la cPLA<sub>2</sub> bêta, qui est la seule des cPLA<sub>2</sub> qui n'a pu être obtenue par RT-PCR. En effet, la cPLA<sub>2</sub> bêta présente aussi une concentration importante de ces bases (61%).

Ainsi, la prochaine étape consiste en la recherche d'informations dans ce domaine assez complexe que sont les séquences riches en bases G et C. Une fois les correctifs apportés aux différents protocoles, il est fort probable que le projet puisse avancer beaucoup plus rapidement.

Un deuxième élément pourrait être la source des échecs rencontrés au cours de ce travail de maîtrise. La possibilité que l'ARNm des deux nouveaux gènes (1 et 2) soit faiblement épissé, tout comme l'ARNm de la cPLA<sub>2</sub> bêta, pourrait aussi expliquer la difficulté à amplifier les séquences codantes de ces nouveaux gènes par PCR. En fait, puisque les séquences des amorces sont déterminées en fonction de la séquence codante, il est possible que la présence de régions non codantes dans le produit de RT (ADNc) nuise aux amplifications par PCR. Il s'agirait donc de redéterminer la séquence de quelques amorces afin que certaines se trouvent dans les régions introniques de la séquence génomique (voir annexe B pour les régions introniques du nouveau gène 2). De cette façon, le problème des bases GC ne serait pas éliminé, mais les chances d'obtenir un fragment par PCR seraient augmentées.

Enfin, la présence de la cPLA<sub>2</sub> gamma au niveau de l'EPR pourrait expliquer en partie les conclusions des travaux de l'équipe du Dr Salesse (Van Themsche *et al.*, 2001) qui veulent que deux nouvelles cPLA<sub>2</sub> soient spécifiques à l'EPR. Ainsi, puisque la cPLA<sub>2</sub> gamma n'était pas connue au moment de leur travaux, il est possible qu'elle enzyme constitue l'une des fractions cytosoliques obtenues lors des manipulations en laboratoire. Toutefois, divers tests devront être faits afin de déterminer les caractéristiques biochimiques de cette cPLA<sub>2</sub> pour confirmer ou infirmer l'implication de celle-ci dans les résultats obtenus.

### 5.2.1 Atteinte des objectifs fixés

Le premier objectif consistant à déterminer les diverses cPLA<sub>2</sub> présentes dans l'EPR et/ou la rétine neurale a été partiellement atteint. Il a été déterminé que la cPLA<sub>2</sub> alpha est absente de l'EPR et de la rétine neurale, alors que la cPLA<sub>2</sub> gamma est présente au

niveau des deux tissus. L'élément manquant dans l'atteinte de cet objectif est la détermination de la présence ou de l'absence de la cPLA<sub>2</sub> bêta dans ces deux tissus.

Finalement, l'objectif consistant à rechercher des cPLA<sub>2</sub> spécifiques à la rétine neurale n'a pas été atteint puisque la présence de l'ADNc de la nouvelle cPLA<sub>2</sub> dans plusieurs autres tissus permet d'exclure cette possibilité. Cependant, le nouveau gène 1 n'a pas encore été cloné (il est possible qu'il ne soit pas aisé de l'amplifier pour les mêmes raisons que le nouveau gène 2) et pourrait s'avérer être spécifique à la rétine neurale. Par conséquent, les hypothèses quant à l'existence de cPLA<sub>2</sub> spécifiques à la rétine neurale et/ou l'EPR restent valides.

### 5.3 Perspectives de recherche

Les perspectives de recherche sont très claires en ce qui concerne ce qui doit être fait à court et à moyen terme. La collecte d'informations relatives aux séquences riches en bases G et C pour la préparation d'amorces permettra de reprendre les différentes manipulations déjà effectuées, mais avec les correctifs appropriés.

Une fois la séquence entière identifiée et isolée, il s'agira de cloner cette protéine dans un vecteur approprié et de vérifier son expression *in vitro* afin de confirmer son activité phospholipasique A2 et ses différentes propriétés biochimiques.

Ensuite, il s'agira de cloner le nouveau gène 1 de la même façon que l'aura été le nouveau gène 2. Ce clonage devra être accéléré puisque la séquence déterminée lors de ce projet de maîtrise a été très récemment mise à la disposition des différents chercheurs sur le site de la NCBI. En effet, une année après que cette séquence ait été mise au point par ce travail de maîtrise (01/12/2001), une séquence très homologue à ce qui est appelé ici le nouveau gène 1 a été déposée sur le site de NCBI. Toutefois, la séquence diffusée diffère légèrement de la séquence présentée ici, plus précisément par l'extrémité 5' qui est plus courte sur le site de la NCBI. Il n'en reste pas moins que plusieurs équipes sont

possiblement déjà en train de travailler sur ce nouveau gène, bien qu'aucun article n'ait encore été publié à ce sujet.

Enfin, il s'agira, via la localisation chromosomique de ces nouveaux gènes (15q13.3-15q15), de repérer les maladies ou troubles liés à ces nouvelles protéines. Ultimement, l'identification complète de l'ADN génomique de ces nouvelles cPLA<sub>2</sub> permettra de rechercher, chez des patients atteints de ces différents troubles, des mutations au niveau de ces gènes.

## RÉFÉRENCES

- Ackermann, E. J., E. S. Kempner et E. A. Dennis (1994). Ca(2+)-independent cytosolic phospholipase A2 from macrophage-like P388D1 cells. Isolation and characterization. *J Biol Chem* **269**, 9227-33.
- Adachi, H., M. Tsujimoto, M. Hattori, H. Arai et K. Inoue (1997). Differential tissue distribution of the beta- and gamma-subunits of human cytosolic platelet-activating factor acetylhydrolase (isoform I). *Biochem Biophys Res Commun* **233**, 10-3.
- Alonso, F., P. M. Henson et C. C. Leslie (1986). A cytosolic phospholipase in human neutrophils that hydrolyzes arachidonoyl-containing phosphatidylcholine. *Biochim Biophys Acta* **878**, 273-80.
- Apitz-Castro, R. J., M. A. Mas, M. R. Cruz et M. K. Jain (1979). Isolation of homogeneous phospholipase A2 from human platelets. *Biochem Biophys Res Commun* **91**, 63-71.
- Arni, R. K. et R. J. Ward (1996). Phospholipase A2--a structural review. *Toxicon* **34**, 827-41.
- Balsinde, J., M. A. Balboa et E. A. Dennis (1997). Antisense inhibition of group VI Ca2+-independent phospholipase A2 blocks phospholipid fatty acid remodeling in murine P388D1 macrophages. *J Biol Chem* **272**, 29317-21.
- Balsinde, J., M. A. Balboa, W. H. Li, J. Llopis et E. A. Dennis (2000). Cellular regulation of cytosolic group IV phospholipase A2 by phosphatidylinositol bisphosphate levels. *J Immunol* **164**, 5398-402.
- Barrett, A. J et M. F. Heath (1977). Lysosomal Enzymes. In "Lysosomes: A Laboratory Handbook". North Holland Biomedical Press, Amsterdam.
- Bazan, N. G., M. G. Murphy et G. Toffano, Eds. (1992). Neurobiology of essential fatty acids. New York: Plenum Press.
- Bennett, C. F., A. McCarte et S. T. Crooke (1990). Purification and characterization of a soluble phospholipase A2 from guinea pig lung. *Biochim Biophys Acta* **1047**, 271-83.
- Berman, E. R. (1991). In "Biochemistry of the Eye". Plenum Press, New York.

- Bernstein, Paul S. (1999). Macular Biology. In "Age-Related Macular Degeneration", pp. 1-16. Mosby, St-Louis.
- Birkle, D. L. et N. G. Bazan (1986). The arachidonic cascade and phospholipid and docosahexanoic acid metabolism in the retina. In "Progress in Retinal Research", Vol. 5. Pergamon Press, New York.
- Borsch-Haubold, A. G., F. Bartoli, J. Asselin, T. Dudler, R. M. Kramer, R. Apitz-Castro, S. P. Watson et M. H. Gelb (1998). Identification of the phosphorylation sites of cytosolic phospholipase A2 in agonist-stimulated human platelets and HeLa cells. *J Biol Chem* **273**, 4449-58.
- Boyle, D., L. F. Tien, N. G. Cooper, V. Shepherd et B. J. McLaughlin (1991). A mannose receptor is involved in retinal phagocytosis. *Invest Ophthalmol Vis Sci* **32**, 1464-70.
- Burke, J. R., L. B. Davern, K. R. Gregor, G. Todderud, J. G. Alford et K. M. Trampusch (1997). Phosphorylation and calcium influx are not sufficient for the activation of cytosolic phospholipase A2 in U937 cells: requirement for a Gi alpha-type G-protein. *Biochim Biophys Acta* **1341**, 223-37.
- Channon, J. Y. et C. C. Leslie (1990). A calcium-dependent mechanism for associating a soluble arachidonoyl- hydrolyzing phospholipase A2 with membrane in the macrophage cell line RAW 264.7. *J Biol Chem* **265**, 5409-13.
- Chen, H. et R. E. Anderson (1993). Metabolism in frog retinal pigment epithelium of docosahexaenoic and arachidonic acids derived from rod outer segment membranes. *Exp Eye Res* **57**, 369-77.
- Choukroun, G. J., V. Marshansky, C. E. Gustafson, M. McKee, R. J. Hajjar, A. Rosenzweig, D. Brown et J. V. Bonventre (2000). Cytosolic phospholipase A(2) regulates golgi structure and modulates intracellular trafficking of membrane proteins. *J Clin Invest* **106**, 983-93.
- Clark, J. D., L. L. Lin, R. W. Kriz, C. S. Ramesha, L. A. Sultzman, A. Y. Lin, N. Milona et J. L. Knopf (1991). A novel arachidonic acid-selective cytosolic PLA<sub>2</sub> contains a Ca(2+)- dependent translocation domain with homology to PKC and GAP. *Cell* **65**, 1043-51.
- Clark, J. D., N. Milona et J. L. Knopf (1990). Purification of a 110-kilodalton cytosolic phospholipase A2 from the human monocytic cell line U937. *Proc Natl Acad Sci USA* **87**, 7708-12.

- Clark, J. D., A. R. Schievella, E. A. Nalefski et L. L. Lin (1995). Cytosolic phospholipase A2. *J Lipid Mediat Cell Signal* **12**, 83-117.
- Clark, V. M. (1986). The Cell Biology of the Retinal Pigment Epithelium. Ruben Adler Debra Farber ed. In "Cellular Biology A Series The Retina for Cell Biology Studies". Academic Press, Los Angeles.
- Clark, V. M. et M. O. Hall (1986). RPE cell surface proteins in normal and dystrophic rats. *Invest Ophthalmol Vis Sci* **27**, 136-44.
- Clarke, S. (1992). Protein isoprenylation and methylation at carboxyl-terminal cysteine residues. *Annu Rev Biochem* **61**, 355-86.
- Cobb, M. H., T. G. Boulton et D. J. Robbins (1991). Extracellular signal-regulated kinases: ERKs in progress. *Cell Regul* **2**, 965-78.
- Damron, D. S., D. R. Van Wagoner, C. S. Moravec et M. Bond (1993). Arachidonic acid and endothelin potentiate Ca<sup>2+</sup> transients in rat cardiac myocytes via inhibition of distinct K<sup>+</sup> channels. *J Biol Chem* **268**, 27335-44.
- Davidson, F. F. et E. A. Dennis (1990). Evolutionary relationships and implications for the regulation of phospholipase A2 from snake venom to human secreted forms. *J Mol Evol* **31**, 228-38.
- de Carvalho, M. G., J. Garritano et C. C. Leslie (1995). Regulation of lysophospholipase activity of the 85-kDa phospholipase A2 and activation in mouse peritoneal macrophages. *J Biol Chem* **270**, 20439-46.
- de Carvalho, M. G., A. L. McCormack, E. Olson, F. Ghomashchi, M. H. Gelb, J. R. Yates, 3rd et C. C. Leslie (1996). Identification of phosphorylation sites of human 85-kDa cytosolic phospholipase A2 expressed in insect cells and present in human monocytes. *J Biol Chem* **271**, 6987-97.
- Dennis, E. A. (1983). In "Enzymes" (P. Boyer, Ed.), Vol. 16, pp. 308-353. Academic Press, New York.
- Dessen, A. (2000). Structure and mechanism of human cytosolic phospholipase A(2). *Biochim Biophys Acta* **1488**, 40-7.
- Diez, E., F. H. Chilton, G. Stroup, R. J. Mayer, J. D. Winkler et A. N. Fonteh (1994). Fatty acid and phospholipid selectivity of different phospholipase A2 enzymes studied by using a mammalian membrane as substrate. *Biochem J* **301** ( Pt 3), 721-6.



- Diez, E., P. Louis-Flamberg, R. H. Hall et R. J. Mayer (1992). Substrate specificities and properties of human phospholipases A2 in a mixed vesicle model. *J Biol Chem* **267**, 18342-8.
- Dufton, M. J. et R. C. Hider (1983). Classification of phospholipases A2 according to sequence. Evolutionary and pharmacological implications. *Eur J Biochem* **137**, 545-51.
- Eisel, Doris (1995). The DIG System User's Guide for Filter Hybridization (R. van Miltenburg, Rüger, B., Grünewald-Janhl, S., Leons, M., Schröder, C., Ed.). Boehringer Mannheim, Germany.
- Evans, J. H., D. M. Spencer, A. Zweifach et C. C. Leslie (2001). Intracellular calcium signals regulating cytosolic phospholipase A2 translocation to internal membranes. *J Biol Chem* **276**, 30150-60.
- Fairbairn, D. (1945). *J Biol Chem* **157**, 633-644.
- Feeney, L. (1973). The phagolysosomal system of the pigment epithelium. A key to retinal disease. *Invest Ophthalmol* **12**, 635-8.
- Fliesler, S. J. et R. E. Anderson (1983). Chemistry and metabolism of lipids in the vertebrate retina. *Prog Lipid Res* **22**, 79-131.
- Fukuda, T., D. K. Kim, M. R. Chin, C. A. Hales et J. V. Bonventre (1999). Increased group IV cytosolic phospholipase A2 activity in lungs of sheep after smoke inhalation injury. *Am J Physiol* **277**, L533-42.
- Gijon, M. A. et C. C. Leslie (1999). Regulation of arachidonic acid release and cytosolic phospholipase A2 activation. *J Leukoc Biol* **65**, 330-6.
- Gijon, M. A., D. M. Spencer, A. L. Kaiser et C. C. Leslie (1999). Role of phosphorylation sites and the C2 domain in regulation of cytosolic phospholipase A2. *J Cell Biol* **145**, 1219-32.
- Giusto, N. M., S. J. Pasquare, G. A. Salvador, P. I. Castagnet, M. E. Roque et M. G. Illicheta de Boschero (2000). Lipid metabolism in vertebrate retinal rod outer segments. *Prog Lipid Res* **39**, 315-91.
- Glover, S., M. S. de Carvalho, T. Bayburt, M. Jonas, E. Chi, C. C. Leslie et M. H. Gelb (1995). Translocation of the 85-kDa phospholipase A2 from cytosol to the nuclear envelope in rat basophilic leukemia cells stimulated with calcium ionophore or IgE/antigen. *J Biol Chem* **270**, 15359-67.



- Gordon, W. C., E. B. Rodriguez de Turco et N. G. Bazan (1992). Retinal pigment epithelial cells play a central role in the conservation of docosahexaenoic acid by photoreceptor cells after shedding and phagocytosis. *Curr Eye Res* **11**, 73-83.
- Gronich, J. H., J. V. Bonventre et R. A. Nemenoff (1990). Purification of a high-molecular-mass form of phospholipase A2 from rat kidney activated at physiological calcium concentrations. *Biochem J* **271**, 37-43.
- Guénard, H. et al. (1996). Physiologie humaine (É. Pradel, Ed.), pp. 103-108, Paris.
- Hanahan, D. J. (1986). Platelet activating factor: a biologically active phosphoglyceride. *Annu Rev Biochem* **55**, 483-509.
- Hanel, A. M. et M. H. Gelb (1995). Multiple enzymatic activities of the human cytosolic 85-kDa phospholipase A2: hydrolytic reactions and acyl transfer to glycerol. *Biochemistry* **34**, 7807-18.
- Hattori, K., H. Adachi, A. Matsuzawa, K. Yamamoto, M. Tsujimoto, J. Aoki, M. Hattori, H. Arai et K. Inoue (1996). cDNA cloning and expression of intracellular platelet-activating factor (PAF) acetylhydrolase II. Its homology with plasma PAF acetylhydrolase. *J Biol Chem* **271**, 33032-8.
- Hattori, K., M. Hattori, H. Adachi, M. Tsujimoto, H. Arai et K. Inoue (1995a). Purification and characterization of platelet-activating factor acetylhydrolase II from bovine liver cytosol. *J Biol Chem* **270**, 22308-13.
- Hattori, M., H. Adachi, J. Aoki, M. Tsujimoto, H. Arai et K. Inoue (1995b). Cloning and expression of a cDNA encoding the beta-subunit (30-kDa subunit) of bovine brain platelet-activating factor acetylhydrolase. *J Biol Chem* **270**, 31345-52.
- Hattori, M., H. Adachi, M. Tsujimoto, H. Arai et K. Inoue (1994a). The catalytic subunit of bovine brain platelet-activating factor acetylhydrolase is a novel type of serine esterase. *J Biol Chem* **269**, 23150-5.
- Hattori, M., H. Adachi, M. Tsujimoto, H. Arai et K. Inoue (1994b). Miller-Dieker lissencephaly gene encodes a subunit of brain platelet-activating factor acetylhydrolase [corrected]. *Nature* **370**, 216-8.
- Hefner, Y., A. G. Borsch-Haubold, M. Murakami, J. I. Wilde, S. Pasquet, D. Schieltz, F. Ghomashchi, J. R. Yates, 3rd, C. G. Armstrong, A. Paterson, P. Cohen, R. Fukunaga, T. Hunter, I. Kudo, S. P. Watson et M. H. Gelb (2000). Serine 727 phosphorylation and activation of cytosolic phospholipase A2 by MNK1-related protein kinases. *J Biol Chem* **275**, 37542-51.

- Heinrikson, R. L., E. T. Krueger et P. S. Keim (1977). Amino acid sequence of phospholipase A<sub>2</sub>-alpha from the venom of *Crotalus adamanteus*. A new classification of phospholipases A<sub>2</sub> based upon structural determinants. *J Biol Chem* **252**, 4913-21.
- Hirabayashi, T., K. Kume, K. Hirose, T. Yokomizo, M. Iino, H. Itoh et T. Shimizu (1999). Critical duration of intracellular Ca<sup>2+</sup> response required for continuous translocation and activation of cytosolic phospholipase A<sub>2</sub>. *J Biol Chem* **274**, 5163-9.
- Hirabayashi, T. et T. Shimizu (2000). Localization and regulation of cytosolic phospholipase A(2). *Biochim Biophys Acta* **1488**, 124-38.
- Ho, Y. S., L. Swenson, U. Derewenda, L. Serre, Y. Wei, Z. Dauter, M. Hattori, T. Adachi, J. Aoki, H. Arai, K. Inoue et Z. S. Derewenda (1997). Brain acetylhydrolase that inactivates platelet-activating factor is a G-protein-like trimer. *Nature* **385**, 89-93.
- Hogan, M. J., J. A. Alvarado et J. E. Weddel (1971). In "Histology of the human eye" (Saunders, Ed.), Philadelphie.
- Hope, W. C., T. Chen et D. W. Morgan (1993). Secretory phospholipase A<sub>2</sub> inhibitors and calmodulin antagonists as inhibitors of cytosolic phospholipase A<sub>2</sub>. *Agents Actions* **39**, C39-42.
- Horton, R. H., L. A. Moran, R. S. Ochs, D. J. Rawn et G. K. Scrimgeour (1994). Lipides et membranes biologiques. Trans. C. François. In "Principes de biochimie" (DeBoeck-Wesmael, Ed.), pp. 243-280. Neil Patterson Publishers, Bruxelles.
- Hudspeth, A. J. et A. G. Yee (1973). The intercellular junctional complexes of retinal pigment epithelia. *Invest Ophthalmol* **12**, 354-65.
- Jacob, M., P. K. Weech et C. Salesse (1996). Presence of a light-independent phospholipase A<sub>2</sub> in bovine retina but not in rod outer segments. *J Biol Chem* **271**, 19209-18.
- Jacob, M., P. K. Weech et C. Salesse (1997). Bovine retinal pigment epithelium contains novel types of phospholipase A<sub>2</sub>. *Biochem J* **327** ( Pt 2), 455-60.
- Jacob, M., P. K. Weech et C. Salesse (1998). Phospholipases A<sub>2</sub> of rod outer segment-free bovine retinae are different from well-known phospholipases A<sub>2</sub>. *Biochim Biophys Acta* **1391**, 169-80.

- Kapus, A., R. Romanek et S. Grinstein (1994). Arachidonic acid stimulates the plasma membrane H<sup>+</sup> conductance of macrophages. *J Biol Chem* **269**, 4736-45.
- Kol, S., K. Ruutinen-Altman, I. Ben-Shlomo, D. W. Payne, M. Ando et E. Y. Adashi (1997). The rat ovarian phospholipase A2 system: gene expression, cellular localization, activity characterization, and interleukin-1 dependence. *Endocrinology* **138**, 322-31.
- Kozumplik, V., F. Staffa et G. E. Hoffmann (1989). Purification of pancreatic phospholipase A2 from human duodenal juice. *Biochim Biophys Acta* **1002**, 395-7.
- Kramer, R. M., G. C. Checani, A. Deykin, C. R. Pritzker et D. Deykin (1986). Solubilization and properties of Ca<sup>2+</sup>-dependent human platelet phospholipase A2. *Biochim Biophys Acta* **878**, 394-403.
- Kramer, R. M., C. Hession, B. Johansen, G. Hayes, P. McGray, E. P. Chow, R. Tizard et R. B. Pepinsky (1989). Structure and properties of a human non-pancreatic phospholipase A2. *J Biol Chem* **264**, 5768-75.
- Kramer, R. M., E. F. Roberts, J. Manetta et J. E. Putnam (1991). The Ca<sup>2+</sup>-sensitive cytosolic phospholipase A2 is a 100-kDa protein in human monoblast U937 cells. *J Biol Chem* **266**, 5268-72.
- Kramer, R. M., E. F. Roberts, S. L. Um, A. G. Borsch-Haubold, S. P. Watson, M. J. Fisher et J. A. Jakubowski (1996). p38 mitogen-activated protein kinase phosphorylates cytosolic phospholipase A2 (cPLA<sub>2</sub>) in thrombin-stimulated platelets. Evidence that proline-directed phosphorylation is not required for mobilization of arachidonic acid by cPLA<sub>2</sub>. *J Biol Chem* **271**, 27723-9.
- Kramer, R. M. et J. D. Sharp (1997a). Structure, function and regulation of Ca<sup>2+</sup>-sensitive cytosolic phospholipase A2 (cPLA<sub>2</sub>). *FEBS Lett* **410**, 49-53.
- Kramer, R. M. et J. D. Sharp (1997b). Structure, function and regulation of Ca<sup>2+</sup>-sensitive cytosolic phospholipase A2 (cPLA<sub>2</sub>). *FEBS Lett* **410**, 49-53.
- Larsson, P. K., H. E. Claesson et B. P. Kennedy (1998). Multiple splice variants of the human calcium-independent phospholipase A2 and their effect on enzyme activity. *J Biol Chem* **273**, 207-14.
- Lecointe, N., J. Meerabux, M. Ebihara, A. Hill et B. D. Young (1999). Molecular analysis of an unstable genomic region at chromosome band 11q23 reveals a disruption of the gene encoding the alpha2 subunit of platelet-activating factor acetylhydrolase (Pafah1a2) in human lymphoma. *Oncogene* **18**, 2852-9.

- Leslie, C. C. (1997). Properties and regulation of cytosolic phospholipase A2. *J Biol Chem* **272**, 16709-12.
- Lin, L. L., A. Y. Lin et D. L. DeWitt (1992a). Interleukin-1 alpha induces the accumulation of cytosolic phospholipase A2 and the release of prostaglandin E2 in human fibroblasts. *J Biol Chem* **267**, 23451-4.
- Lin, L. L., A. Y. Lin et J. L. Knopf (1992b). Cytosolic phospholipase A2 is coupled to hormonally regulated release of arachidonic acid. *Proc Natl Acad Sci U S A* **89**, 6147-51.
- Lin, L. L., M. Wartmann, A. Y. Lin, J. L. Knopf, A. Seth et R. J. Davis (1993). cPLA<sub>2</sub> is phosphorylated and activated by MAP kinase. *Cell* **72**, 269-78.
- Lio, Y. C. et E. A. Dennis (1998). Interfacial activation, lysophospholipase and transacylase activity of group VI Ca<sup>2+</sup>-independent phospholipase A2. *Biochim Biophys Acta* **1392**, 320-32.
- Lodish, Baltimore, Berk et al. (1997). *Biologie moléculaire de la cellule* (U. DeBoeck, Ed.), pp. 971-978, New York.
- Mancuso, D. J., C. M. Jenkins et R. W. Gross (2000). The genomic organization, complete mRNA sequence, cloning, and expression of a novel human intracellular membrane-associated calcium-independent phospholipase A(2). *J Biol Chem* **275**, 9937-45.
- Marmor, M. F et T. J. Wolfensberger, Eds. (1998). *The Retinal Pigment Epithelium*. Edited by M. F. e. T. J. W. Marmor. Oxford: Oxford University Press.
- Matsumoto, B., D. M. Defoe et J. C. Besharse (1987). Membrane turnover in rod photoreceptors: ensheathment and phagocytosis of outer segment distal tips by pseudopodia of the retinal pigment epithelium. *Proc R Soc Lond B Biol Sci* **230**, 339-54.
- Matsuzawa, A., K. Hattori, J. Aoki, H. Arai et K. Inoue (1997). Protection against oxidative stress-induced cell death by intracellular platelet-activating factor-acetylhydrolase II. *J Biol Chem* **272**, 32315-20.
- Miljanich, G. P., L. A. Sklar, D. L. White et E. A. Dratz (1979). Disaturated and dipolyunsaturated phospholipids in the bovine retinal rod outer segment disk membrane. *Biochim Biophys Acta* **552**, 294-306.

- Miller, B., H. Miller et S. J. Ryan (1987). Vitreoretinal junction in infectious endophthalmitis in a primate eye. *Br J Ophthalmol* **71**, 454-7.
- Miller, S. S. et R. H. Steinberg (1979). Potassium modulation of taurine transport across the frog retinal pigment epithelium. *J Gen Physiol* **74**, 237-59.
- Min, J. H., M. K. Jain, C. Wilder, L. Paul, R. Apitz-Castro, D. C. Aspleaf et M. H. Gelb (1999). Membrane-bound plasma platelet activating factor acetylhydrolase acts on substrate in the aqueous phase. *Biochemistry* **38**, 12935-42.
- Miyashita, A., R. G. Crystal et J. G. Hay (1995). Identification of a 27 bp 5'-flanking region element responsible for the low level constitutive expression of the human cytosolic phospholipase A2 gene. *Nucleic Acids Res* **23**, 293-301.
- Moolenaar, W. H., O. Kranenburg, F. R. Postma et G. C. Zondag (1997). Lysophosphatidic acid: G-protein signalling and cellular responses. *Curr Opin Cell Biol* **9**, 168-73.
- Mosior, M., D. A. Six et E. A. Dennis (1998). Group IV cytosolic phospholipase A2 binds with high affinity and specificity to phosphatidylinositol 4,5-bisphosphate resulting in dramatic increases in activity. *J Biol Chem* **273**, 2184-91.
- Murakami, M., R. Matsumoto, Y. Urade, K. F. Austen et J. P. Arm (1995). c-kit ligand mediates increased expression of cytosolic phospholipase A2, prostaglandin endoperoxide synthase-1, and hematopoietic prostaglandin D2 synthase and increased IgE-dependent prostaglandin D2 generation in immature mouse mast cells. *J Biol Chem* **270**, 3239-46.
- Nakatani, Y., T. Tanioka, S. Sunaga, M. Murakami et I. Kudo (2000). Identification of a cellular protein that functionally interacts with the C2 domain of cytosolic phospholipase A(2)alpha. *J Biol Chem* **275**, 1161-8.
- Nalefski, E. A., L. A. Sultzman, D. M. Martin, R. W. Kriz, P. S. Towler, J. L. Knopf et J. D. Clark (1994). Delineation of two functionally distinct domains of cytosolic phospholipase A2, a regulatory Ca(2+)-dependent lipid-binding domain and a Ca(2+)-independent catalytic domain. *J Biol Chem* **269**, 18239-49.
- Nemenoff, R. A., S. Winitz, N. X. Qian, V. Van Putten, G. L. Johnson et L. E. Heasley (1993). Phosphorylation and activation of a high molecular weight form of phospholipase A2 by p42 microtubule-associated protein 2 kinase and protein kinase C. *J Biol Chem* **268**, 1960-4.

- Parthasarathy, S., U. P. Steinbrecher, J. Barnett, J. L. Witztum et D. Steinberg (1985). Essential role of phospholipase A2 activity in endothelial cell-induced modification of low density lipoprotein. *Proc Natl Acad Sci U S A* **82**, 3000-4.
- Perisic, O., H. F. Paterson, G. Mosedale, S. Lara-Gonzalez et R. L. Williams (1999). Mapping the phospholipid-binding surface and translocation determinants of the C2 domain from cytosolic phospholipase A2. *J Biol Chem* **274**, 14979-87.
- Pfeffer, Bruce A. (1991). Improved Methodology for Cell Culture of Human and Monkey Retinal Pigment Epithelium. In "Progress in Retinal Research" (N. e. J. C. Osborne, Ed.), Vol. 10, pp. 251-291. Pergamon Press, New York.
- Pickard, R. T., X. G. Chiou, B. A. Striffler, M. R. DeFelippis, P. A. Hyslop, A. L. Tebbe, Y. K. Yee, L. J. Reynolds, E. A. Dennis, R. M. Kramer et J. D. Sharp (1996). Identification of essential residues for the catalytic function of 85- kDa cytosolic phospholipase A2. Probing the role of histidine, aspartic acid, cysteine, and arginine. *J Biol Chem* **271**, 19225-31.
- Pickard, R. T., B. A. Striffler, R. M. Kramer et J. D. Sharp (1999). Molecular cloning of two new human paralogs of 85-kDa cytosolic phospholipase A2. *J Biol Chem* **274**, 8823-31.
- Posada, J. et J. A. Cooper (1992). Molecular signal integration. Interplay between serine, threonine, and tyrosine phosphorylation. *Mol Biol Cell* **3**, 583-92.
- Prescott, S. M. (1997). A thematic series on phospholipases. *J Biol Chem* **272**, 15043.
- Qiu, Z. H., M. A. Gijon, M. S. de Carvalho, D. M. Spencer et C. C. Leslie (1998). The role of calcium and phosphorylation of cytosolic phospholipase A2 in regulating arachidonic acid release in macrophages. *J Biol Chem* **273**, 8203-11.
- Qiu, Z. H. et C. C. Leslie (1994). Protein kinase C-dependent and -independent pathways of mitogen-activated protein kinase activation in macrophages by stimuli that activate phospholipase A2. *J Biol Chem* **269**, 19480-7.
- Reynolds, L. J., L. L. Hughes, A. I. Louis, R. M. Kramer et E. A. Dennis (1993). Metal ion and salt effects on the phospholipase A2, lysophospholipase, and transacylase activities of human cytosolic phospholipase A2. *Biochim Biophys Acta* **1167**, 272-80.
- Rhee, S. G. et Y. S. Bae (1997). Regulation of phosphoinositide-specific phospholipase C isozymes. *J Biol Chem* **272**, 15045-8.



- Saari, J. C. (1987). Metabolism and photochemistry in the retina. In "Adler's physiology of the eye: clinical applications" (T. C. M. Company, Ed.), St-Louis.
- Saari, J. C. (2000). Biochemistry of visual pigment regeneration: the Friedenwald lecture. *Invest Ophthalmol Vis Sci* **41**, 337-48.
- Salesse, C., F. Boucher et R. M. Leblanc (1984). An evaluation of purity criteria for bovine rod outer segment membranes. *Anal Biochem* **142**, 258-66.
- Sapir, T., A. Cahana, R. Seger, S. Nekhai et O. Reiner (1999). LIS1 is a microtubule-associated phosphoprotein. *Eur J Biochem* **265**, 181-8.
- Sapirstein, A. et J. V. Bonventre (2000). Specific physiological roles of cytosolic phospholipase A(2) as defined by gene knockouts. *Biochim Biophys Acta* **1488**, 139-48.
- Sarna, T. (1992). Properties and function of the ocular melanin--a photobiophysical view. *J Photochem Photobiol B* **12**, 215-58.
- Sauvadet, A., T. Rohn, F. Pecker et C. Pavoine (1997). Arachidonic acid drives mini-glucagon action in cardiac cells. *J Biol Chem* **272**, 12437-45.
- Schalkwijk, C. G., M. Vervoordeldonk, J. Pfeilschifter et H. van den Bosch (1993). Interleukin-1 beta-induced cytosolic phospholipase A2 activity and protein synthesis is blocked by dexamethasone in rat mesangial cells. *FEBS Lett* **333**, 339-43.
- Schievella, A. R., M. K. Regier, W. L. Smith et L. L. Lin (1995). Calcium-mediated translocation of cytosolic phospholipase A2 to the nuclear envelope and endoplasmic reticulum. *J Biol Chem* **270**, 30749-54.
- Scott, D. L., Z. Otwinowski, M. H. Gelb et P. B. Sigler (1991). Crystal structure of bee-venom phospholipase A2: correction. *Science* **252**, 764.
- Scott, D. L., S. P. White, Z. Otwinowski, W. Yuan, M. H. Gelb et P. B. Sigler (1990). Interfacial catalysis: the mechanism of phospholipase A2. *Science* **250**, 1541-6.
- Seilhamer, J. J., W. Pruzanski, P. Vadas, S. Plant, J. A. Miller, J. Kloss et L. K. Johnson (1989). Cloning and recombinant expression of phospholipase A2 present in rheumatoid arthritic synovial fluid. *J Biol Chem* **264**, 5335-8.
- Serhan, C. N., J. Z. Haeggstrom et C. C. Leslie (1996). Lipid mediator networks in cell signaling: update and impact of cytokines. *Faseb J* **10**, 1147-58.

- Sevanian, A. et E. Kim (1985). Phospholipase A2 dependent release of fatty acids from peroxidized membranes. *J Free Radic Biol Med* **1**, 263-71.
- Sevanian, A. et L. L. McLeod (1987). Cholesterol autoxidation in phospholipid membrane bilayers. *Lipids* **22**, 627-36.
- Sevanian, A., M. L. Wratten, L. L. McLeod et E. Kim (1988). Lipid peroxidation and phospholipase A2 activity in liposomes composed of unsaturated phospholipids: a structural basis for enzyme activation. *Biochim Biophys Acta* **961**, 316-27.
- Shakib, M., P. Rutkowski et G. N. Wise (1972). Fluorescein angiography and the retinal pigment epithelium. *Am J Ophthalmol* **74**, 206-18.
- Sharp, J. D., R. T. Pickard, X. G. Chiou, J. V. Manetta, S. Kovacevic, J. R. Miller, A. D. Varshavsky, E. F. Roberts, B. A. Strifler, D. N. Brems et et al. (1994). Serine 228 is essential for catalytic activities of 85-kDa cytosolic phospholipase A2. *J Biol Chem* **269**, 23250-4.
- Sharp, J. D., D. L. White, X. G. Chiou, T. Goodson, G. C. Gamboa, D. McClure, S. Burgett, J. Hoskins, P. L. Skatrud, J. R. Sportsman et et al. (1991). Molecular cloning and expression of human Ca(2+)-sensitive cytosolic phospholipase A2. *J Biol Chem* **266**, 14850-3.
- Shimizu, T. et L. S. Wolfe (1990). Arachidonic acid cascade and signal transduction. *J Neurochem* **55**, 1-15.
- Sierra-Honigsmann, M. R., J. R. Bradley et J. S. Pober (1996). "Cytosolic" phospholipase A2 is in the nucleus of subconfluent endothelial cells but confined to the cytoplasm of confluent endothelial cells and redistributes to the nuclear envelope and cell junctions upon histamine stimulation. *Lab Invest* **74**, 684-95.
- Six, D. A. et E. A. Dennis (2000). The expanding superfamily of phospholipase A(2) enzymes: classification and characterization. *Biochim Biophys Acta* **1488**, 1-19.
- Song, C., X. J. Chang, K. M. Bean, M. S. Proia, J. L. Knopf et R. W. Kriz (1999). Molecular characterization of cytosolic phospholipase A2-beta. *J Biol Chem* **274**, 17063-7.
- Stafforini, D. M., S. M. Prescott et T. M. McIntyre (1987). Human plasma platelet-activating factor acetylhydrolase. Purification and properties. *J Biol Chem* **262**, 4223-30.



- Stafforini, D. M., L. W. Tjoelker, S. P. McCormick, D. Vaitkus, T. M. McIntyre, P. W. Gray, S. G. Young et S. M. Prescott (1999). Molecular basis of the interaction between plasma platelet-activating factor acetylhydrolase and low density lipoprotein. *J Biol Chem* **274**, 7018-24.
- Stephens, W. W., J. L. Walkers et W. Myers (1898). The action of cobra poison on the blood : a contribution to the study of passive immunity. *J. Pathol. Bacteriol.* **5**, 279-301.
- Tang, J., R. W. Kriz, N. Wolfman, M. Shaffer, J. Seehra et S. S. Jones (1997). A novel cytosolic calcium-independent phospholipase A2 contains eight ankyrin motifs. *J Biol Chem* **272**, 8567-75.
- Tay, A., P. Maxwell, Z. G. Li, H. Goldberg et K. Skorecki (1994a). Cytosolic phospholipase A2 gene expression in rat mesangial cells is regulated post-transcriptionally. *Biochem J* **304**, 417-22.
- Tay, A., P. Maxwell, Z. Li, H. Goldberg et K. Skorecki (1994b). Isolation of promoter for cytosolic phospholipase A2 (cPLA<sub>2</sub>). *Biochim Biophys Acta* **1217**, 345-7.
- Tay, A., J. S. Simon, J. Squire, K. Hamel, H. J. Jacob et K. Skorecki (1995). Cytosolic phospholipase A2 gene in human and rat: chromosomal localization and polymorphic markers. *Genomics* **26**, 138-41.
- Tischfield, J. A. (1997). A reassessment of the low molecular weight phospholipase A2 gene family in mammals. *J Biol Chem* **272**, 17247-50.
- Tjoelker, L. W., C. Wilder, C. Eberhardt, D. M. Stafforini, G. Dietsch, B. Schimpf, S. Hooper, H. Le Trong, L. S. Cousens, G. A. Zimmerman et et al. (1995). Anti-inflammatory properties of a platelet-activating factor acetylhydrolase. *Nature* **374**, 549-53.
- Tortora, G. et S.R. Grabowski, Ed. (1994). Principes d'anatomie et de physiologie. Nouvelle édition ed. Anjou: Centre Éducatif et Culturel Inc.
- Trotti, D., A. Volterra, K. P. Lehre, D. Rossi, O. Gjesdal, G. Racagni et N. C. Danbolt (1995). Arachidonic acid inhibits a purified and reconstituted glutamate transporter directly from the water phase and not via the phospholipid membrane. *J Biol Chem* **270**, 9890-5.
- Underwood, K. W., C. Song, R. W. Kriz, X. J. Chang, J. L. Knopf et L. L. Lin (1998). A novel calcium-independent phospholipase A2, cPLA<sub>2</sub>-gamma, that is prenylated and contains homology to cPLA<sub>2</sub>. *J Biol Chem* **273**, 21926-32.

- van Kuijk, F. J. et E. A. Dratz (1987). Detection of phospholipid peroxides in biological samples. *Free Radic Biol Med* **3**, 349-54.
- Van Themsche, C. (1999). Mémoire de maîtrise. Université du Québec à Trois-Rivières.
- Van Themsche, C., M. Jacob et C. Salesse (2001). Human retinal pigment epithelium secretes a phospholipase A2 and contains two novel intracellular phospholipases A2. *Biochem Cell Biol* **79**, 1-10.
- Venable, M. E., G. A. Zimmerman, T. M. McIntyre et S. M. Prescott (1993). Platelet-activating factor: a phospholipid autacoid with diverse actions. *J Lipid Res* **34**, 691-702.
- Waite, M. (1987). The Phospholipases. *Handbook of Lipid Research* **5**.
- Watanabe, M., J. Aoki, H. Manya, H. Arai et K. Inoue (1998). Molecular cloning of cDNAs encoding alpha1, alpha2, and beta subunits of rat brain platelet-activating factor acetylhydrolase. *Biochim Biophys Acta* **1401**, 73-9.
- White, S. P., D. L. Scott, Z. Otwinowski, M. H. Gelb et P. B. Sigler (1990). Crystal structure of cobra-venom phospholipase A2 in a complex with a transition-state analogue. *Science* **250**, 1560-3.
- Winitz, S., S. K. Gupta, N. X. Qian, L. E. Heasley, R. A. Nemenoff et G. L. Johnson (1994). Expression of a mutant Gi2 alpha subunit inhibits ATP and thrombin stimulation of cytoplasmic phospholipase A2-mediated arachidonic acid release independent of Ca<sup>2+</sup> and mitogen-activated protein kinase regulation. *J Biol Chem* **269**, 1889-95.
- Winstead, M. V., J. Balsinde et E. A. Dennis (2000). Calcium-independent phospholipase A(2): structure and function. *Biochim Biophys Acta* **1488**, 28-39.
- Wratten, M. L., E. Gratton, M. van de Ven et A. Sevanian (1989). DPH lifetime distributions in vesicles containing phospholipid hydroperoxides. *Biochem Biophys Res Commun* **164**, 169-75.
- Wu, G. S., A. Sevanian et N. A. Rao (1992). Detection of retinal lipid hydroperoxides in experimental uveitis. *Free Radic Biol Med* **12**, 19-27.
- Wu, T., T. Ikezono, C. W. Angus et J. H. Shelhamer (1994a). Characterization of the promoter for the human 85 kDa cytosolic phospholipase A2 gene. *Nucleic Acids Res* **22**, 5093-8.

- Wu, T., T. Ikezono, C. W. Angus et J. H. Shelhamer (1996). Tumor necrosis factor- $\alpha$  induces the 85-kDa cytosolic phospholipase A2 gene expression in human bronchial epithelial cells. *Biochim Biophys Acta* **1310**, 175-84.
- Wu, T., S. J. Levine, M. G. Lawrence, C. Logun, C. W. Angus et J. H. Shelhamer (1994b). Interferon- $\gamma$  induces the synthesis and activation of cytosolic phospholipase A2. *J Clin Invest* **93**, 571-7.
- Xing, M. et R. Mattera (1992). Phosphorylation-dependent regulation of phospholipase A2 by G-proteins and  $\text{Ca}^{2+}$  in HL60 granulocytes. *J Biol Chem* **267**, 25966-75.
- Young, R. W. (1970). Visual cells. *Sci Am* **223**, 80-91.
- Young, R. W. et D. Bok (1969). Participation of the retinal pigment epithelium in the rod outer segment renewal process. *J Cell Biol* **42**, 392-403.
- Yu, W., P. T. Bozza, D. M. Tzizik, J. P. Gray, J. Cassara, A. M. Dvorak et P. F. Weller (1998). Co-compartmentalization of MAP kinases and cytosolic phospholipase A2 at cytoplasmic arachidonate-rich lipid bodies. *Am J Pathol* **152**, 759-69.
- Zimmerman, W. F., W. Godchaux, 3rd et M. Belkin (1983). The relative proportions of lysosomal enzyme activities in bovine retinal pigment epithelium. *Exp Eye Res* **36**, 151-8.
- Zvaritch, E., G. Lambeau et M. Lazdunski (1996). Endocytic properties of the M-type 180-kDa receptor for secretory phospholipases A2. *J Biol Chem* **271**, 250-7.

## ANNEXE A

### ARTICLE SOUMIS

**Full title :** Cloning of the phospholipases A2 expressed by the human retinal pigment epithelium.

**Short title :** Expression of phospholipases A2

Word count : 3286

Number of table :1

Number of figure :1

**Section code : RE**

**Keywords :** phospholipase, retinal pigment epithelium, cloning, RT-PCR.

**Authors and degrees :** Marc-André Laurin, B.Sc., Vicky Beaudoin B.Sc., Christian Salesse Ph.D.

**Affiliation :** Groupe de recherche en énergie et information biomoléculaire, Département de chimie-biologie, Université du Québec à Trois-Rivières, Trois-Rivières, Québec, G9A 5H7, Canada.

**Correspondence:** Dr. Christian Salesse, Université du Québec à Trois-Rivières, Trois-Rivières, QC, G9A 5H7, Canada. Phone (819) 376-5011 ext. 3308, Fax (819) 376-5057, E-mail ([christian\\_salesse@uqtr.ca](mailto:christian_salesse@uqtr.ca)).

#### Abstract

**Purpose.** Clone and sequence the phospholipases A2 (PLA<sub>2</sub>) expressed by human retinal pigmented epithelial (RPE) cells. **Methods.** Primers of different PLA<sub>2</sub> were assayed by RT-PCR using RNA extracted from RPE prepared from human eye donors to specifically amplify the transcript of these proteins. The PCR products were then cloned and their identity was confirmed using NCBI data bank. **Results.** Human RPE cells express secreted PLA<sub>2</sub>s of group IIA and group V as well as the cytosolic PLA<sub>2</sub> gamma which is a member of the group IV PLA<sub>2</sub>. **Conclusions.** RPE cells express three known phospholipases A2, i.e. two different secreted PLA<sub>2</sub> (sPLA<sub>2</sub>) and one cytosolic PLA<sub>2</sub> (cPLA<sub>2</sub>). Indeed, the sequence of both group V sPLA<sub>2</sub> and cPLA<sub>2</sub> gamma corresponds to the published sequences of these enzymes. However, given that we could not obtain the sequence of group IIA sPLA<sub>2</sub> downstream of its catalytic site using different primers designed to amplified the full sequence of this enzyme, we conclude that human RPE cells express an isoform of group IIA sPLA<sub>2</sub>.

## Introduction

Retinal pigment epithelial (RPE) cells are responsible for the shedding and phagocytosis of the outer segments of retinal photoreceptors (for a review, see <sup>1</sup>). The ingested phagosomes will then be hydrolyzed by enzymes of RPE cells such as phospholipases A<sub>2</sub> (PLA<sub>2</sub>). These ubiquitous enzymes play crucial roles in normal cellular functions by participating in the metabolism and turnover of phospholipids <sup>2,3</sup>. PLA<sub>2</sub>s constitute a diverse family of enzymes that hydrolyze the *sn*-2 fatty acyl ester bond of phosphoglycerides, producing free fatty acids and lysophospholipids.

PLA<sub>2</sub>s should be particularly important for the physiology of photoreceptors since their outer segments contain very high amounts of polyunsaturated fatty acids. Indeed, bovine rod outer segments contain more than 50% docosahexaenoic acid (22:6w3) <sup>4</sup> which are located at position *sn*-2 of these phospholipids <sup>5</sup>. It was shown that the rat retina strives to maintain a constant proportion of docosahexaenoic acid <sup>6</sup>. PLA<sub>2</sub>s from RPE cells could thus be used to specifically hydrolyze and recycle these highly unsaturated fatty acids to photoreceptor inner segments.

There are numerous types of PLA<sub>2</sub>. They have been isolated from different sources and classified into three major groups: secretory PLA<sub>2</sub> (sPLA<sub>2</sub>), cytosolic PLA<sub>2</sub> (cPLA<sub>2</sub>), and Ca<sup>2+</sup>-independent PLA<sub>2</sub> (iPLA<sub>2</sub>). There are six subgroups of sPLA<sub>2</sub> in mammals: groups IB, II (A, C, D, E and F), III, V, VIIA and X. Except for group VIIA, sPLA<sub>2</sub>s are typically small proteins (≈14 kD) which contain multiple disulfide bridges. In addition, they all use a catalytic histidine and need μM Ca<sup>2+</sup> concentrations for hydrolysis <sup>7,8</sup>. cPLA<sub>2</sub>s, which are also called group IV PLA<sub>2</sub>, are subdivided into three subgroups: cPLA<sub>2</sub> α, cPLA<sub>2</sub> β and cPLA<sub>2</sub> γ. They are much larger than the sPLA<sub>2</sub>s (61-114 kD) and use a catalytic serine instead of a histidine <sup>9,10</sup>. These three cPLA<sub>2</sub> isoforms contain two homologous domains named A and B which are interspaced with a sequence specific to each isoform. cPLA<sub>2</sub> α and β contain an amino-terminal domain called C2 that has been shown to be involved in their Ca<sup>2+</sup>-dependent binding to membranes. In contrast, cPLA<sub>2</sub> γ lacks this C2 domain and, accordingly, it is calcium-independent <sup>11,12</sup>. cPLA<sub>2</sub> β has additional domains next to the C2 domain, making this protein the largest one in this group IV. iPLA<sub>2</sub>, or group VI PLA<sub>2</sub>, like the other cPLA<sub>2</sub>s, uses the same catalytic nucleophilic serine but are not dependent on Ca<sup>2+</sup> for membrane binding.

We have previously shown that PLA<sub>2</sub> activity can be measured in bovine RPE cells <sup>13</sup> and that these PLA<sub>2</sub>s are different from known PLA<sub>2</sub>s <sup>14</sup>. More recently, we have evaluated the sensitivity to different treatments of partially purified PLA<sub>2</sub>s from human RPE cells and performed immunoblot experiments <sup>15</sup>. These results suggested that human RPE cells probably contain two novel types of intracellular PLA<sub>2</sub> enzymes, one of them being similar to a bovine intracellular RPE-PLA<sub>2</sub> recently identified <sup>14</sup>, and that cultured human RPE cells secrete a sPLA<sub>2</sub> enzyme. In the present paper, we present our results on the cloning of the types of PLA<sub>2</sub> expressed by human RPE cells.



## Methods

### EXTRACTION OF RNA FROM RPE CELLS AND AMPLIFICATION OF DIFFERENT PLA<sub>2</sub>S

Human eyes were obtained from the National Eye Bank Inc. (Sainte-Foy, Que.). The eyes were dissected to collect the RPE cells as described by Van Temsche et al.<sup>15</sup>. RPE cells were either cultured as described<sup>15</sup> or used immediately. RPE cells were kept on ice prior to RNA extraction. Total RNA was extracted from fresh or cultured RPE cells according to the method of Chomczynski<sup>16,17</sup> (Tri Reagent, Sigma Co., St-Louis, MO). The quality of total RNA was subsequently verified on agarose gel. The RNA was transformed in cDNA using the SuperScript II RNase H-Reverse Transcriptase (Gibco BRL, Burlington, Ont) using 5 µg of total RNA. PCR reactions were performed at 55 °C for 35 cycles. Elongation times were 1 and 3 minutes for sPLA<sub>2</sub> and cPLA<sub>2</sub>, respectively. The reaction mixture contained 2 µL of cDNA template, 2.5 U of *Taq* DNA polymerase, and 10 % DMSO, 1.5 mM MgCl<sub>2</sub>, 200 µM dNTP mix as well as 0.5 mM of each NMT primers in a total volume of 50 µL. The primers used for known secreted and cytosolic PLA<sub>2</sub>s are presented in table 1. The primers were synthesized with a DNA Synthesizer ABI-394 from Applied Biosystem Inc by the Service de synthèse of the centre de recherche du CHUL (Ste-Foy, Québec).

### CLONING OF HUMAN RPE PLA<sub>2</sub>S

The PCR products were then separated by electrophoresis on 1% agarose gel in TBE and the bands of interest were extracted and purified with Ultrafree-DA columns (Millipore Co, Bedford, Mass.). These bands were then cloned into pGEM-T easy cloning vector (Promega Co, Madison, WI) and then transformed into *E. coli* DH5α strain. Plasmid purification was performed with the High Pure Plasmid Isolation Kit (Roche, Laval, Que.) followed by an Eco R1 (New England Bio-Labs, Mississauga, ON, Canada) digestion to select positives clones which were then sent for sequencing (SUCOF, Ste-Foy, Que.) using T7 and SP6 sequencing primers.

## Results

Groups IB, IIC-F, III, VIIA, X, XII sPLA<sub>2</sub> as well as cPLA<sub>2</sub> α and cPLA<sub>2</sub> β could not be amplified by RT-PCR using RNA extracted from human RPE cells. In fact, bands have been obtained, but after cloning and sequencing, they were shown to correspond to other types of proteins which, thus, strongly suggests that these enzymes are not expressed by RPE cells. Figure 1A presents the electrophoresis of the PCR products prepared using primers for group IIA (gIIA) and group V (gV) sPLA<sub>2</sub>s. It can be seen that the strongest bands migrated to a position which is in very good agreement with the number of base pairs calculated on the basis of the designed primers (see Table 1). Indeed, gIIA and gV sPLA<sub>2</sub>s migrated to ~ 300 and 550 bp, respectively. Sequencing of these bands revealed the presence of both gIIA and gV sPLA<sub>2</sub>s. These sequences possess 100% homology with the known sequences for these sPLA<sub>2</sub>. However, only the sequence upstream of the catalytic site (H48 and D49) of gIIA sPLA<sub>2</sub> could be amplified. Other primers designed downstream the catalytic site failed to amplify either part or the full sequence of gIIA sPLA<sub>2</sub>. These results suggest that RPE cells express an isoform of the gIIA sPLA<sub>2</sub>. In contrast, the entire cDNA of gV PLA<sub>2</sub> was successfully amplified with a polymorphism substitution on the allele 140 (T140C)<sup>18</sup>. Figure 1B

shows the electrophoresis of the PCR products prepared using primers for cPLA<sub>2</sub>  $\gamma$ . The strong band located at ~1500 bp corresponds very well to the number of base pairs calculated from the primers (Table 1) and the known sequence of this enzyme. Cloning and sequencing of this band showed 100% homology with the known cPLA<sub>2</sub> $\gamma$ .

## Discussion

The results of the specific amplification of PLA<sub>2</sub>s by RT-PCR show the presence of transcripts of gIIA and gV sPLA<sub>2</sub>s as well as cPLA<sub>2</sub>  $\gamma$  in RPE cells. We have obtained the complete coding sequence for gV sPLA<sub>2</sub> as well as cPLA  $\gamma$ , whereas only a partial clone for gIIA was obtained from the N-terminal up to the conserved amino acid (H48 and D49) in the catalytic region. Attempts to amplify gIIA downstream of this catalytic region failed. The amplification of the complete sequence of gIIA is in progress. We thus propose in agreement with previous results<sup>15</sup> that an isoform of gIIA sPLA<sub>2</sub> could be present in RPE cells.

The results presented here partly contrast with our previous conclusions on the PLA<sub>2</sub> content of human RPE cells<sup>15</sup>. Indeed, on the basis of the sensitivity of different PLA<sub>2</sub>-active fractions eluted from cation-exchange chromatography to para-bromophenacylbromide (pBPB), Ca<sup>2+</sup>/EGTA, DTT, heat and H<sub>2</sub>SO<sub>4</sub>, we had concluded that human cultured RPE cells probably contain two different intracellular PLA<sub>2</sub> enzymes. In addition, we had performed immunoblot experiments where no cross-reactivity was observed between the RPE intracellular PLA<sub>2</sub> enzymes and several antisera directed against sPLA<sub>2</sub> and cPLA<sub>2</sub>, and thus concluded that intracellular RPE-PLA<sub>2</sub>s are different from well-known secretory, cytosolic and Ca<sup>2+</sup>-independent PLA<sub>2</sub>s. However, only antibodies against cPLA<sub>2</sub>  $\alpha$  were assayed as antibodies against cPLA<sub>2</sub>  $\beta$  and cPLA<sub>2</sub>  $\gamma$  are not commercially available. In fact, the most striking difference between these data and the present results is that cPLA<sub>2</sub>  $\gamma$  is calcium-independent<sup>11,12</sup> whereas we have previously demonstrated that intracellular PLA<sub>2</sub> from RPE cells are calcium dependent. It can thus be concluded that other types of intracellular PLA<sub>2</sub> should be expressed by human RPE cells and that these enzymes could be novel types of intracellular PLA<sub>2</sub> enzymes.

We also previously reported an additional RPE-PLA<sub>2</sub> enzyme that was secreted and which exhibited sensitivity to pBPB, Ca<sup>2+</sup>/EGTA, DTT, heat and H<sub>2</sub>SO<sub>4</sub> characteristic of sPLA<sub>2</sub> enzymes<sup>15</sup>. This PLA<sub>2</sub>, with an apparent molecular weight of approximately 22 kDa, cross-reacted weakly with an antiserum directed against porcine pancreatic group I sPLA<sub>2</sub> but strongly with an antiserum directed against N-terminal residues 1-14 of human synovial group II sPLA<sub>2</sub> (which is now called group IIA sPLA<sub>2</sub>), which suggested that this extracellular enzyme was a member of the sPLA<sub>2</sub> class of enzymes<sup>15</sup>. These data are in good agreement with the present finding of the expression of a group IIA sPLA<sub>2</sub> in RPE cells. In addition, our previous observation of a molecular weight of 22 kD for this enzyme<sup>15</sup> agrees well with the fact that we could not amplify by PCR the full sequence of this enzyme when using primers designed on the known sequence of this group of sPLA<sub>2</sub> and supports our conclusion that this sPLA<sub>2</sub> is an

isoform of the known gIIA sPLA<sub>2</sub>. Moreover, the present observation of the expression of an additional sPLA<sub>2</sub>, namely a gV sPLA<sub>2</sub>, by RPE cells could not be documented in our previous report<sup>15</sup> as antibodies against this enzyme became only very recently available.

It is very interesting to note that RPE cells express sPLA<sub>2</sub> enzymes as they could be involved in the phagocytosis of photoreceptor outer segments by the RPE. Indeed, it was previously shown that the M-type receptor, which is closely related to mannose receptors<sup>20</sup>, is involved in the phagocytosis process in macrophages and is also present in the plasma membrane of RPE cells<sup>21,22</sup>. Moreover, it has been demonstrated that binding of different types of sPLA<sub>2</sub> to M-type receptor leads to the activation of phagocytosis<sup>23,24</sup>. It can thus be proposed that gIIA or gV sPLA<sub>2</sub>s or both could be involved in the phagocytosis of photoreceptor outer segments (POS) by RPE cells. In addition, it was shown that pretreatment of POS with gIIA sPLA<sub>2</sub> increases the rate of disk membrane/plasma membrane fusion by five-fold, in a dose-dependent manner<sup>25</sup>. This fusion is critical to the formation of disk membrane packets at the distal end of the POS, prior to their phagocytosis by RPE cells. Furthermore, it has been demonstrated that gIB, gIIA and gV sPLA<sub>2</sub>s induce activation of cPLA<sub>2</sub>  $\gamma$ <sup>26,27</sup> which could also take place in RPE cells and be responsible for POS phagocytosis. The exact role of gV sPLA<sub>2</sub> is not fully understood but it might be related and complementary to gIIA PLA<sub>2</sub><sup>28</sup>.

### Acknowledgements

**The authors are indebted to the Natural Sciences and Engineering Research Council of Canada (NSERC), the Fonds FCAR and the Fondation des maladies de l'œil for financial support. C.S. is a Chercheur boursier senior of the Fonds de la recherche en santé du Québec. V.B. is indebted to the Fonds FCAR for a scholarship. We are thankful to the National Eye Bank Inc. (Ste-Foy, Québec) for providing us with human eyes.**

### References

1. Besharse JC and Defoe DM. Role of the retinal pigment epithelium in photoreceptor membrane turnover. In: M. F. Marmor and Wolfensberger. The retinal pigment epithelium. New York:Oxford Press University; 1998.
2. Waite M. The phospholipases. In: D. J. Hanahan. Handbook of lipid research. New York:Plenum; 1998:
3. Van den Bosch H, Intracellular phospholipases A, *Biochim Biophys Acta*, 1980, 604, 191-246.
4. Salesse C, Boucher F and Leblanc RM, An evaluation of purity criteria for bovine rod outer segment membranes, *Anal Biochem*, 1984, 142, 258-66.
5. Anderson RE and Sperling L, Lipids of ocular tissues. VII. Positional distribution of the fatty acids in the phospholipids of bovine retina rod outer segments, *Arch Biochem Biophys*, 1971, 144, 673-7.
6. Tinoco J, Miljanich P and Medwadowski B, Depletion of docosahexaenoic acid in retinal lipids of rats fed a linolenic acid-deficient, linoleic acid-containing diet, *Biochim Biophys Acta*, 1977, 486, 575-8.



7. Dennis EA, Diversity of group types, regulation, and function of phospholipase A<sub>2</sub>, *J Biol Chem*, 1994, 269, 13057-60.
8. Six DA and Dennis EA, The expanding superfamily of phospholipase A<sub>2</sub> enzymes: classification and characterization, *Biochim Biophys Acta*, 2000, 1488, 1-19.
9. Clark JD, Lin LL, Kriz RW, et al., A novel arachidonic acid-selective cytosolic PLA<sub>2</sub> contains a Ca(2+)-dependent translocation domain with homology to PKC and GAP, *Cell*, 1991, 65, 1043-51.
10. Nalefski EA, Sultzman LA, Martin DM, et al., Delineation of two functionally distinct domains of cytosolic phospholipase A<sub>2</sub>, a regulatory Ca(2+)-dependent lipid-binding domain and a Ca(2+)-independent catalytic domain, *J Biol Chem*, 1994, 269, 18239-49.
11. Underwood KW, Song C, Kriz RW, et al., A novel calcium-independent phospholipase A<sub>2</sub>, cPLA<sub>2</sub>-gamma, that is prenylated and contains homology to cPLA<sub>2</sub>, *J Biol Chem*, 1998, 273, 21926-32.
12. Pickard RT, Striffler BA, Kramer RM and Sharp JD, Molecular cloning of two new human paralogs of 85-kDa cytosolic phospholipase A<sub>2</sub>, *J Biol Chem*, 1999, 274, 8823-31.
13. Jacob M, Weech PK and Salesse C, Presence of a light-independent phospholipase A<sub>2</sub> in bovine retina but not in rod outer segments, *J Biol Chem*, 1996, 271, 19209-18.
14. Jacob M, Weech PK and Salesse C, Bovine retinal pigment epithelium contains novel types of phospholipase A<sub>2</sub>, *Biochem J*, 1997, 327, 455-60.
15. Van Themsche C, Jacob M and Salesse C, Human retinal pigment epithelium secretes a phospholipase A<sub>2</sub> and contains two novel intracellular phospholipases A<sub>2</sub>, *Biochem Cell Biol*, 2001, 79, 1-10.
16. Chomczynski P and Sacchi N, Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction, *Anal Biochem*, 1987, 162, 156-9.
17. Chomczynski P, A reagent for the single-step simultaneous isolation of RNA, DNA and proteins from cell and tissue samples, *Biotechniques*, 1993, 15, 532-4, 536-7.
18. NCBI Nucleotide sequences database. National center for Biotechnology Information web site. Available at <http://www.ncbi.nlm.nih.gov>. Accession number XM\_001841. Accessed October 4, 2001.
19. Jacob M, Weech PK and Salesse C, Phospholipases A<sub>2</sub> of rod outer segment-free bovine retinae are different from well-known phospholipases A<sub>2</sub>, *Biochim Biophys Acta*, 1998, 1391, 169-80.
20. Nicolas JP, Lambeau G and Lazdunski M, Identification of the binding domain for secretory phospholipases A<sub>2</sub> on their M-type 180-kDa membrane receptor, *J Biol Chem*, 1995, 270, 28869-73.
21. Wilt SD and McLaughlin BJ, Is another mannose receptor-like lectin present in retinal pigment epithelium?, *Curr Eye Res*, 1999, 19, 1-3.
22. Wilt SD, Greated CJ, Lutz DA and McLaughlin BJ, Mannose receptor is expressed in normal and dystrophic retinal pigment epithelium, *Exp Eye Res*, 1999, 69, 405-11.
23. Zvaritch E, Lambeau G and Lazdunski M, Endocytic properties of the M-type

- 180-kDa receptor for secretory phospholipases A<sub>2</sub>, *J Biol Chem*, 1996, 271, 250-7.
24. Valentin E and Lambeau G, Increasing molecular diversity of secreted phospholipases A<sub>2</sub> and their receptors and binding proteins, *Biochim Biophys Acta*, 2000, 1488, 59-70.
25. Boesze-Battaglia K, Membrane properties that influence bovine retinal ROS membrane fusion, *Invest. Ophthalmol. Visual Sci.*, 1994, 35, 2139.
26. Hernandez M, Burillo SL, Crespo MS and Nieto ML, Secretory phospholipase A<sub>2</sub> activates the cascade of mitogen-activated protein kinases and cytosolic phospholipase A<sub>2</sub> in the human astrocytoma cell line 1321N1, *J Biol Chem*, 1998, 273, 606-12.
27. Hernandez M, Barrero MJ, Alvarez J, et al., Secretory phospholipase A<sub>2</sub> induces phospholipase Cγ-1 activation and Ca<sup>2+</sup> mobilization in the human astrocytoma cell line 1321N1 by a mechanism independent of its catalytic activity, *Biochem Biophys Res Commun*, 1999, 260, 99-104.
28. Murakami M, Shimbara S, Kambe T, et al., The functions of five distinct mammalian phospholipase A<sub>2</sub>S in regulating arachidonic acid release. Type IIa and type V secretory phospholipase A<sub>2</sub>S are functionally redundant and act in concert with cytosolic phospholipase A<sub>2</sub>, *J Biol Chem*, 1998, 273, 14411-23.
29. Seilhamer JJ, Randall TL, Yamanaka M and Johnson LK, Pancreatic phospholipase A<sub>2</sub>: isolation of the human gene and cDNAs from porcine pancreas and human lung, *DNA*, 1986, 5, 519-27.
30. Seilhamer JJ, Pruzanski W, Vadas P, et al., Cloning and recombinant expression of phospholipase A<sub>2</sub> present in rheumatoid arthritic synovial fluid, *J Biol Chem*, 1989, 264, 5335-8.
31. Chen J, Shao C, Lazar V, et al., Localization of group IIc low molecular weight phospholipase A<sub>2</sub> mRNA to meiotic cells in the mouse, *J Cell Biochem*, 1997, 64, 369-75.
32. Ishizaki J, Suzuki N, Higashino K, et al., Cloning and characterization of novel mouse and human secretory phospholipase A<sub>2</sub>s, *J Biol Chem*, 1999, 274, 24973-9.
33. Suzuki N, Ishizaki J, Yokota Y, et al., Structures, enzymatic properties, and expression of novel human and mouse secretory phospholipase A<sub>2</sub>s, *J Biol Chem*, 2000, 275, 5785-93.
34. Valentin E, Ghomashchi F, Gelb MH, Lazdunski M and Lambeau G, On the diversity of secreted phospholipases A<sub>2</sub>. Cloning, tissue distribution, and functional expression of two novel mouse group II enzymes, *J Biol Chem*, 1999, 274, 31195-202.
35. Valentin E, Ghomashchi F, Gelb MH, Lazdunski M and Lambeau G, Novel human secreted phospholipase A<sub>2</sub> with homology to the group III bee venom enzyme, *J Biol Chem*, 2000, 275, 7492-6.
36. Tjoelker LW, Eberhardt C, Unger J, et al., Plasma platelet-activating factor acetylhydrolase is a secreted phospholipase A<sub>2</sub> with a catalytic triad, *J Biol Chem*, 1995, 270, 25481-7.
37. Cupillard L, Koumanov K, Mattei MG, Lazdunski M and Lambeau G, Cloning, chromosomal mapping, and expression of a novel human secretory phospholipase A<sub>2</sub>, *J Biol Chem*, 1997, 272, 15745-52.

38. Gelb MH, Valentin E, Ghomashchi F, Lazdunski M and Lambeau G, Cloning and recombinant expression of a structurally novel human secreted phospholipase A2, J Biol Chem, 2000, 275, 39823-6.
39. Sharp JD, White DL, Chiou XG, et al., Molecular cloning and expression of human Ca(2+)-sensitive cytosolic phospholipase A2, J Biol Chem, 1991, 266, 14850-3.
40. Song C, Chang XJ, Bean KM, et al., Molecular characterization of cytosolic phospholipase A2-beta, J Biol Chem, 1999, 274, 17063-7.

### Figure Legend

Fig. 1. Migration of PCR products of PLA<sub>2</sub>s present in RPE. (A) PCR products of sPLA<sub>2</sub>s. Lane 1 is the Gene Ruler100 bp ladder Plus (Fermentas, Burlington, On). Lanes 2 and 3 are, respectively, PCR products of gIIA and gV sPLA<sub>2</sub>s. (B) PCR products of cPLA<sub>2</sub>. Lane 1 is the  $\lambda$  DNA- Hind III digest and lane 2 is the  $\phi$ X174-Hae III digest. Lane 3 shows the cPLA<sub>2</sub>  $\gamma$  PCR product.

**Table 1. Primers used to amplify different types of PLA<sub>2</sub> expressed by RPE cells**

Types of PLA <sub>2</sub>	Forward primers	Reverse primers	Length of PCR products (bp)
IB <sup>29</sup>	TGGTCATCTCAGTTCTTTTC	TCAACTCTGACAATACTTCT	482
IIA <sup>30</sup>	AGCCACCAAGGAGGAGCAGG	GTCATGAGTGACACAGCAGC	309
IIC <sup>31</sup>	CTGGACAACCTCCACCCTCA	TGGAGTTTGTCCCTGCCACA	406
IID <sup>32</sup>	TGCCAGCATCTGCCTCCACT	AGGAACAGGGTAGAGGGTGT	480
IIE <sup>33</sup>	TGCTCCTTGTGCACCTCCCT	CAGCTTGTGGGATAATGGG	437
IIF <sup>34</sup>	ATGAAGAAGTTCTTCACCGT	TTGGTGA CTGCAGGTGACCT	480
III <sup>35</sup>	GAGGATGGACCATGCCTGGCACA	TTACGGATGTCTTAGTCGTGCTGA	N/A
V	GTGGATACCAATGTCCGAC	CCTAGGAGCAGAGGATGTTG	549
VIIA <sup>36</sup>	AGATGGTGCCACCCAAATTG	TCCTAATTGTATTTCTCTAT	1325
X <sup>37</sup>	GGAATTCGCCTATATGAAATATGGT	GGAATTCAAGGTAGTCAGTCACACTTG	327
XII <sup>38</sup>	TTTGCGGCCGCATATGGAGCTGGCTGCTGCCAAGT	TTTAAGCTTCTAGAATCTGTCACTAGCTGTGGCAT	717
cPLA <sub>2</sub> $\alpha$ <sup>39</sup>	CATTATAGATCCTTACCAG	CTATGCTTTGGGTTTACTTA	2987
cPLA <sub>2</sub> $\beta$ <sup>40</sup>	CACTTCTACCGGGACTGGGTCT	TCCTGGTTGTTGCAGACATTGT	2855
cPLA <sub>2</sub> $\gamma$ <sup>11</sup>	AGTTTCCATAATTCCTGGGCTC	GCCACGTTTCATCAACTCTCTAA	1564

## ANNEXE B

### SÉQUENCES DES NOUVEAUX GÈNES 1 ET 2

Les séquences nucléotidiques suivantes sont présentées dans l'orientation 3' vers 5' des deux nouveaux gènes, c'est-à-dire selon l'orientation initiale des HTGS trouvés. Le premier acide nucléique correspond donc à l'extrémité 3' du clone RP11-35K2.

La position de l'acide nucléique de la première colonne est indiquée immédiatement à gauche de chaque rangée. Et puis, plus à gauche, les chiffres arabes indiquent le numéro de chacun des exons pour les deux nouveaux gènes. Le nouveau gène 1 se trouve au début du clone HTGS (plus en amont sur le chromosome 15) et comprend 26 exons, alors que le nouveau gène 2 se trouve plus en aval et compte 25 exons.

#### RP11-35K2

1	GAATTCTATC	AAATGCCAAA	TACAATTTTT	CCATCTTTAA	AAGAATGAGA	AATTAAGCCA
61	CCATTTAGAC	CTATGTTTTC	TATGTAAATA	AAACTGAAAA	AAAATTTTTT	TTTTTGAGAC
121	TGGGTCTGCG	TCTGTTGCCC	AGGCTGGCTG	GAGTGCAGTG	GTGCAATCTC	GGCTCACTGC
181	TGCCTCGACC	TCCTGGGCTC	CAACAATCCT	CCTGCCTCAG	CCTCCCAAGT	AGCTGGGACT
241	ACAGACACAT	GCCACCATGC	CTAGCTAATC	TTTTGTAGAA	ACAGGGTTTC	ACCATGTTGC
301	CCAGGCTGGT	CTTGAATGCC	TGGGCTCAAG	CTATCCCCTC	ACCTTGGCCT	CCCAAAGTAC
361	TGGGATTACA	GGCATGAGCC	ACCGCAACTG	GCCTAAAACT	GAAATTGTTT	AGCAACCTTA
421	ACGGGCTATA	TTCTAAACCT	TTAATGATG	TCATTGCTAG	TTTTCACTTC	ATCCAAAGTT
481	TCTACGTTTC	TGTGAAGTTG	TGGCACCCTA	GAATAGATAC	ACCATTAAAA	TATGTCTCCT
541	AAATTTACAG	TAAATCTGAT	GATATCTTTT	ATCTCTAAAT	TTAGAGAAAA	AAATCTGCTA
601	ATAGCACACG	GCATTACTAG	AGTGCTTGCT	ATATTCCTGA	TATGGCGCTG	AGCCTTCTAT
661	ATTGTTTCAC	CTCATGCTCT	GAGGTGGATG	ATGACATTAC	CCTCGTTTTA	TTTTTATTTT
721	TTTTTTTTTT	TTTATTTTTT	TTTTTTTGAG	ACGGAGTCTC	GCTCTGTCTG	CCAGGCTGGA
781	GTGCAAGTGC	GGGATCTCGG	CTCACTGCAA	GCTCCGCTTC	CCGGGTTCAC	GCCATTCTCC
841	TGCCTCAGCC	TCCCAGTAG	CTGGGACTAC	AGGCGCCCGC	CACCTACGCC	AGCTAATTTT
901	TTTGTATTTT	TAGTAGAGAC	GGGGTTTCAC	CGTTTTAGCC	GGGATGGTCT	CGATCTCCTG
961	ACCTCGTGAT	CCGCCCGCCT	CGGCCTCCCA	AAAGTCTGGG	ATTACAGGCG	TGAGCCACCG
1021	CGCCCCGGCC	TACCCTCGTT	TTATTTACAG	GTGGGAAAAC	TGGGGCATCA	GAGAGGCCAA
1081	GCGGCTTGCC	CAAGGTCACA	CAGCGGATGT	TCGAGTGGAA	ATGGAATGCA	AGCATTACAGA
1141	CTCCAGAAGT	TGCACTGTCT	TCAGAAATGG	CCTCAAGTTA	GTGGTTTGCT	CAGGGGTGAA
1201	GAGCAAGACA	AAGTTCAGGG	CCTCATCCCA	GGGTGTGTCA	CTTGGCATGA	GGGACGAGGA
1261	CCCCCATTTT	CTCTCAGCTG	AGGGGAAGAG	CTCTCCACAA	TGTCCCCCTG	CACGGTCCCT
1321	TGGCTACCTT	GACAACAAGG	CCGAGCTCTC	CCTACTCTCC	CTGGAGTAAA	GCTGGGCTCA
1381	TGAGGTGCTA	CCGTTTCAAA	ATGTGTTCAG	CTTTAATTTT	TAAAATATCT	TCAAGAGTTT
1441	CTCCATCAAA	ACCTCTGTAT	CCCTAATAGC	CTAGAAATGT	TCGTGTTGAT	CAATTCTTAC
1501	ACATTCCACA	CAAACTGAG	AAATGCTGGT	AAAAGTTACT	GTGAGGGATG	CCCACTCTGT
1561	GAAACTCCAG	CCTCTATGGA	GTGGATGCAG	AGCGTGTGGA	CTCCATATAC	CCATGGGAAT
1621	GTGGAATTTA	GGTTTTACAG	CCCCCTGCGG	ATCTCAAGGG	TGAAGACATT	CAGCAGCGAG
1681	GCTGGAATTC	AAAGCCAGAT	CACACTCTTG	CCTCCACTCA	CGTCCAGTTT	GATTTTTTACG
1741	CCCTGTGGCC	TGGATGGGGC	TATGTCTGGA	GGCTGGTATG	ACTGAGGGAG	GAGAGACTGC
1801	TGGGGAGGGA	GGGGCTTGGG	AGAACAGCAG	CCAGTGTCTG	CCAGCTCCCC	GCAGCACACA
1861	CAGTTCGCAA	TTACAGAAAA	TGTAGGTGTC	GGGAGTGGGT	TTGAAGCTTG	ACAGGCACCT
1921	TGCTCCAGAA	TCAGTGCTAA	TAGCAACGGT	GGGCTGGCCA	AGACGGCCTA	CAAAAGGTCA
1981	CACACACTGA	TTAAAGGGAT	GAACCAGACG	AGAACTGCAG	CGTGTGCTTT	TTGGGGGAAGA
2041	GCTCGGTGCT	GAGCTCCACC	TTGGGTGGCT	CGCGGCAACA	CGGCCCTTTC	TCGGCACCTT
2101	CACCAAGCCT	GGCCAGGAGG	AGGCCCTTCA	GATTCCGGGC	CAGGGAGGTG	AGCATACGGA
2161	TATCGTAGTT	AAGCAGTGAG	TCCCCAGCT	GGACAATCAC	TCACACAGCA	CAGGTGTTGA
2221	AGTCAGACAT	CAGTGTGAG	GAGTGCCTGA	ATTACTTTCT	GATGTTTTCAT	TCCTGTACAA
2281	GTACAGCAAA	TCCCACTACT	TTTGAAATAA	CAGGGGTTC	AAGGAAAAAT	ATAAATGAAA
2341	AAATTGTTTA	AAGAAGAGCA	GATGTTTCCT	ACATTCAAGT	TTATATTGCA	ATCGAAGCCA
2401	GAGTTTTCTA	AGTAGCGACC	CTACTGGACG	GACCTTTAAC	AATAAATCAT	TTTTAAAGTG
2461	TGCTTCAAGT	GTGAAATAAT	TTGAAGTCAA	AGGTGTCCCC	CTTTGACTGT	GACCTTGGTA
2521	AATCAATCTC	TGGATCTCAG	ACACAGAGGA	AGAACTAGAA	TTAGGAAGAA	TTACTGTATT
2581	ATAAAGCTTG	CTTGCAATTAT	CAATTTGTTT	GGTGATACTT	TCACAGATAA	TACAAAGTTA

2641 ACATTTCAGTG TTCAATTTCA CCATCAAAGA AATACCAGTA AAGACAGAGC ACATTAACAA  
 2701 ATTTTGTGAA ACTCGGAATC ATCAATGCTG GCAAGGGTCC TGTGAGATAA ACACCCTTAT  
 2761 ATATGTTGGT GGGGGAATAA GTTGGTCTAA CTTCACTAAA AAGCAATTG ACACCATGTG  
 2821 TTAGAAGCCT TAAAAATGGG GGTCTTGCTG GGAGCGGTGG CTCGTGCCTG TAATTCCAGC  
 2881 ATTTTGGGAG GCCGCAGTGG GAGGATCGCT TGGCCTAGGA GTTCGAGACT AGCCTGGGCA  
 2941 ACATGGAGAA ACCCCGTCTC TACAAAACAT ACAAAAATTA GCCAGGTGTG GTGGCGTGCA  
 3001 TCTGTAGTCC CAGCTACTTA AGAGGCTGGG GTGGGATGGT CGCTTGAGCC CAGGGAGGTT  
 3061 GAAGCTGCAG TGAGCTGTGA CCATGTCAC TCACTCCAGC CTAGGTGACA GAGTGAGACC  
 3121 CTGTCTCTAA ATAAATAAAG AAAAAATAAA CTCAAATGGG GATTTTGGCG AGTTGACCTC  
 3181 TAACAGCTCG TGCTGGTGTT ATTTGAGGGC CTGTTTCCTC TCCTCTCTTG GGCTGTCTAT  
 3241 TCCTGCAGGG CATGGTCTAT ATACTCAGTA CTTTAGAAAC CTGGAGTCTG GTAGGAATGA  
 3301 TTTGAGTCAG CCTCTTATCT TTACAAGCGA CAAGACTGAG GGTAGTGCA GGAACATGGC  
 3361 TTGTTGCTGG TCAAACCTAC AGGGTGGCCA TGCAGGATCC AGAACCTGCA GCCCCTTTCT  
 3421 CCTGGCCCCA GACTTCTGCT CTTACACCAG GCACAGCACC ATCACCTCGG GGTCTACAC  
 3481 TGCCCCCAGC TCCCTGGTGC TCTAAGCAGG ACCTGGCAGA TCCCTGACTT TACATATGAC  
 3541 ACTCTAAGCT GATTGCCCTG AAGAAGATTT TTCCTGGGG GACTGCAGAA ACTCAGATTC  
 3601 ATCTGTTACA CTGCTGGGTA AACAGGTACA TGTCAGTGCA GTGAGGGAAC TCGGAACAAC  
 3661 CAAGACACCA AGTATCTGAC ATGGAAAAAG GGAGAATGAA GTGATGGATC CACACACAGC  
 3721 AAGCTGGGAG GGAAGAGTCT GAGCTGATGC TGAGGTGCAA TGTGCATAAA TGGGTCAATC  
 3781 ACACAGTCAG ACGGGATAGA CAGCCTGATC AACTGTGTTA GACATTTCTC ACGGTGGGAG  
 3841 TGCCCTCAGAG GCTTTCCCAA AGCTACGAGT TTCAGGGACT CTAGTATTAT CTGCAATAAC  
 3901 AACAAATCAA AGGAACAAGA GGACACAGGA AGACCACCGT GATTCAGAGC TGATTAGCCC  
 3961 CTCCACAGGC ATCCTGTGCT TTGGAGACTT GATTTTGTG CTCCTGGCTCA CTTTGCTCCC  
 4021 CCTTTCTGTT GTTCTCGGAG AGGCTGGCCA ATTCCAGCAA TGATTGCTCC AAGGTTCTCG  
 4081 AGCCAAGCCA GCCACATAAA AAGGTATGGA CATTGGGATT CCAAGCAGCT GGGCGGAAGG  
 4141 GATGGTGGGG TGGATGGGAG ACCTGACATC CCCTGTGCTT GGTCCGATGA GGTCCCCAGT  
 4201 TCAGACTCAT CCAAGTGGCT CTGGGGTACC CCGTGTGCCA GCATTTCAGC CAAGCTCACA  
 4261 GAGCTACTTT GATGCAAGAG TATATATGGA TGGTGTCTTC TTTTTCAT TGCTTAGGCT  
 4321 AAGATTTTTC TTTTCACTTC ATAGAAGTGA ATGTCATTTA TGACAAGGGG AAGAAAAAAA  
 4381 ACTCTCAAGA TATTGAAAAC AGAGATACTC TCAATATGTC CAAGCACCAT CTTCTCAACG  
 4441 ATGAGCTTAA GAATGAATAA AAACCAAAAC AATCTAAGTG GCACATTATT CAAACAATG  
 4501 CTGGAGTAGT GGGCAAAGCA TTACACTTTA AATAAATATA CCTAACGCTT GAGGACTCGT  
 4561 TAGTGCTGTA AGTCTTTACA TGAATAGTCT CATTTTCATC TAACCTACCT ATGGGCGAGG  
 4621 ACTATTTCTA TCCCTATTTT ACAGATGAGG AAATAAGGC TCAAAATGGG TAAGCGGTTT  
 4681 ATTAAGATGA AGGATTTCTT ACTAGTCTTA TGAGGACACC ACCCACCTTA AATCAGAGCC  
 4741 TGAGAGGCGG ACGGGTTCCA CTGGGAATGT TTACAGGAAT GTGTCCTGTG AATCATAACA  
 4801 AATACACAAA TGTAAGATTT TAAAAATAAT ACGGCACATA TATTTTGAGT GTGTATGAG  
 4861 CATTATTTTT AAATTATCAA GCCAAATCTA AAGAAATGGT TTGTGTGGGC TGGGCGTGGT  
 4921 GGCTCAGGCC TGTAATCCCA GCACTTTGGG AGGCCGAGGT GGGCGGATCA CGAGGTGAGG  
 4981 AGATCAGAGC CATCCTGGCT AACACGGTGA AACCCCGTCT CTACTAAAAA TACAAAAATA  
 5041 AAATTAGCCA GGCATGGTGG CGGGCGCCTG TAGTCCAGC TACTCGGGAG GCTGAGGCAG  
 5101 GAGAATGGCG TGAACCCGGG AGGCTGAGCT TGCAATGAGC CGAGATCGTG CCACTGCACT  
 5161 CCAGCCTGGG TGACAGAGCA AGACTTTGTC TCAAAAAAAA AAAAAAAA AAAAAAAA  
 5221 GAAATGGTTT GTGTGATGAT AAACGAGATC TTAATAATAG TTCGTGCATA CCTTGGCATC  
 5281 TGTGCATGGA GTTTGGGGT TCAATTCATT GCTATTGCC TAAACATTTT TAAAAAAGT  
 5341 TTTGGAAAGT TTCAATGTAAT ATATTTTCTC TCAGAATACT TACTTGATCC ACTATTTTAC  
 5401 TAGGAGAGGG ATGAAAAGGA AGAGTTACTA CTTCTACTCT GATTTGATTT GAGCCTGGAA  
 5461 AGATGTTTTA AGAGCAGTTT CTGGCCGGG ACCTGGGCTC ACACCTGTAA TCCCAGCACT  
 5521 TTGAGAGGCT GAGGTGGATG GATCATGTGA GGTGAGGAGT TCAAGGCCAC CCTGTGGCAA  
 5581 ACACCGTGAA ACCCGTCTC TACTAAAAAG CCGGGCGGTG GTGGTGGCGA  
 5641 CCTGTAATCC CAGCTACTCG GGAGGATGAG GCAGGAGAAT TGCTTGAACC TGGGAGGCGG  
 5701 AGGTTGCAGT GAGCCAAAAT TGCACCATG CATTCAGAC TGGGTGAGAC TCCATCTCAA  
 5761 AAAAAAAGAA AAAAAAGGAG TTTCTATTGT TTAATCTATC CTGTGTGTTT ACACCAGGAA  
 5821 AAGTAAGTGG GAGAATCAAG CATTTATTAC TTCCAAAGTC TTACAAGATA GGCCAAGATA  
 5881 TATAAATATT TCCTAGCAAA ACAGTATGAC CTGCACTATC ACTGCTTCTA GATACTAACA  
 5941 AATGACAGGT TACTCTGAAG CCGTACTGTA CAGTTTATAA TGCAATTTAC ACATTGAGCT  
 6001 TATTTTCATC TCACTGCCAT GGGAGCAGGT ATTTTATTC CTAGTTTATA GAGGAGAAAA  
 6061 CTGCTTATGC CAAGTTAAAT AGCTTACAGG AGACCACATT AATACTGGGT AATGACTACT  
 6121 TTTAACAATT TGATGTATTT CCATCCCATC ATTTTTCAGT AAATGTAATT TTCCTTTATA  
 6181 AAATTGAGAT AAAAGTTAAA CAATTTTGCA TCTGATTTT TCACCATTAG TCTTCTCTCA  
 6241 TACCATTAAA AATTCTTTGA AAACATGATT TTAATGGCT GCATACATT CTAAAGTTGA  
 6301 TATGTATCAA TTTATTTTGT CATTTCTGT TAGGCATTTA AATAGCTTTC TATTTTGTG  
 6361 TGTATATAAT AATGCTGAGA TAAACATCCC TGAACAAACA TCTCATCTAC TTAATAAAGT  
 6421 CTCCTTTGAA TTAAGTAAAG ACTAAGTTAG GCTGATCTC ACTTACATTT TTTTTTTTTT  
 6481 TTTTTTTTGA GATGGAGTCT TGCTTTGTCA CCCAGGCTGG AGTGACAGTG CATGATCTCA  
 6541 GCTCACTGCA ACCTCTGCCT CTTGGGTTCA AGTGATCTC CTGCTCAGC CTCCCAAGTA  
 6601 GCTGGGATTA CAGGCGTCCG CCACACATG GGCTAATTTT TGTATTTTGA GTAGAGACGG  
 6661 GGTTTTACCA TGTGTGTCAG GCTGGTCTCA AACTCCTGAC CTCAGGTGAT CTGCCCGCCT  
 6721 CAGCCTTCCA AAATGCTGGG ATTAGAGGTG TGAGCCACCG CGCCAGCCTC ACTTACATTT  
 6781 TAAATGTGAA AACTAAGTTT TATTTTCATT TTTCTTTAG TTAATCATT GACATGTTAT  
 6841 CACCAGGACT GTTTGTTTCA ATGATGTAC CTTTGAAACT GCTGTGCCTA CTAAGGGAAT



6901 GGACATTAGG TCATTGAACC AATAAACTGG AATTTGCTGG GTAAAACTTT TCGTATCTCA  
 6961 GGCTATTTTT CCAGCTATGG ATGGGATTTA TACTCCTCAG GGCTGAGATG TGCTGAAAG  
 7021 CACACTGCTG TCTACCCACT GCCATTACCT GTAAGTTGCT CCACTCCACC TTCCATGCTG  
 7081 TTTTGGGAGG AGGGTGAAAG AAAGGGAGGG GAATTTCACT AGAATCAAA CAGTTTCTTA  
 7141 CAGTTCTATT TGTTCTTCCA GTCTTTCACT CATGCAATGA CTATTATTG AGTGTTTTTG  
 7201 ATGTTGTAAC TTGTAATTTA GGAGTATGAC AACAGACTTT AGGTTGTAAA CTCTTTGATC  
 7261 TGGGAGTACT GTAAATCGTC ATGAGGAGGT TATGGAACAC AGCATATCCC AAAGTTACTT  
 7321 GACTATGGAA CCCTTTTATT TTCTTTGAAC ACCTACGAAT GATCTCCAC AACTGCATTC  
 7381 CACAGAACAT AGGTTAAAGG AAAAGGCTAA AAGGCATCTA ACAAACCTCA TATCCATTTT  
 7441 GTTTTTTTCT CCAAAATAGAA TCATTTAAAT GAAATCCTTT TAAAACTAGT AATAGATAAT  
 7501 TTCTTAATGC AATTAATTTA TAGATAAGAA GCCACCAGCT GACATAATCT CTTAGGGTGA  
 7561 AATAACATAG GCAATACCAT TAAGATTAGA AGCTTGATAT GAATGCCTAC TTCCTGGTAT  
 7621 TATGAAATAT GGAACAGAAG GGGAGACAAA GTAATCATGA AGGTAAGTGA TACAGTTATC  
 7681 TATCTAGAAA ACCTAATCTA CTGACAAGCT GTCAGAATAA CTTCTTGAGT CAAATGACCA  
 7741 GATATAAAAC AAAAATGTAC AACTCAGTAG CTTCTTACA TATCAGCAAT ACCCTGCCAA  
 7801 AAAAAATAG AGAAAATAAA AGCCCATTTA TAAGAGCAAC AGGGCATAAA ATACCTAATA  
 7861 CATTATAAAG AGCTCCATAG GACCTGTATG AGTAAATCAA TCACTTTTCC AAGAAGGTCT  
 7921 GCGTTAAAGC AATGGTGAGC TGCTGCTTGG GAAGAGCAAA TACAAAAATG TAACCTAGAG  
 7981 ACAGAATGCA GTTCCAGGAC AAATGGATCC CAGTGGGATT TTTTTTGTG TGTTTGAAAG  
 8041 AACCTAAGTA AGAAAGTTCT TCTCAAAGAA TGACAGAAAA TGAGAATTGA TAGCATTTGA  
 8101 AAAATAAAGA GTAATGCACG CAGAAGAGTC TTGAGCTAAA ATTATTAAAA AGTATTTCAA  
 8161 AGTGACAATA TTAAAGATGG TATGGTGTGG GCATAAGAAT TTACAAACAG AATAACACGA  
 8221 GACAAAGGTC CTAACAGACT CAATATAGTA TTTAATAACT TAGCGTATGA GGAAGCCTCA  
 8281 TGTGACAGCA GAAAAAGGTA GATTTAATAA TAGTGTTAGA AAATGACAAA AGACAAAAAA  
 8341 AAAAAAAGG CCACAACAAT ATAATTGAGA CAGATTTTAC TTCCAGGTGG GTTAAAAAGT  
 8401 TAAATATTTA AAAAGTAATC CTATTTTAAA AAGAGCGCCC GGGCAGCCCG CCCGTCCGGG  
 8461 AGGGAGGCGG GGGGACGCCC CAACCCGCCC GCGGCCCGCT CTGGGAGGTG GGGGGCGCCT  
 8521 CTGCCCAGCC ACCCTGTCTG GGAAGTGAGG AGCCCTCTG CCCGCGCCCG ACCCTGTCTG  
 8581 GGAGGTGTAC CCAACAGCTC ATTGAGAAGC GGCCATGATG AGGATGGCGG TTTGTGCGAA  
 8641 TAGAAAAGGG GGAATGTGG GGAAGAGAAA GAGAGATCAG ATTGTTACTG TGTCTGTGTA  
 8701 GAAAGAGTA GACATAGGAG ACTGCATTTT GTTCTGTACT AAGAAAAATT CTTCTGCCTT  
 8761 GGGATGCTGT TAATCTATGA CTTTACCCCC AACCCCGTGC TCTCTGAAC ATGTGCTGTG  
 8821 TCCATCTAGG GTTAAATGGA TTAAGGGCAG TGCAAGATGT GCTTTGTTAA ACAGATGCCCT  
 8881 GAAGGCAGCA TACTCCTTAA GAGTCATCAC CACTCCCTAA TCTCAAGTAC CCAGGGACAC  
 8941 AAACACTGCG GAAGGCGGCA GGGTCTCTG CCTAGGAAAA CCAGAGACCC TTGTTACAT  
 9001 GTTTATCTGT TGACCTTCCC TCCACTATTG TCCTATGACC CTGCCAAATC CCCCCCTCGG  
 9061 AAGAAACCCC AAGAAATGATC AATAAATACT AAAAAAAGG AAAAAAAGG GGCAGGGCAC  
 9121 GGTGGCTCAC GCCTGTAATC CCAGCCCTTT GGGAGGCCGA GACAGGCAGA TCACTTGAGG  
 9181 TCAGGTGTTT GAGACCAGCC TGGTTAATAT GGCAAAACCT TGTCTCTACT AAAAAATACA  
 9241 AAAATAGGCC AGGCATGGTG GCCCATGCCT GTAATCCCAG CTACTTGGGA GGTGAGGCA  
 9301 CAGAAACCCG TTGAACCTGGA AGGTGGAGGT TGCAGTGAGC CGAGAACGTG CTACTGCCCT  
 9361 CCAGTGTGGG TGGACAGAGT GAGAGTCTGT TTCAAAAAAA AAAAAAAGG GAATATGAGT  
 9421 GAACATTTCG TGAGTCTTAA GATAAACAAG GATTTTCTAA CCAAAAAATA AAAGATATTA  
 9481 TAAAAAAGGA AAAATTGTGA GATTTGATCA CAAATATTTA ATATATCTGC ATGGAAAAAC  
 9541 TATTACAAAA ATTACAAGGC AAATTTCAAA CTGGAAAAAC ACTTTTACAA CTGTGACAAA  
 9601 GGATTGATAT CTGTTACTGA GAAAAACATT TACATTTATT TATAAGAAC AAGACAGGGC  
 9661 CAGGCTCGGT GGCTCATGCC TATAATCCCA GCATTTTGGG AGGCCGAGGC AGGTGGATCA  
 9721 TGAAGTCAGG AGATCGAGAC CATCTGGCT AATACAGTGA AACCCCGTCT CTCTAAAAA  
 9781 TGCAAAAAAT TAGCCGGGAG TGGTGGTGGG CGCCTGTAGT CCCAGCTACT CCAGAGGCTG  
 9841 AGCAGGAGGA ATGGCGTGAA CTTGGGAGGC AGAGCTTGCA GTGAGCTGAG ATCGCGCCAC  
 9901 TGCACTCCAG CCTGGGTGAC AGAGTGAGAC TCCATCTCAA AAAAAAAGG AAAAAAAGG  
 9961 AAGGACAAGA CCAACACTTC AAATGAAAA TGGACAAAGT TCACAGACAA TTCAGACAA  
 10021 TCACAAAGTA ATATGGAAGA CAACTATATT TTATAAGTTC AACGTCATA GTAAGAAATA  
 10081 AACAAATTA AATAAGTCTG TTCTGGCTGA GTGCCCCATG GCTCACGCCT ATAATCCAG  
 10141 CACTTTGGAC TTTGGGAGGC CAAGGCGTGA GGATCGCTTG AGCCCAGGAG TTTGAGACCA  
 10201 GCCTAGACAG CACAGCGAGA TCCCAACTCT ACAAATAATG AATAAATTGG CTGGGTGTGG  
 10261 TGGTGCAATG TACCTGTAGT CCCAGCTACT CCAGAGGCTG AGGTGGGAGG ATTGCTTGAG  
 10321 CCCAGGAGTT TGAGGCTGCA GTGAGCTATG ATGGCATCAC TGTAATCCAG CCTGGGCAAC  
 10381 AGAGCAAGAC CCTATCTAAA ACACCTTTTT TTCTATCAA ATTAATTTTA ATTAGTTTTA  
 10441 CTGGCAGTTT GGTGAATTAC TGATCCTCAT ATTTCTGGTG GGAGTCTAAT TTGGAATAAC  
 10501 TGTCTAAAAA AAGAAAGATA GATACATAGA TTTAGGAATA TATATAAGCC TGAAAAAAGG  
 10561 AATTTATATA TATATTTTT GAGACAAGGC CTAATCAGT TGCCAGACT GGAGTTCACT  
 10621 GGTGCAATCT CAGCTTACTG CAGCCTCGAC CACCTGGGCT CAAGTGATCC TCCCACCTCA  
 10681 GCCTCCAGG TAACGGGAC TACAGGCATG CACCACCATG TCCGGCTAAT TTTGTATTT  
 10741 TTTTTTAGAG ATGGGGCTCA CCAATGTTGCC CAGGCTGATC TTGAATCCT GGAGTCAAGT  
 10801 GATTTGTCCA CTTGGCCTC CCAAGTGCT GGGATTACAG GTGTGAGCTG CTGCGCCAG  
 10861 CCTATTATCT ATATTTTTTG AGGCACTAAT TTATGCTTAG GCTATTCTAA GAAAAAAGT  
 10921 AGAAATACAA AGATTTATGG CCAAAATGT TCAGTTCATT GCAACATTAT GTATAACAGC  
 10981 ATAAAAACAG AAACAACCTA AATGTCCAAC CTGTACTATA GATTATGATG TAGTCATTGA  
 11041 AGTTATTTTT GAAGACTATA TAATGACATA AGAAAATGCT CATCGTATCA TATTAAGCAA  
 11101 AAAAGAACC CATTATTAA ATTATATTA CACAATCCTC AGTTTTGTAA GCATGAATTC

```

11161 TATGTCTTCA TCTAAGTCAC TGAGAGAAAT GTGAACAGAT GGGGTTTGGA ACGGATGGAA
11221 GGTTCAGCAA AGTCAGCCTC CCTTTTCTGG GTTGAGGCTA GCCCTCTTCC ATTCCTCCAT
11281 TAACCTGAACC TTCATCTGGC AATGGTTCCC AAATGTTTCC TAGGGGAGGG CACATAAGAA
11341 TTAAGTGTGAG AGATGGTTTA GAAAAACAAA AACAGATTTT GGAGCCTGAC CCACAGAGAT
11401 TTCGACTCAG GAGGTCTGAG GTGGGGCCCT GGAATCCGTG TCAAAAAGGCC TCCAAGTGAT
11461 TCTGATGCAC AGCCACGTGC AGAAGGAACC ACGGCCTTCT GGTTCGAGGT TTTGAAAAGG
11521 ACATTGAACA CCTTTTGTAA GTCAGCTTCT GAATGTGAGT TAGGGAAGAT AATGTGAGGA
11581 TGGAAAAAGC CTTCCCCCTC CTACATTTCT CCAAATACTC TGTGTCTGCT TGTGAGGTGG
11641 ATGATGTGAC AACTGGCCCA GCTGCTGGCC TGCAGCGCTT CCCTGCTCGG ACTGCAGCTG
11701 CCAGCGCTCC CCTGCCACC CCCCGCCCA CCTACTACC TCCTGCCTAC CCCCACCACC
11761 CCACCCAGCC CCGCCTACCC TGCAGCTCCT TGCCCTGCTC TTCAGGACTT TTGCTCCAAG
11821 TCTTGCTTCC TCTTTGTCTT CTGAGATGCC AGCTGCTCCC TCTCCATGGC TTCCCTTCTG
11881 TCCACTTTTC ATTCTGCTGA GGTGTGTGCT GTCTTTACCC AAACCCCTGT GATGCTTGTG
11941 GCCAGGCCTC CTTCCCCCAG GCCTAGCTGC TGAGAATTG ACGCCTCACA TCCCCTCAT
12001 TCCTCAACCC GCGCATCTG GCTTCTTCA CCACCACTCA CAGGCACACC TCTTGGCAAG
12061 GTCACCTGGC TCCCAAAGGC TAATCTACTG GTTCTGTTTC AGTCTTTTTC CTGGAAGCAT
12121 CTGACACTAC TGACCACCAC CTCTCTCCCT GGCTTTTGTG GGGGTCATCG TTGCTTTACT
12181 TGCCCTCGTAT CCCCCTTTCT TGGTAGCAGC ATCTCAGGA GCTGCCTTCC CCAGCTCTCA
12241 GGCACGCTGG TTGGTGGGGC TGACCTCCCC ACTCCCCACA TTTCATGCCC CTGCCCCATA
12301 GGATTGGTGC AGGGATGGGC ATATGGTCAG GGTCAAAACA ATGAGACTCA AGCCTGAGAC
12361 TTTTACTGGG GCTGCTGAGC CAGCAGAATT CGAGGCTGGA AATGGAGGTG GCCGCTGCC
12421 CACCCCTGAG AAGGAACACA ACATGAGGAG GGCACAGAG CCAACAGAGT CCTGGTGTGT
12481 CTCCTGCAGT CCTGGGCCAA GCCTGAAGCT TGCACATCCT GGTCTGATGG GTTTCAGGAA
12541 AACGTTCTTA TCTTTTATC TAAGCTAGCT TTGGTTGGAT TTCTGTACT TGCAACCTAA
12601 AGAGATCTGA GACAGCACTG AGCCACTACC CTCCCAGGCA ATTACCCACC TCTTTGGTAC
12661 TTTCCCCGGA GGGTCTCCAC AAGCTCTGCT TCCTCTACT TAAGAGTCTT GTGCAAGGAC
12721 TTCTACTTCC CAGATGCTGG CTACTAAATG CTGACACTCC CAAGTCAACA GACCAGCCCA
12781 GATATCTCCC TGAGTCCAG ACCATGTCT TACCCCTCAT GGACATCGCC ACCTTGGTGT
12841 CCACTCAAAC ATGACAGGTC TAAATCTAAA CTTATCATTA TGCCCCATCC TTCCTCTTCT
12901 TGGATCCCCA TCTCAGTGAA CAGCTGTACT CCCCAGTTAC CCCAGAAACC GAGGCCTCAT
12961 GGGTGACACC TCCCACCCT CACTCCCTGG CCTCTTTGG GCAGCTAATT CTCTGCTCAT
13021 TCCTCAGGAG TCAGCTTTTA GGCTGCATTC TCAGGGAGGC AGCCTTTTCC AACCCGCAAG
13081 CCTGAATGAG GGGTCCCTC CTATGCCCCC AGCCACCCTG TGTCTTAGT GTTACAGCCC
13141 TTGTCACTCA GAGCTGTGGC CTGTTGACTC TCTTCTGTGA TGGCTGGCCC CAACCTGAA
13201 ATCAAGATCA TTTGAATAAA TGAACACTAG CATCTGTAGG TGTGATCTTA TATTGTATAT
13261 TGAAAATATA GTCTAATTTT CTGTAAAAAT TAAATATCTC AATATTTTAG AGAGATCTGT
13321 AAAGGGTATA ATATTGGAA CAGAAGATTG ATTTTCAGGA GCCCAGTAT CTGGATGAGT
13381 GTGAGAACCA CTAAATAGGA CCCATAGTTT TTCTGTGTGA GATTTTATTT TTTTAAAAAT
13441 AGATGAACCA GGCTGCCTAT TAGTGAATGC CTTCACCTCC CTTTTCTTA ATTACTGAGT
13501 CTAGAAACTT ATCAAGGGAA TGAAGCAGGC CTTTGTCTCT ATTCAGAAATC CATTTCCCTG
13561 CTTTTTTGAA AATGAGAACA TCTGCTGCT TCCAGTGTTC CAATACTTTT CATATCTTTC
13621 ACAATTTCTT AAAAGCCACC AAATACAGTT CTTTGAGCAG AAATGCAAGT TCTTTCATCC
13681 ATGTGTCCGG CCAGAGATT TCTGAACCTC GCCTGTGGCA GGCCCTGGGA GAGCTCCTCA
13741 GCCTGTGGTT TTAATCCCTT CTTGAACACA GCTGGGCTCC CTTCTGCCAT CTCTGGGCA
13801 TTTCTTTTGA GGAAGGAG GAAGCTGAGT AGTTCTGCTT TCTGTTAATA TTGACGCAAG
13861 AGCCCATGAG TGGGCCGATC ACTTCCTGAT TCTTGTCTT GTTCTAATCA TACCTAAAAA
13921 GGCTTTTTTA ATGTCCAAGG CACTTTTACC TTCTCAGCCC ACTCCAGCTT CTGTATCCTT
13981 CATACCGCCC ATAAACCTTG TCTGCACCTC TGCACATGCA CCCCTGCCCC TTTTTTCCCC
14041 TATTAATATC TATAGTTAAA TGTTAATATC TAAAGTTTTT AAACACAATT ACGGTATCAA
14101 GCATTTTATC AAAAATTTT ATCATACTTT GATTTTACCA ACTTGAGTGT TATGAACAAT
14161 GAACATGTTT ATACATTGAC AAGAATACTC ATACAGGAT TGTGGGATGG CTACTTTCAC
14221 TCCCTTTTCC CTCATCTTTA CTCTTTTTTT TTTTTTTTTT GTAGAGATGG GGTCTCACTA
14281 TGTGCCCCAG GGTGGTCTTG AACTCCTGAG CTCAAACGAT CCACCTACCT CAGTCTCCCA
14341 AAGTGTTAGG ATTACAGGTG TGAGCCACTG CACCTGGCCC CTCATCTTTA CTCTTAGTT
14401 CCTTTGCCCT TGGGTTCTTC TAGCTTATTT CTAATTTGTT TTCTTTCTGG CATCCATTTT
14461 TATTTTAAAT TATTACACAT AGATGCAGAA ATTACTTGGC TGGCGCTAGT TGCAGAAATCA
14521 TGAAAATTA ATCAGATGAT GCTTAGTGTT GAAAGACTGA ATTGTTTTGG TTTTTTATTG
14581 AGATATAAAT CATATACCAG AAAATTTATT CTTTACAGG GTGCACTCCT TTAGTATATT
14641 CACAAAGTTG TGCAACCAAT ATCATTATCC AATTTTCAGAA CATTTTCATC ACATCAAAGA
14701 GAAACTCCAC GCCCACTAAC AGTCATTCTC TATTTCTGTT GCATGCGGCT ATCCAGTTGT
14761 CCTAGCACCA TTTGTTGAAC TATAGTTTCC CCCATTGTTT TAATAGTTTT TAAGTGGTTT
14821 CTTAGGATTC TCTACATAGA AGATCATGTC ATCTGCAAAAG AGTTTTACTT TTTTCTCTCC
14881 AGTCTGGGAG TTTATTTCTT TTCTTGCTT ATTTCCCTGC TCTCTACCCT TTTAAAGCCT
14941 ATATAGTTTC TTGAGTGACT TCCTCTTTT TTCCTCATT GAATAAGTTT TCTGGGGAGA
15001 GGGAACAGAG CATATGAAGG GAGAAAGTGG AGCTCTTAAA AATGTGTAGT TTGGGACGGG
15061 CCGCGGTGGC TCACGCTGT AATCCAGCA CTTTGGGAGG CCAAGGCGGG CAAATCAGCT
15121 GAGGTCGGGA GTTTGAGACC AGCCTGACCA ACATGGAAAA ACCCATCTC TACTAAAAAT
15181 ACAAATTAG CCAGGCGTGG TGGCAGATGC CTGTAATCCC AGCTACTCGG GAGGCTGAGA
15241 CAGAAGAAAT GCTTGAACCC GGGAGGGGGA GGTGCGGGT AGCCGAGATC ATGCCATGGC
15301 ACTCCAGCCT GGGCAACAAC AGCGAACTG TCTCAAAAAA AAAAAAATTG TGTAGTTTGC
15361 ATTCATTCA AATGGACTTC ATACTTTTTT AGTGGAAATG TCTTCTTCA CAGCACCTGG

```

15421 GGCCAGGGAT TTTCTGGAGG AATCTTTGAA ATCTGTTTTC CTCAAGTTTT GGACACAAGC  
 15481 CTGGGTGAGC TCCATATCCA TCCTTCCTCG GCTAGTACAG AGTCTAAGAT GACATGGTCA  
 15541 CTTTCCCCCA TGCCCATCTC CCTCCTTCT CTGGCCAATT CTTTATTGCC AGTCAAAGGC  
 15601 AGAAGAGCCA ACAGCAATAA GAGAAGATTC CCACCATGTC AGCACAGGTT CTGTGTCATC  
 15661 AGTGACCTCT TGGAAATGTGC AGGAAGCCAG GGAGATAAGA ATCTATGGCT ATGTAGATTA  
 15721 GGTACACCCC ACAGCCAAGT ATGAGCTTCT AGGGCCCAAT CGCTAGAAGC AAACAACCCA  
 15781 GACCACTAAC CAGGCTCTAA CTAAGGCAGG ATCCAGGATG CTCAGCACCC TGGGCCAACC  
 15841 CTTCCGCTCA AGATGAATCC ATGGTTCATT TACAGTGTCT GGGGTGATAT TTAATTCCAA  
 15901 GTTCACAAGT CTCCAGGGTA GTCGCCCTTT TGTTTGGGGA TTTCTCCCG GGTGCTAAAT  
 15961 CACTTCCTCC TAGCTCCAAT CTGAAAAGCA ATTCAGGTCT CCTGTGGCCT CACCAGCTT  
 16021 CCTGCCCACA CTGATTTTCT TCTAACCTCA CTGGTCTACC TTGTGTACAT CACCTAACAA  
 16081 CAAAGCCACT TAGCTCAAGG AAAGCCTCAA GAACAAATGG CTTCCTAATA AGGTGATTCC  
 16141 TCCCTCACCC TAGGGTGGCC CACAGCTCCA GCACCTGCTG AGAGGTGGCC AATTTGGATG  
 16201 AGGTAGCTGC CAGGAATCAA ATCTCAGAAC CCTTCATAAA ACTGCGCTTA GCTATTCCAC  
 16261 AAATTTCTCC AAATAATGAA ACAGAGAGAA GAGTGAGAA ATAATGAGAG AGAAATCTGG  
 16321 AAAAATACAA AAGATAAATT CATGTTCTTG TGAGTTGTAT TCATCTTCCT GGGAGTTGAT  
 16381 AATCAGGCTG GCTAGCTATC CAGTCTATGG GAGCAGAGAG AGGGCTGCAT TATCTCCAGG  
 16441 TTGGCACAGC ATCTAGCAGG GGCTCCCTGA TGACACCTGG GCACAACAAG GACCTCACAG  
 16501 CAAAGACAAA CAGGAGGAAA TGCTAGGGT GTGGGGACAT GCCACCACTA AATTACAAAA  
 16561 ATATATACCT TTCTTCTTGA TAGTTCCTCA CAGTACCGCT TCCCACTCTT GTACCACTTG  
 16621 GCCATAATAC TTTATTATCC CTAGTACCTA AAAACTAGAC AGTATTACTG AGACCTTAAA  
 16681 TGTGACATTT GTCATCTTTC AAGGTGCCAG TTAGTGTAAAT ACAGCTATGG TTCATTTTGA  
 16741 GTGTCTGGGT GACGGTGACA TATTTAATTC TATTTTGTG TTGTCCATAA GCACCAATA  
 16801 AAGCAATTTT GAAATTACTT TCTTCACACT CAATTTATTA TCTAGAACGA CATTTCAATA  
 16861 CACAGTATAT TTGTTTTGTG GAATTCATTC GCAGTGTATC TGGAGGCCTG GATTTAACAT  
 16921 ATTTCTGTGA GTGATAGCTA CTCTATTGTG AGGAACATTT TAGGTCACTG GAGTGGAGAA  
 16981 AAAGAGGTCG CTACTTAACA GAAATGGCAT GTATCCAAAT ATAGATGCAG TTTTCTCAAC  
 17041 ACCTGTTGCA CATCAAATGA TGAAGAGATT GTTACAATTT CCAAAACCCCT GGACCAAAAC  
 17101 CTTAAATGAT CAGCACTTGG TTTCAACTGA TACACGGCTT AAGTCAGAAA AATAAGTTCT  
 17161 TCTTACCAA AACAGTGAGC CTAAATATGC TGGCTGCTTT AGAAATAAAA CAGGTGATCA  
 17221 ACAAAGCAAT CTGTTTCTGT CTAAATAATT TCCCAGCAAA AGTCACTTCT TTTGAAACTA  
 17281 TATGATTGA ACCACAGGCT ATGCACAGTG ATTGTGTGCT TTCAGAAACA CGGTAAGCAT  
 17341 GACAAATATT TTCAGTTGTC AAGGTTAAGC GAGGAGGGAA TGACCCACCT AAGAAATGGA  
 17401 AACTAACAAA TAGTAATATG ACACTCCCCC TTAGTTCCAC AGGCTGGTGC GGCTGCTACA  
 17461 CGCAGCTCAC TTCTACGGG CGGCTAAAGG CGGTGCATGA ACTGAACGTA GGCGGCCTCT  
 17521 TTATCCCCCT CCCACAACAC TGTGGGGTGG CACGTAAAGA CCTCAAGGGA AGGGGGCCTG  
 17581 GCGGATACGC TCTCTGCTG CCATTCAATT GTTCAAGAAA TAGTTGAACA AATGAATAGT  
 17641 TAAACAAACT CTCCAACCTAG GACTCTGTCC CATTCCGATA CCCTACGCCT CTCCCCCAT  
 17701 TTCCATCCCT AGGCATCACT CCCTCGTCAA GAATATGCTG ATTATGGCTC GCGCCCCCTC  
 17761 TCTCCCTTGG ACCCTCCCGC CAGGGCGACC GGACGAGGCC AGCGGGAGCA ACGCCTTCCG  
 17821 AACCGCAAGG AAGGAGGGCT CGGAGTCCGC AGGGCGCGGG TTCGCGGGAC CCTCCAGCGC  
 17881 CGGGGGCTGG CCAGCCGGGG CGGGGCGAG GGCACCGGGC GCCGAGCTCC GGGAGGGGTC  
 17941 CCCGAGGTCG CGCGGGCCGG GAGGGGCAGC GCGGGGAGGC GGCCGACTGC GGCGCACACG  
 18001 TGGGGAGGGC GGAGCGGGGC GGGGGCCGGG CGAGGGGCGG TGACGGTACC TGATGAAGGT  
 18061 GGTCTTGCCG GTGCTGTACT GGCCACACAG CAGGATCATG GGCTTGTCTC CGAAGTCGGC  
 18121 TCTCTCCAGC GCAGGCGAGT GGAACCTCGT GAAGCGGTAC GCCTCCTCCA GCGGCAGCAC  
 18181 CTTGCGCAGG TAGAGCGAGC GCAGCCCGCC CGTCACCGTC TGCACCGCT CCGCGCCGCC  
 18241 AGCGCGTTCG CGCCCGCCCG CCTGCGGCC CATCCAGCTG AACATCCTGC CGCCAGTCCA  
 18301 CGCTCGGATG GGACCTTGCT CCGGGTTCGA CTCTCCCCG CTCGCACTGA GCCGCCCGGG  
 18361 CGCGCTGGC CGCCGCCCCG GAGGGGCCCC GCCTTGTGCT GCCACAGCCA GTCTTCGGCC  
 18421 GCTACGTCAC GCGCGCTCGC CCACTATTCA GACGAGCTCT CTCAGGTCTT TCCTGATTGG  
 18481 ATCTCTGCCA GAGAGGGCTG AAGGCGTCAG CCAATTGAGA GGAGGTCTT AATGAATTAT  
 18541 TGATGACGGA GTCATTCTAG AGGACTGGGC GGGGAGATGG TGTCTTCCCG GCCGGGTGGC  
 18601 ACTGACCGGT CATGCTCAGG GCCCCTTCT GGTGCGCGGT GCCTCACCTC TGGCCAGGTG  
 18661 GCATCGGCTT CCGGGTAAAT CTGAGCTGCC CAAAGCGAGC CTCCGGAGGC TGCTCCATCT  
 18721 CCTGGAACGT TTAATCATTC AGGCCTCTGG ACGTCGATAC TGGGGATGAA AAGTTCAGAA  
 18781 AGTGACAGG CAGGTTCTGT CTTCCCGGG GCGGGTGGTT ATGACTGCAT TTTACAGGCC  
 18841 AAGAGACCCA GGCTCACGGG CATAATTGAC ACAATCTGA GGGGAAGTGG ACTGGAATCT  
 18901 ATTTGAGTAG CCGTTTCTCC TACTTCTCCA CATTGCCCTC ACTAAATTCT GGATCAGCTA  
 18961 AAGGCTTTTT CCCCTTAGTT TTCTTTATGT GTTTTCTGT AAAGCCGAAA ATTTAGTCAT  
 19021 TACTCTCATT ATTACATTAT CCTCATTTAT ACAACTTTAT TTCCACATAT TGACATGTGA  
 19081 GTCATTGTCA CCTCCTAGCT GTTTTTGTG GTTGTGTGT TTGTGTGTG TTTCTGAGAC  
 19141 GGAGTCTGCG TCTGTGCCCC AGGCTGGAGT GCAGTGGCAC GATCTCGACT CACTGCAACC  
 19201 TCCGCTTCCC GGGTTCAAGC AATTTTCTTG TCTCAGCCTC CTGAGTAGCT GGGACTACAG  
 19261 GCGCTCACCA CCATGCCAG CTAATTTTGT TATTTTGTAG AGAGATGGGG TTTCAACATA  
 19321 TTGGTAAGGC TGGTCTTGA CTCCTGACCT CAGATGATCC GCCTGCCTTG GCCTCCCAAA  
 19381 GTATTGGGAT TATAGGCGT AGCCACCGCC CCTGCGCTGT TTCTTTTCTA GGTAAAGAGG  
 19441 ATACTCACTT ATCTAAGTTG AGATTGGGG AAATAATCAC CTTTTGTGTC CTTGATTGTA  
 19501 GCTTAGGAAT ATTTCTTAGT TGTTTAGAG CATTCTCAA GTGTGATCCA TGGGGCCACC  
 19561 TGCATATATG CATTATGAG CTGGGATGCT TGTTAAATG CAGATTCTG GTCTTCAGTC  
 19621 CAGACCTACT GAGTTAAAT ATTGAGGAGA GATGGAGCTC AACAGATGCC CTAGGTGACT



19681 CTGATGTGGC CTAAGTCTGG GGAAACAGTA CACCCTGCAG TACTTCTCCA ACCTAACAGA  
 19741 AATGCGGATT CTGATTCGGT AGGTATGGGA TGGGGCCCAT GACTCTACAT TTCTAACAGG  
 19801 CTCTCGGGTA TTGCCAATGC TACCCATCCC TGGACCATAC TTGGAGTATG AGAGAACAGG  
 19861 ACTTTGCATC AAAAGATCTG GGTTTGAACC TTGACTATGT AAATTAATGG CTGTGTGATT  
 19921 TTAAATGTGA TGAGTTACCT CTCTGCCCTA GTTTCTTAA GAGCAAAATG GAAACTCCCT  
 19981 TGGAGTTGTT GTGTGGGTCA CATAGTTGGA GGGAATTGTA TGTAACTAGG GAAGTAAAA  
 20041 GAAAGCCAAG GGAAAATAAA GTTAGACATC TGCCTTGTGT CAAAATAAAA CAAAGATTGA  
 20101 AATGTAAAAA GAAAATGAAA CCACTGCACT ACTAGAAGAA GGTACAAGAG ATTTATACAA  
 20161 TTTGGCGGTG GAAAAGCCTT TCTAACCAAC ATTAAGGACA GAAATATGA AAGCAAAGGC  
 20221 TGACATATTT GACAACAAAA GTAAAAGCTT CGATGTATTT TTTTAGAATA CATAAAAACA  
 20281 AACGGATATA AATACAATCT TACCATATGA TGCAGAAATC AGCTCCTTAG TATTACCCA  
 20341 AATGAATTGA AAACCTTATAT CCACAGAAAA GCCTGTACAT GGATGTTTAT AGTGGCTTTA  
 20401 TTCATAAGTG CCAAAAATTG GAAGCAACTA ACATGTCTTT CAGTAAGTGA AAGGATAATC  
 20461 TGTGGTACAT CCAGACAATG GAATCTTATT CAGCACTAAA AAGAAATGAG CTATCAAGCC  
 20521 ATAAAAAGAC ACAGAGGAAA CTTAAATGCC TATTACTAAG TGAAGAAGC CAATCCGAAA  
 20581 TGACTATGTA CTGTACGATT TCATCTTTAT GACATTCTGG AGAAGGCAAA ATTATGGAGG  
 20641 AAGTAAAAAG GTCAGTAGTT TCTAGGTGTT AGGGAGAAGG AAGGAATGAA CAGGCAGAGA  
 20701 ACAGAGGATT TTTATGGCAG TGGCTCTACT TTTTATGGTA CTATAATGAT GGATACATGT  
 20761 CATTATATAT TTGTCAAAAC CCACAGAATG TACAACACCA AGAGTGAATG TAAGTGATAA  
 20821 TACTCTGTTA ATTTTCAGAT AGTTTTAGGT TCACAGCAAA ATGGAGCATA AAGTACCGAG  
 20881 AGTTCCCATTA TACCCCCACC CTCCCGGACA CACAGCCTCC CCCACTATCA ACATCCTGAA  
 20941 CCAATGGAAC ATTCGTTACA ATCAATAAAC CTAACCTAAA TAGATACATC ACTAAATTA  
 21001 CTTAAATGAG AACTAGATAT AAAACTAATA TAAATGTAGA ACTAAATGTA AAACCTAAGA  
 21061 TTATAAAACT TCTAGTAGAA AACACGGGAG AAAAATGTTG TGAACCTGGG TTAGGGAAAG  
 21121 ATTTCTTAGG TGACACCAAA AGCACAATCC TTAAGGAAA AAAAAGTTGG TAAATTGGAT  
 21181 TTCATTAAAC TTGAGAACTT CTGACCTTGG AAAGTACTGT GAAGAGAATG AAAAGACAG  
 21241 ACAAGGCACA GAATATACTT GCAGATCACA TATCTGATAA AGGACTTGTA TCCAGAATAT  
 21301 TTAAAAACTT CTTAAAACTT AATAGTAAAA ACCCAATTT TAAATTTCC ATTTAAAAAA  
 21361 TGGAAAAAAG ATCTAAACAG ATATTTTACA AAAGAAGATG TATGGATGGC AAATAAGCAC  
 21421 ATGAAGACAT GTTCAACATT GTTGTGTTGT AGGAGAATGC GAATTA AAC CACAGCAAGA  
 21481 TACTACCACA TACCTATTAG AATGTCTAAA ATTAGGAAAG ACTGATCATG CCAATCATTTG  
 21541 ATGAGGATGT GGAGCAACTG GAACTCTTAC ACATTGTTGG TCAGAATGTA AAATGGCACA  
 21601 ACCACTTTGG AAAACAGTGT GGCAGTTTCT TAAAAAGTGG AACGTACACT TACCACATGA  
 21661 CCCAGCAGGG GAGGAGCAGA GGGAAGGATT GCAAAAGGCA CTGGAAAGCT TTTGGGGCTG  
 21721 AAGGGTTTAT TCATTACCTT GATTGTGGTT ATGGTTTCAC CATGTGTATC ATATGTAAAA  
 21781 ACTCATGAAA TTATGTACGT TGGATCCTCC ATATCTCTGT GTTCTGTATC CATGGATTCA  
 21841 ATCAACCCACA GATGGAAAAT ATTTGGGGAA AAAAATGGAT GGTGTCATCT GTACTGAACA  
 21901 CATACAGACC TTTTTCCTT TTCATTATTC CCTAAGCAAT ACAGTATAAC AACTATTTTC  
 21961 ATATCATCTA CATTAAATAAT AAGTAATCTA GGTAGGATTT AAAGTATATG GGTTCCTGAT  
 22021 GCATGGCGTT TTAAGTTTAA AAGAAAGTAT GTGGAGGATG TGCTTAGGCT ATATGCAAT  
 22081 ACAATACCTT TTTTTCCTT TTTTACATA CCATTTTATA TAAGGGAGTT GAGCATCCAT  
 22141 GGTTTTGGTA TTCCCAGGGG TCCTGGACTC ACGGATAAAT GTACTTTAAA TATGCACAGT  
 22201 TGATCATATG TTTATTACAC CTTGATTAGT TGTTAAATAT GGTATACATT TTTGACTCAG  
 22261 GAATTTTAT CTTGAGGAAA TTTATCCTGA GGGTATTTAT ACTGAGGAAA TGATTAATGT  
 22321 AAATTTTAT CTAAGACGAT GGTGTGCAGC ATGGAGCTGT CTATCATCAT GGAGGATTAG  
 22381 TTAAGCAAAAT TATGGGTTCC TATGTGGTTA CTGAAAACCTA TGTCAATTTA CAAGAATATT  
 22441 AATGATGTGC AAAAATGCTC AGAATATATT TTTAAGTGGA AAAAGGCTAG ATTACAAAAG  
 22501 TGAATGTACA TGCAGCTGCT AGGGAAAAACA ATATGGAGGT TCTGCAACAA TTTAAAAATA  
 22561 GAACTACTGG CCAGGTGCAG TGGCTCACGC CTGTAATCCC AGCACTTTGG AAGGCCAAGG  
 22621 CAGGCAGATC CCTTGAGGTA AGTAAGGAAT TCAAGACCAG CCTGGCCAAC ATGGCAAGC  
 22681 CCATCTCTAC TAAAAATAAC AAAAATTAGC CAGGCATCAT GGCGGGCACC TGTATCCCA  
 22741 GCTACTCGGG AGTCTGAGGT AGGAGAATCA CCTCAACCCG GGAGGTAGAT AGAGGTGCGA  
 22801 GTAAGCTGAG ATTGTGCTAC TGCCTCCAG CCTGGGTGAC AGAGCAAGAC TGTCTCAAA  
 22861 AAAACAAAAC AAAAATCTAC ATATGATTTA CCAATCCCCA ATTTGGGTAT ATATTTAAAA  
 22921 GAATTAAGAG CAAGCTCTCA AAGAGGTATT TGTCTACCCA TGTTACAGC AGCATCATTC  
 22981 ATAATAGCCA AGAGGTGGAA GCAAGCCAGT GTCCATCGAT GGATGAATGG ATAAACGAAA  
 23041 TGTAACACAC ACACACACAC GCACACAGAG GAATATTATT CAGCCTTAAA ATGGAAGAAT  
 23101 ATCTGGTATC ATGCTACACC CTGGAACACG GATGAACCTT GAGGACATTA TGCTGAGTGA  
 23161 AATAAGACAG TCACAAACAG CCAATGCTG TATGATTCCA CTCATATGAG GTTTCTAAAG  
 23221 TAGTCAAATT CACAGAAGCA GAAAGTAGAA TGGTGGTTAT CAGGGGCTGG CGGGAGGGGA  
 23281 ATGGGGAGTT GTTCAATGGG TATAGAGTTT TAGTTTTGCA AAATGAAAGA GTTCTGGAGA  
 23341 TCTGTTGCAC AACAGTGTGA ATATGCCTGA AACTGCTGAA CTTTACACTT CAAATTATTA  
 23401 AAACGGTAAA TGGTGTGCGT GTTTTCTACC ACACGCTTTT TAAAGTGAC GTACAGGATT  
 23461 TTTGTACTTT TGTTTTAAAC AGTTTAAAGT ATCTAAGAAA AAGATGGGAA GGAAAAACCC  
 23521 CCAATAATA ACGATGTCTG AGTTTGGAC TGGTAAATGA TCTTAGTTTC CTCTTTATAA  
 23581 CCTGTAGCAT TTTTCAAATG TTACACAGTA ATCAGTTTCC TGTATATAAA GCAGAAAAAA  
 23641 TTAGCAGGAA GAAAAAATG TACTGTCTTT AATCTTAAAA AAAAAAAGA AAAGAAGAGA  
 23701 AAGTCCTAAA AGCGCTGGC CCAGAGGAGG CACAGTCCC ATCTTGCTTC GCTCCCTGCT  
 23761 TGCCTTCATG CAGTTGCAGC TTTCTTATTT ATTATTTTGG TAGGCACATA GTAGGTGTGT  
 23821 ATATTCATGG GGTAGCAGCA GCTTAATTTA TTTTGTGGG CACATAATAG GTGTGCATAT  
 23881 TCATGGGGTG GTGACAGCTT TTTTCTTAT TTATTTTGT TTTGTGGGCA CATAGTAGGT

23941 GTGCATATTC ATGGGGCAGT GACAGCTTTA TTTTATTGTT TTTGTGGGCA TATAGTAGGT  
 24001 GTGCATATTC ATGGGGCAGC GGCAGCTTTC TTATTTACTT TATTGTTTTT GTGGGCACAT  
 24061 AGTAGGTGTG CATATTCATG GGGCAGCGGC AGCTTCTCTT TGCTTGCTCT ACCCTCAATG  
 24121 AAGCCAGACA GTGGCAGATC CTCCCCGCCA TAGTCATGTT GCATCTGTTT GCAAATCATT  
 24181 CTGCTCTTTG GTTTTCTTTC TTCAGACTAA ACAGTCCCAA TTGCAGTCAC TCTTTCATCT  
 24241 GAATCCAGTT GTTTTTTCAG TCAAACTCTA TCACAGGCAG TTCTACAAGG GAATGCCATT  
 24301 AGTTGTTTCG AAATCTATTG TATTGAATGC ACGCTTGAAA TTAGGAAGCT GGTAAATAATG  
 24361 GTCAGAAGAT TTGCTCTAGA CCATGTTTCT TGGTTGGCTT TAAATGGTGA TCAAGTCCAG  
 24421 TGTTTTACCC TGAAAACCTG ATTGCAGAGT GGTAAAGGC CTGGGAGACT CGTGCTGAGA  
 24481 CCAGAGGCAG TGCTGTGGCC TGGTGTCTCT ACACGCTGAT GGCCTTCTC TGTGGCCACT  
 24541 CACCCAAGTC ACTGCCACCA CCCAGAGGAG GGGCTTATAG GACACACAGG TGGAGATGGC  
 24601 TAGGAGGTAC TTGCAGGAAG TCTGCTAGCT GCGTATAAAT TGTCTTCTT TCGCAGCAGA  
 24661 ACTCTGATGT TTAGGTGGGC ATATTGCCAC CCATGCAAAA AACAAAAACT TCCAGCCTCC  
 24721 CTGCAGTTAA GTGTGATGAT GAGACCAAGT TCTGGTTGGT GATATGTTAG CTGCAGGATG  
 24781 TGTTGGGACT CCAGGAAGTC CTTTAAAGAA GGAGTTCAAT CTTTTTCTC CTTCTTCTG  
 24841 GTGGCCTGGA ATATGGGTGT GATAGCTGGA GCTCCAGGGC TTACCCTGGA CCATGAAGCC  
 24901 AAGGACCACA TAAGGGTGGA TAAAGCAAAG AGCCAGAAGA AGCAGGGTCC CTGACGAATC  
 24961 GTAGGTGTGC CTTGAACTGC CTTCTGTGG ATGTCTTTTA CATGAGTGTG TTTAAGCCAC  
 25021 CATTTTTTCAG AAATAAACAT GCAATTCCTA ACCAATAATC AATTGTGTAC CAGAAATGGA  
 25081 CTGTGCCACA ACGGAACCTA AAATGTGACA TGTGCTCAGC GGACAAGGGG GAAGGTCAAT  
 25141 ATTAAGGATT CAAATATTGT AGCTGGAATT TAAATATGG CTGCACAGT CTTGACATGC  
 25201 CCTCCATCCA GAGATGGGTT CATTTGCCCC TCCCCTTGAA TCTGGACTGG CTTGTGACTA  
 25261 CTTCTGCCAA TAGATTTCCA TGGAAAGTGA GCTACGTAAC TTCTGAGGCT AAGTCATAAT  
 25321 GGGCCATGCT GCTGCCACTG TAACAGACAC TCTTGGAACC CTGAGTGCCA TTTAAGAAGA  
 25381 CTGACCATCC TGAGTCCATC ATGCTGGGGT CACCAAGGTG CTCTGATTGA CACTCCTAGC  
 25441 TGGGATTGCC TTTTCATCTT CACAGACTTC CCAGCCAGG CACCAGACAC AGGAGTGAAA  
 25501 GAGTCGTTGG ACGAGGGCCC TTCAGCCCAG GCTGTTCCCA GATCCCAACC ACTGGAGTCC  
 25561 TCACCAGCAG TTCAGATCAT CTGACCTGTG GCCCAAAACA TGGAGCAGTT ATGCCCTTTT  
 25621 CAAATTCCTG ACCACAGAA TCTGTGAGCA TAATAAAATG GTGGTTAATT TATGCCATTA  
 25681 TACATTGAGG GCTTTTGTAT CTCAGCAGTA GATAACTGGA ACAGCTGGTA ATTCTTACTA  
 25741 TGTTAAGATA AAAAGTTAGT TCTTTTTCAC AGAGAAGAAG TCCAGACATA GGCAGTCCAG  
 25801 TGCAAGTTTG GTCACTTTGT AGTCAGCAAG AACGCAGATT CTTTATCTT CCTGTCTGCC  
 25861 TGGAATGTGG ACATGATGGG TGGGGCTCCA GCAGCAATAT TGGACCAGAT GGTAACTCTG  
 25921 AGGATGACAG AAGAGAGCAG AAAGGAGCTT TGGTCTCTGA TAATAGAGAA ATAACCTCTT  
 25981 CTCTTGTTTA AGCCACTGTA ATTTTAGATC TTCAGGTTAT ACGGCAAAAG CATTGCTAAT  
 26041 AGGCACAGAA TTGGAATTC TGGGTTTAGA ACCAGGAGAT TGGCTTGCTT AAAGCTAAAA  
 26101 GGTTAAGATA TTTCAAATTA GAAGCAGTAG TTGAAGTCTT GGGTATAAAT GAGGTTGCTT  
 26161 AAGGAGAGTA CAGAGTGAGA AGAGGAAAGA GCTAAGGAGC CTCGAGGAAT GGCAGCAGTT  
 26221 AAAAGAGAGG CCAGAGGAGG GGCTCCACCC TTTCAGGGG AAGAAACCCG AGACAGAAGG  
 26281 TGTTCACTGA ACACGACTTC TTCCCCTCTG TGTCTGGGTT GGGAGCCCTT CCTATGCCTC  
 26341 TTCTGGACAC AGCTAACATG CAGCTTTTCT TTCAATACAC AAGGGCTCTG TTTTACATGA  
 26401 GGGCTCCCA CATGGGACCA GAAGGAACCT CTTAAAGTT TGGAAATCAAT TGTAAATTCT  
 26461 TCTTGTGGCC CCAGTGCCCTA CAACAGTGAC TGGCATGTAA TGGGCAGGCA GTAGTTATTT  
 26521 GCAAGCTGAA CTGAATTGAA CTGCTTTTGG CAAGGAAGTC CTGGGAAGGG AAGGTTTCCA  
 26581 GAAGGAGGGA GTGGTCTGCG GTGCAAAACAC TGAGAACAAT CAAGTCAGAT AAGGCCTTTA  
 26641 CAAGACTGGT GGGTTTGGCT CAGTGCATCT GGCAATGAGG AGGCCATTGG TGACCTTGAG  
 26701 GAGAAGTGTT TTCGTGGGGA GGCAGCGGGG ACAGGGTAAC TGGTGTGTGG GAGCGGGTGG  
 26761 GGTGCTGGTT GCTCTTTTGG GAAGAAAGGG TGGAGACAAG GAGGGACGGG TTTTATTCTT  
 26821 TTTGTCAATT TCATCATCTT CATCATAGGC AGAGAGCAGG GAACTGGGAA GGGAGCCTCC  
 26881 CTTGACTTCT GCAGGATACC ATGTGCTTGC CTTGTGAGAG TTGATGACAC TTTTGTCAAT  
 26941 TCCTTAGCCA ACTCACCATT TAACTTCTCA TGTATGCATA ACAATCTTTT CATGGTTTTC  
 27001 TCCCAGAACAA AAATGTAAGT TCCTGGAAAG CAGAGACTGT TTACAGCTTG CTGTGCACAA  
 27061 AGTAGGTGTT TCGTTGGTGT TAACTGCCTC TGTTTCTCT CAAGCAATGT TTCCTGAAAT  
 27121 ATACCTGCGC TTGGAACAGC ATGGATCTGT ACGAGCTGGT GCCCACTAG TTAGTGACAT  
 27181 GGCTTCATGA AAACAGAGGA GCACCTCACC ACTAGAAGAG ACCCTTGCAG TAGGAAACAA  
 27241 AGCTGCATGA GAAAGGTGG GATGGGATGG AAACGGCCTA CTCACTCCCC CAGGAGTGGA  
 27301 GAGGAGGAAG CCTGAAGGGA TGAGGGCAGG GTTATGCATA TTTTGCAGCA TCAGATAATG  
 27361 AAAGCAGATG AGTCAGCAGC ATTCACCCAC ATGGCCATTT TATTCCATGC ACAAAATAA  
 27421 TGATGATTTT ATAAACCTCG TTTTCAAAAC TTTTGTGATA TCCAAGGCAT GGCTGCCTTC  
 27481 GCCTTACGAT AAACAAGTCC TGTCTGTTGC TTGGCTCACT CTTTGATACT GGAAGGAGAG  
 27541 TCTGTTGCAG TCAGAAAGAA AAGATGTCAA AGTGGAGGCA GCCAGTTGCT CTGGTGGAGG  
 27601 GAGGGCTGCG TGTCCCAGGT TAGCATTACC TGGATGGGGA CCAAGTTTGT AGGTCCAGAG  
 27661 CTCAGTTTAT CAGTGGCTCA AGGGAGCACT GATTATACG TTCCAAGCTC ATGGGAGGGA  
 27721 GTGGCGTAAA TCCAGGCCCT TCAAACTTCA ATATGCACAG GAATCACGGG GGGCGCCTGG  
 27781 GGATGGGAAT CTCCTGAAAA ATGCAGCTTC TAATTTAGTA CAGCTGGGCT AAGGCCTGAG  
 27841 ATTCCGCATT TCCAACCAGC TCCCACTGCA ACTTTGAGTA GCAAGGATGT ACATCGGAGG  
 27901 CCAATTTTTC TAAACTTTTG GTACCTGTCA CCACTGGTTT CCTCACAACC CAAAGACTGC  
 27961 TTAGGGAGCA TGGTGGCCCA GAAAGGGGAC TCCGCTTCC AGGGTTGCTG GGTTCCTGCG  
 28021 AGACCTCATG TCTTCTGCTT TTTCTGCTTT CACCTTCTAT CAGCCTGGGC CTGTCACTGG  
 28081 CTCCGGTTAT ACATTTCTTT GGTATGTTT CCAGCCTGAA GGGAGACCTG CTGACTTGGG  
 28141 GTTCTCTCTG GTCCCAGGGT TTTCCAGAG ATTCAATTAG AAGCAATCTT AAAGGGCGGG

28201 CTTCTTCCTC TGACGGCTAA CCTAGTCCTT TTTCACTAAT TCTTCCCAGT TTCAGTCTCT  
 28261 GTGGGTGACT CTCCAAGAGT TGGAAATGCAC TGGTCACGGT GTTGACAACA CAGATAAGAG  
 28321 GAATCACATT TAGAAATAAG TAATCCCATG AGGCAGCCAG AGAGATGGGC AGTAAGAGGC  
 28381 TGAGAGATGC TCTTGAGCA GGGTCGTCAG TGAGCACCAG CAGCTGGAGT GGGCAGCGCT  
 28441 GGATGCATCG CCTGCCTGGC CCCCACCTC ATTTACCCAC AGCCTGCCTG ACCTTGCTGA  
 28501 ACTCCATCCC AGATCCTCAT TTGGTGAACC TGGGCTATAG ACTTAGGCTT TGCTGGCAT  
 28561 TTGTTTGGAG CTGGATGCTT GGGCATGGCT GTCTGGTCAT GATGACATGG ATCACCAGCC  
 28621 CTGGAGCTCT GGGTAGCTGG AGAGAGCCAC TGAAACCGGC CACCTGTTTG GCTTTGTTGT  
 28681 GGCAGAGACT GGCTTATTTC TAGGGCTCCA TCTAGATCTT CTAGCTTCAT GCTTCCTATT  
 28741 AATACTCCGA TTCATAAAAA TAAAAATCAT GAAATATAAT TGCTATATAA AATAATGAGA  
 28801 CTAGGGCCCG GTGCGTTGGC TCATGCCTGT AATCCCAGCA CTTTGGGAGG CCGAGGCAGG  
 28861 TGGATCACCT GAGGTCAGGA GTTCCAGACC AGCCTGTACT AAAAATACAA AAAATTAGCT  
 28921 GGGCATGGTG GCGGGTGCCT CTAATCCCGG CCACTTGGGA GGCTGAGGCA AGAGAATTGC  
 28981 TTGAACCTGG GAGGCAGAGG TTGCAGTGAG CCGAGATCGC GCCATTGCAC TCCAGCCTGG  
 29041 GCAACAGAGT GAGACTCCGT CTAACAAAAA CCAAAATCA CAAGACTAGA ATTGCCTTAA  
 29101 ATTTTGAAGT GTTTTCCAG TGAAACGCCG AGTGCTCTC AGGTCCCTTC GTGCATTTC  
 29161 CTAGGCCCGC CACACGAGGT AGAGGACAGA GAGAGGAGGG CTGCCAGCCC TGGGCTCTGC  
 29221 CTCTGAGTCG AGGTTTCTGG GGGCCAGGCC TCTTCCCAGG GCTTTTCTC TCCACAGATG  
 29281 TGTGGGAGCC TGTGGCTATG AGACCACCTT CTCTATAGGG AAGCACTTGA AGAACTCTA  
 29341 GTGGGTATAA AACCCCAACA GAGCTCCTCT CTCTCTCTT CTGACTTCT CAACTTTTTA  
 29401 AAATTTTTTA CTTTTCGTGA GAAAGCCAG GCCAACCTGC TCACACCCAT TGCCGGAGAG  
 29461 CAGAGCTGGC TGCGTAGTAA CCTGGTCAGC CCTGAGGGGT CCTGCACCTC TGATGGTAGA  
 29521 AGCTGACACA GAACAGGGGA GGCTCCCTGG GGCCTAGGAG GGACACTGGC CCTTCAGGCG  
 29581 CTTCTTCTTC TCCACTGCGA GCCGCAGAGC CTGGAGGAGA GTGTCCTTAT TATTCAGGAT  
 26 29641 CACTATCTCT GAGAGTTTCA CCAGCTTGTC AAAGGTGGCC TCTGTGTATG TCAGCTCCTT  
 29701 GGTGGCATAG GGAGTTTGG GACCATAAAT GTCCACCTGG CCCTGCTCCA GCTCCTCAGG  
 29761 GCTTCCCTCT ACACCTGCGG GCAGCAGGAT GACTTGTAT AGACTCTGTC ATGTTAGGTT  
 29821 GGAATGACTT GTTTCACAT CTCTCCCCAA GGCCACCAA CCTGCACTCA CGGACACACA  
 29881 CACTTGCATG AAAACCTGTA CACACGCACA CCTCACACA CGCACCTGCA CACCTGCAAG  
 29941 CTCTCAGGG CGTGGTGCAG ACTTGCCCAT TGGAGGATTG CCCTAGCACC AAGCACAGTA  
 30001 CTGGGCACAT GGCTGAATGT CAGAAAATGT AGAATGACTG AAAGAATGGG GGATGACACT  
 30061 ACCAGCCCAT GGGGGTGAGG GAATGGGAGG CGCCCTTCTT TGTGGAAGCC AAGGTTGGCT  
 30121 CAGGCTGGAG CTGAGCCAGG TTCCCCGTT AGTGCCCTTA GAGACTCTAC AGATCTAAAG  
 30181 GCATTCCCGG GCCAAGGGCA GCAGGTGCTG GGAAGTGAAC ACAGAGGCAG CTCACCTGGT  
 30241 GCCTTGATAT TTCGAAAGT GTCATTGATG AGTGGGAAGA AAGTCACGAT GGGGGCATCG  
 25 30301 GGTTCCTGGG GGTTCCTCAT CAGGTAGCAT TCCTTGAGAT TTTCATTCTC ATCTGGCAGC  
 30361 TCGTATTGGG GGAAGGGGAT GTTCTGCACA GTGCAGTACT CACAGGTTTG TTTCAGGGGC  
 30421 TTGAACAGCA CAGAGGGCGT GTTGTGACAT TAGGAGGAGT GGGAGTGGAG GAGAAGCACA  
 30481 GAGTTCTTAG CCCCAGCTTC CTGATTTCCC CAGCTAACAA CTCAGCTCCC CAACTGCTTT  
 30541 GCCAGTGTAT GGTATTTTGA GACTAAAGCC AGGTCAGGAC AAAAGGCCCG AGGCAGTAGG  
 30601 TGGATGCACC TCCAGTTGCA AGCTGCTTCC CCAAGTAAAA TGGACAAAAG CCCTTTCTGT  
 24 30661 GGTGTTTCGA TCCAAACACA GACTTGGGTG GATGTGGTGG GTAGCCTGGA AGGGCGACTG  
 30721 GAACCAAGCT ATACACAGCC AGAGCTGCCC AGGCCCTGT GGGATGTGCA TGGGCTTTGG  
 30781 AGCTGACAGT AGTACAGGGC TAGAAAAGAC TCTGCAAGGG CATCTCTAAA AGCATGATGT  
 30841 CGCAGGAAGA GCACTGCCTT AGAGTGTCAG AACCCCTCA GTTCTGCCCC TCGTTCTCTCT  
 30901 GCTCACTAGC TGTGTGATCT TAAGCAAGTC TCTCTGTGTC TTGGTTTCTT CATCCATAAC  
 30961 ATGAGAAATG ATAGCAGTTG TGCCGTCTCC CTCCAAGAGC TGGTGTAAAG GCCAAACAAG  
 31021 ATCACACGTA GGAAAGTGCA TTGTACCCA TTACAAACT TAGATTGTTT TCATTGTGTC  
 31081 CCTGTTGGTC TGTGTACACT TCACTTCGCC TGCTTCAGG GTAGAACTGT ATACACATAG  
 31141 AACTTTCCAA AGTAAATTAA AAAAATAAAA ACTTCTGTTT CTGACAAGTG CCCTCTCTC  
 31201 TCGCCAGGCT TCCCTCCAC CTGAGAGCAC TTGAGTGTA TTTCATGGAG GTCATATACC  
 31261 CTGGGGGAAG GGAGAGTGTG ATCATTTTGA AAATAGGATA TTACAAGCTC TTGAATGGAA  
 31321 TGGGCTCTAC TAAGGTTATT TTTAAAAACT CATTTTTTGT AGCTGAAGAA TTGCCCAATA  
 31381 TGGCAAATGC TCCAAGGCAA GAGTAAAAAT CGAATTTGCT AAGTCAACAT GGGCACCATG  
 31441 GTGTGAGCAG CAGCTGGCTG CCCGTGAGTT CTGGGTGTCT GACTTCCCAT CATGACTGCT  
 31501 GCTTGGCTTC CTCTGATGCT CTCTGGGGGC TGTCCAGAGC ATGCCTCTTC CCAGTTCACA  
 31561 GCGAGTGAGG CCACTGACTG CCGGGCCCAT CTTAGGAGAG GCTGCAGCCC ATGCTCAGGA  
 31621 GGGAGGCTGG GCACCTGCTT TTGTCTGGGA CCCAGCACAG TAGTTGAGGT GGATGATGAG  
 31681 GTCGGCTTTT CGCTCTGGCC TGAGGAGGGG CGGGTAGCTG GAGTTGACAA AGAACGCAGT  
 23 31741 GTCCAGCAGG CACAGGTGGT TCGCGGACTC GGTGAGCTGG TTTGGGAAAC CATCTAGCAC  
 31801 TGTGTCTGAA AGCCAAGCCT TCCCGTTACC AACACACCTG GTGCCCTGAC AGCCAATGGA  
 31861 CAGAGGCACT TGGGGGGACG GAGCGCCCTG TCTGGAGCAT CCTTCGCCTC AGCCTGACCA  
 31921 GGCAGCCCACT TGAGTTCCAG TGCCCTCTGG GTGGCTGAGG AAGAACAGCG TTATCTTGGG  
 31981 CTTTTAGCTA GTCTTTTATG GGATATTGAT ACTTGTGGGG TATAAGGGGC TTAAGACCAC  
 32041 CCCATAAATA GTGTAAACCC TGGGCCCTGA GTGAGCTGTC CCATTGGCCC CAGATGGTGA  
 32101 GGATGAGGCT TAACAGGCCA GGCTGAAGCA GCCATTTGCA GGCCACGGCC AATTAAAATA  
 32161 TTAAGAAGGC GAAAGTCAAT GCCTTAGGGC CTCAGCAGCA GAGGAGGCGG CACTGAGCTG  
 32221 ACGCTGCAGG ACCACAGTTT AAGGCCTGGC TCTGCTGCTT ACCACTTGGG GGTTCCTGAG  
 32281 CTGGGCCATC CTGCCCCAGT TCCAAGCGGG CACCCTCAC ATTTTTTTTT TTTTGTAGAC  
 32341 AGGGTCTTGC TCTATTGCCC AGGCTGGAGT GCAGTGGCGT GATCATGGCT CATTGCAGCC  
 32401 TTGACCTCCT GGGCTCAAAAT GATCCTTCTA TCTCAGCCCC CAAGTAGCTG GGACTACAGG



36721 TCTGACCCCA GGCCAAGCAG CTGACTCCTG GGTGACCAAC CCTGGTCACC CCAGTGTGCA  
 36781 AAGGCTACAG AACCCGTGTG GAATGGCAAA CACATGTATA TGCCCAATGA CATGTGGGTC  
 36841 TCCTTTGAAA GAAAACAGGG TCATAACTGG GCCCAGAGAT CTGGGCATCA CAGAGCCAC  
 36901 TGTCTTGAA GTAAAGCTGC AGGCTCAGGG AAGAGCCGGT CCACATTCAT TGTCAAGGTG  
 36961 GCCCAAGGGA AAGGGGCTCC TGTCTGGGCC AGAGCGCAAG AGCAGAGGCC TCTCAACAAC  
 37021 AACCTGCCCC TTGCTTTCCT GGGAAACCCAG CTCTTCCCA CCAAACCCCC AGAATGTGTA  
 37081 GACTCCTGAT GTCGGAAGAC TCGGTGCTTC CTCTCATCA TATGCGGCAC TCATGATTGT  
 37141 TACAATATTG TTGAAAATGT TAACTAGCAC TTATTGATGT AATTCCTACT TAAGGAATAG  
 37201 TAAAGAGCAG GGCCTGAATT TCAAACCCAT CCTGACTGAA TCCCAAGGCC GATGCCCTTG  
 37261 GCCTCTGCCT GCTCTGCTTC ATAGGGAGGA CAGGTAAAGT GGTGGGTAT CACAGGGGCA  
 37321 GGGTGGGTGG GATCCATCAG AAGTGGCGAG GGCTTGGGT TGGAAGCCAT GTCGATTCC  
 37381 TTTCTGAATC CCCATTGACC AGCGCAGTTC CTGGCATGCA GGAGGTACTC AAGTAATTTA  
 37441 ATTGAAGGAA TGAGGGAAGA TTGAAGATGA GGGTCCGGTC ATTCTCAGGT GCTCAAACATG  
 37501 TCTGGCTTTC TGGAAGATGG AACCTACTAT TATAAAAAAT AATTGCTGCA ATGACTTTCT  
 37561 GGCAAGTTCT GATTGAATT CAGAACTCCC CGGACACCTT TGGAGTGGAG CAGGGAACCG  
 37621 TGGGCTTGGG CAGAGGTGGG AAGCTCTCCT TTGGGGCCAG GCAACAACAA AAGTTGTAA  
 37681 TCTCGGCTTG TGCAACCTAA AGGGCACAGA TGTCAGCTGA GCCCACCCTG ATGTCCAGGA  
 37741 TCGCAACACC CCTTGAGAT CCCTGTGCTG TGCCATGTGT CCAGGAAAAG GACCACCCTG  
 37801 TGCTGGGTCG CAGCTCTCAC AGCAGCAGCC ACCCAACACA GGCTGGTCCT CTCTTGGGC  
 37861 ACGTATTAGT CAGATCGCTG GATCTCAGAA TAGGAAGGAA CCTCATTGGG CCCACTGGTG  
 37921 GCCATGAGAA CGGGCCCTGC AGACCTCCAA CTGCAGAGAG TGTATTGAC TGAGGACCCC  
 37981 AGCTGCTGTG CTCTGAAATC CACAGCTGAG TCTGTGCTGA GGCTGCATT CTGTGACTG  
 38041 AGTGGGGCAG GGATACTGAG GCTGGCCCGT TCTGGGAGAC GCGGGACTCC TGTGACAGCC  
 38101 AGCTTTGGCT GGAGGACCTT CCAAGGGCCT TGCGAGACCAC ACGGTGGTCT GGGATGCTGC  
 38161 CGCCTGACCT TTCTTCCTCT CCCCACCCCT AGGGTCAGAC TCGCTTCACG GCGTCACTGC  
 38221 TCTCCAGTCT TCCCCGGCTT CCTCCCCGTT TTTCCCTCAC AGGTGTCTTG GCGTGGCTCT  
 38281 CAGAGGACGT GCTGCCCTAG GCAGGTTTGC CTCCAACACG AGGGTCTCAG CCATCCTCTA  
 38341 AAGGGTACAC GGGAGACAAT CCCACCACCT TCCTCCGTCA TAGAGTCCGA CACCTTATCG  
 38401 GTCTTCCAGC CCAAGGATAC TCCCGCAGGT CTAACTAAT CCCTGTGTG CAATGGAAGT  
 38461 CCTGACTCCC CTGAGGAAGA AAGAGAGCCA GGCATCAGCC AGGGCAGGGT GGGGCCAGG  
 38521 AGAAGCCGAG CACCTACCTC GTCCCCCAGG CAGGTCTCTA TCAGCAGGCC CCAGAAGTCT  
 17 38581 GTAAAGGTGA CCCTGTAGCC TTCTTGGCTG CGCTGCCGGA GCTCCTCCTG GAATTTGCGG  
 38641 AGCTGTCTG GGAACAGGGA GGGTAGCTTG TCCTTTACCA CATGTCTCCG AGCCTCAAAG  
 38701 ATAGCAGGCT CCAAGTTTTT GGAGGACCAG TCAGGGTCAC GGTACAAGGT AGCCATGGTC  
 38761 CTGGAAGAG ACAGGCACAG CTGATAGGGC TGACCCGGCC CTGTCTGAT GAGTCTGGGA  
 38821 CCCGGGCAGG AGGGAAGGTG GGCAGGCCAT TGAGAACCCT AACCTGCCAG GAGGGGAACC  
 38881 TGGAGATGAG GGGTGAAGCA GAGGCTGGC AGACAGTGAG GAGCAGGGGA GGTGAACATT  
 38941 ATTACTCATA ATCACAAGTG CTACCAGTTC TCAAATGCCC ACTGTGTGCC AGGCATAACA  
 39001 GTGGTTAAGA TGAAGATTCT GGATCCAGAA CTTTCTGGAT CAAGCCAGGC TCTGTGCTC  
 39061 ACTAGCTGTG TGATCTTGGG CAAGCTACTT AATCACCCTT CCCTTCAGTT TTCCTGTCTG  
 39121 TGGAAATGGG AGAATTACAA TAATTGTACC TGCTTCACAG AGTTCTGGTG AAGATTACAT  
 39181 GAGTTAAGGT GATGTGGGGT AGCTAGAACA GTGCATGGCA CAGAACAAT GCTGGAGCCC  
 39241 ACATTAAAC CCAGGCTCGA CTGATTACAG GTCTGGCTG TTCTACTCT GCCATGTAA  
 39301 TTCCTATAAT TCTGATGTTG ACTGAGGATT TATTTCCCTT TCTCTTCTTA ACTTACATGT  
 39361 AGTTGTGCGG GGGGAGCAAA ACAAGAGGT GATCCCACTA GCCCTGGGTG TGGCTCAGT  
 39421 CTGCCTCTTC TTGGCCTGGA TCCAGCTTTT CTGCGTCAGG ACAGGAAAGG GCCCTTTTCA  
 39481 TTCTGCGCCT GAGAATGACC TCCCTGGCAG TGCAGTAGT CTGCGGCTGC ACATTTAAGG  
 39541 AAACAGGGGC CTCCTCCCTC ACAGGACTTG CTTTGGGAA GGTGGGAGT GGTGAGGAGA  
 39601 CTTTTCAACA AAAAGCACCC TGCTCTCCAT AAGCTTCACC CATTCAATTT ATACTTTAAA  
 39661 AAAGCTTCCA GAAGCCATTG CAAATGAGTT TTTAAAGAA TACGACTTGT GCTTTCCCA  
 39721 GGCCAGGGGA CCCCACCCAT CCCACCCCGG CTGGGCTGTA GTTCTCTCTG CTATCTGTCT  
 39781 ATCGCATAGT AGTCCGGGAG GACGCAAGTG CCCCTCAACC TGCTGGAAGC AAGTTACCTT  
 39841 CTTAACCTAC TACCGGGGGC AGCCATTTAG CATGGAGGTT AAGGATGTGG GCCCGGGACT  
 39901 AGGTTCAAAT TCTACTCCAC TGCTCATTA CTGAGTGTCC CTGGCAAAT ACTCTGTTT  
 39961 TCCAGCTGAA AAATAATGAC AATGCCTCCC TCATAGTTTT TTGTGAGGAA GAAACAAGAT  
 40021 GAGCCATAGG AAGCAGTGAC TCCAGTGCTT AGCAAGAGGC AAATGTCTCC TTCTGTCTCA  
 40081 CCTGGGCTGG CGCCTCTGCC CCCCAGGGGG GGTGTGTGGC ACAGGAGCCC TTGGCTTTCC  
 40141 AAGGCCCAT CACATCCATT TGCTCACACG GGGGCTACCC AGAGCTCCTC CAGTATACAA  
 40201 GGCACAGCAA GGGACCTCAG GGCTCACAGG CCCACCAGTC ACAGACAGGA CACGGCCCTG  
 40261 CTTGATAGGC CACCCTGAAC CAGGGCCACG CAGGGTTTTG ATGGGCCCCA CAAAAATATG  
 40321 TTATATGTTT TTTATTTTTG TTTTTTTTAA AGACAAGTCC TCCCTATGTG GCCCGGCTG  
 40381 CTCTTGAGCT CTTGGCCTCA AGCAAGCCTT CTGCTCTGCT CCCCCAAGT GCTGGGATTA  
 40441 CAGATGTGAG CCACCATACC TGGCCACATA AAAAATATTA AAATTTATAT TTCATGCTGG  
 40501 CCACATAAAA AATATTAAAA TTTATATTTC ATGACTGTTG TATAAAGACA AATGTATTCA  
 40561 TATTATTTAT TAAAACATTT TCTTCAACCT ATAAGTTTGT TAATGAAATG ATCCTGATTT  
 40621 TAAAAGAAAT TAACCTATTT TTGTGGGCCC CTAAAGTAC AGTGAGCCGC AGGCACTGTG  
 40681 CCTTCTGTGG GTAAGCTGAC CTCACCCTGG CCTCCCTGTG GGGGAACCCA GTGCTCAGCC  
 40741 CCTGGCCCTG CTCACCAGCA GAGGGTGAGA AGGAGGGAGC TCAGCTGCAG GGCCCTGCCC  
 40801 GTTGCAATGTC CCCAGCAGTA TGGGGTCATG TTCCTGGGTC CATACCACAG CCCAGGGAA  
 40861 ACAGTAACAC CAGGAAAGGC CAGGGCATAC AGTGTCTCTG GCATACTTAG TAGCGGGAGA  
 40921 GCGTCTTACA CCCACCGCAT GAGGTCAGAG CTAAGTTTAT CCGCACTGTG CAGATACACA

40981 AGCCGAGCCT CCAAGGAGTC AAGCTGCATG CCCAAGGCCA CACCACTGCA TGTGGGGCAC  
 41041 TTCCCAGGCT TCAGGAGTCA CTTTGTCTTT GGGTGGCGGG GAGGAGCTTC ACCTGAGACC  
 41101 CAGGGCCAGC CTGTTCTGTGG CTTGCCCTGG GCTCCACAGA CAGTGGCTCA TGTCTCCTTA  
 41161 CCAGGTGGCC CCTGATAGAC CAGTGATGTA GCTGGCACAG TCCAGGAGGT TCAGCTTCTG  
 16 41221 CAGCCCCAGC AGGTGGCCAT ACATGGAGGT CATGGATCTT GTTCCACCCC CAGTGGCCAT  
 41281 GATGGCTATC AGCGGCACCT GGGGTTACAC AGAGGCAGGC GGTGGGGGAA GGCTCCAGAA  
 41341 GCAAAGCTTG GGAGAAGACT TAGAATGACG GCAACTTTGA AAGAACAATG GAGGAGGCAG  
 41401 AGCTGGCTAC AGAGGGCAGG GAGTGC GGCG TTGAGCCACC TGACACCTGC CTTTCCACAG  
 41461 AGCAGTGGGT CTAGAAATAG ACCTCAGCTC CACACCCAGA AGTTGGTGTG AGTGAAGGAC  
 41521 TCTGAAGCCT AACACAGGCT TGGTAAATCA CAGAAAGCTG TGGCTAGCAG AGTTCCTAGC  
 41581 ACTGTTAGTC TAGCCCTCGG TTAACAGGTG ATAAACTGG GGTCCGAGGG TGGTGGGCAG  
 41641 CTTGTTCAAG AACCCCTACTA ATGTGCGATC CCTCAATGGG ACAAAACCAAC ATCTTCTGTA  
 41701 GCTGGATCTG AGGGAAGGCA CCAGGCCTGG CACATGGTGC CTGCTCAGCA GGTACTGTAC  
 41761 TGGCAAGGCC AAGAGTCAAT GGACTGGTGA TGTGGGAAGG AGCAGGGTTG CCAGCTCTGG  
 41821 GTCAGCAGAG GTTTAAGAGG CCCCTGGAAT GTTTCCTTAA AGATAGGGCC AGGTACAATG  
 41881 GCTCACACTT GTAATCCAG CACTTTGGGG GGCCGAGGTG GGCAGATTGC TTAAGCCCAG  
 41941 GAGTTCAAGA CCAGCCTGGG CAACATGTAG AAACCCATC TCTATAAAAA AATACAAAAA  
 42001 TTAGCCAGGC ATTGTGGCTC ATTCTGGAG TCCAGCTAC TTGGGAGGCT GAAGTAAGAG  
 42061 GATTGCTTGA GCCTGGGAGG TGGAGGCTGC AGTGAGCTGT GATGGTGCCA CTGCACTCCA  
 42121 GCCTGGGATG CAGAGCAAGA CCCTGTCTCA AGAAAAAATA AAAAAAATA AAAAAAATA  
 42181 GGAGGGAGGA AGGGGCTCTT TGTCAAGACT CATTCGTCTT GTTTCTCTT GAGTGGGGGC  
 42241 TGTAAGAACC TTCCATTGGC TAAAGCGCTA TTACAAGTCC TCCTCGGGCA CCCGCCAGGT  
 42301 CTCTGTCTCT ACCTAGGGTC AGCCTCAGAA AGCAGGCGCT TGGAGGGCTT GCCTGCTGCT  
 42361 CAGGCGAGTG TGGGCTTTGG TCCACACAAG GGAGGCAACA GAGGGCTGTG ACAGAGGGCC  
 42421 ACCGTCTGCT CCCGTGGGAA GAAGGGGCGG TCTGGCCCTA GGAGTGCAAG TTGCTTCTCT  
 42481 CCAGCTTCTT AGCGCTGTTA TTTCTGGGAG TGTGTTTACC TGTGCGGATT TGAGAGGAGA  
 42541 GACAATTGAG GAAAGAAGGA CCTAAGGCCT TTTGGGTGGA GGTGTGGGGT GCAGAAGGTG  
 42601 GTACAGCTTA AAATGTTTTC CTGGCTGGGG TGAGGAGGGC ATTTTATCCC TGAGTAAGGC  
 42661 CCCTTTAGCC ACAGCAGGAA CTCTTAGTGC TTGGGAGGAG GCGCTGGGGC TCTGTCTTT  
 42721 CTGTCTCCAT TTTGTCTCTC CTCTCTCTTG TGAATATAAA ACAGAATCAC TTCCAGCACA  
 42781 GCTCTACCA GGGCGCCTGT ACCTCAGCAG TAAAGCATC CATGGAGGTG GCGCCTGGGG  
 42841 TAGGGCAGGG CCTGGGTTGG CTGGGAGGGC ATGGTGTGG AAGGACCAGC TGGGCCAGG  
 42901 CTGGCCACAT CCCTGATTAC CCACCTCTGC CTCTCTCAGG TCTTCTCCA GCTGCAGCAC  
 15 42961 CTGCTTCAGG GCCTTGCCA CCACGACCTT CCGCTTCTGC AGAACTCCA GCTCTGCTGG  
 43021 GCACAGGCTG AAGCCAGGCC GCACGTCCAG TGTCTCAGGG CTGGAGGGAG AGAGGACCTT  
 43081 GTATTGGGCT TGCAGCCTCT ACCTCCCTGG AGTAGACACA GCTTTAGGGT CCATTGGACC  
 43141 TAGGCTCAAA GCGTACTTCT CCGACGACCT ACTTCTGTGA CCTTGGGTCT CCATTCTCCC  
 43201 ATCTGTAAGG TGGGGCGAAT CGTGCTGAAC TCTCAGCACT GTCTTCAGCA CTAAAGAGA  
 43261 TAACAAATGC AAAGCCCTCA AAGCCTTAGC TGGCACAATG CGAGCTGGTA CTCAACAAAA  
 43321 GATAACTGTT GGCAGTGTGT TAGGGACGCT TTGAGTCTCT TGTCTCTCTG GCCAGATCTG  
 43381 GAAAGTCAAG AAAGTCTTCC CCCAGGTGGC TTCTCTAGCC CAGGGCCAGG CATGGGTGAC  
 43441 AAAAGGCGCT GAGATCATTT TGCCAAATCC CTCTGAACAA GGCCCCACGG TATGGACTCT  
 43501 TCTGTGCTAG GTAAACAAAC TCCCTGTTCC TATGGAGCTC ACGTTCTTCA ATGAGCTCTT  
 43561 CCTGCTCCTA GAATAACAAT TATTCTGTGT GTGTGTGCAA TTGACAAAAA CCATCCCATT  
 43621 TTGCTGACAG TTGTGTACCT TCGATGCTG CAGGCAGGTG GGGCAGGCAC CTCTGTGACA  
 43681 TGCACAATGC TTGTCTCTTT TTTTCTTTT AATCACATCC CTGCCCCAAA CAAGTCCAGG  
 43741 CTCCTGGCCT GGCATCCAAG ACCCTGCCTG ATCTTTCCAA GCTCACCTCA AACTACCCGC  
 43801 GAAACTTCTG GGTCAAGTCC CAGAGCTCCC AGAAGCAGTC CTGTTTCCGG GTCCTGGCAC  
 43861 AGGCCCTCTG CTTCAACCAG AGTGTCTTCC CTTTGCCAC CTCCAGCCTG TGCATCTCT  
 43921 GAGGCCACAG TCCCAAGTGA CTCCAGATGA ATGTACCTTG ACCATTTCCT ACTGAGCCTT  
 43981 GGCTCACTGA CAAAGCCACA TCTCCTTGAA GGCTGGGCCT AGGAGATTTC CATCATGTCT  
 44041 CCTCTTTGTC CAGCACAGGG CTGCAGACAT GGCTGTTAGA AAATAATTCT TCTTTTGGCC  
 44101 TCAACAGGGA ATGAATAATA ATAATAAAC AACAATAATA ACAATAACAT TAGCTGACGT  
 44161 TTGCTGACAG TCCCACTTGC CAGATGGATT ACTGAGTACT TCGCACGTGC TGTCTAGCTG  
 44221 GATCCTCAGA GCAACCCAGT GATGCACAGG ATCCTTGTTA TTCTCCCAAT GTACAGATGA  
 44281 GGAAACAGAG GACTCAGTAA GGGCAAAAAG CTTTTCCAAG GGCAAAATAGC TAATCAAAC  
 44341 GTCATAGAGG ATCATGGCCA CTTGATTTCT GACAATCAAG ATTTCTGAAA ACGATGTTGT  
 44401 TTTTCAAGAG TGGTGTCTGAG ACATCTGGAT ATCCACATGC CAAAGAACGA GGTGAGCCC  
 44461 CTCTTTTACA CCATGTAGAT TGATTACCTA AAAATTGATC ATAGACCTAA ATGTGAGAGC  
 44521 TAAACTATAA CAAGTCTTAG AAAAAACAT GGAATATAT TTTCAGGACT TTGGTCGGGC  
 44581 GCGGTGGCGC ACGCTTGTAA TCCAGCACT TTGGGAAGCC GAGGAGGGTG GATCACTTGA  
 44641 GGTTCAGAGT TCGAGACCAG CTGGCCAATA CGGTGAAACC CCGTCTCTAC TGAAAAACA  
 44701 AAAATTAGCC AGGTGTAGAG GTGTGCATCT GTAATCCCAG CTACTGGGGA GGCTGAGGCA  
 44761 GGAGAAATCG TTAACCCAGG AGGCAGAGGT TGCAGTGAGC TGAGATCACA CCACTGCACT  
 44821 CCAGCCTGGG CGACAGTGAG ACTCCATTTC AAAATAAAAT AAAATAAAAT AAAATAAAAT  
 44881 CAAATAAAAA TTTCAAGACA TTGGATTAGG CAATGATTTC TGTTTCTTA GATATGACAC  
 44941 CAAAAGCACA AGTGACAAGA TAAAAAACA AATTTCCTCA AAAGTAAAGT CCTTTGTGCT  
 45001 TCAAAGGACA CCAACAAGAA TGTGAAAAGA TACGCCATAG AATAAGAGGA AATATTGCA  
 45061 AATCATATAT CTGACAAGGG ATTTGTATCT GGCATATGTA AAGCACTCTT ACAACTTAAC  
 45121 AATAAAAAAA CACAATAAAC CCACATAAAA AGTGGGCAAA GGATCTAATA GAAATGTCTC  
 45181 CAAAGAAGAT GATATGCCAG CCAGCCCATG GAAAGATGTT CAACATCATT AGCCATAGGG



	45241	GAAATGCAAA	TCAAAACCAC	AGTGAGATAC	CACTGCACAC	CCACTGCTAT	GAATAAGTGA
	45301	GAGGCACGGA	CAGAATGCGG	GTAATATACC	AGGGTGTAGA	CTTCATGTGT	AATTCATAAC
14	45361	TCTCACCTTG	GGGAGGGAAA	GATGAAAAGC	TTGTGAGAGG	TGACAATTCA	GAGCCAAGTG
	45421	ATCCCAGATC	CACACTGTCC	ACCCAGACCC	CACTCAGCAC	ACACGTGCAC	ACACAGGCGC
	45481	TCCATCCCTG	CCTCTCTCTT	CTGGACACAA	TCTCCCCACC	ATCTTCAACT	CAGCCTCCAT
	45541	ACACTCTTTG	GACCACACTG	TAGCTCATCA	GTCTCCTGCC	CCACAAGACT	CCTCCATCCT
	45601	GCCTGAGAGA	GACCCCTCAG	AGTCACATTG	TCTCTGTATC	CTCCCCAGGC	TCTCCCCACC
	45661	TCCCTACCAC	AGGAGCTCCA	GGGGAAGTAA	GTCAGGGGCC	CTTCCCAGAC	TCAGGCACTC
	45721	ACCACAGGCA	GGGTCAATC	CTGACCATCG	GAAAGGCAGT	CGAGGGGCTG	GCTGATGGGG
13	45781	CCCTTCTTGC	GAGAGCGGCA	GCAAAGCTGG	AGGGATGGGT	GGGTTCTGTG	AGAGGGGCTG
	45841	GGGAGCTCCT	CTGGCACCCC	CCACCTGTCC	CAGTCCCCCA	CACCTCCCCC	AGTGTGAGCC
	45901	GCCTTGATAC	CCACAGCTCC	AGTGACTCTT	GGGCACTTCC	ACGTGCACCT	GGGACTGGAA
12	45961	GTAATTTGGG	TAGTGGGAGC	AGGCAGCGGT	TTGGAAGCAG	GCAGAGGTTG	GGCAGCAGGG
	46021	TTCCAAGCAG	GGCCGGACAC	GCTGGGTGTT	CTCAAAGGAT	TCGTTACCCA	TCACCAGGAG
	46081	GTCTTCTACT	GGGAGAGAGG	ACAAGAGGGT	GAGACCTTGG	GAGCCTCTCA	CTACCCACCC
	46141	GCAACTTGGT	CCAGCAAAAA	ACCTATTTGG	AGGTATCCAG	GAGGAGTTTA	GAATCACCTG
	46201	GGGCACCACT	TCAAATTCCC	CAGAGATTCT	CAATCTATAG	GGTAGGGCAG	GGCCAGGAA
	46261	CCTGCATTTT	AATGAGGTTT	ATGAGGATGC	AGACAAACAT	AGTCCCCAGA	AAACCTTTTG
	46321	AGATGCTGAG	AAGGTCTCTA	GGACCACTCC	TTGGCTCAGC	TTTGCCAGAA	GATGCCCCAA
	46381	AGTAACCTGC	CCACTGTAGA	GGAGGCAGAC	CCTCCGCACA	CTGTGCTTTT	TGATGTAGAC
	46441	GCCTCTTCTG	TCTCCCTCTC	CTCCCTCATG	TCTGGTAGCC	TTGCATGGGC	AACATGGGGT
	46501	GGGGTGGGGA	GTAAGGTAGG	GGTATGGTAT	TCATGGATCT	CTTAGGTCTC	TGCCAGCTAT
	46561	AAAGTGCTAG	GACCCCTCTG	GAGCTGTGGT	CTGCCTAGAG	GAAGGGGCAT	AGTTAGAGGG
	46621	GCCTTGGGCT	TCGGGTGCTG	ACCACCTCAC	TGACCTAGGC	AATCTGAGGT	CCTGGCATGG
	46681	CACAGCACAG	ACAAGCGTGG	ATAGAGGGAG	CAGCACTGGG	ACCCAGAGCC	TGAGATACTC
	46741	CAAGGTTATG	TGGGTGCTTG	GGCGAACACG	CCCTTTACAG	AGCAGGATGG	GAGGCCAGGG
	46801	GCTCAGGACC	CCTGGGGGTG	AGATGATTCT	GAGGGGTTTC	AGCCGAGCTT	TTGGTCCCCC
	46861	TGAGCCGAGG	GAAGCCCTGG	ACTAGGTGGA	GTCTGATGCT	GATGGGATTT	GGAAGCAGCT
	46921	CCCCCTTGTC	CCCCAGTAGG	CAGAAAAGGC	AAAACAGCTC	ACTTTTCTCC	CTCTTCCTCC
	46981	GCCTCCCTGA	TTGTGCATGA	ACCTCCAGGC	AGGAGACTTG	TCGAGACTGT	AAAAAACAAA
	47041	GGAGGGTCAC	AGAGTGTCTG	AGCACCAGCC	ACCTCCTTGC	TCTTCCAGGC	TGAAGGGAGA
	47101	TGAGGGACCA	GAGGATACAT	TCAGCATTGG	TCTGACTGGG	CTGATGCTGG	GGAATGCAGA
	47161	ATGTGTTGAC	CACATTTTCT	GAACCCAATC	AAGGCATGCC	ATCTCACTCA	GAGATAGGAT
	47221	ACTTGACCAC	GAAGCTGTCC	CAGGATGTCT	AGGACCCACG	GTCACAGTTT	ACTAGGGGTC
	47281	AGGTCTCTAG	CTGCAGGTCG	GGAATCTGAC	CTGAGGGAAT	GTTTCACTCT	GGGCTTTCAA
	47341	AGTGAGGGCT	TCCGACAAAA	TAGCCTTGTC	AAGATGCTGA	GGCAGATCAT	ACTCCTTTGC
	47401	CACACATTCC	CAAGACTCTT	TCTTTTGGAG	GGAGAGGTCT	CCTTGAACCC	TGAGGCACCTG
	47461	GATGCATTTA	ATTAAATAG	CTGTACCAAC	AGCCGAAATA	TGAATTTTTG	TTTTTGGCAC
	47521	AGGGTCTCAC	TCTGTACCCC	AGGCTGGAGT	GCAGTGTGAT	AATCACAGCT	CACTGCAGCC
	47581	TTGACCTCCT	GGGCTCAAGC	GATCCTCCCA	CCTCAGCCTC	CTGAGTAGCT	GGGACTACAG
	47641	ACACATGCTA	CCATGCCTGG	CTAATTTTTG	TATTTTTTTG	TAGGGATGGA	GTTTCACTAT
	47701	GTTGCCACAG	CTTGTTTTGA	ACTCCCGGGC	TCAAGTGATC	CCCCAACCTT	GGCCTCCCAA
	47761	AGTGCTGGGA	TTATAGGCAT	GAGCCACCAT	GACTGGCTAT	GAAACAATTT	TTAATACAGA
	47821	AATTGTGCTT	GTAGGCCAGG	CACAGTAGCT	CACACCTGTA	ATCCCAGTAC	TTTGGGAGGC
	47881	TGAGGCTGGC	AGATCGCTTG	AGCTCAGGAG	TTTGAAACCA	GCCTGGGCAA	CATGTTGAAA
	47941	CCCCGTCTCT	ACAAAATACA	AAATAATAGC	TGGGGGTGGT	GGCGTGCTTG	AGCACGGGAG
	48001	GCAGAGGTTG	CAGTGATCCG	AGATCGCAGC	ACTGCACTCC	AGCCTGGGCG	ACAGAGTGAG
	48061	ACCTTGCTCT	AAAAAAAAAA	AAAAAAAAAA	AAAGGTAAAC	TGTGCTTGTA	GAATTACTGG
	48121	GGTTGTGGGG	TACAGGACCA	AGACTGAGTC	TGAGAGGAGG	CTGCACTTAG	TGGGTTGTTT
	48181	TATTGGGGAG	GCTCTTAAGG	GGAAGGGTCA	CTTTTGGGAT	ATGAAGTGCT	CGGCCAGGGT
	48241	GATGTCCAG	CTACCTCCTC	CCTCTGTACA	GAAGCCCCAG	CTCGTGCCCC	AGTGCTGCCT
	48301	GTGCCCCATT	TGGCTGGCCT	CGTTTAGTCT	AACCTCTGCA	GAGGTGAGGA	TAGGACAGAG
11	48361	ACCAAGCTGG	GCCCACTGTT	TTCTCCCTTG	GTCTGTGTTT	CCTGAGAGCA	GCAGCCAGCC
	48421	GTGGCCTCTG	GAGCTACAGG	AACCCAGGAA	GATAATTACC	ACCAGCACGC	CATTGGTGAC
	48481	GAGGGTCTCA	GGTGGAGAGG	GACTGAAAAA	CAAAAGGAGA	ACTCCTGGTC	AATTCTAGCA
	48541	TTCTTGACTG	GTTTCTTCTA	ATGGCGACAA	AGGGAATAAA	TAGGGTCAAT	AACAAAATAT
	48601	CAGAGCCTTG	TTTGTTCTCA	TAAAAAATAC	TAGCTTCGGA	AACCTTAGCC	TTATAGGTTG
	48661	TCAGGGAGGA	AACCTGGCCTA	CATTGGAAC	AGAATGTCCA	GGAAAAAAT	AGATTTATAG
	48721	GTTAGAAGAG	TAGGAGAGTT	CTATTTTTTT	GAAAGGATAA	CTGTTTTTTA	TGTCTAACCA
	48781	GATGCAAAATG	CATGGGGTTA	GAAGAATAAA	TAGGTTGTAA	GAAATGGCTG	CAACCTGCCA
	48841	GCTAGAGCAG	ACATAAACAG	TTTTGGTAGT	GAAAATGATC	CATTCTCCAT	TCTGTGCTTT
	48901	AGCTTAACCT	TGCATTTAAC	AGTCTGGCCC	AGCTGAGAGG	GTGAATGTTA	AATGCCCTTG
	48961	TGTTATAAAG	TTGAGAACAT	AAGCCTCCAA	GGTTTCAATT	AGCCTGGCAG	AGTCCACAGC
	49021	ATCCCTTGTC	AAAATTAAGC	CCAAGGTGAC	TGTGTCTCCA	GATCAACTTT	ATTGTTGTTG
	49081	TTGAGAAGGA	GTCTCGCTCT	GTCGCCCAGG	CTGGAGTGCA	ATGGCACGGT	CTGGGCTCAC
	49141	TGCAACCTCC	ACCTCTGGGG	TTCAAGTGAT	TCTCTGCCT	CAGCCTCCTG	AGTAGCTGGG
	49201	ATGACAGGTG	CGTGCTGCCA	TGCCCAACTA	ATTTTTTGTA	TTTTAGTAGA	GACGGGGTTT
	49261	CACCAGATCA	ACTCTTATTA	CCAAGCATGA	CATGTTTCAT	TTTGCAAGCC	ACTTTCTAAT
	49321	TCAGCCGAGG	AGTTTTCCCC	GCCCTCTAAA	AATTCTGTCA	AATCAAGAAT	TGGAGGGGAG
	49381	AGCTTCAATG	TTCTTCAATG	TAAAGTAAAT	AACGTTGATT	TTAATTTGCC	AATCAGCTGT
	49441	TGATGGGAAG	TAAGTCCCT	CACATAGTTT	GAAGATCCCC	TAGACAGAGG	ATTCTCCAAG

49501 TGGCAGTTGA CGAGCCCTCT TGTCTGTCT CCCACCTGTC ATCCTCCTTC TGTTCCTGGG  
 49561 CTCATCACAA GCACAGGGGG CTCAGCCCTG CCAGGTGAAG AAGAGGGAGG AGTGAGGAAC  
 49621 AAGCTGCCCA GTTACAGGGC ACTCTGCCCG GCCTGAGGCG GGGGTTGCCT TTCTTTCTTT  
 49681 CTTCTTTTCC TTCTTTTCCT TCCTTTCTCT CTTCTCTTTC CTTCTACCT TCCCTCTTTG  
 49741 TTCCCTCCTT CCAGACCTCC AGCTCTCTGA GTCTCAGCA CAGTCATGTC TCCAGGATCC  
 49801 TGGATCGTTT TCCTAATTGA AAATGATTTG GGCTGGGCGC GGTGGCTCAT GCCTGTAATC  
 49861 CCAGCACTTT GGGAGGCCGA GGCAGGTGGA TCACCTGAGG TCAGGAGTTT GAGACCAACA  
 49921 TGATGAAACC CCATCTCTAC TAAAAATACA AGAATTAGCC TAGCATGGTG GTGTGTTCTT  
 49981 GTAATCACAG TTAATCGGGA GGCTGAGGGA GGAGAATCGC TTGAATCCAG GAGGTGGAGG  
 50041 TTGCAGTGAG CTGAGATCGC ACCACTGTAC TCTAGCCTGG ATGACAGAGT GAGACGAAAG  
 50101 AAAGAAAGAA AGGAAAGAAA AGGAAAGAAA GGAAAGAAAG GGAAGAAAG AGAGAGGGAG  
 50161 AGAGGGAGAA AGGGAGAAAA AGGGAGGAAG GGAGGGAGGG AGGGAGGAAA GAAGGAAGGA  
 50221 AGGCAGGCAG GCAGGCAAGA AAGAAGATTG GAAAGAACCC TGCAAGGTGG GCCTGAGCCT  
 50281 CCCATTGACG TGGTTCCAGA GTTGGCTGCA GAGTTGGCTG CTTGCGGCCA GGGTTGGGGT  
 50341 GTTTGTTCTA ACAGTTCTGG TCAAGGAAGA GACCCTGAGG CTTGACCTT TTTCTCAGCT  
 50401 GGGCCTCAAC AAGGCCCCCA TCGCCATGGT GGCAGGAACT CTTCCAGAA GCCACCTTGG  
 50461 GGAATGTAGG GGGCCTGAGG CAGGCTATTT GGCCAGAGTA CTAGATCTTT AAAGCAATGC  
 50521 TGAGGGCAGG AGAATGGGT AAGTCAGATG GGGCCAGGC TCCTACACTA TGGAGAAGCT  
 50581 CCTGGGTGAT TCTGATGCAC AGTGAAGAGT CAGGGTTAGT GCCAGGAGTG AGACTGGAAG  
 50641 AGGGTGAGG AGCAAAAAGG CCGACCTGCC TCTGAAATGC TCGTTCCAG AAGCTACTTA  
 50701 ATGGCCAGGA CCTTGGTGGG CCCCAGGGCC TGGGCTACT CACCTCTCCT CCAGCAGGAA  
 10 50761 CTCCACCTCC AGCTCTTCCA TGCCCTGGTA GGGGAGGGAG GGAAGGCAGC TGTGAGTGGC  
 50821 GTCTGCACAT CTCTGCTGAC CAGAACCTTT TGCCACCCTG GGAGCCGCTG TGGACCTGCC  
 50881 TCCGTCTCTT CCCCACAGCT CCGTCTGCC ATGGAGTGTG CTTGTGGTCT CCTCTGCCAA  
 50941 TCTCTGTCCC CAAGTCTTTG ACATCAGGAG CATCACTGAC AGCCCAGGGG AGGTCAGGGT  
 51001 GTGGAGCTGA CCCCAGCCCC CAGCATGTAC ATTCTCTGCT CCGAGAGCAT CTCTTTCTG  
 51061 TCACTTCCCT GGCTTCCAC CGCTCTCCTG CCTTCCCAAG GGACTCAGTG CCTGTGGCAA  
 51121 TGGAGCGCTG AGATACGCC TACAGAAAGCC CATTCCTTCC TTCACTGAGT CCTCACACC  
 51181 TCTTGCTGTG TACTGGAGGC AAGCCAGGGC CTCCTGGGC CTTACTTTCC CCACCAGAAT  
 51241 GTAAGCTAGG TGAGGGCAGG GATGCTAGTT CCCAAGGGGA CCCAGTGCCT AGAACCGTGC  
 51301 CGGCCCTCTA CAGGCACTCA AATGCTGGCA GAATGATTGA ATTGACAGCA TCCTGGGGGC  
 51361 TCCAGTCTG ACTGTCTCTG ATCTGCAGAT TCGCCACCT TTACCCTCAT CTGTCTCTG  
 51421 TTCCTGGGGC CATTCCAGCA ATGGGACTCT GAGGGAAGAA AAGGAGGGAA GGAACCAAG  
 51481 ACTTACTACC TCTCGTTGT GTTCTGTGAA CACTGCTAGT TGCCTGAATT ATGGTTTCTC  
 51541 ATTTAAGTCC CACAACAACC TTGTGTGGTA GGTATTATGA TCACCATTTT GTAGATGAGG  
 51601 AAACAGAAAC TTAGGGAGTT ATTTTACGA CATCACATAG TGAGTGACAG TGCTGGGATT  
 51661 TAAACAGACA GCACCATCTT TTCTCAACTT GACTGTGTAC ACAGAGAAGA CCTTGGGGAA  
 51721 TGGCTTTGTG GGAATCAGCA GCAGCTATGA GGTATCATG AAGCCGGAGT CTGAGGGCCT  
 51781 GTACCTGCA CCCACCTGCG GGTGAGTGG AAACCTCACG TGGGTTTTCT TTCGGAACA  
 9 51841 GAGCTTGGTG AGGTCATAGA GAACTGTGAG GAGATGGTCA TCTGGTGTCA CTGTGTCTTC  
 51901 ATCAGACAGA CTCAACTCTA GCACGTTCTA GGGGAGAAGG AAGGATGCCA GTCTGGTTAC  
 51961 CACGGTTGCA TTGACCCCTT CCCAGAACAA GGGGCTTGAC CTCTGTCTTA GCAAGGAGGT  
 52021 AGGTCTGGGG GATGGGAGAC AGCCAGAAAG TCCTGTCTCA GATGTGACAG ACATTTGTAG  
 52081 GGAACGAAT TAAGGAAGGC AAGTGTCTAT TCCTCCTTT CCAACTCCCT GAAGCCTTCT  
 52141 TGGCCTTCTG TTCTCACTTT CCTGCTAGAG AATTTTGCTT CTTTACAAG CTGTTAGAG  
 52201 ACCTGTGAAG TATGGAAGGA TTTAGTTTCA CCTGAGGGGA AGGATGGATG GGAAGAGCA  
 52261 AGCAAGACCC CTGCCAGGAT ATAGAATGTG GTGAGCACTG GGCCCTGGG CACTGAAGAC  
 52321 AGCCCTTTG GTGCTAGAAA ATCCAGATTA GGAATCGAA CATCTAACAC TCCAGGGAGT  
 52381 GATGTGGCAA AGCTTGTAGG GGTCTGATA ATTGGTGGG TATATGAACT CTATCCTTTG  
 52441 ACTCATCAGC CCCTAGTAAG TTCTCATGGA TGGTCCCAA GCAGCTACCT TGTCTCTGG  
 52501 CTGAACCTG AGGTCACTGT TCTCAGGTGA CATGGACTAA TCTCTGAGGC CTCTTGTTC  
 52561 AAGGATCCTG GGGGATGTGG GTTGAAGCTG CCTCTTAAAT ACATCTAGTG GACATAAGGT  
 52621 GGTCTTGGAG GTTCCCAACG GGATGTGATT GTTGACCTTG TAATTACTAG AAACACAGGT  
 52681 ATGCTCCAGA ACCAACCAG CAGCTCCACA ATCCAAAGGA AGCTCAAATT CTGATTTCAA  
 52741 GATGGCAGAC TGAACCCACA TCAAAATTCT CCTCCCTTAC CAAACTCACA TGTATGATAG  
 52801 AAAAGGTATT TTAATAATAA TAAGTATAAA GCCCTTCTTG AAAAGATACG TCACGTAACA  
 52861 AAAATAAATA AGTATCCCTA AAACATGGAG GCAGTAAGGT GGAGCTGTCA AGCCTCAAGC  
 52921 AACTCCAGAG TGGGGCTGAT GCCACTATGA TGAGGCCAAA TTTGTTGTAA GGGATCTCTG  
 52981 CTGAGGCTGT CACAACCTAG GATGCTCAA GGATAATTAT CCCATGGGAG ATGTGAACTC  
 53041 AACTTAGAGT CATCATATAT CTGAAGAAGA CCAATCCCAT GGAACAACTC AACAAATAAA  
 53101 ATACACACCC AAAGAAATGG AAATAATAAA GCAATATGAA AAGAATTAA GAGGGAGAGA  
 53161 ACCCTCTGGC TGGAATTAGA CTGGCCTCAG ACCTAGTGAA TAGAATGGTT CTTTCACTCT  
 53221 GTGTATTCCC CTGCTTGGAG AAGGCTCTGG AAACACAATC TGTTTTGGAT TAAGAGACAC  
 53281 CTAGTTTCAG CAACAATTAG GACAAGAGAG ACCAGGAGAA AAAGCTAGAA AGAGAATAG  
 53341 ACTCAGAGAG GGGCAAACAT GAGACAGGCA GAAGGAAGTA GTGGGTGTGA GTGGGAAGAC  
 53401 AGGTTGTAG GACAAGAAGT AGGCAGAAAG GAAACAAAG AGGCATAAGA ACACAGAGGC  
 53461 AAAGAGAAGG ACAAGTGGGC AAGAGGGGAG TAAATGCAG ACAGGGAAGA CAGGTGCAGT  
 53521 CCTCCATGTC TCACCTTAC TCGGCTCTGG ATCTGGAAGT TGAAGCTTTC ATTCCACTCT  
 8 53581 GGATTTGGGC AGTTGGAGAT GGTCTTGTG CTCAGCTTCT TCTGAGAGGC GGTGGGCAGC  
 53641 CAGAGGCTCA CAAACAGTC TGTCTGGCTC ACTGTCCAAG AAAGATAAGT GGAACCAATC  
 53701 CAGGTTACAG GAGAACATGT GGAGTCAAAA GGAACCTTCA GGCAAGCCTC TTCCAGAGCC



53761 AGGCCCCATC TCAGCTCCAG TTCCTCTCCT CATTTCATGGG TCTAACTTTC CAGAGACTCA  
 53821 GATCCCCAAG AAGTCACAGA TTGCAGGCTG TGGGTCTCGC AGAGGGCGAC CCAATCCCAC  
 53881 TGGCCATTCT CATAGACACA GATGGCCTCC GGCTCTTATT AGAAGCACCC TTGTCCCTGA  
 53941 GCAAACACCT GCTGGTTTGG GACAAAGTAT GGTCACACAG TTTTCCCTCT GCCAAGGTAA  
 54001 GATCTGGGAG ACCAGTCCCC AAGGGTCCGG ATTACTTCTT GTCTTAGCCC AGGGGACCAG  
 54061 GAAAACCAAA GTAGGGCCAA GCGCTCCCTC TGGGCTGAGG AGAAGTGAGG ATGAGAAGGA  
 54121 ACACCTGTGCT GTGAGTCGGC CAAGAGGAAG GGCCCCAGTG GGCAAGGACT TGAGTGAAC  
 54181 TAGCTATTCA GCAAGATCTA GGAGCAGGAA CCCCCTATCT CATATGGCTT AATGCCTCCA  
 54241 CTGCATCCAG CCAGGCTCTT CTTACCTATC CTGATCTGCA GCATCAGGTT ATGGCCCCAAG  
 54301 AAACCTTCCC CTCTGTGGCC ACATACTTCT TTTTCATTCT AAATGAACAA CCAAGTGTG  
 54361 ATGTTTGTGG ATATGCTCTG GAGAAGGGAC AGAGGCTTTG ACCCCAACAA GCCCTCCCAA  
 54421 AATGACAGAA GGTCTCCTAA GTCTTCCACC TGGAGCAACA GAAGGGATAG GGTGAAAATA  
 54481 GTAGCCACAGA CTGGCTCCCT GCTCTAGTCT ACCTGGGAGG AGGGAGGCAG AGGGGCCATG  
 54541 ACTGCCCTTG TCACTCTCCT GGCAGTTGGT GAGGGTGAGG CCCAGGGTAC ACACCAGATA  
 54601 CATAGGAAAT GGTGAAAGCC AGCAGTTTCC AGTGTACAG AGTCCTCTCG ATGGTAAAGC  
 54661 TGGCAGGAGG ATGGGAGTTC TACACTGGTA GGACTGATTC AAAGTAGTTT TTAATTATTA  
 54721 ATAGTCACTG ATGAGAGTGG AGCCAGGGCT TTCTTCGCCC TGGGTTGTCA CCACACTGAA  
 54781 AGGTTTATCA GAATCAACAG CCCTGATCTA AATCACAGGT GTTCCTAGCT TCTGATTCA  
 54841 TCCCTGAAGT CTGGTTAATC TTTGAGGCCT CATCATCCAG TGAATTGGAA TTCACAGCCC  
 54901 TCCTCTCTATT GAAGAAAGGT GAGATGACAC TCTCTTTTAG ATGTCATGAT AATCACTTAA  
 54961 TCCAGACAAA AAGAATGATT GAAACCCAAC AGAATCAAAA GAAAATGATG ACCTCATCAA  
 55021 CCTACTCTTC CTTCTATGCT TCAACCTGCT CTTCTTTATC ATCACTCAGG TGTCAGGGC  
 55081 AGAAATCAGT GTCATCCTTG ACTCTCACAT TCAGTTGATT TAACCTTTG TCTCTCTCAG  
 55141 GACCCCAAC TTCTCATCTG CCCTGCCACT CCCTGGTGT AGACCACCAT AATCTCTGAG  
 55201 GCCACTTCCC AGATAGGCTT CCCATCTCCT GTCTTGCCCC ACTCTAATCT TTTCTCCAAA  
 55261 GAGAGGCCAC AAGGAGCTTT TTAGATATGA CCATGCCATT CCCAGATTG AGATCCTCCA  
 55321 GGGGCTCCCT TCTGCCCTCA GGAAGGAGTC CCAGCCACTC AAAGTGGCTG ACCCTGCTCA  
 55381 CTCTCCCAAG CTCTCCCAAG TCTCTCCACC CTCCCTAATA CCCACCCCT GCCCTGTGTG  
 55441 TGCGAGCATG TGTATGTGTG CACACATGTA CATTACACA CATTTACACA CATAAGCATT  
 55501 CACTCAAAC TACACACACG CACTCACCGT TTAGACTCCA GCCAAACACA GTCACTTGTA  
 55561 TCCAAAACCT CTTTGCACA TGTAGCTCCC TCTGCCAGA ACACCTCTCC ACTTTACCC  
 55621 ACCCTCTCTC TTGCTAGCCC CTTGGGACTT CAGTAGGAAC CTCCTCTGGA AGCTTTCCCT  
 55681 GAGCACCCAA GTCTGGACCA GGTACCCAG AGCCAGAGT CCTCTGCTCC TCTCCTGGG  
 55741 AGCACCCCTA TGCTGCTCTG CAACAACAGG TGGACTCATC TGTCTCCATC AACTGTGGC  
 55801 CACAGGATAC AGGTACAGAT ACATCACTGC AGATACAGGT CAGAGGTGAG TGTGTCTCAT  
 55861 TGCTCTGCTG ATAGGGAGAC TCACCAGGAA CCTGCTCAGC CCTCTTGAAC CCTTCCCTTG  
 55921 GCTTTTCCAG AACACTGGCC CCCCACCCCC GGGCTGGGAC CACACCTTC AGGTCTGTCT  
 55981 CCGGTACGCT GGGAGCCTGC AGGCAGGGTG CCAGCTCTAT ACCTACTGGT TCTGTAATGC  
 56041 ACTTGGCGCG CATGATTTGA TGGACCCTGT AAGAGCCCTT TCCTCTAAA ATTGTTTTC  
 56101 TGTCAATCGT CCATGCCTAA AATTCTCCGA TAAAGCCAT TAAATAAAAA CTCAGTATT  
 56161 TTGGCCTAAG GCTTAAATGT CTAGTTAAGA TAAAGAAGTA TTCTTGGCAG AGACAAGATT  
 56221 GATCAGACAC CCACATTCCA AGCATTAGTT TTTCACTTCT TTTGCAGAGA GCTGCTTCCC  
 56281 CTGTCCACCC CCAGCAATGA ATGGTTAGGC AACGGTGACC TGGTCTGAG AATTTATGAG  
 56341 ACTATGGGGA TTTCTAGGC ACAGATAAGT TCTAATCCTT GTTCCTAGTG GAGTGACAGA  
 56401 GGGCTTGGTC ATCTGAATGA CATGTCTCCT TTCTTGAAT GTGGGTGGCA GAGATGACTC  
 56461 AATCATCTCT ATAGATGTTT AGAAGTTTGT CTCTCTAATA CCTCTGGAGA CACCTGTCTG  
 56521 ATATCCCTTG GGGTGTCTGT TTAGTTTATA ATTTCTCAG AGAAACAAAA ACTCCTGACA  
 56581 ACTGGGAGGT TTCTAGGCCT GATTGTCTCT CAGTCTGGC CCAAGGCTCT TTCAGTTATC  
 56641 GAGGAGTTCG CCTGAATCT TGAGTATTCA CTGCCATTG GCCTCTTGA AACTTTAGGA  
 56701 GACTCAGGCA CAAGTTACAA GACTAAATGG AAGTGCTGTT TATTTTGGAT TCAAACATAGA  
 56761 AGGAAAAAAT TTGGGACAAG GGTAAACCAG GACCTAAAAG ACATCGGGAT AGTCTTAATC  
 56821 TAACCCAGGC AAGGTAGCAA CACAGCCAG CCTGCAGCAA CTGTGGAATC CAGACCAGAC  
 56881 ATGACTGGGG TTGGCACCCC GAGATGCCAC ATGTGCCCTC CGCCTTGTTC TTCGGGGAAA  
 56941 TGAGGAAAAA GTAAGCAACA GGGGGGTTTC CAAAGAGGGC CAGGTATCGC CCAAAGGCAG  
 57001 CCTGTCTGCC CAAGGGGCAT GTAGGGAAGG GAGGAGGCTT GTAGAGCCAG ACTACAGCCT  
 57061 GTTCTCTTAG CTGCCTGCTC TTCTTCACAC CCCGTACCTT GCTGAGTGAC AGGTGTGGCA  
 57121 GGAAGAAGAC AGAAGAAAGG ACCCTGGACA GTGGGAAGTC TGAGCATGCA GGTGCAGGCA  
 57181 TGCATGGGGG CATGCAGCCA GGTGACTGGA GAGGCTGGGA CAGACACACC TCAAAAGGAG  
 57241 TTGGAGGGCT GGGCATAGTG GTTCACGCCT GTAATCCCAG CATTTTGGGA GACTGAGACA  
 57301 GAGGATCACT TGAGCCACGA AGTTTCGAAC CAGCTTGGGC AACATGGTGA AACCTGTCT  
 57361 CCACAAATAA TTTTTTTTTT TTAAGTTAG CCAGGTGTGG TGGCCCATGC CTGTAGTCCC  
 57421 AGCTACTTTG CAGGCTGAGA TGGGAGGATC ACCTGAGCCC AGGAGGTGCA GGCTGAGTG  
 57481 AGATGAGATT GAGCCACTGC ACTCCAGCCT GGGCAACAGA GTGAGACTCT GTCTCAAAAT  
 57541 AAAATAAAAA TAAAAATAA AATAAAAAAGC AGAAGGGGTT GGAGGAAGGT AAGCACTGCT  
 57601 GTGGATATGT GGGAGCTGGG TTGTCCCAAG GATGAGGGTG AACTCTGCCC AAGAGTGGGA  
 57661 AAGTGTCTGA GTGACCACCG TGGCATGGAC AAGAAAGGCA AGAGAGGTGG CCACGTCCCT  
 57721 GACCCAGAGG AACTCCCTG GATCCTCAA TCATTACCG ATAGGCTGGG AAAGACCCCA  
 57781 AGAAGCTTGA GGAAGGTCC CTGGCCCTTT GGTGAGCCAT TTGGCAGACA CAGACCCAAA  
 57841 CTGGAGGCTA CTGGAGCCAC TCCTGTCTGG GCAGGCTCTG GACACATTGC CCTCCAGAGG  
 57901 CTCTGTCTCT GGAATCCTGC TTGGTCCCTT TGCTGTCTCT GGGGCTCAGG GCCCTCCAC  
 57961 ACTTATCCAT TCACACTCAG CCCTGCATTG GAGTGACAGC TTCCTACAAG TTCCAGACT

58021 CTCCGCACCT GCACTCTAGC ACAAGCTGCA TACAAGGCCT CAGGAAGTCA TGCAATCTCT  
 58081 TCCTCAAGAC CCCAGGTACC CCAACACTTT GAAGGCATCC AATCCCTCCT CTGCCTCACT  
 58141 CCTTGGTTCT CCATACCTGG CTTGTCTCCT CTCTTGCCAC ACTCTCCACA CCTGGGACCA  
 58201 GACCCTACTT CCCTCTTAGG CAATCATTAA CCTCTTGTG CAAAATGGGC CTCCAGGGCC  
 58261 TCTCCTCAAG TTGCTGGGAG CATCGGGTAG TGTGTGGATG TCTGGCTACA GGTCAGCCAG  
 58321 GCAGGTGCTC AGGGGGTGTG GCTCCCTCC CTCAGGACAG CCTCTCCCA AGATCGGGCT  
 58381 CCCAAGGCCT CTGGATACTC AGGTCCTGCC CCTCTCAGAC CCCAGGCTCT CTGTGCAGCC  
 58441 TTTCTTGGCC ACTCCCAGCT CTCATTGTTC TCCTCTGCCT CTAAACTCCC TGGTGACAAG  
 58501 ATGGGAACGT TGTATTCTA TAACTGCTTT GGACATAGAT AGCTCCTCTT GTTAAATGT  
 58561 ACACTCAATC TCAGGGCTCT GTGTCTTCTC TGCCCTCTC AGGCCCTCCC TAGCCTTTCC  
 58621 GGATCCCTGT GGCGCTTGAG TCATTCTGT GTGGTTACCG GTGAGCAGG ATGCTCCCA  
 58681 CCTGTGGGGC CTGGATATAC AGCTTTTTC CCCTTGGGG CCAACTTCCC CTCTGGATGG  
 58741 GGCCAGGGGT GGGCAGTGTG GGAGGGGCTT TCAGTCTCCT TGATCTCTTA AAATGCCACC  
 58801 CAAAGACAGG GGTGGGGAGG ATTCTGCCCA CTGAGATGCC TCCATTGGCC CCCGTGGCCT  
 58861 GGAATGCAG CAGGCTCTGG GATGAGGGGG GAGGTAGGG CTGCCAGGAA ACAGCACTGC  
 58921 ACAAGACAGT CTGCTAAGTC AAGCCACATC CATGGGCTGA ACTGTAGAAC TGGGTGCTGT  
 58981 GGAATCTGGC ATGGTCCCCA GAGCCAGTT TCAGATCTTT GTCTTGAGAA GGGGCTCTGG  
 59041 GAGTCACCTG GCAGCTGGCT CCACGATGCC TAAGTGCTC TCCTTCCACT TCACCCCAAC  
 59101 AGGCCCTAAC CGAAAGGAGG AGCAAGGAGT GCTCCGTGTG GGCAAAGCGG GTGCTCTTGG  
 59161 AAATCGATT CAGGAAGAGA GAGTGGGGAG GAGACAGAG GCGAACAGAG CTCTTGGGTG  
 59221 CAGGCACAGA CTGCGGGGCA GGTAGCGCCG AGGGAACAGC ATGAACAAGG GGCCTGGAAG  
 59281 AGTGACAGAA ACCCCAAAGT CTGAAATACA ACTTGAGTAT AGAACACACA TTCGTGCACA  
 59341 TACATGCACA CACATACACG TATACACCAT GCACGTATAC GCGTCCGCGC GCACACACAC  
 59401 ACACGGTACC ACCTGACGAT TTCTCCAGCC AGGGCCTCTT CCATCCAGAT CGGTAAGCA  
 59461 GAGGAGAAGG AGAGAAGTGA GGTGGATACT CACGCATATC AGCCTGCCGG ACATTTTTC  
 6 59521 TCCGGATGAC CCTCACTGTC AACAGGTGGC ATGGAGACAG CCCCTCCTGT GGAAGGAGAG  
 59581 GACAGAGGAC TCAGTCTTCA AGCATTACCC CAAGGCCATT GCCCAGGGG AGACTTTCCC  
 59641 GCTATGCTCT CTCAGGCCGG CCCAGTTGCA TTACAACAAC CCCTAGGCTG GTTTTTTTT  
 59701 TTTTTTTTTT TTTTTTTTAC AGAGTTTAC TCTTGTGTC CAGGCTAGAG TGCCATGGCG  
 59761 CGATCTCGGC TCACTGCAAC CTCTGCCTCC CAGGTTCTAG TGATTCTCCT GCCTCAGCCT  
 59821 CCCATGTAGC TGGGGTTACA GGTGTCTGCC ACATACCCAG CTAAATTTT GTAGTTTTC  
 59881 TTGAGACAGG GTTTCACCAC GTTGTCCAGG CTGGTCTTGA ACTCCTTATG TCAGGTGATC  
 59941 CACCCACCTC GGCCTCCCAA AGTGCTGGGA TTACAGGCAC GAGCCACCAC GCCTGGCTCC  
 60001 CCATGCTGTT TTGTCTTGT TTGTGGGGT TTCTGTCTGC TCTTCAGCCT GGAAGGTCTT  
 60061 GGAAGCTCCA GAGAGCCTGG GACCTGACT GGTTTTTCAT CTTGTACCT CCAGTACCTA  
 60121 TAGCATGCTC ATATCTAGAA GGTGCTAATA CATGTTCAAT AAGTGAGTGA ATGAATGTTT  
 60181 CAGCCCCTAC CACAGAATGA ACGCCCATGG CAAGGGAGTC CTGAGGTGT CTCCCTCTCC  
 60241 TTTCTTCACT GAATAGCTGC TTATTGGCAC TCTCTCCTA CTGTCCCACA TGCAGATATT  
 60301 TGCTCTTGCC CAGGCAAGAT GGCAGCCTGC TTGGGAGGCA GGCACAGTGT GGCTGCGGGA  
 60361 AGGGGGGTGA CAGACACTGA TCACGCACGG TGACATCTAC TCACTGGGAG ACCATGTGCC  
 60421 AGGGGCTCAC CCCCCTGGT ACTTACTTAG CTCCAGCCCA GGCCTTTGCA TCTCCCTGTC  
 60481 CCTTTGGGCC TCTTCCCAG GGTATCTCCC CCAGGTCCTT CTGGTCACTC TATTTTCCCT  
 60541 TCCCAGGAAA CAGTCTGCCT CTCTCCACT CACTGTTTCT CGCAAGAGGT CTTGTGGAG  
 60601 TCTTGCTGG TCTGCGATGA CCTCTCTTC CTGATGGCAA AAGCTCAAGT AGTCCTTGGT  
 60661 TGTATGACTG GCTCTGTGTA GATGCCCTG CGGGTCATGA TAACCCAGG GGACAGAGC  
 60721 CCTGAGGTGT CCCCCTTCCC TAGCACCTTG TCCTAACAA TGGGGTGTGG GAACACTCCC  
 60781 AGCCCTGCTC ACGCGGTGCT GATGGCTCCT CCTCCGTCC TGCCCTCTT CTAGTTCTGA  
 60841 TCACCTCCTG GGGGCTCTGC CATCATCTCA AGCCAGGTC CTCAAGGAAC ACCCAGAGCC  
 60901 CTCCCCTGGG TGCACTGTCT GGGCAGTGG CTGTGGTTCC TTAAGGCCAT ACGTCCCTCC  
 60961 CGAGCTGCCC CTGCACTCTG AGTCTCGCAG CAGCCACTAC TCTACTTAT CCCTTTTGA  
 61021 CCCATTACTT CATTACTTTC TGCCCTCCAG AGGCCCTGAA ACCCCCTCTT CCAGCCGCAC  
 61081 CAGCTTCTCT GGACCTTACA GCAACAGAGT GTCGCCGGAT CATCCTTCTT GCCTACCGAG  
 61141 GGACCCGAGC GCTCTCCCAG TGCCCTCTG CCTGTTGCC TTGGCCCCAG GCGTTAAGCC  
 61201 TGTGGGAGCC GCTCCCGTTT CACCCTCTC CTGGCCAGCG TCTCTTCCAT GGGCTGTTT  
 61261 CAGAGCTCAG GGTCTGTGTG CTCGCTCGTA GGGCTCCCTG CCTCTGTTGC TCTGAGAGTG  
 61321 CCAGCCCTCT GATTGTTCTT CTGAATGAAT GAACCATCTT CTAGAGTATT ATTTTTCGAGC  
 61381 ACCTACTGTG TTCCAGGCC TATGCTAGAT GCTGGGGGGC CAGGGGTGAA TACGAGGCAG  
 61441 AGACAGAGGT TAAAGGCAAC AGCGGCCAGT GGCTTTGATG GGGAGAGCAG GTCAAGCTGG  
 61501 CTACAGGTTG TTGCTTTCCA GGCTTGGGGC CTCTGGCTGG CTCCAGAAGC TTCCCGGCCA  
 61561 TCAGTCATGC TCCTGGACTC TGGAAACCTC TTCTGGTCCC TAGGGGCCCA AGGCTGCAGA  
 61621 GGTGAGGCGC GTTCTACTC TGAGCCCGCA CACCCGGGAA GGGGCAGCCT CGCGCCGGTT  
 61681 CTCTTTCAGT GTCTGTGTT TGTCTTGGT TCATGCCCTT CCGTTTGAA CTCATGACTT  
 61741 GAAGCCCTT CGCTGAGAGT TTTACCCACG CTTAGAAGAG TTTTGTGTGT CGTCCATTCC  
 61801 CAGCTGAGC TCACGGTAAA TGCTTACCAA CCGAATGTCT GTATTTATAG AAGCGTCTGA  
 61861 ACCGGCACCC TCGACATGCG ACCCTGAATG TTCTACTCCG TGTTGCCAGA ATCACTTTAT  
 61921 ACCTTTTAA GAGGATTCAT TTCTTCTCTC TTTAAATACT ATTGGGATTG GGGGAGGAG  
 61981 AAACAAGATA CGAAGAAAA AAAAGTTTCT ATAATCCAC CACTCAGAGA CCATCTTTT  
 62041 ATCTTTACTT TTTTTTTTTT TTTTTTTTTT GAGACAGGGT CTCACCTGTG TGCCTTGGCT  
 62101 AGGGAATCTT TTGATGCTGG TCCTTTCAGA CTTTCTGGG GCTTAATTCT TTTTCTTTG  
 62161 GAGACGGAGC CTCACCTGT CATATAGGCT GGAATGCAGT GCGGGGATAT CTGGCTCAC  
 62221 TGCAACCTCT GCCTCCAGG TTCAAGCAAT TCTCCCTCT CAGTCTCCTG AGTAGCTGGA

62281 CTACAGGTGC TCACGACCAC GCCCGGCTAA TTTTGTATT TTTAGTAGAG ACGCAGTTTC  
 62341 ACCGTGTTGG CCAGGCTGGT CTCAAACCTCC TGGTCTCAGG TGATTCTGCC TACCTCAGCC  
 62401 TCCCAAAGTG CTGGGTTTAT AGGTTTGAGC CACTGTGCCT GGCCCTGGG GCTTAATTCT  
 62461 GATCAGCAAC AGAATGCTGA TTCTGGGAAG GTCACGTTAT TTAGCCCCCT TTAAGGTCTC  
 62521 TTCTCTCTCT CTGAACAGCT GCCTTAAAC CAGGACAAAA GTCAAACCTCA CAGCCAATTA  
 62581 TAGGTCTGGA GATTATATTT GTGAGATCAA CCTGAAATA TTGAACATCT CATGGTCAAT  
 62641 AACTTGCTCT AACAGCATTG TGTGGGAGC CGGAGACTTG GGCCAAAAGC CATGCACACT  
 62701 GCTTCTCTTC TGATGCTGGA GACACGGCTG TGGCCTCTGG CCAGACCTCC AAGGTGCACT  
 62761 CCTCTTTTCT CACTCTGCCC GAGTGCAGCA TGACCCTGG GCTGGCCTGG TGAGGAGGCA  
 62821 GGGTCAACTT TCACCAGGCA AGCAGTTTGG GCAAAACCTG TACCCACCGG CCCACTTTCT  
 62881 CAGAGGGCAA CTAGAGTCCA AGCTGTGGAG GCTTCAGCCC AGGCAGGAGT CCGTCCTGAG  
 62941 GGATATTCCA CAGTGTTCCT GGGCTCTGAA CCCAGGTCCG GTGGTCTCT CGGCCTCACA  
 63001 ATCCTTCTCC TGCCTCCAGG AGGGGGTGCC TGTGGCCCAG CCTGCCTCT TCAATATCTA  
 63061 CTCTGGCAGG GAAGGCGCAG CCACACTCTC CCTTGGGACC TCACTCCAGG TGTCTGCTT  
 63121 TCCAGGACAA GTTCACTCCC ATCATCTAAC CTCAGTGCCT TCCAATGCAG GTGACCATTT  
 63181 CTTCCTGCTC TAACCTCAGCA ACAATAACAG CTAGCATCTG CTGCATGCCA GCCATGCGCA  
 63241 GGTGCCATGT ACAGGTCTCT TGTATATTAT CTGTCTTAAT CCTTACAACC GCCCCACGAA  
 63301 GCAGTTATTA TTAGCTTCGT CTTTATAGAT GGAGAGTCAG AGGCTAAGTC CTTTGACACA  
 63361 GGGCTCACAG CTAGAATCAG AGACATCCAA TGCTACTTCC CAAGCCTGTG TTTCTCCAC  
 63421 TGAGAATCTC AGTTTACTTG GCAAAACAGGA AGACAAATGT TGCTTGCCTT CATCTCCTCT  
 63481 TCCTTGCTTT TCTTCTCTTT AAAACAGTTT TTCTGACAGA TTCTTCTTTT CACTCAGCGC  
 63541 TCCCCCAAAG CCAAAAAATA ACCACCAAAC TCTGACAAAT CTCTATGGC CTGCACCTTT  
 63601 ACTGGCTGTT TCTACCAACA CATTCTAGGA TATCCCTCTC TTTTAAACA TCTAGAGCAC  
 63661 ATCGAGGACA TAAGGCTCTT TGGAGAAGTT CCATTTTCC AGTAGCCACA GATTGTCTCA  
 63721 CTTGAGCCTT AGATCACACC CTCCAGCTCT CGAATTTTCA GGAGAGACCC AGTTTTCAGG  
 63781 AGACAACCAC ATTTTGCCAG CACAGAATTG GGACTGGGAA GGGGGTTTGA GAGACATCCC  
 63841 AGCCGGATCA TGCTGGGAGG AAGGTGGCAG GAATATGGGA CTAGCTCTG GGGACCAGAT  
 63901 GCATTCTCAG AACCAAGTCT ATGAGAAATG GCCTTCTTGG GTTGCAATGA ATGTACCTAT  
 63961 TTTGCTTCTG ATATAATCTG TTTGATCAGT GTTGCTAGAG ATGAGCAAAA ATTCATGGCC  
 64021 TGGAGGTAAT TAATCTACAG TGGAAAGATG CAATCGAGCA AAGCACCGTG TCTCTGATGT  
 64081 GCGTGTGGGA ACCAGAAACC AGGGAGGCTG TGGAAACCAA GTTTACAGGG TTGCTTATGG  
 64141 TGTGGTGAGG GCCAAAGCTT TAAGGACTCA GATAAGCTCA TCCCTAGAAC AGGATGCCCT  
 64201 GGGTGGGATT GGGTATCAGG GTGAAGAGTC AAAGTGTTA AGAACAGACA TGATCAGGCA  
 64261 CCATGCCAGC TGGGCTGCAG CCCTGGAGTA GGGACAGGCA TTCATCCTTC ATGGGGGCAT  
 64321 TGGGAGGGCC CTGGGGTAAG GACCAGGGAG CACTTGAAGA CTGGAGCTC AAGGCAGGAA  
 64381 CTGTTTGGCA AACAGTACAG AGGTATATTT AGCCACTTTG CCAATCAGTA CAGCTGGGCC  
 64441 ATCGAGGACA GGTGGACAG CAGTTTCTG GGAAGCCAGG GGAGAGAGAG GGCCCATCTG  
 64501 CTCAGGCTGC CGACACACCA TCCCAGCCCC CTCCCCAGCC TGGATTCTGG GGTGTATCTT  
 64561 GCCAGGACTC CACTGCCCCC TGGGGACCTA CTGGGGACTG GGGGACCTTG CCAGCTCATG  
 64621 TGGAAAAGAG GCAGCCCCCTT GAATCTCAGG ATTGGCTATG ATGCAATCCA ATGCTGCCCC  
 64681 TTCCCATCTC TACTTTTATA GATATGACTT TTGGCAGGGG TGAGGGTGGG GAACAAATG  
 64741 TATGACTCTG CTCTGGCAGG CCCAGCTGGG CTGTGCTCTA GCTGCTGTGC ATATGCTCTT  
 64801 ATCTAACCTT CCAACCACCG CCAAGGAAGG CATCATATC CTGCTTTAC AGAGGGGAAG  
 64861 GCCGCGTTTA GAGAGATTAC ATCTCAAACA GGCTTACACT GCCCAGGAGT GGCAGAGCCA  
 64921 GCGACTTGGT GCAGGTCTTC AGGGCGCATG CACTCCATCC ACCTGGCTCT CACTATGAAT  
 64981 GTACAGCACC AGCTAGCAGC TTGCTCCAGG CCATGTGCTG GACTAGTCAG TAGCAGAGAC  
 65041 AGGAGGAAAA GCCTCAGCCT CCGCATTGGA GCTCTCCCTA TACCCCAAGA ATCTAGCCGT  
 65101 TACCACAAAT CAGTCAAAAG AATACTTAAT TACCAAAAGG CTCCCAGGGT CAGGCTCCTG  
 5 65161 GGTGGGAGTG GGGTCTGGGG AAGGGTCTGA CAACCACATT TCAGAGGTGG TGGATCCGAT  
 65221 GCCCATCGCT GGGCCAGACA TACCTCAGC AAACCATCCC ATGAAGTCTG GGGAGCCTGG  
 65281 AGAGAGGGGA AGATGGCACC TGGGGAGCCT GCAGGAACCT CCTCCCCAGG GCTTCTCAGG  
 65341 TCAGGCCAGG CAGGTGATGC CACCTCCTCC TCCCTAGAGA GCTGGGTGTA CCCTTCAGTT  
 65401 CCATGCAACA CCTGTGTCAG AGCTTGGACT AGCCAGGTGC TGTCTTCCAG AACATGTGGG  
 65461 CCTCCTTTA AGGAGACCAT TGGGAGGGGA CATGATGGGA CCTAATACA TGTCCAGCC  
 65521 ACTCTGAGCA ACATCCCAGA CTAAGAGTCA CCTGGGAGGA GCTACCATGG CTGTTTCTCT  
 65581 TTTTATCACC AGGCTAGTTG AGTGGTCAGG GCAAGAAATC CCGGCACCAA CATGGACCCT  
 65641 CCAGGCCTGA CAGGGGCGA TGCAGCTCAT GGAGAGGCTC CTGAAGCCAC ATTTGGGGTC  
 65701 CCGGCTCTTT AACTGCTTCA CTGCCCCCTT TGAGGTAAAC TCTGTACCT CTGTGGAAGA  
 65761 CCTCTCCTGG CCAAGCTCT TGGAGGCTCA AAAAGATTGG CAGGAACAAT AGGTGGCCAG  
 65821 ACGCCTTTC AGCCACCCAA GTGCTGTGGG GAAACCAGG CACACCCAG GAGGAGTGT  
 65881 CTACCTGGCT GGGCAGCAGG GGCATTGGCC ACCACCCACC ACGGAGCTCC CTGACTTGGG  
 65941 CCTCTAGGAG ACTGAACAGG GTCTTCATCT CCCCCACCAA GGGGCAAGT ACTTTCAGGG  
 66001 CATCACCAT GTCAATCAAT GATTTAACTT GAGCCGGGGA GGTGGCCCTG GGAAAGGTAA  
 66061 TGAGGAAGAA GTGCTGCCTT CGAGGAACCA ACTAATGATT CTTCCAAGG TGAGCCTACA  
 66121 GCAGATTGAA GCATTGAGGT CATCTCCCTA TTGCCTAGAT GGGGAAACTG AGGCAGAGTT  
 66181 GGTCTAGAGC CCTGGCAACC TGATCCATGA ACCCAGAGGT CTCCCTGTGA GGCTGTGAGG  
 66241 AGTTGGATGG ACCCATGAGC CTTGACTGTA AACCAGGTG GGCAGCTTTG TGGCTGATGG  
 66301 CAACAGATTG ACAATGCCAC CAATTCTCAC CATACTACT ACCATACTCT TGCACAGCAT  
 66361 AGAAGCTTAG GATGTTTGCA GGGTTAAATC CTGCTTGCCA TTAGGGCAGG CCAATGTGG  
 66421 ATCCCCCTGG CATACTCTCT ATCTTGGGTG CTAGCCTGGC CCCAGGATGG CCCCACGCCA  
 66481 CTTCAATGATG CCTTGAAAT CTTGAGGGTA CTCTCAGCAG GGCCACCCAA GCAGGAGCCC

66541 TGGCCCCAAC TGGGATAAAC CACCTCCCC AGGTTTAAAG ATGAGGCTGC CTGAGCCCAG  
 66601 GTCCTTGGGA TGTCCAATGC CTACCAAGCCC TCCCTGAGGC CCGTCTCCTG CTCTCAGCTC  
 66661 CCTGCCCCGTG CCCCTCACTT CCTGGCCATG CCTCCCTGGC CCTCAGCCTC TGGCCCCAGC  
 66721 CACTGAGCAC CCCTCCCTCA CCTTCAAGGC ACAGGTGGCC GAGGGCCTGC TGTCTTCTGC  
 66781 TTGGCATGGG TCTGGCAGGG AGTGGGAGAC CCTGGCGTGC ATCTGCCTCA GGAGCATCAC  
 66841 GGCAGCTCTC CAGCTGGGTG GCAAGATCAG GAACTTATAC TCTGACTCAG AGCTCCACCT  
 66901 TCCTGGCCAC AGCTTGGCAG TAGCTGGGAT TGCTCTGGGA GTGGTTTGAG TGGGTAGACA  
 66961 GATAGGGTGC ATGGGCCTAG ATCTCTTCAT ATCAAAGAAA TCACAGGTCT TCAGGTCATA  
 67021 GAAATACAGG GACCCGAGG GCTCTGGTCT TTGGGGTCAG CAGTCTGTAG CCCCCAGCCA  
 67081 CTCAGAAATGA CCAGGCCTTC AAGGGCAGAG TGGGTGGCAA GAAGAGGACA CTGGACCCAC  
 67141 ATGACTGTAG GACTAGCCCG CCATAATTG GGAAGCTCT ATCTCTTCAG CTTATAGAGG  
 67201 TCATGGCTCA TAGATGCCCC TCTAGGAGGA TGCGACCGTC GCTCCCAACT CCACTTCAGT  
 67261 TTCATGGCTT GGTACCCAG GAGCTGTGGC CAGGGACCTT CATGGGGTCC CTGGCCACAA  
 67321 CTCCTCAAGA GACCAGGGTC AGTGCTCCCA AAGCTAGAGG GTGAAGAAAG AAAATAATTG  
 67381 CCCAATACTT CGTGAGAAAG GGTCTGGGGC AACCATCCCC TGTCCTCCCT CCTGGGAGGC  
 67441 CAGGTGGAGA GAGAGCGTGT CAGCCCTCG CTTGGTCTCA GGACAGCTTG GAGGATGCAC  
 67501 CCGCCCCGCC CAGATGATTC CTATCTCTAG CCTGGTGGCT CACTGAGGGG CCTGACAAGG  
 67561 GCAGCTTTGC CTGCTCTGCT GCCTGACTCT GTACACTGGG AGTTTCCTGC CTGAGCTATT  
 67621 GCCTGGGGCC CCACAACCGG CTTTTCTGGC CCATAGATAG AAGTGCTTTT TTGAAATTGA  
 67681 GGATGAGATC TGAACATGGG ACAGACTGGG CCTGAATGCA GGGATGGAAC CTGGTTGGAG  
 67741 GTACCTGTGG GAACTCTTCA CGTGCTGGC CTTCCAGAAA GTGGAGGGGC CTTAGCCTGC  
 67801 TGCTCTCCTG GCTTCCTTCC TTGCCCTGCC CTCCCTCCT CTGCCAAGGA TGGGGACTCA  
 67861 GCACCACATC CCTCAGTAAA CTCCACAAAT TCAATTTTTT TTTTAGTTC ATGTACCTAG  
 67921 TAAGCATCTG CTGTATACAT CCTACTGAGA GAGGAAATCC TAAGCTTGAA TTGTCTGACA  
 67981 GTCTGGCTGG AGAGCCAGAT GCATCCAACC CCTCAGAGGC TCAGGGATCC AAGTTTCCCT  
 68041 GTGGTGAGAG GGGGCATCTT AGAAGGATCG ATAACAATCC TTCCTTCACA TTTGTGAAGA  
 68101 TCCATGGTTT ACAGCAGTAG TCCCCATCCT TTTTAGTGC CAGGGACTGG TTTACCCGAG  
 68161 GTTAATTTTT CCGCGGGTGG GCGGAGATG GTTTTGGAT GATACAAGTG CCTTACATT  
 68221 ATTGTGCACT TTATTCTAT TATTATCATA TTGTCATATA TAATGAAATA ATTATACAAT  
 68281 TCACCATAAT GTAGAATCAG TGGGAACCTT GAGCTTGTTT TCCTGCAACT AGATGGTCCC  
 68341 ATCTGGGGGT TATGGGAGAG AGTGACAGAT CATCAGGCAT TAGATTTTCA TAAGGAATGT  
 68401 GCAACCTAGA TCCCTCGAGT ATACAGTTAA CAATAGGGT TTGTCTCCTA TGAGAATCTA  
 68461 ATGCCACTGC TGATCTGACA GGAGGTGGGT TTCAGGTGGT AATACAAGTG ATGGGGAGTG  
 68521 GCTCTAAATA CAGATGAAGC TTTACTACT CGCCCACTGT TCACCACCTG CTGTGTGGCC  
 4 68581 TGGTTCCAAA CAGGCCACGG ACTGTACTG GTCTGTGGCC AGGGGGTTGG GGGCCCCTGG  
 68641 TTTACAGAGC TCTGCCACAC AACTAAGAAC TGGCTAGGGA ACACAGAAGA TGGAAATG  
 68701 ATTCCAGCTG GAGAGGAAA GAATGGATCT GAAAGGCCCT TACAGAGGAG GTGGCATCTG  
 68761 AATTGGGCTT TAGACGAGGA ATGGGATTTT AACAGCTAGT GAGCCTAGGC ACCACTAGGA  
 68821 GGAAGAGAGA GGCATCCTAG GAGGAAGAAA CAGCATGAAC ACAGGCAGGG AGGCCAGAAA  
 68881 GCAGGGCAGC CTTATGTATG CAGGTGGTCT GGTGTCTGTG CAGTGAGTTG TGTGGGAAGA  
 68941 GATACCTGGG GTAGAGAGTG GAACACAGGG TTGGTAGAAA TTGGGCAGAA AGAAGGAGGC  
 69001 TTTGTGAAAG AACAAATCTG GAAACACCCC AGGTAAAGGCC CCTCTGCTCT CTGGAGGGCC  
 69061 AGGCACATGA AGAGGTCCCA CAGATACCAC CAACTTGGTT CTGGCCCTGC ATTTGGACCC  
 69121 AGTCTGCCCT GTGTTCAAGG GTCACCCCTT ATTCCATACT AGATTAGGGT CCAAATAATC  
 69181 ATGACCCCTA AAGGCCATAC TTTATATTTG CAGGTAGTCT CTAAGCAGTA GGGGGCTATT  
 69241 GATGGTTTTT GAGCAGGAGA CTGGAACCTT AAGAAGGCAA GGATGATTGA TGTGAAACA  
 69301 AGGGATCAAT TAGGAGGTGA GATGTGAACC ATCTGGGATA GGGGAAGTAG GGTGGAAAGT  
 69361 TAAGGGAGAG ATAAGGCCAG AACAGGCCAT ACAAGCTAAG TCCAGGAGTT TCTGTGGCAA  
 69421 CTCACCCAC AGCACTATAC TGTCTGAAA GTCATGGGGT CCCATCTGTC TCCCCATCCC  
 69481 TCCCACTCCC CCACCCTTCC CAATCAGATA GGCCAGGTTG AGCCCCAAAC ACCAGGCAT  
 69541 TCCTCAAGGC CTCCTATTTG GCCTCTTCTC ACTGAAGAAG CCCAGCCACA AACCCCCACT  
 69601 CCAACTCATA TGTTTCCTCA CTGTCTATCA ACTCCTCAGA GCAGGGTCTC AACTCATTTT  
 69661 AGTTGAATTA AATCAAATTG AATAGGGGGC TGATGGGAGC TTGGTGGAAG TTTCTTCCCC  
 69721 CACATCTCTT GTGGTGCCTC TGTCCCAGGT GGGAGCCAGT GTGACCTGA ACTAAGTGGA  
 69781 CTGATCTACT GGAAAGTATT TACCGACTGC CTGCTGGGTG CCCATCTCTG TGCTGACCAC  
 69841 TCAGGTGGGG TCTGGAATAA AACAACCAAG AATGTAAAAT CTAGAAAGGG GAGGAAAGGG  
 69901 GGTAAATTA AATGGTAAAT TGGCCAGTGT GGGCTCTAAG TGGGATGGGA GGGGAGCAGA  
 69961 GCAGAGGACT CGGGGAGTTG GAAGTGGACT GGGACAGTGA GTAAGGAGTG TCAGGCCAAAC  
 70021 AGTGACATCC CCACGCCAGG GTGCGGGTGC CAAAGCAGCA ATCTTGGTAC ACAACTGCTG  
 70081 CTCGAAGTCA AGTGCTCTTT TGGTAGAGTT CCTTCTGTGT GGAATTTCTG AACTTGTCTC  
 70141 TGCGGCAAGG GGAGGGAACA GTTCTTTTCC ACCCAGCACT GATGACCAGG GAGCCCCAGG  
 70201 ATCCTCAGAT CCCCTGGGCT ACTTGGGCCC AGCCTTGCTG CTGTGTGCGC TGAGGCCCTC  
 70261 GACTCCAGGC CTGTGCGCCAG GCTTGTACC TGACACTTCA GATGGGGTGA CCAGGCACTG  
 70321 CCTCTGCCAG CAGACTCTTG GCCTCAGGTA ACCCAATATT GAAAAGCACA TTCCTCCAG  
 70381 AAGCCCAAGT GACGCCACTC CCCTGTGGGC CCAGCAAGGA TGACTCTCA CTGTCTCTCT  
 70441 TACAGGTGGT TCAGGGCCCT TCCCTTCTC TCCACCTCC TGCCAGCAGC ACCTGTCTCC  
 70501 AGCCTCCTCT CTGCCTCCTA CAGCTTGTGA GCTTTAGAAA CACTCCCTTC AACCTCTCGT  
 70561 TCTTCTGCTG AAAAGCACCT AACGCTTTTC CAGTCTTTGC CAACAAATC TGGAATCCAA  
 70621 TACCTTCCAT AAACCAGGTA GGACCCAGGA GCAGCTCTAC GTATATACGG AATTTTAGCA  
 70681 TACAATAAAG GAAGTATGTG AAATCAGTCG TGAACTGTA AGAATTAAG AGGAGCCAGG  
 70741 CGCGGTGGCT CACGCCTGTA ATCCAGCAC TTTGGGAGGC TGAGGTGGGC GGATCACGAG

70801 GTCAGGAGAT CGAGACCATC CTGGCTAACA CAGTGAAACC CCATCTCCAC TACACAAAAA  
 70861 ATTAGCCGGG CATGGTCGCG GGCACCTGTA GTCCCAGCTA CTCAGGAGGC TGAGGCAGGA  
 70921 GAATGGCATG AATCCGGGAG GCGGAGCTTG CAGTGAGCCA AGATTGCACC ACTGCACTCC  
 70981 AGCCTGGGAG ACAGAGTGAA ACTCTGTCTC AAAAAAGAAA AAAAAAGAAA GGAAAAGAAA  
 71041 AGAAAAGAAA GAAAGAGGAA AGAAACATGA AAGGTGACTC GCCAGTAAGA CAGGTTTACT  
 71101 TAAGGGAAAA CAAACCTGAG AGGGGCTTCT GGCCAAGTTA GGTGAGAGGC ATACTCTCTT  
 71161 ACAGACTAAG AGTTTTTAAG GATTGAGAGT GGGAGAGTTT ATCAGAGGCT TGGACTGCTT  
 71221 CTGTGTCTCT TTGTTGTGCT TATCTGGGAG GGAGAGTTGT GTGTCTGTTT CCATACATCT  
 71281 TCCTGTAGCT GCAGGCATAC CTCCCCCAC CCCCCTACTG AGTCTGCTTT TAGCTTCCCT  
 71341 ATCTTAGTGC ACCTGAAGGG AAAGGAATGT GCTTATTAAG GCCCAATGTT TTAGTGGGAC  
 71401 CCATTGTATG CAGGCAGAGT TTGGCAGTTC CAAGAGACGT TCCCCCTACC TCCCTCTGTA  
 71461 CCCAAGCTAT CTTATCTGTG TTTTACTGCC TGCTCTTTCT GGCTGCTTGT AGTTAGAAGA  
 71521 GAAGTGATTT CTTTGAAATG CATGAGGCTA GAAAGGGAGC TGGAACTTAG AGTGGTGGTG  
 71581 TTTGTCTGAG ATGACGGTGC TCCTGCTCTG TCAGAAAGGA TGGACCATTC ATTATTGGAA  
 71641 AAGTAGCTTT ACCATTGGA ATAAATGAA ATTAACCTCA CAAATAAAT TCCAGATAGC  
 71701 TTAATATCTT AACCATTAAT ATAAACTAT ATATAATTTA AAATGAAATG TGAGAGTACA  
 71761 TTTATATAAT GTTAGGTGGA GGAGGATCTT CCTAAACAAG GTATAAAACC TAGAAGCAAA  
 71821 TAGGGAAATA TTGCCAGTTT TGAATACACA GAAATTTTAA ACTTCGGTAA AGCAAAAGAT  
 71881 GCCATAAGCA GAGTTCCACC TCTACAATAG ATCTCACGTT TCCCCTCTTA ATAAATATTC  
 71941 ATGCATAACA AAATCACTTT AAAAAAAGAG ATATAGAGGA GAGTCTTGCC CTACCAAGTA  
 72001 CTAAGATAAA TTACAAAGCC ATACTAATAC AAACAGTATA TTAGTGGGAA AAAATGGAGG  
 72061 GAACAGAGAC AGACCCTGCG TTTTCAAGGG AACTTAATAT ATGATAAAAT TGGCACCACA  
 72121 AGTCAATGGG AAAAGAATAA ACTGTTTAAT GGGTGACATG GCAAAAGTGA CTTACTATAT  
 72181 AGAGTAAAG AAAACAGCAT TCCCATTTCA TACCATAATA AAAGGCAGGC TGTAAAAAAT  
 72241 TAACACAGGG AAAAGCCAGG AAATGCAAAAT TCATCAGAAA TGCAAAATAA AATGAAATG  
 72301 AGATTTTACT ACAGTACTA AATTGAAAAA AATTGAGAAA GTTGGATAAT GGCAAGTGT  
 72361 GGTGGGAATA TGGGAATAGA GTCACCATCG TGCTCTGCTG TGGGAAGAGGA GGCTATCAGG  
 72421 TCAGTGAGGA CACTCTGTCG TCACCTAGTC AAAGTAAGAA TATCAGACAG CAATTCTGAT  
 72481 TTGGATATAG AAAGATATGT ATGAGGGTCG GGTGTGGTAG CTCATGCCTG TAATTCCAGT  
 72541 ACTGTGAGAG GCCGAGGTGG GAGGATCATT TGAGGTCAGG AGTTCGAGAC CAGCCTGGCC  
 72601 AACATGGTGA AACTCCATCT CTACTAAAAA TACAAAAATT AGCTGGGCGT GGTGGTGGGT  
 72661 GCCTAGAAAT CCAGCTACTC AGGAGGCTGA GGTAAAAGAA CCGCTTGAAC CCAGGAGACA  
 72721 GAGATTGCGA TGAGCTGAGA TCACACCACT GCCCTCCAAC CTGGGCGACA GAGCAAGACT  
 72781 CTATCTCAAA AAAAAAAGAG AGAAAGAAAG AAAGATGTGT ATGAGGGCAT TCACTCCACA  
 72841 TCACCTACAG CATGAGAAAT GTGGAACAG CCCATGTGAG AGCATGGCTG GAAGCATGAA  
 72901 TTGGGAGTGA AGGATGTGCA ACCTGCTGCA GTCATTCAA GCAACACACG AGACGCACGC  
 72961 AGAACCAACT GATGGGTCTG GGACAATCAG TGTGTACAGG AAGAGGCAGG AAACACTGGT  
 73021 AAATCTATCA AGGACATCTA CAACAACCTC AATATGGCTA GGAATATAT TCAAATAAAA  
 73081 TATGCATTGA ACACAGTAAA ATGATTGCCT TTCTGCTGGG AGGGGATGGG GAAAGTTTTG  
 73141 TGGGTATAGG GAAATAGCAA TCAATCAAGA AAGCTGGAAG GGGCCCCCA GGGACCAGTG  
 73201 AGGATATTGG GCCATGTCTT GGAGTCCCTC GCACCTCAGG TGAGATATT TTTTCCCTCC  
 73261 AGCTTTTCTC CAGATGTTGC TGCCACTGCT CTTGTCTAAG CCATTATCTC CACTTAACTC  
 73321 CTATATTGCC TAACGGATCT CCCTGCTTCT ACTTTTGCTT TCCTATAAGC CACTCTCTAC  
 73381 AGAAGAGCAA AAGTGATCTC CTGAAAACAT AAGCGAGATT ACCTGACTGG CCGGATTCAA  
 73441 CTGTGCCCCC CACTTCTGGC GGCTGACTTC AGACCTTCTC TGAAGGTCTG TGTGGTCGGG  
 73501 CTGCTCACGT TTGTCTACT TTTCCCTTCA CACCCAAGCT CCAGCTACAC TGGTCTCTCT  
 73561 TGCCAGGCT CTTCCCATG TTGGGAATTT GGCCTGAAGT CCTTGACCGC ACTCTGTTTA  
 73621 TAGCTGCTCT ATCTCATCCA TCACAGCCCT GCTCAAATGC CATTTTCTGT GGCAGCCCTC  
 73681 CTGGGCCACC CTCACCTCT CATCTACCTA CTGCTTCTG TGTGCTCTA TTAACATAAT  
 73741 TCATCATATC CTCATAATC TTGTTTTCTT ATTCTTTTGG CCTTTTCCA CCAGACTGTA  
 73801 ACCCTTTTCA GGACATGGAC ATTGCCTGTC TTGTTTCATCA GCTGGAAATA CCTAATGCCC  
 73861 ATCAGTAAAG GAATGGCTAA ATTAGTTACG ACACATTAC ACACGTTACA CAGATGTCAA  
 73921 CAGGAGCAAA GTGGCTCTCA TTTATTGAAT GTTACAAAAG TAGCAGGGAC TGGTCTAAGT  
 73981 ATTTTATATT GCCTGCCAAG CAGATACTAT TATTATCTCA TTTTAAAGAA GCAGAAATTGA  
 74041 AACACAGAGA GGATAAGAAC CTTGCTCAAG ATCACGATGG TAAGTGACAG AGCTGTGACA  
 74101 CTAGCTAGTT TGGATTTAAC CACTATGATC CCAGGTAGAA AGCTGCTGAC CTGGAAGAACT  
 74161 GAGATTCTGT ACTGCGTTTT TATTTTATTT ATTTATTAC TTTTATTATTT ATTTATTCTT  
 74221 TTATTATTAT CTTTATTATTT TTATTATTTT TTGAGACGGA GTCTCACTCT GTTGCCAGG  
 74281 CTGGAGTGCA GTGGTGTAAT CTCGGCTCAC TGCAACCTCC ACCTCCTGGC TTCAAGCAAT  
 74341 TTTCTGCGCT CAGCCTTCTG AGTAGCTGGG ATTACAGGAG TGAGCCACCA CACTGGGCTA  
 74401 ATTTTGTAT TTTTGTAGA GAAGGAGTTT CACCATGTGG GCCAGGCTGG TGTCAAACCTC  
 74461 CTGACCTCAG GTGATCTGCC CACCTCGGCC TCCCAAAGTG CTGGGATTAT AGGCATGAGC  
 74521 CACCGTGCCC AGCCTGTACT CTCTTTTTAA AAGGCTTCTT CTATGTGCTT ATGTTTCTTG  
 74581 ATAAATAAAA TCCAATTTCC TAAACAAAT GTACAATGAA GAATATAATT TTCATGACAT  
 74641 AAAGTAAAC AAGCCTTACA GCTTCTCTGG GTAAGATGAG ATTAATAAAT CAACCATGGA  
 74701 AGTTAATATT TTGAGATGTC CCAGAACCTC TGCCAGCAA ACTAGATGTG ACTCTATTTC  
 74761 ATGATGACAA GATAACAAAG TATAAACATT ACAAGAAAGG ACCTTCACTC CAATGTAGGT  
 74821 AATGTTTGTA TTCTTAGTGT CATCAATGTA TCTGGCTCAT GAATATTTTC TGCCCTGCCT  
 74881 TATCTTTACA AAATATTAAG GGCCAGCTTA GTGCTTTTCT TTTCTTTTCT TTTCTTTTGA  
 74941 AGACAGTCTC GCTCTGTTGC CCAGACTGGA GTGCAGTGGC GCGATCTCAG CTCATTGCAA  
 75001 CCTCTGCTTC CCAAGCTCAA GCGTGATTCT CCGATGTTAG CCTCCGAGT AACTGGGATT



```

75061 ACAGGCACAC ATCACCACGC CTAGCTAATT TTTGTATTTT TTTGTAGAG ACAGGGTTTT
75121 GCCATGTTGC CCAGGCCAGT CTCAAACACC TGGGCTCAAG CAATCTGCCC ACCTCAGCGT
75181 CCCAAAGTAC TGGGATTACA AGCATGAGCT ACTGTGCCTG GCCCTTTTCA TTCTCAAAGC
75241 ACTTGATATCC ACAGCTCAAT ATTTCTCTTC CTGCCATTAA AACTTGTGGC CTCTGGTTTC
75301 AGCTAGATGC CAGGCTTTAT TTCCCTGCTC TGTACTCTCT TGCTCATACA GTGAGGTGCC
75361 AAAAGCCTTC TGCACACTGA GTTTACTCCT TATCACCCTC CACTTTGGTA AGGATGACAT
75421 CGTCATTGTC TCAAGACTGA GTATGGTATA GCGGTCTAGA CTTTAATAGA CCTGGGTTTG
75481 TACCTTCATT CTGTTACATA TAAACTACGT GGACTTGGGT AAGTTTATCC CTCTAAGCCT
75541 GAGCTTCTCC ATCTGTAAAG TGGGAAGCAT AACAGTTTCT ACCTCATGGG ATGGTTGTCC
75601 AGAATAAATG AGGTAAGGA CTTAGCACAT GACGGGCACG TGGCTGATTC ACTATAAGTG
75661 ATCCACAGTA CAGTAACATC AGGGATACAA AGACATCATC GCTGCCCCAA GGACCTCCCA
75721 ATTTAAGTTG GATGAGGCAG GAAGGTACAA CAGATAGTTT CAGGACAATC ACCTCATATA
75781 TGAGAAGAGA GGTAAAAATA AGCTGCTCAG AGAGCTTGGC CACTCAGGGC TGGGAGATC
75841 AGAAGAGGGC CTGGAGGAGG AAGAAATTGA ACTTGCCCTT GAAGAAGGAA GGATATGGGC
75901 CTGACAAATG AAGAGGAGGG GAGAGGACAT TTTTCTCAGG TGAGGTGGCA TGAGTAAAGG
75961 AATGAAGGTT GGAGAAGGGA ATTAAGGAGG TTCTATGGTA TGGAGTGAAA GAATCACAAT
76021 CTTTCTTCAG TCAGCAGTGG AACTCATGGC AGGTTTTTGA GCAGAGAGAG TTGAAGAAAG
76081 AGAGGAAACC CTGCTCCTAC AGTAGAACAA GTGACTCTCA GCACTGTGTG TGGTTAGGGT
76141 GAGCTGGAAG GCCTTCATCA GCCATCATGG GGCTCCTGGA GCAAGTGGTA GCCTGCAAAA
76201 CGCCCCAGAT CAATGCCATAT CCCTCTTTGT CTGCGAGGAA CAGCCACCAG GTTTAGTTTT
76261 CTCCAAATAC ACGATAGGCT CCCAGAGCAG GGCACACTTT GCATCAGGCT CTCCAATAT
76321 CACTCTGGGC CCCGAGGGGA TGGAGCTCAG GAAAGAGAAA TCTGCACGTT TGTGGAGGAG
76381 GAGGAGAAAT TTGGAGCAGG GCCTGTGTGA GAGGGAGATG GTGGGAGGCC AGGTGGCAGC
76441 CTGCAAGGTT TTCTTCACTT TCCCTCTGGC CATCCCCCAT TAAGCTAGAA CAGAGTCTTG
76501 CTCCATCTGT GTTGAGGTGG TCTCACAGCC TTCTGCTTCC CCCTCAGCCC AAGTAACCAC
76561 TCATTTGCCT CCTGTTGCTG TAGACTAGGT GTGTCTTTTC TGGAATTTCA CATAAATGGA
76621 ATCATGCAGT GTTTGCTCTT TTGTGTGTGG CCCCTTTCAT TTAGCATAAG GCTTTTGAGA
76681 TTGGTCCAGG TTTTGTGTG TAACAAGAGT TTACTTTTTT ATTGCTGTGT AGCATTCCAT
76741 TGTAATAATTA TTACAATTTG TTTATCCATT CATCAGTTGA AGGGCAAAAT TTTTCTTTT
76801 CTTTCTCTTT TTTGAGACAG AGTGTGTTC TGTACCCAG GCTGGAGTAC AGTGGCACAA
76861 TCAGGGCTCA CTGCAGCCTT GACCTCCCTG GCTCCAGTGA TCCTCTCACT TCAGCCTCCT
76921 GAGTAGCTGG GACCACAGGC ATTTGCCACC ATGCCAGCT AATTAAGAAA AAATTTGGTA
76981 GAGTCGGGTT CTCACATATG TGCCACGCGT GGTCTTGAAC TTCTGGGCTC AAGAAATCCA
77041 CCCGCTCTTT CTTCCCAAAG TGCTGGGATT ACAGGTATGA GCCACCGCAC ACAGTCACAA
77101 AGTTCTCTCC CTTCTATGCT AGCAATCATA AATAAAGCTG CTGTAAACAT TTGCACACAG
77161 GTTCTGTCAT AAATATACAT TTTTCTTTCT CTTGGCTAAA TCCCCAGACG TGAATGGCT
77221 GGGTCTGATG GATGTTTTTA CAGAGTGATG GCATGTCTGG GCTCGTCTAA AAATCCAGC
77281 CTCTGCACA GAGCCTCCTC TCTTGGCAC TTTCCCATTT TTCCTCCTCT CATTTTCTCT
77341 TCTTCTCTCT CTTTCTCTCT CCTCTTTCTC CCATAGGCTT CCCTTTTCCC TCCTCCATCC
77401 CACAGTCCTT TCCTGCCAAG CTCCTGGGTT CTCCACAGA GGGGGTATTG TAACAGGCAG
77461 CCCTCAGCTT TTCTTGGTCC AAGGCCAGCT GACCAAGCAC CCACAAGTAG ACTCTGCCAA
77521 GCTTCTTGAC ATCAGCTCCT GAGCCAGGGC ACCTCTCGGA GACTCCAGCA CTTACAGGG
77581 AGAGGGAAGG GGAGGTCTTC TAGCAAGTTC AGCCTGCAG TAATTGAGCA AGGCTGGGTT
77641 TCTCCTAGTT CCACTGTCTT AAGAACTGTG TCAGATGGTT CTAGTCTACC TGATATCATT
77701 GCGTGTGTTG TCTCTCCAA ATCTCATGTT GAAATGTGAT CCCCAGTGT GGAAGTGGAG
77761 CCTAGTGGGA AGTATTTGGG TCATGGGGAT GGATTTCTCA TGAATGGCTT GGTCTTCTG
77821 GTGGTAATGA GTGAGTTCTC GCTCTGAGTT CATGCAAGAT CTAGTTGTTA AAAAAGAGTT
77881 GACCTCCCAA CACCTCCTTG CTTCCCTCTT TGCCACCCAA CATGCTTGCT CCCTGTGAGC
77941 CTTCTGCCAC GAGTAAAGC TTCTTGAGGG CTCACCAGAA GCCAAGCAGA TGCTGGTGTC
78001 ATGCTTGTG CGCCTGCAGA ACCTCTTTTC TTTGTAAATT ACCTAGTCTC AGATATTCTC
78061 TTATACAATG CAAAATGGAC TAATATACTA CTCCATTAC TGGTGCTCTA TGTTTTGACT
78121 GAATTGAGTT CCAGCCAAGT TCTAACAGAG AAAAATAATT CCCCATGGGC ACCATTCACT
78181 TCAAACCCCG GAGTCAAGGG TCCTGCCCCA CCACCCACA GGAAGGGGA GGCAGAAGGG
78241 GCAGTGTGAG GCCTGGGAGG GGCATGAGCA GGAGATGCCT TCAGATCACG TCCTGGGTTT
78301 GACAAGGACA GCACAGCTGA GGAAGGCCTT GGCTCCCTAC TGACTTTAGG TCAATGTCCA
78361 TTTGCATTTA GGAAGGCAGT GTGGGGGTGG ACAGAACCACC GGGTGGGGAG ACTTCTGTAT
78421 ATTTTACCCA TCTCAGGTGA GTCTGAGAGG CAGCCGTGGT GTGGACCATC CAACCAAGTG
78481 GCCTTAGAGG CTTTCTCAC TGCTGATTTT TGGGGCCTGG CTTGTTCTTT GCTGCCTGGT
78541 GAGGCTATGC CATGCTTGGG CACAGAGAAG AGAATAAGGT GAGAGATTTG TGGGGTAAGA
78601 CACAGACTAA AAATGTGTT TGATGTCAGA AAGGAGGAAC ACAGCCAGGA AAGAGCTTTA
78661 AGTAACTAAA AATAGGCTGA TATTTAGTGA GGGGACCAGG ATTTGGGTTA CGTGAAGACA
78721 GAATTTGTGG CCTGTCTTG CAGGATCCA GGGCAGGGGA CTCCCCCACC CCTCCCCACT
78781 GCATGTGTGT GTGAGTGTGT GAGAAGTGTG TGAGTGTGAG AACAGGTATG TATGTGAGTA
78841 TGCTGTGAGC ATGTATGAGT GGTATGTTGT GTAAGTGTGC GAGTGTGTGT TGTATGAGTG
78901 TGCTGTGTGA ATATATTGTG TGAGTGTGTG TGTGTGTGA GAGTACTTG GAGCGGACGC
78961 TTATGTGTTT TTGCTTGAGG AAGAAGGCTC CTGAGTCTT GCTGGAGAA TCTGAGAAAT
79021 AAAGTCCAGT TGCACCGAGA TGGATTAAAG TCAGCAGCCC TGTGGTGCTG GGCCTGTGGT
79081 CCTGGAGCCT GGTTCCTGCG GTGCCGATG AGCCTTGCCC CTGCTTGCCC TGCCAGACAA
79141 GAATGAGTGC AAGTCTTGTG GCCAGTAAAG GCAAAATACC CTGGGGATTT TTTCACTGCC
79201 ACAGTGTGG TTTAAGGTT TGAGTTAAAG AATTTTCACT GAGGAACACAG AGGAGTGGCT
79261 GAGCTGGCCC TTATGACACT GGGAGGGCTT CTGGGAGGGG CCCTGCCACG CCTCTGCACT

```

79321	CTGTCTGGTT	GGTGACTAGT	GGGCTCCAGG	GCCAGCATGG	TGAGATGCCA	GGCACCTGGA
79381	GGAGGCATTG	GGGCAAAGGA	CATCCTGGGT	GGGTGGGGTG	GGAGGGGCAG	GTCCTTTGGG
79441	GTGACTCTGC	AGTCCTGGTT	TCTGTATCT	GGGCAGCACC	TGGGTGAGGC	TGAACACATG
79501	GATGATCAAG	CACCAAAGGC	AGAAGAGGGT	CCTCTGGTCT	GTGAGGAGCA	GGAGAGGGAC
79561	AAGAGCCACC	ACCTGGCACC	CTCTCCCTGC	TGCCGCCCCG	CAGCCTGTGA	AAGCAGCCCC
79621	TGGCCCTGGG	TGGAGTGGGC	CCTGGTGGGG	ATGGGCTCAA	GCCTCCTGTC	CTGTGTAGGG
79681	GAGCTGGCAG	CATGGCAGAG	ACCCCTTGCC	AAGGTGCAAA	TGACACTTCA	CCTCTCTCAA
79741	CTGCTATTTT	CCTCTGTGCA	CATGGATGTG	TAGGGATTGG	GTGGAGTGGG	AGAGAGGAGT
79801	TCAAGTCCTT	TTTCCGGTGA	TGTAGAAATT	CAGAGGGAAG	GACTCTGGGC	TCCAAGAGCA
79861	CAGACCTGAC	CCCTGTAGT	ATCTGCCTCG	TAGCCCTCAG	CAACTGTCTG	CTGCAAAACG
79921	TAACCAAAGA	CTCATTTGAA	AGTAACCAAA	CATCCTGGTT	TGTCCAAGAC	TAAAGGGTTT
79981	CCTGGGACTG	ACGACTTTTA	GTGCTAAAAC	TGGAAAGTTC	CAGGCAAACT	GGGATGATCA
80041	GATCACTTAT	TTCTAATTTT	CCTTGAAAGT	ATTTTGTGAA	GTAGAGACAT	ACCTGATATG
80101	CCAGACCCGG	AGTACTCCTA	GGCTTTGTAA	TTATTTGAGG	CAAAATAATT	TGATTACTTT
80161	CTGATTTTGT	TTTGGAGAAT	ATGAAGAAAA	TCTGTATGTG	CTGATTCTGA	GTACATAAGC
80221	CAGTCTGGCT	GTTAACGGTT	GACTGTGTCT	GTATGTGGGA	TGCTGGTATA	GTGGACACAT
80281	GACCCTAGAG	TCCCTTTCCC	CTTCTCTGTA	TAAGAAAGCA	CCGTTTTCTT	TAGGATCTTT
80341	ATGGAGCTGT	CATCAGGGAA	TCATCCTCCT	TTCTCAAAGG	TGGGCCCCAA	CCAGTCTGGC
80401	CAGTCTCATT	CTCATCCTCT	TAGACTCAGT	GATTGGTTCA	GGAACAATTA	AATCTCTCTC
80461	TCTGCTCCTT	CCTCTTCTAT	GTCTCTTTT	TAAGGATTAT	ATGAATATTG	AGAAAGAGAG
80521	AGAAGAAGAG	GTCAGCTTTG	AGTTATTGGT	GGCCATCTTT	GTCACCATGT	GGAAGGACTC
80581	AGCCTTAAAA	AGTTGAGAGA	AGGAGAAAAA	AAGACAAAT	CCTGAGGACA	TGATCTGGAC
80641	TGCTAGATCC	AGCTATGCCT	GAAGCCTATA	TGCCCCTTGG	ACTTCCCAGT	TACATGAGAC
80701	AAATAAGGCT	TTTTATTTTC	TTAAGCGTAT	TTGAAATCAG	CTTTAGTTAC	TCTGCCACTT
80761	GCGCCTGAAC	AGAATCCAC	ACAGACTCCT	GATGTGCCTG	GCTGCCACAT	CTGCAGGAGG
80821	GCTGTACTCT	GGGCTCTGTG	CAGGGGATCA	CAGGGGTGAT	GCCTTTCACT	GTCCTCTGAA
80881	CTGAGTCCCT	TATGCCTTTA	AATTAGGGCA	TATGCTCCTC	CAGGGCACAT	CTACGGTGGG
80941	AGGCACATGT	CTGAACATTT	GCCTTCAGGA	GCCCTCTTCT	CCAAATCGTG	GGGGAAGAGA
81001	TCTTCCAAGT	TCCTTCTCTC	ATCTTCTCTG	TCTTCTCGCC	TCATTTCCAG	GATTTGCTCT
81061	GATCTCTCAT	TTGTCTTTGA	CTGGAAGTCA	ACACTGTCTAT	ATGTCTACCC	ACCACTCACA
81121	TACCTTACTG	CCTTGGTGAC	ATCCCTGCA	TAGGGAAAGC	CTCTCCCAT	CCTGGGAGGA
81181	GCAACTAGCA	CCTGAAACAT	TACATTTCAA	CACCTGTAAA	CTGGGACAGA	TGAAAAGCGT
81241	TGGTTATTGT	CCATGCCTAG	GGGCTAGAGA	GTGGGGACTC	ACCCTCGTTC	TCCTGCTGGG
81301	AATGGGAGAG	GCAACCATGC	TTTCCCAGCG	AGTAAGGCTC	AGTTGAAGGA	AGGGAAGGCG
81361	AGGCAGGATG	GGGCAGGCC	AGGGGATGTT	GCAAAACAGC	CACAGGAAAT	GGAAATCATG
81421	TCATATGCAA	ATCAGGATAA	GAATTTGCAG	ATGGGCTCGG	CACTTGGGCT	TACAGTTAAT
81481	CTTCTTAACT	AAGGTTAAAC	ACTATCCCTA	GGAAGAGAGA	GAGGGGAAGT	GATTTGGGTAA
81541	GGAGCACCTT	GGGGCCTTTG	ACTCTGTAC	CAGTGTGATT	TTTTAAGGCT	GGGCAGTGGA
81601	GTCATAGGTA	TCCATTTTAT	TATTCTCTAT	ACATTTTGTA	TATCGTATAT	ATTGCTTAAT
81661	ACTTTTTTAA	AGCAACCTGA	ATTCATGTCT	ATCACTCCCA	GGCATTTCCT	TACTCTTTCT
81721	CTTACCTATG	GTGTGTTCTT	GAGCTATACT	TGGAATTATT	TCACGTTTTT	AAACTTTATA
81781	TAAGCTAGGT	AACATACAAA	TATCCATTCT	GCTATAAACA	TGTTTTTTTT	CTCTACTTTA
81841	CATTAGGCTT	GTAAGATTCA	CCTATATTGG	TTAACAATAG	ATATGGTTCA	TTCATTTTCA
81901	CAGTTGTGAC	CTATTCCACA	ACTTATGATT	TTGTGGGTAA	GACATTTAAG	CTGTTCTTAA
81961	TGATGGCAGC	GGCGGACCAT	CTGGAGCGGC	CGTGCCATC	ACCGGCCGCA	GCAGGGAGGG
82021	TGCGGGGAGG	AGGCCGACAG	CCTCTCTGTC	AGCCACCCGC	CCTGGGGGCC	GCGGCGATGG
82081	GGCTGGGCCA	GGTCGCCCCG	TGGCAGAGGA	GCAGCGCAGC	TGGGGAGCAG	CGCGACTGGG
82141	GAGAGCAGGG	CCCTGAAGTG	GAAGTGGGCC	CAGAGCGGTG	CTGCTCTTGC	ACTGGGAGCA
82201	GGAGCCCCAG	GAGGGGGAAC	TGGGGATGCG	CTTCTGGGGC	CCGGGTCTGA	AAGTGGGAGC
82261	GGTGCCGCTT	TGGGGACACG	CCAGCGGGGC	AGCCATTGCT	GCCATCGCCG	AGGATGCCCG
82321	GTTCTCTGTC	CTCGGGTGGG	GGCTCTGCTC	GGCGCTGCTC	AGGGCCGCGT	CCCCAGGTCC
82381	ACCCTGCATC	TGGGTGACCG	CCTAGCCTGA	CACTCCCAAA	GGCTGGGCCA	GGGCTGCCCA
82441	TCCACACCAC	CCTGGACCCG	CCCTGGTTGC	CAGGGCAACA	GAAGGGAGCT	CATAGCCGTG
82501	TCGCCCTCTG	CCTGGACACT	GGCCGGGGCC	CAGCCAGGAC	CTGGAGCCCC	TGCCCCAGGC
82561	TGCGAAGGAG	GGCGGATGG	AGCTGCATGC	TCCGCGGAGC	TGGCGGGAGC	CGGGAGCCGG
82621	GGAGCAGGCA	ACAGCCAGG	TGGGGCTGT	GGGCCCTGGC	CTCTGTGTGC	TCTTGGGGGC
82681	CCTGGGAGTG	CCCCCGACCC	CCCGGCCCTG	CAGGCTTGGA	AGTGCTGTCT	CCCGTTGCCT
82741	GGCTTCTCCC	CGCTGCCGCG	GCCCACTCCG	ATCTCAGAGC	ACCGTGGGGC	CGAGTCCAGG
82801	TGCTGTTGCA	GCCTGGCCGA	GGGTGCACAT	ATTTGGGGCA	GTGCTGGTAT	GCCAGCCCCC
82861	TGCCGCTTCA	GCCCCCTCCA	GACTTTGGGT	GCCAACGAGC	ATGGGGTGGG	GGGTGAGGCC
82921	AAGGTGGGGG	ACTGAGGGCA	GCTCCGTGCT	GGCCTGTAGG	TGCCCTTTGG	CACCAGCAGC
82981	CTGGCCACCA	TAGGTGGCAG	CAGAAGGCAG	ACAAGCTCCT	GGGTAGAAGG	GGGTGGGTCC
83041	CTGGTGAGGC	CCACCTTCA	GGCCAGGCTG	CCAGTCTCTG	AGATCAGAAT	GGGAACCTGT
83101	GGTACCTTTT	CCAGGCTCAC	CCAGGCTGTC	CCACGGACCA	ATGTGCGCGC	ACTTCTCTCC
83161	CTCGAAGCCC	TGGGCTCAGC	CAGAGCAGAG	CGGTCTGTAG	GATGACCAGC	TGCTGAGAGG
83221	AGCTTCCCCC	TCCAGGGCCT	CCTCTCTGCT	AGGAGTTGAA	CACTAGTAGG	GACGCCCTGG
83281	CTGGGAAAAA	AGCTACCTAC	TGCGGGTCTC	CTCTGAGCTG	TTCTATTGCT	GGATAAATCT
83341	CCTCATCTTC	TTGCTCACCT	CCCACTTGTC	TGTGTACCTT	GTTCTTCTTG	GTCACAGGAC
83401	AAGAACCTTG	TACCTGTCTA	ATGGCGGGTC	TAAAAGATCT	ATAACACAAA	CAGGGCTGAA
83461	ACATGCCCCC	TGCTCACTTC	GTTGTGGGTG	AAGAGAAGGA	GAGAAGAGCT	GCGGCCCTTC
83521	AGGGAACCCA	AACCTGGGAG	CTCATGGAGC	CACATCTGTG	ACTCCCTCTT	CGGGGCCCTG

83581 TGGTTCCCGG CATCTCCAAG CTCTGGGCA CCACCACATT CCCTGGAGCC AGCTGTGGAA  
 83641 GCTGCTTGCG GTGCACCTGG AGCTGCCAC CTCACTGCGG CAGCTGGCAT GTAGCCAGAC  
 83701 CCCACGTTTG CTCACACACC CATTGCTGCT CTATGCAGTC TCCCTTGCCA GGGATCCAGG  
 83761 CTGGTAGTGT GAGCTGAGTG CAGCCTGCCA GGCTGAGTGG GCAGAACGAA CCCAGCGGAC  
 83821 CCAAGCAAAA CTGGGGCAAA GGTGCCAATA GCCACAGGTT TCTGGCCAGA AAAGTGACAC  
 83881 CCCAAAGATC CCATAACATT AAGTTTTGCT GTTACTGTAT TTCATCAGTT CCAAGGTACA  
 83941 CATTTTTTCC CATTTTAAAA TCTGTGAAAT AAGAGATACT TTTTAAAAAC TTATAATGTG  
 84001 TAATATTCTG ATGTGTAATA TCTAAGGTAT TGGTATATCT TAGATCAGTG GTCCCCAACG  
 84061 TTTTGGCACC AGGGACTGGT TTTGTAGAAG ACAATTTTTC CATGGATGGT GGGGGCAGGG  
 84121 ATGATTCAAG TGCTTTACAT TTATTGTGCA CTTTATTTCT ATTATTATTA CACTGTAATA  
 84181 TATAATAAAA TAATTATACA ACTCACCATA ATGTAGAATC AGTGGGAGCC CTGAAC TTGT  
 84241 TTTCTGCAA CTAAATGGTC CCATTTGGGG GTGATGGGAG ACAGTGACAG ATCATCAGGC  
 84301 ATTAGATTCT CACAAGGAGC GTGCAACCTA GATCCCTTGC ATAGGCAGTT CAGAATAGTG  
 84361 TTTGCAC TCC TATGAGGATC TAACGGCATT GCTGATCTGA CAGGAGGCAG AGCTCAGGCG  
 84421 CTAATGTTCA CCAGCCTGCT GCTCACCTCC TGCTGTGTGG CCCAGTTCCT AGCAGGCCAC  
 84481 AGACCTGATC GTACCGTGAG CTTCCTTAGA ATTGACTATG GTAAGGGAGA ATAAGAGGCT  
 84541 TGAAAGCAGT GTGCGGGCAG GGAGTACAAT GCTACTGTGC ACTTTCTTGT ACATGTCATG  
 84601 TTTCTTTAGT GTCTATATTC CTAGGAGTAT AAATTGCTGG GTTATGAGGG ATATTATCAC  
 84661 CAATTTTACT AGATATTACC AAAGTGTCT CCAAAGTAAT TGTACTGAT TCACACTTCC  
 84721 AATACTATGT ATGAGAATTC CCATTGGGCT ATACGATGGG GTATCATCTG ACTTTTGTAG  
 84781 TTTTGCTAAT CCAGTGGATA TAAAATGGTA TCTCATTGTG GTTTTCATAG GCATTTTCCT  
 84841 TATTATTAGG GATCGTGGGC ATCTTTTCAA GATCTTTTAC TTTCACTAAT CCATGTTCTT  
 84901 GGGACTGACT TTTCAATTGAC TCCACCCACT TCCCCACCC ACTGCCTTCA AGCCTCTTAC  
 84961 CCTAGTCAAA TCTAAGCACC ATCAACACTT ATGAATATA ACAGTAGTAT ATCCTAAAGT  
 85021 ACAAGACTAT TTATGCCCAA GTAATTACAC TCTTGGAAAC TCACACCGAG GAATACAAAT  
 85081 GCTATCTGCT CAAAAATTTC ATTTAAATTT TTTGATGGTG AAGAATTGAA AATCACATAA  
 85141 TAGGCTCAAT ATTGGTAGAA ATATTAGTAA ATAATTGCAT ATTTCTATGA CTGAATTCTA  
 85201 TGCAGTCTTG GAAAGAGAAC TATAGAGAGA ATGTACCAAT ATGGAAAAAT ACTTTGTGTA  
 85261 ATGGTGTGAT ACTGGTATAC TGATTAGCAG ACTGATCAAC AGAATAGAAT GAATATCTCC  
 85321 AAATAGAACC AAGTATACAT ATATACATAC ATATGTATGT GTACGTATAT ATATATATGA  
 85381 ATACTTGAAT TTGCATATGT ATGTATATAC ACATATACTT GGATTTAGGA TATAAGAAGA  
 85441 CTGGATTATA AAGATAAATA AGTAAAAAAT ATTAATTTTT CTATAAAAAAT CATAGCCACA  
 85501 GAAGCAAAACA AGATCACATG CAAAATAGCA ATGAAAAATA AAATATATAA ACCTAAGGAG  
 85561 AAATAAATAA GAAATTCACA GACACTATGT AGAAAGCTAT AAAACTCTCC TAAGAGACAT  
 85621 AAAACAGCTT GAAAAATTAA CATACATACC TTGTCTTAA ATGGGAAAAAT AAAATGATCA  
 85681 ATTCTTCTTA AATTGATCTA TTTATTTAGT ACAATAACCA ATTAATTTTC CAACCAGAAG  
 85741 TTTTGGGGGA ACTTAACCAA AAGTTGTAAA ATGTATGCAG AAAAATGAAA ATGAAAAAAT  
 85801 GTTCAGTGAA TTTGTAAAC AGATAAAAT AAGGAGAGAA AACTAATTCT ATCAATAAAT  
 85861 GTACTTTAAA GCCACAGTAT TTGGCCAGTT GTGGTGGCTT CTGCCTGTAA TCCCAGCACT  
 85921 TTGGGAGGCC GAGGTGGAGG ATTGCTTGAG GCCAGGAGTT CAAGACCAGT TTGGGTAACA  
 85981 TAGTGAGACC CCTCCACCA TTCTACAAAA CTTTAAAAACA CACACACACA CGCACACACA  
 86041 CACACCTACA CACACACTAC AGTATTTAAG GCCTAGAGCA TCCAGAGACA AATCCAAGTG  
 86101 CATACAGGAT ATTTACATGT ACAAATGACA ACATTGCAAA TTAACAGGGG AAGGACAGGT  
 86161 TATTTGATAA ATTATGTTGA AAAAAATAAG TTTCTGAAAA AAATAATCAT TTTGTCAATT  
 86221 GGGGAAAAAT AAGCTGGTTG CATATCTCCT TGTATTAATA TAAATTCAG ATGTACTGGT  
 86281 ATTGTTTTAT CATTATTAGT GAGAAAGCAA GATATGAAAT TGAGATGCAA TAAAGGAAAA  
 86341 TGAATGGATA ATTCTGACTG TACAGTTTTT AGATATTAGG TTTCTCTATG TCAAACACG  
 86401 TACACAATAA TACAAGACTA ATTTGTAAAA ATGTTTACTA AACTTATGAT GAAGGGTTAA  
 86461 TTTTCCCAAT AAACATAAAT CTCTTTCACA TAAATACTGA AATGCAACA ACATAAAAAAC  
 86521 TGACAAAAAG TATAAACAGG CAATTAATGG TAGTAATGTG TATGTCGCA GTTCCAGAAC  
 86581 TTCGCTGAAC ATTGGAATCA TATGGAGATC TTTAAAAATA GTATTGTTGC CTGGCTCTCA  
 86641 GATACTCTGA TGTAATTGGT ATGGTGTATA ACTTGGGCAC TGGAAATTTT AGAAGCTTCT  
 86701 CTGATGATT CTAAGTTCAA GAAAGGTTGG GAATCAATGG ACTATGCAAT AAGTCTATAT  
 86761 AAAAGATTAT TGCTAATAAT TTAAGACTG TAAATTAATA CAAAAAGAG ATGCCACGTT  
 86821 TCATTTATCA GATTGGCAAG GATAAAAAAT TGATAAGACG TAGTGTGGT GAGGGTGTGA  
 86881 GGAAGAGTCA GTAAATGTT GGTGATAGGA TGATAGATTAG AGCAAAGTTT TTGGAGGGCA  
 86941 ATTTGGTAAT CTTATGAATA TGCAACCTT GCACTGCCTT TGACTCAGCA ACTGTATTTT  
 87001 TAGGAATTTA TTTTACACAA ACATTGCTC TAAATCTCAA AGATATTTGA TTTGGTATTA  
 87061 TTCCCATAAAT ATTTTAATTT TTGTTGTTGT TATTTGCCTG TTTGGTTTAA TGAGGGATAG  
 87121 GGCTTAGAAG TGGTTGATGG GTAAAAGGAA AAAAAAAGTT GTTGATGAAT GGAAATAGCA  
 87181 GTGCCCCCAT TCCACTCGCA CCCACCGATG TATCCATGGC ATCTAGGCGT TGGATACTGG  
 87241 TGAGTGGAGC TGGGCAGCAT CTTCCCTTAC GAATCCAGTT GCATCCATAA ATTTACATAG  
 87301 GAGCAGACGA GCCTCCTGGG GTTGGGGAAA GAGGTGGGGA CTCGGACTCC TCATTGAGGA  
 87361 TTCAAGTTAT TGCAGTGTAT ATGATGAAC AAAGAGACCT CCCAGGAGA CCCATTGAAG  
 87421 TCAGATGACA AGGAAGGACA CTGGGAGTTC CAGAGAGGAA ATATGCCCTA AAGAAGAGGG  
 87481 AAGTTTAGAC CAGTGTCTG TACTGAGTAT TGGGGTTTAC CTTATCAGAA CCACTCTTTC  
 87541 ATGCTTATTC ACTTTGCTAC CTGCGGTCTC TTTGAGGGGA ACTGGCTCTG TGTGTGTGTG  
 87601 TGGAGGGACA GAGCTGCATC ATTATCAACA GACACTTGGT CACCAGGTCC CTTTGCTGAG  
 87661 GGACCTCTCT CATTGCTCCT GACAGGGCAG ATGCATTCTT TAGCCTGGGC AAATCAGCAG  
 87721 AGAGGCTTTT GTAAAGTTTC AAGTCAGCAG ACAAATGATG GGGGTCTGTC TCATGGACGG  
 87781 GGACAAATGG CATGCGAATT CTCTGGGCTA TTTGGAGGGA AATCTTGACT TCTCTGCCAA



87841 CTTGACAAGG AGTGGAGAGG AATCATTGTA ATTTGCATAT CAATGACATT TTCTAGGCTT  
 87901 TTGGGTGTTT ATAGCACACT CCAATGGTTC ATAGCATCAG AAATTGGAAA CAACCTAGAT  
 87961 GATATGATGT TATGTTATGT AGCCATTTTC AAAAAATCTGG TTTTAAATAG CAGAGAAAAA  
 88021 TTAATCACTA TTCACTGATA TGTAAGAAATG ACAAGTTTTA CAATACGAGC TTAATTTTAC  
 88081 ATATATCAAT ATACGTATAT ATAAAGAAAA TTTATTTCTA AATATATGTG CCTATAAAAA  
 88141 AGACTGGAAG GAAGTATATT AAGGGTTCAG GAAGTTATTT CTGGATGATG TTATTGGCAA  
 88201 TGTTTATTTT CTCTATTACA TTTTCAAAT TTTGTAATGT CTATACTGAA TACATATATT  
 88261 ATTTATAATC AGAAGAACAA TTTTGTAAAA AGAACATTAT TATAAATAAA GTTGCTCTTT  
 88321 AAATAAGACA AAAAATGCAT TTAGAGACAG GGAAAGAAAG CTAACATTTA TTTTCATGCCT  
 88381 ACCATGTCCC AGGCCACCAT ACCAGATATT TATATATGTC AGCTCTGTGA AAAGTATCTT  
 88441 CTCCCATGGC CATCCTCTGG GAGTAGCAGG AGGTGTAGCT TGAGGTGGGG TCAGGCAAAAC  
 88501 CAGCCAGGGG AAGGAATTTT TTCTCAAGAG CTGGGAAGCC CTTTCTTCCT TCTACTAAGC  
 88561 AGCTGGCCAT CTGCCCACCA ACCAGATCCT TCTCTTTGCC TTGGGTGAAA CATCCACTCA  
 88621 TTGGGCAAGG AAGGGGGCAG AGCCGTGGCT CACTGAAAAT GTGACTGAAG CAGATACTGG  
 88681 AGCCCAGGCC AGCAGGGGCT TGAACATATA GAGCAAGGAA GGCAACTTCA GGTATGAAAA  
 88741 AGAGACGTCA GCACCAAGAA GGGTTAGACT GATATTGTGC CTGACCTGAT GGTGGAATA  
 88801 GGGCCTTTCT GAGTGGGTAG ATTTGACCTT GGTGTAAGAA AGAAGTTAGT CACTTGGAAG  
 88861 GTAATAAGCT TCCATTCTTA GAATTATGCA AGGGGTTCCCT GCCTGGAAGG TGCAGAGGAG  
 88921 ACTCCTACAC TGGGCAAGCG GGGGTGGCTC TGGAAACCTA ATGATCTGTT CCAGTTTTGA  
 88981 GAGCCTGTGA CTCTCATCCC CACCATGCCC CAGATTGAGA AGACAGCTTG CTACCCCTGT  
 89041 CAGCTGCTGG AAGCATGTGG GAGAGGCCTT CCGGAGGCCA GGCCACCTT TGTGCTTAA  
 89101 GAGGTAGGTG GCTCTGACGG TGGGAACCTT ACTAGGCAGG GAGGAGGGAA CAGGGAAGGG  
 89161 CCAGCCTGGC CTAGTGAGGG AAATATGTGA GGAAGGTGGG CCTGTGGGCA CCCGTAGGTC  
 89221 AGGACCTGTG GGTGGGCATC CACTCGGCCT TTCTCAAGC CAGCCCTGAA ACTCATCTG  
 89281 TGGGGCAAGT CTGATGCGAA CATGGAGTCT CTTGTCCTGG GTGGCTCTCA CCTGCTAATA  
 89341 GACACATCCT CACAAATGGT AGGTATTTAG GGAGCATGGT GATGGAGGAT AAAAAAATA  
 89401 AAAAAAATA CAACACAAAA CCAAGAGAGC CAGTGACAAA AAGCTTACAA CAACAACAGA  
 89461 AGTTCACAG CAAGACCCAT AAGACAAATG TAGAGGAAAG AGCAGGAGAT CTAGAGTTAG  
 89521 AGAGACCCGA GTGCGTATCT TGGCTCTGGC ACTTACTGGG GCAGCAGTCA GCAATTACTT  
 89581 TATCTCTCTG AAACCTACTG TTCACTTTCC TCATCTGTAA AATGGAGTTA ATAGTACTGA  
 89641 TTTACAGGGA TTGTTGTAAA GATTATATAT TATATATACA TAAACACCGT ACATATACAT  
 89701 CTGGCACACA GAATCTGGCA TGTCTATAAC CTTAATAAAT TGTATGGTT ACTGTTATTC  
 89761 TTTTTCTATC AGGACAAATA AATTGGAAG AACTAACCCT AAGTGTACTT GGACCGCAAA  
 89821 GAACCTCAAG GTTGGCATGG GCTGGAGAAG TCAGTGAAAG CTTTGTTTAC TGATGACAGA  
 89881 GACATTTGAG CACCTCCTAA ATGCTAGGCA CTGTGCTACA GGTTCAGAGT ATCAAAATCA  
 89941 ACATGGTGCC TGCTCTTGGG AAGCTTCTAG TGAAGTAGCT TTGAAAGAAA ATAATTTCTA  
 90001 ACCATCAACT GCGATAAGAA TCAGGAAGGA AAGTAACAGG GTGCTACGAA AGACGGGAAC  
 90061 GGATTCTGAT TTGGGGAGAC GTCAGGGATG ACCTCCAGGA GGAAGCAGCC TTGAAGCTGA  
 90121 AACCTGATAG AAAATCGGGG TTAGCTACAT GCGGGCTGCG GGGAGAGTCC CCCAGGCTGA  
 90181 AGGTACAGCA GTGCAAGGCG TTAGAGGATG GAAAAGGCTG GCGTGTTCAG AAAGTGAACA  
 90241 CTTGTGCTG GGGCACGGTG AGGAAGGAGT GATGAGTCA GAGAGGCTT GGAAGCCATG  
 90301 GTCAAGATTT GGGATTTTAT CTTAAGAAAC TGGATGCCAC CGAAGGGTAT TAAGCAAGGA  
 90361 CAAAAACCA AACACCGCAT GTTCTCACTC ACAGGTGAGA ACTGAACAAT GAGAACAATT  
 90421 GGACACAGGG TGGGGAACAT CACACACTGG GGCCTGTTGT GGGATGGGGG GAGGGGGGAG  
 90481 GGATGACATT AGGAGATATA CCTAATGCTA AATGATGAGT TAATGAGTGC AGCACACCAA  
 90541 CGTGGCAGCT GTATACATAT GTAACAAACC TGCATGTTGT GCACATGTAC CCTAAACTT  
 90601 AAAGTATAAT AAAAAAATA GAAAGGAAGT GATGTGAATG AGATCATCTG GCTGCTGGGT  
 90661 GGGGACAGAA TTGGGCAGGA CAGGGACAGG AAGATGTGTT GTAGTGCTC AGATGGGAAG  
 90721 CAAATGGTCA CTTGGGCCCG GGTAGGGGCG AGAGATGGAA GCAGTGGTGG ACTCGAGAGG  
 90781 TATTTCTGAG CAGAACCCTC ACCGCATGGT GATAGATTAG ATATGAGGGG ATCCAGTGGG  
 90841 CTTCACTGAC CCGGCAGGAT AGATTAGATA TGAGGGCATC CAGTGGGCTC CAGTGACCCA  
 90901 GCAGGTGAAG GTACACAGGG CTGGTCTTGA AGGGGAGGCT GTCAAGATCC CACGTGGAGG  
 90961 CCGGTGGGGA GTCTCAGGGG GCTGAGCCTT GAGTTAGGGG CAAGGTTTGG AAAGGTGGCA  
 91021 CTTGGGGAAG GGAGACCCTC TACGATGGGA GGTAGCAGAA CTGGGGAAAG AAGTAGGAAT  
 91081 GAGCTTGCCCT CACTCCACGT GGGTGGCACA GAGTAGGGCA AGCCCTGCTG AGAGGGAAGG  
 91141 GCATGAGCTG GGGGCCAAGA GTGTGAAACC GGGAGGATGG GGTGGGGTGG GGCCCTTGGA  
 91201 GTGGGCGGAG GCCTCTGCCT CTCATGAGAT GCAAGAGATA GTAGGGAGCT GTTAGGTTCC  
 91261 CAGGCCAGCA CCGAGCCACC CAGGCAGCAC CGGCACTCCA CGGCAGTGCC GGGGATGAA  
 91321 AAGGCATGAA CAAGCAGCTG CAGAGCTTGT GCTTTCTGAG AGTATTTACT GACCACCTGC  
 91381 TCTGTGCTCA GAAAGATAAC GGGATGTTTA TGGAAAAAGT CTCTTGCTC CCAGGACACC  
 91441 ACTTTCTCAT GGGTTGCTTC CAGCCTCACT GCCTGCTGCT TTGCATCTAC TCTGATGGCT  
 91501 CTTTCATCCTT GCCGGGCATC TGAAGGTGG AAAACGGGAG AGCAGAGACT TCTTCCCTT  
 91561 CGCTTCTGTC TCTTTATCCT TTAGAGATGT CATCCAGCTC CATGGCTTAA AAACCTTCTA  
 91621 CATTCGTGCT TATGATTTCA AACAGATTTC CTCTCAACT TGTCTCTGA GTTCCAGACT  
 91681 TGTATTTTCA ACGGCCACCC TGACATCTCT GCTTAGATGT GAATCTCAAG TTTAAATGCG  
 91741 CAAAGACAGA GGCACCCAGA CCGCTCCACC CCAACCAATT CTATCCTAGC TGAGTAAGGG  
 91801 GCATCTCTCA GTTACTAAGG CAAAACTCA CTAGGCATCT GTGATTCTC CTTCCATATC  
 91861 CAGTGTGTCC CAAGTCCTGT AAATTCGACC TCCAGGAAAC CCAGAATCCA TCCACTCCTA  
 91921 CTGTCCAATG CTACCAACCC TAACCAAGCC ACCTCCATTT CTCTCTGGAC CACAGCAAA  
 91981 GCCTTCTCAC TTGGCCCGCA CTTCCCTCTT GGCCCCAGG AAATCCATGC TCCACAAAAC  
 92041 AGACAAAGTG ATGCAGGTAA AACATAAAAC AGATCAAGTC AGTCCCTCAGG TTAAAGCCCT

92101 CCATGGGCTC CCCCTGTCTT TTGGAACAAA ATCAAAGCTG CTTTCCGCAG CATCCGGTCC  
 92161 ATAAGGCCCC GACTTCACCT CCCAATGCTT TCTTCCTTGC TTTCCCTGCT CTCTGCTCTC  
 92221 CAGTCACCCCT GGTTTTCTAC CTGTTTCTAC AAAAGGTCAA CCTTGTGTCC ACCTCAGGGC  
 92281 CTTTGCACCTT AGAATGTGCT CTTTCTGGAA TACTTGCCAT CAATCTTCTC ATCATTCCGA  
 92341 TCTTAGTTCA AATGTCACCC TCTTCTGATG TCCCTCCCT GCTGCCTGTC ACTCGCTGTC  
 92401 ACATTTTGTC CTAATAGCAC CAATATCAGA AATAGTCAGA TATTTATTTA CTGTCTCCCA  
 92461 CATTGAAATG TAAGTTTCGT GAGAGCTGGA TCTTGTCTGA CCTGCTCAGT GCTCTCTCTC  
 92521 CCACTCCTAG AACAACTACTA GGTGCATAGT GGGAGCTCAG TAAATATTTG TTCAATGAAT  
 92581 GAAAAAATGC TTTGGATGTG GTCCCAGGTC TTCTTGAGGG AGTGTGACAA GCCACATGAG  
 92641 AAGTTATCTA AGAACACAGG GCAGGAGATG GTAACCTCTC CACAGGGCAG GAGATGCGGT  
 92701 GATTCAAAGG TAATCTGATA AGCTGATATT TGCTTTAGGG TTAAGTGGATG ACAATATGCT  
 92761 AGACTGTAAT GGGAGCCTCC AACTGGCCCC AGAAAAGAGG AAGGCTGAT CCATAGGGGT  
 92821 GGTCTGGAG GGAATCCTTC AAGGACCAAA ATGGAGGGCA GATGAGGAGC TATTTGAAGA  
 92881 CTGGTGAGTG TCTCAGTCCA TTTTATGCTA CTATAACAGA ATATCTGAGA CTGGGTAATT  
 92941 TATAAAGAAC AGAAATCTAT TTCTTTAAAA TATGTTTTTG GCAGGTAGAG GTGGGGTGTG  
 93001 CCACTGTTGA CCAGGCTGAA TCCCTGACCT CAAGTGATGA TCCTGTCTTG GCCTCTCAAA  
 93061 ATGCTGGGAT AATGGGTGTG ACCCACAGCA CCTGACCAAG AACAGAGATT ATTTCTTATA  
 93121 GCTCCAGAGG CTGGGAAGTC TAAGGTCAAG GGGTCCACAT CTGGCAAGGG CCTTCTTGCT  
 93181 GTGTCATTCC ATGGTGGAGG GCAGAAGGGC AAGAGAGCAC AAGAGAGAGC AACAGGGTGC  
 93241 AGCACTCAAT TTTATAACAA TCCCCTCTC AGTGACTCCC ATGAAAACGA CATGAATTCA  
 93301 TCCATGAGGG CAGCAGCCTC ATGACCGAAT CCCCTCCCAA ATGTCCCACC TCTCAACAGT  
 93361 GTTGCCTTAG GGATTAAGTT TCCAACACAT GACCTTTGGG GGGCATATTC AGACCATAGC  
 93421 GGTGGGTCTT TGCAGTTCCC CCAATCTGCT CACGTAAGTA ATCGAAGATT CCACTGTTAC  
 93481 CAGCTGTGCA CACATTCCCC AAGCCTACAT TCCCCCATT CATACATATA TACACACCTT  
 93541 CTTGCAGAGC TGGTGGTGAC TCTTTTATC ATGTTTCAAG ATTTTACAAA GGTCTCAGGG  
 93601 AATGACCAAT TGGCCAGTCT CTGAGAAATT CCATGGAAAT TTGCAAAAAT AAAACACCTA  
 93661 TAAAACTGA CATTTCATCA CAAAAATCCT AATAAATTTT GTTGCAATACA TATGCACAA  
 93721 AGAATATTAT TTAGCCACTA AAATGATACG AAGTTTAGGT AGCAAGATAG AAAATATTAA  
 93781 TACTATATAA TTTAGGCTAA CCCCCCTCCC GGGCATAGAT ATGGTATAAA TATTATGCCA  
 93841 TATTGTTATG CTCTGCATAA TGACGGTTTG GTTAACAATG GACCACATAT ATGATGGTGG  
 93901 TCCTGTAAAG TCAGAAATATA CTTTTTTTAC TGTAGCCTCT CTATATTTAG ATACATTAG  
 93961 GTATGCAAA ACTTACTGTG TTACAATTGC CTACAGTGT CAGTACAGTA ACAGGCTGTA  
 94021 CAGCTTTGCA GCCTAGGAGC AAAAGGCTAC ACCATATAAC CTAGGTGTGT AGTGAGCTAT  
 94081 GCTACCTAGG TTTGTGTGCA CACACTCTGT GATGTTTGCA AAATGATGAA ATTGCCTAAT  
 94141 GATGCATTTT TCAGAACATG TCCCTGTGTT TAAGCAACAC ATGACTGTAG TTATCGTATT  
 94201 GGGTTAATGG GATTATAGTG GATTGGCCCT TTATGTTTTT CAAACATTTT TAGAAATTAC  
 94261 TTAACATACA TCAAAATTAA AATATTAATAA TTAAATACAA TGTCTTTTGT TGGTCTGTG  
 94321 TGTCCCAGTC ATGTGGGCCT TTCTGCTCCC CCTCAAGCAG CTGGAGAAAG ACCAGGGGGC  
 94381 CATGGGTACA GCCTCATATG TGCACTGTGA CAGGCATGGG GATGACGGGG CGGAGGGGAG  
 94441 AAATCACGGT AAAGTCATAG TGATAACAAT GACAGCGAGC CGCATTGCTC CAACTTCATA  
 94501 TGCCCGCAGG ACTGTGCTAA ACATGTTGTT TACTCCGCAC ATCATTTTCC CCATAGAGGG  
 94561 AGAACTGAA GTTCAAGAGC GCTAAGGAGC TTGTCTGAGG TTACAGGGCT GCCAAATGAC  
 94621 AGAGCAAGGG ATTGAAGCTG GTATAACTCC AGAGCCCGTG CTCTTCACCA CTGAATGAG  
 94681 GCCAATAAAT TCACAAATAT TAAGCCATTA TGTTTTCGTT GCTCCCAAAA CCTTTCAGCT  
 94741 GCCACTTCTA GGTGTGAGTC AGGAGGCTGA AGATGCCAGG GCAGGCAAGG GCAGGCCATT  
 94801 AACCTCGTTA GCAAGACAGC AGGGGAGACA GCTCGCCCCA GTTCCCATCC CCTCCAGCG  
 94861 TATCAACAGC AGGCCTAGCC GATGAGACAC AGAGACTTCT AGAACGGAGG CAGAGTCAGG  
 94921 AAGAGGAAGG GGAGAAGCAC CTGCTCACCG GAGACAGGGA GACTGCAGAA GATGAGCTCT  
 94981 TGTGGGAATG GTGGGATAGT GTGGGGCAGA GACATCAAGG GAAACCTTGG GCGGGGGCTC  
 95041 TTGACCCCTC TACATGTG ACTCTGGGAA GTGCCACCAG CAGAGGGCCG AGCCTCCAGT  
 95101 CCACCTGGGT TAAGGGATGG AGGCTGAAT CACTTTTATA ACAATCTCAC TCTCTCAGTT  
 95161 ACTCACTCCC ATGATCATGA CATGAATTCA TCCATGCGGG CAGCGGCCTC ATGACCAAA  
 95221 CCCCTCCTAA GGGTCCCACG TCTCAACACT GTTGCAATTG GGATTAAGTT TCCAACACAT  
 95281 GAATGCTGGG GGACATACTC CCCCAGTATG GAGGATGGTC CCCCATATG GAGGACCGGG  
 95341 TCATCCCCAA GCAGCCCTCA GTTACCTGTG TGATGGATGC TGTGGTAGGT ATGAAGTGGG  
 95401 ACCTTCTGGA TCCAGACAAG GTGACTCCGT GGTGGGGCT CAGGATGGAC CCTCCAGCC  
 95461 AGGGGCCAGA TAGAAGGAGG CAGAGACCAG AGGGGTCCAT AGGAGGTGAG GGGATCTTGG  
 95521 CAGTTCTTCC CTGTAAGCCA GGCTCGGCT GGCAGGAGA AGCGAGGCTC ACTCCCTGCA  
 95581 GAGCTCAGTC TGGTGCCACC GCTAAGACCA TAGCAGCCAG AGGGGCAGAG GAATGTGTGTC  
 95641 ACCTCCCCAA ATCACTGAGA GCCTCCCAAG AGCTCACACG CCCGCCAGA GGAGAACTG  
 95701 CACGAGGCAG ACCTGGATCA CCTGATAATT ACCTTCAGAT GCTGTTTGAT TTCTGTAGCC  
 95761 CACTCAATTT ACCAGATGGG AATTGCGTAT TTGCTCATAG GTGGGGTTTG CGAGGTAAGC  
 95821 CAGATCTATG TGTCCCTCC AACTCTCTCT GACTCGCATG GAATTACACT TACGACCTG  
 95881 CCAGGCCACT ATGCTAGACT GCATGGTTCT AGAGCCTATG TGACAGGCAC TCTGCCAGGG  
 95941 GCTTCACATG CATGGTCTTG TTCACTCTTC TCAGCATCGT GTTGACAAAT ATCTTCCATC  
 96001 TCAGGTGCCC ATGAACTAG CTATGTGGCC TTGGGCAGGT TCCTTGACCT GTCTGTGCTT  
 96061 CAGTTTCTTC ATCAAATGCA GATACTGATT CCTGCCTAGC AATATCAGTG CATAGAATGC  
 96121 ACTTGATAAA AGGTCACCAC TGTGTGTAAC AGTCTTTTGT AGCACAGAGG GAGGTAAGGA  
 96181 CAGATAGGTG ATGGCAGCAG GGTAAAGGGA CTCTCCGGA GAGAAAGAAA TGAACAAAAA  
 96241 CCACAACCTC TTAATCTGTC AGCAATGCAG GTGATCTTAT TCCGAGTCAG GAAAGTGAGG  
 96301 TTAATAATTTA AGGGACCAGA CCCCCCTCCC AGGAACACAG ACATACCTCA GAGCCCCAAG

96361 GAGCCATAGG ACAGAAAGGT GGGAGACCAG GAGTGTCCAG GTCCTGTGTA TCCTCGAACT  
 96421 CACTGAAACT TCTTCCTGAC CTGCTGCCTT CTTTCATCCGT TTGTGGGCTC TGTGGGACAA  
 96481 ACACATTAGT TCCCAGGCCA GGACAGCCTT CCGAGGCCTG GAGACTCATC CCAGCCTCCT  
 96541 CTCTGGGACT CAGGACGGGC TGTGTGGAGC TGAGGCTGGT CCTCAAAGTA AGGTGTTGTT  
 96601 GGAGCCCTCT GGGGTAATTT CAAGGGAGTT TTGTGTGAAA GGGGAAGCAGT GCCACCAGAC  
 96661 TCCTCCTGAA GCTAAGCCAG GATGAATTAT TAAGGAGCTG AAGGAAGCCA GGGCTTAAC  
 96721 AAAGGAGCCC CACGTCTATC TCTGTGAATA TTTCAAACCA TCTGGACCTA TTTGTGGAG  
 96781 AGAAATTCTT TCGCTGCCT GAAAGACAAG CAGGGAGCTA GCATTATGCT CACATAGGAC  
 96841 AGTTCTCCGT CAAGGGAGGC CTGCGTTTCA GGGCTGGGAT CCACGCTCAG AGAAAGCTTC  
 96901 TTAGTGCTGC AGAGCTATTT TTAAAGTGCG AAATCAATA TGTCAGTGGG AGAGGGAGGG  
 96961 AGAGAGGGCT GGGGGTAAA TGGCTGAGGC CATCCAGGCC CCTGAAGGAG CCTCACTTC  
 97021 CTGGGTGGGT GAAGCCACTC ATTTGGGTCT TGCCGTGACAC CCCAGATCT CTATTTTAAAC  
 97081 CCCTGAAAT AAATACCTAG CCTTGTGTTT AGAGCTGATG GTGGGGCCAC AGGCCAGAAT  
 97141 AGAATTTAAT CATTTACTCT GGCTCTCTTT GGTTTAATAG ATGCCATATC TGGCAGCTTA  
 97201 ACGGGGATGT GTCCTAGTTC CATGTTAAAA ACAAAAATTA AAAAAAATTT AAGAACGCTT  
 97261 CTTTATGTTT TGTTTACAAC TGTTTTTAAC ATACATTTGA AATGTTTCATC TTCCTGTCTT  
 97321 TCTGTGTGCT GGCTCTGAGG CACCCAGCAT CCTCTCTAGT TGGGAGTATC CCGCTGTCCC  
 97381 CCTCTCTTTC TGAGGCTGTG GTTTTCCAGT GGGTGTCTCT GGCCCAAATA GCCCTATCAG  
 97441 AAATGAAGCA ACAGGCTGGG TGCAGTGGCT CATGCCATA ATCCCAGCAC TTTGGGAGGC  
 97501 CACGGCCAGT GGATCGCTTG AACCAGGAC TTTGAGACCA GTCTGGGCAA CTTAGCGAAA  
 97561 TCCTGACTCT ACTAAAAACA CAAAAATTAG CTGGGCATGG CGATGCGTGC CTGTAGTCCC  
 97621 AGCTACTTGG GAGGCTGAGG TGAGAGGATC GCTTGAGCCC AAGAGGTTGA AGTTGCACTG  
 97681 AGCTAGATTT GTGCCCTGTC ACTCCAGCCT GGGCTACAGA ACAAGACCCT GTCTCAATCA  
 97741 ACTAATAAAC CAAGCAACAC CCAGACCCAC ACTTAGACTG ACCAGTCATG TGTTTGGGTA  
 97801 GGAGACGGTT ATAAAGTGTC ATTTTCCGT TTGTGCTATA TAGGAAAAAA TAGCTTGAAT  
 97861 GGCAGTGCTA ATAAGGATAA AAGGAGTTGA CATTGCTCAA GCACACACTT GGTTCCCTTA  
 97921 CAAAGCAAAG GCAGTTTGGG ATAAACAAT AAAAAGTAAA CTTTGAATA AAATAAATGG  
 97981 AAAATGCATT GTAAAAATAT CAGTTTGGAA TTGACTTGAA GAGTTGCCAG ATCTCTTCCC  
 98041 TGCACCACAT TTCACAGTTT ACAAGCCTTG TTGGTTTACG TTTTCCATTT CTTAGGGGCC  
 98101 TTGTTCTGGA GGTCTTTACT AGAAGCCAGT GCCGGCAAGT GACTGGGGAA AAGGCAACTC  
 98161 TTCATGGAGG GAGTGTGCAG ATGCCAGACA TGGCTATAAT ATTTTACTTT TTATTTTGCC  
 98221 ACAAGGGAGA AAGGAAGCTT GACATTGGAA GAAAAAAGT TATGTGGATG GAGATATGTA  
 98281 GATACAAAGG TGTTTGAAAT CAAGCTTCCT TGAAGTCAA GAGAAAGTTA TGTGTTAAGA  
 98341 CCCAGAATAA TTTCTGATTC AAGGTGAAGT TTTAGAATGT GAGAGGGAGA CCTGGATTAG  
 98401 ACTAAATATT GATAGGTGAT GGTTTATTGT GCAATTAATA AAAACACAAA TTATCAGTTC  
 98461 TTTATCCAGT TTCCAGTGAG AATTACGTTT GTGAATAATC GGTGCAAAAT AAAACCAACA  
 98521 CAGCTTAATA TTTCAGTGG TAACTTTGA AAGAATCGTA TCTTGGGGAT AAAAAACCAA  
 98581 CCTGGCAAAA AAAAAAATAA AACCCAAAAA CAAACAAACA AACAAAAAAA CAAAATTAGA  
 98641 CTTGCTAAGC AATAGTAGCC ACAAAAGAAA ATATACGAAG GTCAAGCTAT AATGGAAAC  
 98701 TAGTGAGGAT TAGTGAGTGG AATAACCAAG TGTACCTCGG TATTCTCATT TTTCTACATC  
 98761 TACCTTGTAT TGTAATTTAA AAAATAAATA CAGAGATCAG AATGGGGTCG TTCAAGATTT  
 98821 CTTTTGGTAG ATGTGTTTTG TGCTTAGTCA AGGTGAAAGG ATGCCCAGGC ACTGCCACTC  
 98881 AGAGGTCTCT GCCCTGTCCG GGGCATCTGG GGGGCACCTG GAAGAGGTCT CTTCCTATA  
 98941 GCTCTGCCCA CCCCACTGTG CAGGTGAGAC CCAGAATGGA CGCTGCTTTG AGGCAGCTGA  
 99001 CCAGTGGGAT CTGACCTGGA GCAGCTGGTG GTGGCCAGGG CTGAGAGCCA GGATGAGTGT  
 99061 GGAGGTGGTA GCACCTGGA GAGTGAGCAG AGAGGGCAGC TCGCTGGTA GACTGTGTGT  
 99121 GGGGCCAGG CAGGGAGGAG CCAGGCTTCC TTTCCCTGTT CTTGGGACAT CCTCTCTCCT  
 99181 GTTCTTTCTT CCTGCCCTTG AAGACCCTCC TGTGCTTGG GTTATTTTTC TCTCCGTGG  
 99241 TCTGACTCCA GGACTTCTCC AAGCCAGGAA CTATGGTGAC AGATAGACAG GCAGACAGGG  
 99301 TGAGAAAGGG GACCTCTCTC GCAGACCTCT AGGAACCTAG TGCCAGGGGA GGGGGCTCTG  
 99361 AAGTCAAGCT TCCTTGAAGT CAAAGAGAAA GTTATGTGTA AGACCCAGAA TAACTTCTGA  
 99421 TTCGAGGTAA ACTTTTAGAA TTTGAGAGG AGACCTGGAT TAGACTAAAT ATTGATAGGT  
 99481 GATGGTTTAT TGTGCAATTG AAAAAACAC AAATTATCAG TTCTTTATCC AGTTTCCAGT  
 99541 GAGAATTACT TTTGTGAATA ATCAATGCAA AATAAAACCA ACACAGCTC ACATTTTACA  
 99601 TGGTAAACTT TGAAAGAATC ATATCTTGGG GATAAAAAA GCAATCTGGC AAAAAAATAA  
 99661 TTAGACTTGG GCACAGGGGC AGGAGCAGAG ACGAGCCCTC ATGTCAATTC CAGATTGGTG  
 99721 TCAGGGAGTA TCTGAATGTT TATCCTTTTA AAATGTTTCT AGGGATGTG TCCAGGCCTT  
 99781 GTTGAGCTTA GTCTCCTGTA AAATTCACCT CCTACAAAT AAAAAATAAA GCCAAACAAA  
 99841 GAATTTTAACT CTCAGCTGAA ATAACAGCTC CAACAAACAT TCTTGCACTG AGCACCATGT  
 99901 GCCCGAATC CCTCAAGGGC ACCTGGCCAC TGGTCTGCC AAATTCCTCA CTTCCTGCC  
 99961 CAGCCAGGCC CTGGGTGGCC CCACAGTTGG CCACTCCTAG AATAGGCTCC ACAGGACTAT  
 100021 GTGTATTTT TGTATTTTGT AGTTTTTTTT TCTTATTTCA TTCTTAAAT GGAAGATCTA  
 100081 CAGGTGGATA GAACGCTCAC TAAGTAAGTC CTCAGCTTCC AAAACCTAGA GGCACACTAT  
 100141 TACCCTACAC ATTTAGTTAA GGGAAATGG AAAGGAGAAA AGGAGGGAGG GAAGGAGGAG  
 100201 ACAGAGGAAA GAAAGAAAAG CAAGGCAAGG CTGGAAAGA TTTGCAGAAG TAATGACAGC  
 100261 TTTTTTCTTT TTTTTTTTGA GACATAGTCT CACTGTGTCA CCCAGGCTGG AGTACAGTGG  
 100321 CATGATCTCA ACTCACTGCA ACCTCAGGT GCCAGAGATT CTCCTACCTC AGGATCCCGA  
 100381 GTAGCTGGGA TTACAGGCAC ATTTTATAT TTTTAGTAGA GATGGGGTTT TGCCATGTTG  
 100441 GCCAAGCTGG TCTCGAATC CTGACCTCAA GTGATCCGCC TGCCGTGCACC TCCCAAAGTG  
 100501 TTGGGATCAC AGGCGTGAGC CACCGCACCT GACCAGTAAC GACAGCTTTT AATGGTGAAA  
 100561 AACAGAACAA GCTGAGAAGG CTACAGTGAG GACTGGAATG TTGGGCTCTG ATGTGTGGCT

100621 TTGATGTGTA TATGTTTGAG CTTCTCCCTT TTTTACATCC CTGGGGAAAT TTTAACTTCA  
 100681 TTCTGTAAAG CTTAAATGCT TTTATGAGAA GTATCCTAAT TTAATAAAGT CTTATCTAGA  
 100741 AAACGTTTTA TGAAAAATAC CTAAGATGTA GAATAGGGAG TAGAATAGAA TCTGGTGGGT  
 100801 AAAAGTGGGC TGTATACATA AGTTTAAAGA AAAGAAAAAG AGAAAAATGT ATTTTCTGGA  
 100861 GCAAAATCAA TTAAGCCAAT TTAATTTCTT GATCTTCTTG ATTATTTTTA ATAGAATAAT  
 100921 ACTTAGCAAT TTCCAGTTTT GCTGCTTATC TTCGTGGGA CCAAGTCTTT TCCACACACC  
 100981 ATGAAATCAT ACCCTTTATG ATGGGTTTTA CCTTTCTTT GTGGCCGCTA TCTCTTTACA  
 101041 AGTCTAATTT TTGCCAGGAC AATTTTTTTT TTCTTAAGAT GTGGTCTCT AAAAGTTTAT  
 101101 CATGTGAAAT TTTGGTTGGT TTCATTTTGC ATTGGTTAT CACAAAAGTA ATGTTTCATTG  
 101161 CATTGGATAA AGGACTGATT ATTTGTTTTT ATTTTTTATT GCAAAAAGAA AACATAAAGA  
 101221 AAAGGCATTT TCCATGGTAT GGTAAAAATG TCAGGAAGCT TTGTTACATT AAAATAGTTG  
 101281 TTTAAATGT GGGAAAGCCT ACAGATAGAA AAAAAACAAA CAAAAACAT CCTTAGTTTC  
 101341 CAGGGGCAGG ATCTTTAATA AGGAAAGATT TTCTCCAATA AGAGAGGCC TCCACGGGAA  
 101401 GAAGAGTGGG GAGATCACAT AGGGAAAGGG GCTGTTGGAC ACACCAGAGA TCTGAAGCCA  
 101461 CATGGGAAGA ACCATCCAC ATGGGCCTAA ATTAAATGT CCAGATTGGC TCTGGGCACA  
 101521 TTTAACTAAA TGTGTAGGGT AATAGTGTGC CTCTAGGTTT TGGGAAGCTGA GGACTTACTT  
 101581 AGTGAGCGTT CTACCCACCT GTAGATCTTC CATTTTAGGA ATTGAATAAG AAAAAAAAT  
 101641 TAACAAATAC AAAAAATAC ACATAGCAAT TAATGTAGGT TGTTTACACA TGATAATGTA  
 101701 TGAGACCATT ATAATGGTTA TGGTCTCATA TCATTAGGAG GACTCTATTC AGAAAGGAG  
 101761 GATTTAACAG AGTATAGCTG GGGACAGTTA GTAGAAAAA TAAATTTGTT TCATGTTTGC  
 101821 CCCCAGAGTG AGCTGAGGCT CTTCAATGAA CTGGTTCAGG TTGTAAAGGA CATGGAGTAG  
 101881 AGACCGTGCC TGGCAAGACC CCAGGGACAG CAGCTGCTGC TACCACCTT AGGTACAGC  
 101941 AGGATCTGCC CAAGGCATT GCATTATGT TTCAGGGAGC AAAAGTCTAT TTAATTTTGC  
 102001 TCTACAAAAG GGAGATGGGG CTTTATCAAA AGGAGCTGAT TACCTTCAGC CCTGGATCCA  
 102061 GGTGTATCTC CTTCCAGGTC CATGGAAGAC ATACTGGGTG ACATCAGGAA GACATGAGAG  
 102121 GCTAGGCTCA GTGCCCTCTCC AGTGGTTCCT GATTCAAAA CGGATGGTGC CCACCCACCC  
 102181 AGGAGGTGCT GGGGTGCTGT GGGAGGGTGG GGCTCCACA GCTCTAAGA GTATTATTGG  
 102241 CAAGTTATCC AGTCTTTCTG AGCCTTAGTT TCTTCCTTGG TAAACAGGG ATAATAATGA  
 102301 TACCGACTTT ATAGAAGTGC GAGAACTAAC GCCTGCCAAA TAGTAGGGCA CAATTAATGT  
 102361 TAGTGCTTTT TCTGCCAAT TGCCAGTAC AGAAAGTGT GCTTCAGTAA CCTAGGCCAA  
 102421 TGCCAGGACT GATTTCATAG AATATCATAG ATATATTGCG AACGTCTAAA ATGATCTCGT  
 102481 TATAACAATG ATCGTTTTTA CAAAAATAAG ATATATTGAC TGTTTTGTAT CCCATGAAGG  
 102541 TGAGGGGAGG GATTTTCCTA ACTTTCTTTC CAGAATCCGC TATTAGGGTG ATGGACTCAG  
 102601 CACCTTTCTG TCCTCCCTA CCCGTCTTAA GAGTTCCTCG TCACAGAGTC GGAAATGGGG  
 102661 ATAGAGCTTG GCCCAAGTTA GAATTTCCAC CACTCCAGGG GCAGAATGCA GAAAGAAGAG  
 102721 AAAAGTAGCG GGGGGCCTGG CGCAGCTGCT AATGCGGTGC AGAACTCTT TGAGCTGCCA  
 102781 CCCTGGCAGG CCGGCCTTGG GGGGATGTGC CTGAGGATGG TGGGTGAGCA AAGCCACCTG  
 102841 CCCCAGCACA GAGCCTGCGG GTTGTGCTCT GCAAAGCCCT CCCGGTTTCA TTCCAGAGA  
 102901 CCGGGCGAGA AAGGCAAGC CAGGCAGCAG ACCTGGGGG GCGCGTGGCC AGGCTACCCG  
 102961 GATCGTCCAG CGCGACATGG AGTCTGTGTG TCCTCACCAG TGCTGCGTGA CAGGAGCCGG  
 103021 GCACCAACCA CGCCATAGGC TTGTCGGTGT CTCAGCACTG ACACCGTGTG GCCCTGAGAA  
 103081 AAAGCAGTGA GGGGTCCGCA GAAAGTCTTG GAGACGGGAG GTTGTCCCAT TTCCTTTCTT  
 103141 CGGGATGTTA ATCAGCATCC CTGGGACTTT CCTGGTTATC GCTGATGGCA GAAGTCCAGG  
 103201 AGGGAAGGA ATGTGTCTGA GCTCAAAGGC TGAAGGCTTC AAGACTTAGA AGTCAATTCC  
 103261 ATCATAAAG AGAAAGAATG TAATCTCAAG GTGAACCTCA GAGTAGGGAG TGAGTGTGTA  
 103321 TCGTGACCTT GGGGGCTAGT CATCACCCCA GCCTAGACAA AGATGATGTA AAACACCCGC  
 103381 GTCAGGGTTG TGACATCCTC CTGACGTCGA GGAGCGACCG CCCTTCCCG GATGATGTGG  
 103441 CTCGCACAGC TTACACTTTC TCCACCTCCG CTGGCCGTAC CTGCTTCAGT AAGTGGCCCT  
 103501 GCTATCCACG CAGTTGCGTA ATGGATTAGT CAAGACTCTC TTAGTCATAG GTAGGAGCCA  
 103561 AACTGAACAA GCTGAGGCAT AAAAAGCAAT TTAGTGCCTC ACCTAGCTGG GCAAGAGGCC  
 103621 GAACTAACTC ACAGGTTTGA CACAAGGAAG ATTCTGGGAA GTTGGAGAAG AACAGAACCT  
 103681 GGGATTGGCA GCGGCCAGGG CTGTTTGTCC TGCCCTGTGC TCCTCTGTCC CCCTTACCC  
 103741 ACTGGCCTCA CTCTCTCTT CTGTAGACAG CCTCCTCCA GTGGTGGGA TACAGCCCTT  
 103801 GCGAGCCAG GCTCAGCCAC CCCAACATCA AGGGAGAAAT TTTTCAACA GGAATCTTTC  
 103861 TAATAGGGAA GGGCTCAGAT GGGCTGCTC AAGTTAGGTG TCTGCCCCC GAGGCAATCC  
 103921 CCAGTGGCAA GGAGATGACA CACTGAGATA GGCCAGTCTT GGTCACACCC TCTCTCCTAG  
 103981 GGCTGCTGGA CAGGAGTTGT CAACAGAGAA GAGTGACAGG AACACTTACT TGGCAGGCAC  
 104041 AAATCACAGG TTGCTGTGCT CCACTATGCA GATGATGACA CCTCAGATT TTTTCTCTTT  
 104101 AGTAAACTCT TAAACTAGA CATATACACA GAAACTGCA TGAATTATGA GTGTGCAGTG  
 104161 TTTATTCACA AGAGCAAGGA CATAGAATCA ACCCAGGCAC CCATCAAAG TAGATTAAAG  
 104221 AAAATGTGGT GCATATACAC CATGGAATAC TACACAGCCA TAAAAAGAA CAAAATCGTG  
 104281 TCCTTTGCGA TACCATGGAT AAGCTGGAG GCTGTTACCT TAATTGAATT AACACAGAAA  
 104341 CAGAAAATA AATACTACGT GTTCTTACTT GTAAGTGGCA GACACAAAGA TGGGAAAAAT  
 104401 AAACACCGGG AATTCACAAA GCGGGGAGGG AGAGAGTGGG AGAAGGGTGG ACAAACTGCC  
 104461 TATTGGGTAC TATGTTGCTT ACTTGGGTAG GGCCCAAACC TCAGTACCAT GCGATATACC  
 104521 CATGTAAACC CAAACCTGTA CATGTACCCC CAGAATCTAA AATTTGTTTT AAAAAAGAAA  
 104581 ATAAATATGT ATAGAAATAG CAGAATGTAA GATACCTTTA AGTTCACCTT TCAGAGATAA  
 104641 CACTGATGAT GTTATCAGCA TTTATTATTT TCTTAAAGGA TCTCATTTCT ATTTATTAT  
 104701 TTATATATAT ATATTTTTTA GAGACAAGGT CTCACTCTGT TGCCAGGTT GGAGTGCAGT  
 104761 GATGCAATCA TAGTTCACTG CAGCCTCAAA CTCCTCAGTT CAAGTGATCC ACCTGCTTTG  
 104821 GCCTCCACCA GCGCTGGAAT TACAGGCATG AGCCAGCACA CTCGGCCAAG AATCTCATTT

104881 TTTAGCCTCA CTTTTAAAAA ATAAAAATAGA AAGTATACTG CATAAAGAAA AAAAATGTGC  
 104941 AGCTTCATTA AGTTTCCCTT TTGAATACAC CCATGTAACA AGCATCCTGA TGGAGAAACA  
 105001 GAACGTCACC AGCATCCTAG AAACCTTCCT TGGCTTCCTT CCAATCATGG CACACCTGTC  
 105061 TTCCAAAGGT AAATGCTATC CTGATTTCTG TACTGTAAAT TAGTTTGATC AGTTTGGCA  
 105121 CTTTATACCA ATGCAACTGT TCAATATTTT ATTCTTCCAT GCCTTGCTTC TTACATTCOA  
 105181 CATTATGTTT ATGAAACCCA TCCATACTGT GTGTAATACA GAGTGTGTTGT TTTCACGGCT  
 105241 GCATAATATT CCATTGCATG AATACTCCAA ACATTATTTG CCCATTTTAC TGTGTATAGA  
 105301 CACCCTGGCC GTGAAGAGGA AAGTCCCAAG TGCCATTGGG TTGTGACAGT ACTGTGAGAG  
 105361 ACTTTACCTA GGCATTATAT TCTTTGAATT TTGCTTTGAA TCATGTTCTT CTTTGAATAT  
 105421 ATGTGTAAT GCTGGCTGCT AATTAATATA TAAGGTTGTC ACATGAGTTT AAAAGCATTAA  
 105481 TTTCAGCCAA GCGCAGTGGT TCATGCTTGT AATCCCATCA CTTTGGGAAG CTAGGCAGAA  
 105541 GAATTGCTTG AGCCCAAGAG TTCAAGACCA GCCTAAGCAA CATAGCAGGA CTTCTATAAA  
 105601 TAAGAGCGTA TTTTGTGGAA AGTTTACGG AGTGACGAGT ATGCTCTGTG CTCTTGGGGA  
 105661 CCACTACTCC AGGCAATTAA TTGTCTGGC AACACCGTCA CTTGGCAACA CAGCCCCCTG  
 105721 GTGGCAGCCA CTTGAAGTGT AATCTTACCT CAGAAAAAGT TAATTGCTGT TTTTCAAAACC  
 105781 GGGCATTATA GCAAGTGTCA CAACCCAGGA CATAATTTAT ACTTAGACCA GTAATCATTT  
 105841 ATGTGAATAA AAGTGCTCCA TCTCCATGGC ATTTGATCTA ATCAGGATAC AAGTCCATCT  
 105901 CTAAAAATAA GATAAAAAAT TTTTAAAGAA CTACCATTAA TCAAGTACTT CCTGTGTCTG  
 105961 GTACCGTGCC ACACCCCTTG CATACAGTAT TCAGCGTCAT CCCTCCAGCA GTGCTGCGGA  
 106021 GTGGCCATTG TTGTACCTT TTAGAGACAA GGAAGCAGAG GCTTGCAGAT GTTCAGGAAT  
 106081 CATCTCAAGG TAAATGGAGC CAGGATGGCT TCTCCACAA TCATGATCTC CTCTGCCCCC  
 106141 TGCATTTGGA GGCACCTGGC TGTTCACGCG TCACTCAAGG TTCATTCCCA CAGACCAGCC  
 106201 CCCACCCCGA GGAGGGCTGG AGTCAGAGAA GAGCTCACTT TTTCTCCTGT GGCTCTCTGC  
 106261 TGCGAGGCAC GTACTTCTAG TTCTCTTTCC TCTGTCCTATA CCGCCCAAGT GGACAAGTCG  
 106321 TTTTCTGCT AGTAGAAGCC CAAGGCATTG TTCGGGCGAC ATTGTACTAG CGCTTGGGAT  
 106381 TCATGTGGGT GTCTTCCCCA TCTTCTGTG CCCAGGCAGT GTCAGGTTGG CCTGTCTCTC  
 106441 GCCCCCTCCT GACTGCTCAT GACCAGGTGG CAGGTAGAAT AACCAACTTG CCCCACCTTG  
 106501 CCCAGGACCT TCTCCATTTT TGCCTGAAA ATCTCCCATC TCAGGAACCC CCTTGGCTCC  
 106561 AAACAAGCTG CCTTCCCCC TGGATCTTGG GAAAGTCAGG CCCAAATTGC ACCACAGATT  
 106621 CAGCTGTCCT AGATCCTGTG AGAGGATGAG AAACCTCGTA GGCTCCTGG CTTATGATGC  
 106681 ACTGGCCAAT CCAGCAAACC AGAGCTCTGC GTTGGGATTG AGGATTAACA GGTGCTGTGC  
 106741 CTGCTGTGA AAACAGTCTT TGGCAGCATG TAGGGTAGTC CAGCATGTGA CTTCTGCAGG  
 106801 CAGGTGGCCT GGATTCAGAT CCTGGCTCCA CCTCCTGCTG GCTCTGTGGC CTTGAGCAGG  
 106861 TTATTTCACC AACCTGACCT CAGAACCCAA CGCTTAGCAC AGCACATACC TCACTGTGGC  
 106921 TGTTGTGAAA ATTCAATGTG AAAATGCAGA TGCTTGGGAG CATTCAGCCC ACAGGTCTGT  
 106981 TGTCCTGTGA GACAGCCCTT GCCATTCCCT TTGCTTCTTC AGCTGCCAAA CTGCTGCCCA  
 107041 AGTTCTGATC CTAAGACTTA CTTATTTTAA GTATGAAATG CTTCAATAAT ACAAGGAAT  
 107101 ATATATCACA CATTATATA TAAAATGTGA AGAGTAACAA AATATATACC AATGTACTCA  
 107161 AACCTCCTCT CCTTCAGCTT AAAATAGAAT ATTAAGTACC AGTCATAGTG GTTCACACCT  
 107221 GTAATCTAGC AGGTTTGGGA GGCCAAGGCA GGAGAATCGC TGGAGGCCAA AAGTTTGAAA  
 107281 CCGGCTGGG CAACATGGCA ACACCCACC TCTACAAAAA AATACAAAAA TTAGCAGATG  
 107341 GTGGGGGGTG GTGTGCACCT GTGATCCCAG CTACTTGGAA GGTGAGGTG GGAGGATCGC  
 107401 TTGAGCCTGG GAGGCAGAGG TTTCAAGTGA CTGAGATCAT GTTACTGCAC TCTAGCCTGG  
 107461 GTGACAGTGA GACCCTGTCT AAAAAAAGAA ACAAAAGGAG AAGAAGATTA CCAATGCTTT  
 107521 TGTATATCCC TCTAATCATA ACCTTCCCTG TAATGATAAT GAATCTGAAT TTTGCCATTT  
 107581 CCTCACTTGA ACGTATCATT TCTTCACTTT TTGTTATATG TTTACTACAT ACGTATGTAT  
 107641 TGCTGATAAC GTATTGTTAT GTTTGAGATA TCTGGGTCTT GATGCAATG CAATCATACT  
 107701 TATTTTCTTT TTTCACTCAA CATTTGTGTT TTGAGATTCA TCCATGTTAA TATATGAAGC  
 107761 TGTAGTTTCT TATTTTCTCT TGCTGTACAG CATTCCACTG TTGAATGTG TGGTCACTTA  
 107821 TTGATCTACT CCCCTATGCA AACAGTGCCA CCACAAACAC TTTCAATAGG TTCCAGGTGC  
 107881 ACGTGTGCAC AAATTAATCT ACAGTATCCA TCTAACCCGA AAAGTCTTCA CTGATGGCTC  
 107941 AGTGGCAATG GCCTGCCCTC TGGTTTGGTT TTCATTGTGA CCAGAATTCT GAACCCAAACA  
 108001 CTCTCTGCTC CATTAAAATC TTTTCTGTGA TTTTAAAGTA GGATTACAGGA TTATGAGTGA  
 108061 CAAAAAGCTC AACTCAAGTG ATATCAGGCA AAAAAAAGAA AAAAAAAGG TGGGGTAGAG  
 108121 AGAACTGTAT TGGGATGAAT GATGGAGCAG TTTAGAGGGG ACTGGTTGAA TCTTAGACTA  
 108181 AGCACTGTCC ACACAGCCCT GTGGATCTCT CGGTACAGG GCTTGGCCAC CCCCCTTTTC  
 108241 TTTCCATCCT TGGACAGACT CTTCCCTCAT CGGAAACAGA TAGTGGCAGT CACTTCTGCC  
 108301 TGCAGTCCCC AGGTCTGAGT CCAGTGGGAG GGAGCATCCA TCTCTGCAGC CTTAAAAAGC  
 108361 CTCCTGGTGT GTCCTGGTCT ACCTGCTCAT CCCTGTGCTA ATCTCTGTGG CCGGAGGAAC  
 108421 GCAGTCCATC CCTCGAGTTA GGGTGCAGTC AGCCACTGCC CTAAGCTTTC GGGCTGAGAG  
 108481 TGAAGTCAAG GGGATTCTT CCCAGGGAAA TAGGGGTACT GTTCTTAAAA GAAGGGTAAT  
 108541 GACCACAGCG GGGCCAAAAC AGTAAGTGTG ATACAGGCGT GGTGGCTCAC ACCTGTAATC  
 108601 CCAGCACAAAG GAAGCCAAGG TGAGCAGATT GCTTGAGGCC AGGAATTTGA GACCAACCTG  
 108661 AGCAAAATAG TGAGATCTGT TCTCTACAAA AAATAAAATT AGCTAGGTGG GGTAAATGCAC  
 108721 AACGTGTAGT CCAGGTACTC TGGAGGCTGA GGCAGGAGCA TCACTTGCAC CTAGGAGATC  
 108781 AGGCTGTCAG TGAGCTATGA TCATGCCACT GCACTCCAGC CTGGACAATA GAGTAAGAGT  
 108841 CTGTCTCTAA CAAAAACAA AAAAAAAGTG AGTGTCCAAC GCGCCCTGGG CTTGCTGTGT  
 108901 GACCTTGGAT AAGAGACTTC ACATCCCAAA TGCCCAAGTC TTCATCATTT ATGAGTATGG  
 108961 CAGAGCTAAA GGTCCCTGAG ATTCCATCTG ACCTTGACCT CTGATTGTGT GGGCTCTGGC  
 109021 TGCCCTGTCA CCTCTTCTCT TGCCACTCCT TGCCCTTTA ACATCATAGC TTGTCCAGTC  
 109081 CTCCAGATCT TTCCACTATA ATGTGAGTGT CATACTTCTG GGATCACCTC TGTGAGCTCA

```

109141 TTCTCTCTGC TGACAACATG GCTTCTCAGA GGTAGCACAC AGGACAACAG CTCCTCAAAC
109201 CGCACCCCTCT GCCAATACTG GGAGCCTCTC CAGAGCGCCT CATCATGCTC GCCACCACCG
109261 GCCTTTTCAGG CCACAGCCTT CAGAGTCTTC TGTGGATCTC CATCTTCCTT AAGCCCCTGA
109321 ACCTGCTCTC TTACGGAATT TTGGCCCTTG TTTCTCTCTC TCAGAGTTCT TCTTGCTTGC
109381 TCACATCAAT GAAGCCAGCT GCCATATTGT GACCACCCCG TGGATAGGCC CATGTGGCAA
109441 GGAACGGAGG GTGGCCTAGG GACACATGGC CCACGAGGAG CTGACCCCTG CCAGCAGCCA
109501 CATGCATGAG CCTGGAGACA GATGCCTCCG CAGCTGAGCC TCGGGGCGAG GTCAGTCTCTG
109561 GCCAACACCT TGATTGCAGA TGGCCAGTTA AGTCATGCCT CAATTCTCTGA CTCACAGAAA
109621 CTGATTGAGC TCCAGAAAAC CATTGTCAAG AAACAGAGAT AGTGGGCATC CTGATCTTGT
109681 TCCTTAGCTG TAAGGAAATG AATCTAGGAT TTCACTTCAC TATTCAGGAT GACAGGTTTT
109741 GTTTTTTCAT CAAATATATA CTGTATTTGT CAATGCTCTG CAGAGAAACA GAACCGAGAA
109801 AGAGAGAGAG AGAGAGAGAG AGAGTGTGTG TGTGTGTGTG TGTGTGTGTG TGTGTGTGTG
109861 AGTGAAGCGA GAGAGAGACA TTTTGTTTTT GTTTTGTGTT TTTGAAATGG AGTTTGTCTG
109921 TTGTACCCCA GGCTAGAGTG CAATGGCGCA ATCTCGGCTC ACTGCAACCT CCACCTCCCG
109981 GGTTCAAGCA ATTCTTCTGC CTCAGCCTCC CAAGCTGGGA TTACAGGCAC ACACCCACAT
110041 GCCCCGGCTAA TTTTGTATT TTTAGTAAAG ACAGGGTTTC ACCATGTTGG CCAGGCTGGT
110101 CTGTACTGCC TGACCTCAGG TGACCCACCC ACCTTGGCCT CTCAAAGCAC TGGGATTATA
110161 GGCGTGAGCC ACTGCACCCA GCCAAGACAT GTATTGTAAG GAATTAGCTC AACGTGGTGG
110221 AGGCTTTGGT AGTCCAGAGT CTGATGGGGC AGGTTGACAT GCTGGAGACT GAGGGAGGAG
110281 TTGTAGTTGG GGTCCAATGG CACTCTGCCA ACAGAATTCC TTCTTGCTGC GGGGAGGTCC
110341 GTTTTTGTTC TATTAAGGTC TTCAACTGAT TGGTTGAGGC CAACCCACAT GGTGGAAGAT
110401 AATCTACTTT ACTCAAATC CACTAATTTA TTTATAATTA TTATTATTGT TATAGAGACA
110461 GGTCTCTACT TTATTGCCTA CTCTGGTCTT GAACCTCTGG CTTACGCGA TCCTCTCCCC
110521 TCAGCCCTCCC AAAGTGCCGA GATTACTGGT TTGAGCCACT GCATTGAGCC CCAAATCCAG
110581 CAATTTAAAC GCTCACCTCC CTTACAGAA ACACCTAGAA TACTTTTAAA CCAAATCTCT
110641 GGGCACTGTA GCGCAGCCAA TTTGACACAT AAAATGAACC AGTACAGTTA CCTACATCA
110701 GGTATGTCAG ATGCCCTAAT TTATTAAGAG CTGTTTATAA CTTCCGAGTT TGTGTTGAAT
110761 TTTATCAAGC AATTACTTTT ATATTGTTGT TCTCATTTGT TTATACATTG TTTTAAAAAA
110821 TTGTAATTTT TCTAATATAA AAGCATTCTT GCATAGTGGA GATAAATCCA ACTTGATGAG
110881 GAAAAGGGAG CATCTCTCAG TTCATCTTGC TCGCCTTTCC CCGGCCACA GCCTGTGGTG
110941 AGCAGAGGGC AGAGCGCCAG CCTTGTCTTG GGAAGTGGTA CAGACAGCCT GCTGTCTTCC
111001 TGTCTCTGAG CTGGCCCAAG GAAGGCCCA CAGGCTGCAG GAAGAGCACT GTGTGATTTA
111061 CATAGATGCC TTTGTGTGGG GGCCTGACAG CTGCTGGAGT CACGGTCCCT ACCCGGGTCA
111121 GGGTGGGGCT GAGAGGGCCT CCGGGTGTGTT GGGTAGTCAG GGGCTGGGGT GACAGCAGCT
111181 AGTGGGGGGC GGGTATATGA GACAGATAAA CATCCACACT CACATGCCAG GTGTGGCTTT
111241 GAGCCTCAGC CCAGCCCACT CTCCTTGGCT GTCTGGCTGG CTTTATTGAA AGTTCGCCAA
111301 ATGGAAGGGG TCTTTTTTTT TTTCTACCAT GCCGGCTCAC TCTCCAGAGC ATCCGTGCCG
111361 GCAGTTGGCC AAACCTGAG GTTGGGAGCA CAGTCTCCA GACCACCACA TCTCTGGGA
111421 CGCCTGACCT CTGCAGCAGA GGTACCAAAA CTGTCCAGTC CGAGGGAACA GTCCCCACAA
111481 GACTGCCCCA ACTTCAGCCA CCAGCGCAA GTTCAGAGT CCCAGGGCCA CCTCATTCTC
111541 AGACCACTG GCTGCAAAAT TGGGGGTGCT TACAACACC CTGAGCTTCA GTAATTTTCT
111601 AGGGTGACTC ACAGAACTCA GGAAAGCACT AACCTTATGA TTACAGTTTT ATGATCATGA
111661 AAAGATACAA ATTAATAACCA GCCAAAGGAA GAGAGGCTTA GGGCAGATC TGGGAGGGT
111721 CCAGTTGCAA GCTTCCAGGG AGGTGTCCTC TCTCTTGGG GTTCTGGACA GCATTGCCTC
111781 CTCTGGGCCG CAATGTGTGA GGATACGCAC AGGGTGTGTG CCACCAGGAA AGTCTACCTG
111841 AGACTTTGGT GCCCAGAGAC TTTACTGAGG CTTGATCAGA TGCTGCCTGT GTGACCGACC
111901 TTTAGTCTCC AAAGCCCCC AGCAAACAAA GACACTCCTG TCAGACAGGA CATTCACGGA
111961 CCCAGAGGTC ACCTCCCACT AGCCACGGGC AGAGACTGGA CCTCTCTTTG GTGAGGTTAA
112021 TTTAAACTA CGAATAGGGC ACGTTACAAA GTCAGTGCCC AAAGACACAG AGAGGTGTGG
112081 CTGTCAAGAG CTCCAGTCCT GGATTGCGCA CCTTTGAGGT TAGCAAGTTA CTTAATCCCT
112141 GTGGAGCCTC ACCAGCCAGA GGCCAGCAGG AACAGCCCAG TAGCCAAACA AATTTGAGTT
112201 AACTGGTACC CATTGCACTA AGAAGACTGT AACTAGGAG GCAGTGCAGT CTCAGTAAGT
112261 GTGTGTGTGT GTGTGTGTGT GTGTGTGTGT GGTGGGGGGG TGCTTATTGA AGGATTTGGG
112321 CGTTTGTAGG GTGACATTAC GAAATGTTCA AGGAAGCAAA GTGTTGCTGT GGATTGGGTG
112381 CTGTCAAGGA GTGGGCGGAG TCTATAATTG AGGATCTTCA CAATTTTAT CCAGGAGGTG
112441 GAAGAACGGG AGGGACGAAA AGCCTGGTAG AGAAACCGCA GTCTCTCACG TGAACAGGA
112501 GAGGGCACTG TTTGGCCATT TTTATGGTTT CACAGTGCCC TCATTTTTTG TCCATGCTCA
112561 GGCATATTTA TGGAGTGGCC TTGATTCTGT CCTGCTGCGT CACGGGGGCG CGGTGACCCT
112621 GCCTGACACA GGTGCTCTGT GAAGTTTGTG TGCATCAGGC GGACACCAGG GACCTGCAGA
112681 CGAGTCACAG GATGCTTTGT TCTGTTCCTT TCCCAACCTT GAGTTTCATC ACTGATAACA
112741 AGGAGACGGT AATTCTACCT AACACATTAG GCAAAATGTA GAATTAAGGG GAATAATGTA
112801 AAATGCCCAG CATAAAATAA AACACAACAG CTAGTCAATT CCCATATCAC TGTAAAAAAA
112861 ATAAGAGTCC ATGATAAAGG GAAAAAGTGA AGAATCAAAA TAAAGCCAAA TGCAGAAGAG
112921 GCACAGGGAA GGAAGCTGAG AGAGGACATC AGATGAGGGC AGTCCCTGAG CCAAGGTCTG
112981 TCCCGAAGCT CTCACTACAA CCCCCACCTG GGGCCAGGTG TGTGTGTGTG TGTATGTGTA
113041 CACGCACCTT TGCATGAGGG AGGTAAGAAA TGTGCTAGCA GCTGGCCGTG GTGGCTCACA
113101 CCAGTAATCC CAGCACTTTG GGAGGCTGAG GCAGGAGGAT TGCTTGAGTT CAGGAGTTCA
113161 AGACCAGCCT GGGCAACATG GTGAGACCTC ATCTCTACAA AAAACCATTT TTTAATTTA
113221 GCCAAGTATG GTGTTACATG CCTGTAGTCC CAGCTACTTG GAGGGCTGAA GTGGAAGGAT
113281 CACTTGAGCC CAGGAAGTTG AGGATGCAGT GAGCCAGATT GCACCATGTA ACTTGGCCTG
113341 GGTGACAGAG TGTGATGAGA TTCTGTCTCA AAAAAAAAAA AAAAGAAAGA AAAAGAAAAA

```



113401	GAAATGTGCC	AGCAATAAAC	AAAAAGGTTT	ATATACCATA	ATCAAATGGA	AACTGCCCCA
113461	GGAATGCAAA	GTTGGTTTAA	TATCCAAAAA	CCAATTAATG	CAATAGAGTA	CATTAATAAA
113521	ATACACGACA	AAACCACATG	ATCATGCCCA	TAGACACAGA	AAAATCACTT	GACAAAACCT
113581	AACACCGTTT	CATCATAAAA	CACCTCAATA	ACTAGAAATA	AAAGAGAACT	TTCTCAACCT
113641	GATAAGGATG	TCCAAGAAAA	ACCCACAGCT	AAAATCACAC	TAAATGGTGA	GAGACTTATG
113701	CTTTCTTGTC	AGAGTCAGGA	ACAAGACAAG	ATGCTGTGTC	TTGAAACGTG	TGTTTCATCAT
113761	TGTACTGTCT	AGGGCAATTG	GGAAAAATAA	ATAAAAGGCA	TTAGATCAG	ACAGGAAGAA
113821	GTAAGACCAT	CTCTATTTGG	AGATGGCATG	ATCTGTCTAC	AGAAAATTCT	AAGAAATCCG
113881	CAAAAGTATT	AGAGCCGGTA	AACAACCTTA	GCAAGATTGC	GGAATACAAG	ATCAACATTT
113941	AAAAATCCAT	TGTATTTCTA	TAAAAATAGCA	ATGAATAATC	TAAATGAAAT	TAAGAAAGTA
114001	ACTCCACTTA	TCAAAAAGAA	TAAAAATATT	AGGAATAAAT	TTAAACAGT	ACAAGACTTG
114061	GATACGGAAG	GCAGGAGGAC	AAAGAGGAGA	AGGACAAGGA	GAGGAGGAGG	TGGAGGAGCA
114121	GGAAGAGAT	GAGGAGGCAA	TGTGTTTCAG	ATGCTTGTG	ATCCCGAGG	CCTGAAGCCA
114181	CCTGTGAGGT	CAGGCATCCC	CTGTCTTCC	TGTCCCTACT	CCTGATCTGC	TGCCTCCGGC
114241	TGTCACCTCA	AATCTATCCT	GCTCAGGCAT	GGAAGTAATT	GTGGTGTGAA	AGTCTGTCTG
114301	GTTGGACCAG	CTGTTCTACA	CACTGGCCTG	GCGAAGTTAT	TTTCAACTCA	TCTCAAGCCA
114361	GCCATGAACG	TAAGAGAGAA	GTCTGTGTGT	GCAAGTCCCT	CTGGGCAGTG	AGACCAGCTA
114421	CAGAGGCCAC	CCAGGCCTGG	GAGAGCCCAG	AACTCCAACC	AGGAAGGCCT	GTGGCTGAGC
114481	CTCAGCTATA	ATCAGGTTAT	GCCCGCAGGC	CCTGGAGGGT	CCTGCAGCCT	CTGAGCAACC
114541	<u>TCAGGCTCTGT</u>	<u>GCCCTTGGAG</u>	<u>GCCTCGCCTC</u>	<u>TAGAGTCCGG</u>	<u>TGCTTCAGCG</u>	<u>CGGTCTCAG</u>
114601	<u>GGCCTGCAGG</u>	<u>ATGGCACCCCT</u>	<u>GGCTGGTCTG</u>	<u>CACGTTGTAG</u>	<u>TCAGTGCAGC</u>	<u>GCAGCAGGCG</u>
25 114661	<u>CTCGAAGTCT</u>	<u>TCCTCCTTGT</u>	<u>AGGTCATGTT</u>	<u>GGACAGGGTG</u>	<u>TAGGGGAGG</u>	<u>TGGCCCCGGT</u>
114721	<u>GAGATCCACT</u>	<u>TGGCCACCCCT</u>	<u>GGAGCTCTGC</u>	<u>GGGGCTGCGC</u>	<u>TGGACACCTG</u>	<u>CCCAGGGGTA</u>
114781	GGAGGGGTGT	CAGGAGCAGG	ACGCCGGGGC	TTCTACAGCC	TGGCAGCGGC	GCACAGGGAG
114841	GGCTGCCCCC	GCTGGAGGGG	CCCCTGCTTT	CCTCCCAGCT	CTCTGTGTCC	TTTGGCCAGG
114901	AATGTAAATG	TGGGTGGCAG	GGTGCTCTTT	GGATGGGTAC	TATTTGGTGG	TAGAAAATAC
114961	AATTTTCATAT	TTCATTGCCT	TGTGGCGTGA	CACCCCTTAG	GCTATTTCAGG	CAGCCCCACA
115021	TGTCTCTACA	TCGCTTCGAG	CAAGGTGGA	AACAGCCAGA	AGGGGGAAGT	CAACTGGGTT
115081	TGCGAATGGA	TTCTTAGAAC	TCACGGTGA	AGCAAGGGGG	CAGCGAGAAC	TTGCTGAGAG
115141	ATGAAGGTTT	TTTAGAATTC	TCAGATGTGG	GCTGAGCATG	TTAGAGGAA	AGCTGGGAGA
115201	TCTAGAATGA	TGGTAGGAAC	TGGAAAGCTG	GGAATATCGC	ACAAAGGCAG	CAATGAGTTT
115261	TTCACACGAG	AAGGAAATAA	AATTCTGACC	AGTTGGTGA	AAGAGAAGAT	AGGAAGGAAT
115321	GTGATGGGGG	TTTAAATTTC	GGGAGCCCTT	GAGGAGAGGC	CCAGCTCCCT	GAGAAGGGGC
115381	CCAGCCTGGG	CTGGGGAAGG	ATGGTGGAGA	GGCAAGTGA	GTCCCCACA	GAGCAGGAGA
115441	GCAGAGGAGC	CCTGGGCAGC	TGGGGGTACT	AGGGCAAGTG	GAAAGGCTGA	GGAGCCCGGC
115501	ATCATGACAA	AGGGGACCCC	TGCATAGCGA	CCCTGCAAA	AGAGGGCCTG	CAGGGATGTG
115561	GGCATCAGCA	AGGAGGCCAG	GGTGGGCTCC	AAGGCCTGAG	TGGGCTGGA	TGGGCTAGC
115621	CTGAAGCCTT	CTCCTGGCAG	ACTCATTTCC	CTTCCCTCTG	CTGGGCAGCC	GGGGGGCCCA
115681	GGGAGGCCCT	CTGAGGGGCC	AGGCAGCCTG	GGTGCTTGGG	AGGGGCTGCC	TCACCGGGGG
115741	CTGAGTGGTC	CTTGAAGGAG	GCATTGACCA	GCGGGAAGTG	CAGCAGGATC	GGGGCCTCGG
24 115801	<u>GGCAGGCGGG</u>	<u>GTCTGAGAAG</u>	<u>AGGTGGCATT</u>	<u>CCCTTGGCTG</u>	<u>GTGCTGGTCC</u>	<u>TGAGGGCTGG</u>
115861	<u>GTTCCACCCG</u>	<u>GGGGAAGGGC</u>	<u>AGCCCCCGGG</u>	<u>CCCGGCAGTA</u>	<u>CAGCTCCGTC</u>	<u>TGCTGCAGTG</u>
115921	CCTGGTGGGG	AGAAGGTGGC	CCGAGAGAG	CCAGCCCGGC	CCAGGTCAAT	GCTGGGCCAG
115981	TTCCAGCCCA	CACCAGCTGC	TTCAGTGTC	AGACAACAGC	CTGGCTCTGT	CCTGTTTGTG
116041	GTCGGGCCAA	GCTGCAGAGT	GACTCGGAGG	GAGACCCAGG	GGGAGTCCAA	GGACTCTGGT
116101	GGTCTGCAAT	GTTGTTTGGT	TAAAAGCTCT	AGGTGGGGTC	AGACCCATCC	CCCAGCCCCC
116161	CAAAACCCCA	ATTCTGGATT	GCTATTGCTG	GACTGTTCAA	ATCCCAGAGC	CGAGAGGAGA
116221	CCCACATTGG	CACAGCGTCT	CAGGGGAGGT	CTGGGGGATG	GACATGGAGA	TGCTAGGGTC
116281	TCCCTCATAG	TCCGTCCAGC	TGTTCTTTGG	AGTAGGTGTG	GGGAGGTCCC	TCCCTGTTCA
116341	CTGTCTCTGG	GAGCCCAAGG	CCCCTGCACC	ACCCCAGCCT	TCCCACCTAA	GGACCTGGC
116401	AGGAGCCAGC	AGATCAAGGC	CTGACACCTC	TCCTTCCCTT	CGGGACCCCG	GCGGTGTGGG
116461	GCAGGGGCTC	AGTGGCCCTG	CTGCTCGCCT	CTGGAGCTTG	GCCTGGCTGG	GCAGAGGGGT
116521	TCCCTTGGGG	TGACCAGGGT	<u>ACCTCGAAGG</u>	<u>GCGCAGATAG</u>	<u>GGAGTAGTCG</u>	<u>AAGGAGAGGA</u>
116581	<u>TGAGGTCCAG</u>	<u>CCTGCGGCCT</u>	<u>GGCCGGAACA</u>	<u>TGGAGGGAGA</u>	<u>GCTGGTGTGG</u>	<u>ATGAAGTAGG</u>
23 116641	<u>CGGCGTCCAC</u>	<u>CAGGCAGAGC</u>	<u>CGGGGCTCCT</u>	<u>TGGGGGTCAG</u>	<u>CTGGCTGGGC</u>	<u>ATGGAGTCAA</u>
116701	<u>GCTGTAGTCT</u>	<u>TGGTGGGAAT</u>	<u>CGCAGGATGG</u>	<u>TCAGAGGCCA</u>	<u>CCACCTTGCC</u>	<u>ACAGGTCTATC</u>
116761	CATGTCCCCCT	TCTCTCCTCT	GGGTCACTCC	TTAAATGCCA	CCCAGCCCTT	CCCCAGGACC
116821	TTCCAAGGGG	ATCGTCTCCA	GGCCACACCC	CAGAAGTGGA	TGCTGACCTC	CTAGGCCTCA
116881	GCCAGTCCCT	GCCACACCAC	CCAGAGGCC	AACCTTGTGA	CCTAGGGACC	CCTGGCCACG
116941	GCCCTGACCT	GCCCAGGTGG	AGAAGTCTTT	GTGGCTACAG	TAGTCTTGGT	GCAGCTGGAG
117001	CGCCTGGAGG	AAGTTGGGGC	TGCGCTGGTG	GAGGGGCTTG	CCTGTGAGGA	AGCCTTTAAA
22 117061	<u>TGCTGGGGCC</u>	<u>AGCGCCGTGC</u>	<u>CTGGCTGCAG</u>	<u>CCACGAGGCC</u>	<u>TCCAGCCGCG</u>	<u>AGGAGGTCCT</u>
117121	<u>CGAGGTGGTC</u>	<u>AGGGGCTCCT</u>	<u>TCTCTGAAGC</u>	<u>CAGAGAAACA</u>	<u>GAAATTGGTC</u>	<u>CCTTCTTCCC</u>
117181	CTTACTTCTT	CCCCTCAGAC	ACCGCACACC	GCAAGCATTC	TAAAGGAACA	GCGTCATCCC
117241	ATGGCCCTCA	GTGCCCTTGC	CCTTCCCACC	<u>TAAGCTCCTG</u>	<u>GTCTTGTCTT</u>	<u>TGATGTGCTG</u>
117301	<u>TTTCCAGGAC</u>	<u>TCCCAGAAC</u>	<u>TGGTGAGGTC</u>	<u>ATACCAGGCA</u>	<u>TCCAGCAGGT</u>	<u>TCAGGAGGAA</u>
21 117361	<u>AATGTTGCTC</u>	<u>CAGATGGCTG</u>	<u>AGGAACACCA</u>	<u>AAAGAGATCA</u>	<u>GCCTCAGGCC</u>	<u>CCAGCCAGCG</u>
117421	GCCCCACCCCT	CAGGTACTGA	TGGAAAATGG	GAGTGGGGCG	AGGGGTACCC	CTACAATGCA
117481	TCCCCCCCCG	CACCCCTCCC	CCTTGTCTTG	AGGTTCTGCC	CCACCTGAGC	TACAGGACCC
117541	ATGTCATTC	AGACACCTGG	CCAGGGGCTG	CCAGGGGCTG	TCAGAGCCAC	CACCCCTACC
117601	<u>TTCCAGAAAG</u>	<u>CAGATCCGGG</u>	<u>GCTCCGGGAT</u>	<u>CCTCCTCATC</u>	<u>AGCCGTCCCA</u>	<u>TGAAGAAGTC</u>



20 117661 GGAGCCGAAG AGCTCAGGAG GGACGAAGGC CCCGTAATTC AGGAAACCGA CCTCATAGGG  
 117721 GGAGAACTCA ACCCACTCTA ATGGGGTGGG AAGGAGAGGC AGGAGTGTGA GGGGAGTGGG  
 117781 TTGCGGGAGT CCCAGCTCTT GGCAGAAAGAC ACTGCCCTTG CCCCAGCAGC GCCTGAGCCC  
 117841 TGCTGTGCCT GGCACCAATG CAGCAGATGT GAGGAATGGG CAGTCCTGGA CTGACCACAC  
 117901 CCCTCCCAGC CTTTGGTCCA CACCTCAAG CACTACCCAC TCTCAAGCAC AGAGAGATGC  
 117961 CCTGAGCCCC TCAGCACCAC TGCCCTTCCG GAACATTGGG CAATCTGTGC GGAGCTTCCA  
 118021 GCATGAATTC TCTGGGGGCT GTCGGGGTGC AGAGGGTTTG GGGGGTGGAG TATGTGGGAG  
 118081 GGGAGTAGGT AGAAGTGTAC CTTGAAGTC CAGTGTCTCC AGATTGTCTT CTTTGACATT  
 19 118141 GAGGCTCAAG TAGAGGGGCA GAGGGTTCTG ACCCGTTCC AGGGCGGCTC TCTGTCTGA  
 118201 CAGCTTCTGA TCCATCACCT GGGGCCAGAG GGCATCAGGG CCTAAGTGAG GCTGGGAGTA  
 118261 CCCTGGAGGC AGTGGCTCCG CCAGGACAGG GGCAGGATGG GGGACCTGGC TGGGGGTGAG  
 118321 GAGGCAGCTC CTCCCGTGGG GTCTGGGGAC CATATGTGGG TCAGTTCCCA CTCCTGCCCC  
 118381 TGTGTCTGTA ACAGCCCTAG ATGCAGGGTC GCTGGACAGG GCATGCAGGC TGTGTACTAT  
 118441 GAATCTCCAG CGAGGGCCAT TCATAGGGG CACAATGCCA GCGGTGCTTC CAGGCCCCATC  
 118501 TTCGTGGAGT TAAGGGTGAA TTGGGTGTGG ATACTGACAC ACAGACACAA GGACACATGC  
 118561 ACACATCCAC GTATTGCTTT GTAGAAAGGA TGCAGCAATG CCCTTTTAGC ATATACTGCA  
 118621 ACCTATGCCA GGGGATGGAC TGCCCTCTCT TCTAGGTTCA CCTAATGCA GTTCAAGTTT  
 118681 CAGCTCCTCC CGAATGCTTT CCTGACATCC TTGCCTCCCC CTCATCGTCC CTCTTCTTTT  
 118741 TTGCTTTGAG TAACATGAGG TTGTACCCCA TTGATTATTC TTGCCTTCTG TCTTAAATTT  
 118801 CCTCTCAAT AGAGCTCCAA GCTCTGCCTT TGTTTTATTT TATTTTGAGA CAGGGTCTCA  
 118861 CTCTGTCAAC CAGGCTGGAG TGCAGTGGTG CAATCTTGA TCACTGCAAC CTCCACCTCC  
 118921 CAGGCCCAAG TGATCCTCCC ACCTCAGCCT CCCGAGTAGA GTAGCTGGGA CTACAAGTTA  
 118981 CACTTGGCTA ATTTTGTGAT GTTTTGTAGA GATGGAGTTA CATCATGTTG CCCAGGCTGG  
 119041 TCTCAAATTC CTGAGCTCAA GCAATCCACC TGCTTGGCG TCCCAAAGTG CTGGGATTAC  
 119101 AGGCGTGAGG CACCATGCCT GGCCCAAGCT CTTTGATGGC AACGACTATA GCTTGCCTTT  
 119161 TGGGGGAGCC TCCACAGAGC CTCCCTGAAT TGTAAGGGAA CAGAAAGGCC TCATTCTATG  
 119221 CTGAGCTAGA ATTGCAATA GAGCTCTTCC CAAGCACCAG CAAGGCGAAT GGTGGTGCTT  
 119281 GCCTGGAGCA CCACCTTGGG AAGGATTCTG CCGCCAGAGC TGAGGTTGAT GCTGTACGCC  
 119341 CTGGGCTGTT TCTACTGGGC AGTATGTGTA CCATGAATAC GTCATCCTTC CTTTGGGCTC  
 119401 TCCAAGATGG AAGAGTGTCT AGTCTAGAG GCTCCCAACC TGCCGCAGCA GCCCTTATCA  
 119461 TCAGAGTGAC TCAGTGAATG TGTACACGTG AGAGCATGTG AAGCATGCGG ACTTCTGGGT  
 119521 CTCCCTGAGC AATCAGCCTG TCAGAAAGTCA CTTTGGAGTT TGCTTTTAA AAACCCCTAG  
 119581 ATGATCCTGA TGCCACGCCA GGTGTGAGGA TTCTACCCA GCCTACCCAA GGAGGGTCCC  
 119641 TCTCCACAT CCATGGCGAA ACTCAACACA CCCAGTGTGT GTGTCTCTCT TGCTGGCAGC  
 119701 TCTTCTGTCC CTGACGAGAA CACTCATGTG TCAACCACAG TCATCCTTCC ACCCATTTT  
 119761 CCAATGGTCC CCAGTATGAA TTACAGGGAA TTGACAGGAA ATCTCTAAAT AGGGCTCGAT  
 119821 CCCCTGCCAC TGATGGTGGT GGTGTACGTA CGAGTCAGTC CTTAGGTACA AATGTGTGTA  
 119881 CGTGACCCCT TCTCAAATC TTTTATTTA GAAACTCTGT TGCTCTTTTG ATACATATGA  
 119941 CTAGGAAAAA TGTGTCTTTT TACTATAGTT GCACAATTTA TGCCCTAGGG AAAGTTCTGG  
 120001 AAAAAGGTCA GCACACCCCT CCATCTACTG AAACCCACCC GGTATATCTA ATGGCTCACC  
 120061 ACTAAGCTGC GGCAGGACTC TCAGTCTCTA ACCCCACAGT TTCCAAGCAA GTAACCCCAA  
 120121 CCCTGCAGAA CACCATCTG ATGATTGAG GGTCCCATGG GAAACACTGG AGCTTGGTGC  
 120181 AGTACCGGCA ATGCAGACTA AAGGGAAGGA AAGTGTGTCA GCTCTTAACT CAGGAAGGA  
 120241 TAAGAGAACC ACTCTGAGTT TCTCCTTGAG GAATGATAAG CAGCTTCTAA GCAGTTTCACT  
 120301 AACTTGCAGC AAAAGACTGA AGAAAAAGT TTGCAAGAGA TGAGAGTTTG AGAAGCAGTG  
 120361 ATTATAGCTA GTGTATTTTA CGAGCGGGTT TTTATAAAAA CCTAATCAAT ATTAAGATG  
 120421 TATGTAGACT AATAAAGTCA TTCTGAGCAA AGTTTCAGGG TATGATATTT ATTAATATTA  
 120481 ATGAAATAGT GATTGGTCAAT AAATGGAGTT ATATATCTG ACATTGTAAG TGATTTTGG  
 120541 TATTTTAGAT TATTACATTA TAGAGGTAAT AAATACCTGT TTTTATCTAA AGTCTTTAAA  
 120601 AAATTAGAAA TTGACACTAA AATTGCCCGA AGAAAAACATT ATTTTGTAAA GTCTACAATG  
 120661 AATGTTGGCT TTAAGAGGTA AGGTATGATA ATCAGAAAGG AAAAATAATA AATTAATAAC  
 120721 CTTTCATATT GGGACAAAAA TTTAAAAATA GCATAGGTGC TTTTGTGAG ACAGAGTCTC  
 120781 ACTCTAAAGC CCAGGTTGGA ATGCAGTGGG GTGATTACAA CTCACTGCAG CCTCAACCTC  
 120841 CCATGTTCAA GCTATCCTCC CACATTAGCC TCCTCAGTAG CTGGGACTAC AGGTGTGCAC  
 120901 CACCATGCCC AGCTAATTTT TTAACTTTTT GTAGAGATGG GGTCTCTACTA TGTGTGCCAG  
 120961 GTTCATTTTT AAGAATATAT AGGTGTCTAA TAATACCACA AGCATTCTTA TTTTGTGAA  
 121021 CAATTAATAA ATGTCTTCTA AAAAATGTCT ATAAAGCTTG CACATAAAAC CGTGCAATTA  
 121081 GTCATTACTG GAGTATCCAA TGGTAATTAA ATAAAAGTCT GCTTTACAGC CTAGGATCAA  
 121141 ACTGCGTTTT CTAGCACCTT GTAAAGTGAA CTGTTAAAC AGCTTCTCCC ATTCTGTAGT  
 121201 AACTAATTGC TGCCACTCAG CAGGAAAGTC AGTTATTAAA ACACAACACC ATCAAACTG  
 121261 AATATGAAT TAAACAGTAG ATCCAAAGAC TACATAGCAC ATAGTAATAA AATGAAAAA  
 121321 TTTCTAGTAA TTAAGTTGGC TTCATAAATT CAATGCCATT TCAATAAAAA TCCCAACAAC  
 121381 ATTTATTGCA AAACATGACT AGGTGATTCT ATCATATAAT TACAACAGCA GGCCAGTCAT  
 121441 AGTGGCTCAC ACCTGTAATC CCAGCACTTT GGGAGGCTGA GGCAGGCAGA TCACTCAAGG  
 121501 CCAGGAGTTC AAGGCAAGCC TGGCCAAAGC AGTGAACAC CATCTCTACT AAAAATACAA  
 121561 AAATTCGCGC TGCATGGTGG TGCATGCCTG TAGTCGAGG TATGCGGCAG GCTGAGGCAC  
 121621 AAGAATCACT TGAACCCAGG AGGCAGAGGT TGCAGTGAGC CGATATTGTG CCACTGCACT  
 121681 CCAGCCTGGG TGACAGAGCA AGACTCTGAG GAAAGGAAGG GAGGGAGGGA GGGAGAGGGG  
 121741 AGGGGAGGGG AGGGGAGGGG AGAGGAAGAA GGAAGGAAGG AAGGAAGGAA GGAACAGCA  
 121801 AAGGGTCAAG TACAGCCAAG AGAATTCTGA AGAACAATT AAGTGGACTG GTAACCTGGC  
 121861 CTATCAGATA CCAAGATTTA TCACAAAAC TTAGTCAGTG TGCTGTTGGC ATAGGGATAG

121921 AGAATCTGTC CTGTTGGAAC AGAATAGGGC ACTCAGGAAG AGTCACAGGA ATATAAATAC  
121981 ATGAAAAATA TATGAAACCT GATACATGAC AGAATTAGCA ATGCAGATCA GTGGAGAAAA  
122041 AATTCAATAG ATGAGACTGG GAAAAATTGT TATTAATGTG AGAAAAACA AAATTGGATC  
122101 CCTATCCACA TGATACCTGG AGATAAATTC TGGGTGGATT AAAGTCTTAT TAGTAAGAAA  
122161 TAAATCTTTA AAAATTTTAT AAGAAATGTAT AGGAAGATAT TTTATAACTT TGGGGGAGTG  
122221 AAAAGGAAGA CTCCCCAAA CACAAACCAT AAAGGAGAAT GATAAATTAG ATTATATTGA  
122281 AATTTCATG TAAAACCATT AGAACTGGTA AGGATTCACA TCGAGAATAC ATAAAGAATT  
122341 CCTACAAAAC AATATGAAAA AGACAAATTA AACAGGAAAA TGGAAAGGGG TTTGAATAGG  
122401 TAATTTCATG AAGAGGAAGC CCAAATGATA AATATACTTA AAAACATCAA CTGTCTAACT  
122461 AAAAATCAAA GAAATGCAAA ACTATGATAA ATGTTTCCTTT CACACCCCAT AGTTAAACAA  
122521 AACTTTAAAA GTCTGCCCAA GATATAGAGA AAAAAGCAAC TATTATACAC TGGTAGTGGG  
122581 AGTGTGAATT AACACAACCT CTTTGGAGAA TGGTTTGGCA ATACCTAGGA AAGTTGAAGG  
122641 TGCACAAACC CAGTGGCCCA ACAATTAAAC TACTGGGTAT GTATCTTGGA GGCTTAGAGA  
122701 ATCTTTAATG CAGGTGTGCA AGGAGACTTG TACAAAATTG TTCATGCAAA TAATATCATA  
122761 TTTTCTTAA GGATATATAT GTGAAAATAT TGTGAACCTG AAAACCTAGA AGAAACAGAT  
122821 AAATTCCTGG ACACATTCAA CCTACCAAGA TTGAACCATG AAGAAACAGA AAAGCTGAAC  
122881 AGAACAAATA CAAGCAATGA GATTGAAGCA GTAATAAAT ATCACCCATC AAAGAAAAGT  
122941 CAAGGACCTG ATGGCTTCAC TGCTGAATTC CACCAAAACAT TTAAGAAGT AATACCAATT  
123001 CTACTCAAACT TATTTCAAAA AATTGAAGAG GAGGGAATAC TTCCAAAGT ATTATCCAAAG  
123061 GCGAGAATA CATTGATACA AAAGCCAGAT AAGGACACAA ACAAAAAAG AAAGCTACAG  
123121 GCCAAGATCC CTGATGAACA TAAATGCAAA AGTTCTCGAC AAAATACTAT CATACGAAAT  
123181 TCAGCAACAC ATCAAAAAGA TTATTCACCA TTGATCAAGT GGAATTCATC CCAGGAACGC  
123241 AAGGATGGTT CAACATACAC AAATCAATAA ATGTGATACA TCACATCAAC AGATATAGTT  
123301 GGGCTCAAGT CCCTGGTTAC ATAATCAAAA ATGCCTTCCC AAGTCTCACC TCATTTCTAC  
123361 CTCTCCTAGG ACCTTTGTGA GCTATTTTGT TATTACTAAT CTTTGGCCCT TGTGATTGA  
123421 ACCTCTTAGT AAAGTTTGTG TCTTTTAGAT TACAATAGTT CCAGGTAAAG ACAATGCTGG  
123481 CACAAGGCTT CCAATCCATC CTGCCTACTG ACCAGGAGAA TTAATATGTT CTGCCTATGG  
123541 GCGCTTAGA TCAGGTATCC AGAGATTTT ACTCCTCTGT TGCTTGCCA GGGCTACGC  
123601 CCATAACATA AGCAGGAAGC AGTTACAGAA GACAGATTTC TACCCTTCTG CAGCCCTGTT  
123661 AAGATTAAAG AGGAGTGTCT AATCTCTGAG ACAGGAATGA GGTAGGAGGC AGGGCTTGAC  
123721 TCGAAGTACA TTGGAGACTG GCTGAAACAG GGAAGGGCTG AAAGCACCTC TCTGTAAGGC  
123781 ATGCTACAGA GTATCATGAC AGTTTACTCA TTCCCTGGCA ACAACCTGGA AGTTACGACC  
123841 TTTTCTAGA AATTTCTGAA TAATCTACTC CTTAATTTGC ATGTAATTAA AAATGGGTAT  
123901 AAATGTGACT GTGGAACTGT CCCTGAGCTG CTGCTCTGGA CACACTGCTG ATGGGTGGCC  
123961 CTGCTCTGCA GGAGCAGTCA CAGAGCTGTA AACTGTGCC CTTGATAAAG CTATTTTTTC  
124021 TACCACCAGT TAACACTTGA ATTTCTTCTT GGGGACACC AAGAATCTCC ATGGGTGGAA  
124081 CCCAGTTTTT GAGGCTCCCT TGCCCTGCAA CATGATGATA AACCTAAAT CAGGATGTTG  
124141 TTTGCCCTTG GGGCAGGAGT GAAATGAGAC TGGGGAGGGT ATACAAAGGG CTTAACTGT  
124201 ATCTGTAAAC TTTCAATTTCA TAGGCTGGTT GGTAAAGTCA TGCATATTCA TTATGTCATT  
124261 ATGCTTTATT CAATGTCTTA GGTGTGAAGA TTTTCATAAT AAATAATAAA GATTGAGATT  
124321 TGCATAATGA ATATGTATAG CTAATAACAA ATAAGGATGG TAGTTTCTAA GCAGTGTGT  
124381 ACTTAACCGG TAAGAATAAT ATTTAAAGAA AAAAGTTGAA GGCAAATTCA TTTTGGTTA  
124441 GTTTACACGA TAATCATTTCT TTTCAAATTC CTTCTATATA AATTTATTTT TATACCCATT  
124501 AACGAATTTT TCTTCTCTCC CTCCCATGAT AATCATACTT GATTAAACCC CTGACCCAAA  
124561 GATAAAGATG AGTCCTTCTG CTACTAATAT ATTCCATACT TTGCTTTTTC ATTGTTTCAA  
124621 ACGCTTTTGA AGCTTTGTTC AGATAGTTAT GTGTGGTAAA ACAACAGCGT GAAAAAATAG  
124681 CTGATTTTAA TAATTTGTATC TGTGTAAAGT TCAGCAGTTA ACTTACTAAG CTTTCTCTCA  
124741 AGCAGCATTG TAATTTTTTT AAAAATGTA TGTTTAGTTT TGTCTTGATG CTGTTAGAAAT  
124801 ACTAAAGAAA TATAATAGCT GACATTTAAT GCTTATTCTA TGGTGAGGTA TCACTGAGTA  
124861 TTCAGAACAG CTCTATGACT TAAGAATCTAT TCTTATCCCC ATTGTACAGA TCAGAAAATG  
124921 GAGACACAGA GAGGTTAAGT AACTTGCCTA AGGACACACA GCTAGTAAGA GGTACTGGTG  
124981 TTAGAACCCA GGCAATCTGG TTACAGAGTT TGTGCTTTTA ACAGCACCTT GAAGATGAAG  
125041 AAGAACAGTA AATTTCATAT TGCAGGATAA ATAAAAACGG TTCTAGCAAT CTACCCATT  
125101 CAAAACAAAA GCTTTGTTTG CTTAACTGAC TTTGACATCG ACTCATTTCC AGCTTTTGGG  
125161 AGATCCAGAA ATTTCTGTTT TACAGCAATT AATTAAAGAA TCTGCTAACT CAGAAAGGCA  
125221 GCATTCCAAA GTGAGGCAGA CACTAAAACA CTGGCTCTTC CACGCTGCAA GAAATACGCA  
125281 CGCGCATGTC AGCCGCTTGG GAGCCCTGCC TTCGCCCCG CGGTGCACAC ACGCGTCTCC  
125341 CCCACGTGCG CAGGCGCCCT ACCTGGCCGT GCAGCATGGA CTCCAGCACT AGCGCCCAACA  
125401 GGTCCACAAA GGTCTGTGGG TGGCCCTGCT CAGCCCGCAG CTCCAGCTCC CGGCGGTAGC  
125461 TCGCCAGGCG CTCTGGGGAA AAGACCTCCA GCTTGCTCTT GGCCAGGTGC TCCCGGGCGT  
125521 ATCTGATAGG TCCCTCCAGG TCCCTCTGCG ACCACTCAGG GTCCCCGTAC AGGTGGGCCA  
125581 TTGTCTTGGG AGAACGAGGG GACAGAACAG TAGGAGCCCG AGGCCAGCA TGGGGCCCTG  
125641 CAACTCTGCG TTCCCACTCG GCAGCTTCTC CTGACACGTG GCTCCTGCCA CCAAGGCTCT  
125701 GCCGATCTGG TAAATGGGTC TCAGGGCTTT CTCTCTTCT AGATCAGCGC TGTCCAAAT  
125761 TAATACGATG TGAGCCACAC ACGGAATTTT AAATTTCTTA GTAACCACAT TAAAAAGATA  
125821 TACAAGCCTG TAAAAATTAA TTTTAATATT TTTTAAATTT CGCTATATTC AAAAAATACA  
125881 TTTCAACAAG TAGTAAATAC AGTTATTGAT GTAATTTCTA TTTTTTTCAA ACTAAGTCTT  
125941 CCAAATTCAC TGTGATTTTA TAACACATCT CAATGTGGAC TGGCTGCATT TCAAATGGTC  
126001 AATAAACAACT GTGGCTAGTG GCTACCATAG TGGCTTACTG TACTGGACAG CCCAGTTCTG  
126061 GAGCCCACTC GTATCTTCTC GGTCTGTGCC TGAAGCCTTT TCACACTCTG CTTGTCAACA  
126121 GCTGTCACCT AAATCCTCTT GAGGCGTTTC CTCTCTTTT ACTCCTCAA GTCGCATTTT

126181	AATAAGACTC	AAAGGAATGA	ATCTCTAACA	GCGGGATGCT	GCACCTCGTC	ATGCTGCGGG
126241	GTGAGGGCGT	AGGGAGGTAA	TGCTAGCTTT	AGGTAGAAAT	AGTATCTCGA	CCTCAATGTC
126301	TACACAGTCT	TAAATGGACA	ATAAGGGAAT	CTACCTCGAG	TGTGGTGAAG	ATTAAATGCA
126361	ATAATAACAA	TGAAAGATAA	AAAGAGCTAC	CAATTAGGAA	TTCTTGCGTG	TGCTGGGCAT
126421	TAAGTTAAGT	GCTTTACTTG	GGTGATCCCC	ACACCCAGT	CCCACGCGGT	GTGTGCTAGG
126481	ATTGTCCGTT	TGGAGCTGCA	TCAACTGAGA	ATCAGAGGAG	ATAAGTTGCT	CAGTGTGCTC
126541	CAGCCAGTAA	GTGGTGCTGC	CAGGACTCAA	ATCCAGCTGC	ATCTGAGTCA	GCTACCTCAC
126601	AGCCTTTTCAG	TAACATCTGG	AATGCACCTC	CTGCGCCGCC	TCGCTCACGG	CAGGGCTTCT
126661	GTAAACGTCT	GTTTCCTTCC	TTGTGGGAAA	GCTGCTTGGC	GCCCATCAGA	TCACAATGAT
126721	CACACCTACT	TGGCCTCCCC	ACACAAAGTG	TCCCTCTGGT	GCCCACTCTG	CCCCGCTATG
126781	GCCCCGTGATT	GTCTCTGCCA	CCCAGTCCCC	CAGGATCCCTC	ACCACGTAGA	GCCAGAGATG
126841	CCACTGAAGT	AGGTACACACA	GTCTAGGAGG	CCCAGCTTCT	GCAAGGCCAA	TAGGTGGCCG
17	126901	TAGAGTGAGG	TCATGGCCCG	GGCACCTCCT	CCTGTGGCCA	TGATGCCAC
126961	TGGACCACAC	ACAGACAGGC	TGGATTAACA	AGGAAAGGGG	CCAGCTGCCT	CCTGTCTCAC
127021	CAGGGGCCCT	GCCCCGCTCC	TTGTCTCCAG	GGAAAGGACC	GCATAGTTGG	GTCCTGATAG
127081	AGTCAGGGTT	AGAGCTGGGT	CCAGCCCTCC	TCTGGAAGCA	ATGGCTCTCA	GCTCAGGATG
127141	GAGCTCCCAA	GGGAGTTAGA	ACCACAAAGC	CCACTCACGC	TCCCATGTCT	CCACACACAT
127201	GCCCACTCTC	CACTGACTGG	AGTTTCTGGA	GAGGCATAGG	AATCAGGTAC	TCCCTTATGG
127261	CCCCACAGG	ATATCTGCAC	ACACATCCCT	CTGCACCCCC	AGACCTCATC	CTCCTGACGG
16	127321	TCTCTGTCCA	GCTGCAGGGC	CTGCTTCAGG	GCCTTGGCCA	CCACCTGCTT
127381	AGGAAGGCCT	GCTCCTCTGC	ACAGAGATTG	AAGCCCAGGT	GCACGGCCAG	CTCCTCAGGG
127441	CTGTGGCAAT	GGAGGATCCA	GGGGTCAGGG	GACACCTGCA	TAGCTCAGCC	ACTGGCTGGC
127501	CAAGCCCACA	ATGAGCTGCC	AAAGAAGCCC	TCTCCTCCAA	TCTCAAGGAC	CACCTTTGTC
127561	CATAGCCCTC	CCCTCCTGCA	TGCCCGCCTC	TCACTGTCCC	CACAGATGTG	AATGTCCCTT
15	127621	ACCAGCCCTC	TGCCTTGAGC	TGCAGCCTCA	CTCCTGGGGC	CTGAAATCAA
127681	TCTGCTCAGA	CCCTCCTAGA	CCCTAAGCGG	CACATGGGTA	TCCCTGGGGA	CTTGGTCCCC
127741	TTCACTTTCA	TTCTTTTTTT	TTTTTTTTTT	TTTTGAGACA	GAGTTTCACT	CTTGTGTCTC
127801	AGGCTGGAGT	GCAATGGTGC	GATCTTGGCT	CAATGCAACC	TCCGCCCTCT	GGGTTCAGGT
127861	GATTCTCCTG	CCTCAGCCTC	CTGAGTAGCT	GGGATTACAG	GTGTGCGCCA	CCACACCCGG
127921	CTAATTTTTG	CATTTTTAGT	AGAGATGGGG	TTTGTCCATG	TTGGCCAGGC	TGTTCTCAAA
127981	CTCCTGACTT	CAGGTGATCC	GCCTGCCTTG	GCCTCCCAAA	GTGCTGGGAT	TACAGGCTTG
128041	AGCCACAGAG	CCCAGCCTTA	TCTTCATTCT	TACTGGCATT	TCCCATCCAG	TTGAGGCCAC
128101	TTACATTTGG	AGCAGGAACA	TCCATAGTCA	CCTCCTTCCC	AATGGTCAAG	GGCCTCAGGG
14	128161	GCACAGTGAG	GTACCCAGCT	GAGTTGTCCC	CATTCCAGCC	ATTGCTTCTG
128221	GGGAAAGCAT	CATTCAATCA	TTCATTCAAT	CATTCAACAA	ATCCCTACTA	AGTATCTAGA
128281	CTACAATCCA	GACAGACACA	TTCTCCCTGA	TAATACAGAA	TATTATAATA	TTGATAAGTA
128341	TCTAATTTAT	AATTATAAAA	TTATAAGAGC	TAATATTTAT	GTAGTACGAA	GGCATTGTTC
128401	TCACGGAATA	GCATGTATAA	ATTTATAAAG	ACATGGGTAA	TGAAGTATG	TTAGTAAGTG
128461	TTGTGAAGGA	AATAAGTGGT	GTGATATGAC	AGTTTCTAAT	GCAACCAAAG	GTGGAGTGCG
128521	TGGGGTTTTAT	TTGAGGTGGG	GGTCAAGGAG	GACCCCTCTG	AGAGCCACCA	TTTGAGCTAA
128581	GATGTTGAGG	AGGAGGAGGA	GGAGGCAGCC	TTGGAGGAAG	CCAGCTGGGG	ACAGGAGAGC
128641	ACTTGGCATA	TTGGAATGA	GAGAGGAGAC	AAGGATGGCT	GAGGCATCGT	GCATGGGGGG
128701	AGTGTGGAAC	AAAATGAGGC	GGCAGAGGTG	GACGGGTGCC	AGGCTGTGCA	GGGCCATGTG
128761	GGCACAGATG	TCATGTCTTA	TTCTAAGGGT	ATCACTGGAG	GGGCCGCTGG	AGGATTGAGA
128821	GACTGATATG	GTGCAGCAGC	CCCATGCCCG	CTCTGGCCAC	TGGTGGTCTC	CTGTGTCTGC
128881	CAAGAGTGAG	GTGCCTAACT	AGGGGAGGCT	GACAAGGCCA	CAGGGGCCCT	GGCCCTGGTG
128941	GGCAGGGAAG	GCAGGGAGGG	GCCCGCTGCA	GCTGGAGCTG	CCCAAGCCCTA	GGATTGCAAA
129001	CGAGGTCTAG	AGCCAAGACC	CACCCCTCAGG	CGCCCGCTCA	GCTCTGTCTC	TAGGGCTGCC
13	129061	ATGTAGTTGA	AGCGGAAGGC	AGAGGCTGTG	CCCAGGAAGG	ATGTCTGTGT
129121	GACCCCTTCA	GCACCAGCTC	CAGCTCCAGC	TTGTCTTGAT	CTAGGGAGAG	AGTAGGCGGG
129181	TTAGCATGAG	AACAAGAGCC	CGGAACCCCA	GCTCGGACCT	GGGACAGCAG	CACCACCACC
129241	TGGATTGAGA	GGATGCCTGG	GAGGCTGGGC	CATCGTCAAC	AGCATTCTGA	AAAGGGCCCT
129301	TCCAGCCCA	AGGGGCCCCA	CAACCCCCCA	TCCCGACTC	CTGCCCTCTG	CAACATGCTC
129361	TCCCGCTAGT	CAGGGGACCC	TCCTGCCCAG	CGGAGATGCG	TGGCCCTCTG	TGCCATGTGG
129421	GTGAGAAGAG	AGAGGCAAGC	TCTACCTGTC	TGCTGGCCTC	CTGCCCTGCG	CAGGCTTCCC
129481	AGCAATCCTC	CTCACCCCA	CATCCAGGCA	GAGTCCAGGC	CTGGTTCTAA	GCTGGTCTCA
12	129541	CCTGCAACCA	CAGCGGTGCT	CCCTGTGCTG	TCCAGATGCA	CATCCAGGCA
129601	CGGGCCTGGG	GAAGACCACA	TCATGGGCAG	GGGCCTCTGC	AGGGTAAGCT	GAGGTCCCTT
129661	TAACTCAAC	CCTCTGAGAC	CCACTGTCCC	AAATCCACCA	CACACACAAG	GAGTTGTGTG
129721	GGCTCAGGAT	GGGATTGGCT	GCAGAGGCCC	TTCATCCCAT	TGCCGATCCT	CTGCCGCGGG
129781	TGAACCTGAT	ATTCCGCGAG	AGCCAGGAGG	GCCCTGGCTC	CACACCTCCC	CTCCAGCTCC
129841	CCTGGCTTTG	CCAAACTCCC	GAAACCTCCT	AGAGCGCCCA	GCCCTCAGAC	ACAGCTGTTT
129901	CTTCTCTTGT	TTCCCTGTGG	CTTGGAGCTG	GAGGAGGTAT	TCTACCACAA	CTCCAAATAA
129961	TCTAAGCAAG	CCGCCCATCC	ATAGGAGGGT	AGACGGGGCG	TCGGACAAAC	CCTCAGAAGC
130021	TCTCTTGCTG	GTCCTCTGTG	GCCTAACTGG	GGGACAATGA	CAAGGCCAGG	CCCCCGTGGC
130081	CTTGCCAAAG	GAGCCTCTGC	ATGCTGTCTG	TGAAGCATAC	CAGCAGATGC	CCGCCACCCA
130141	CGCTCCACAC	AATCACTGAA	TTTGCCCAAC	TGCTGTGTGT	CTCCATCACC	TCCTAAGAGG
130201	CAGTGCCATC	GGCTTGACTT	TATTTTTCATG	TGTATTTTAA	AAAGTAATAC	CCGTCCATCA
130261	AACAGCCCTC	AGGAGCTTAA	ATGCTGGGCT	CTAGGTGCCT	TCGCCCCGTA	CCCTGGACCC
130321	GCCTGTCTCC	AGGTCTGTGC	TGCTGGTCTG	GGTCACTGTG	ATAATCGAGC	GCGTCGCCCG
130381	CTCGCCACCT	TGCGGCCATT	TCTGTCCACT	GCAGCTTCGT	GAGCCCGTGG	TCAGGAGCGC

	130441	TGACCAGGAC	TCTCCCCAGT	CCCAAGGAGG	AGCTAGCTGC	CTGAGCCTCT	GATCCTCCTA
	130501	GGTTGGCAGT	ACACCGCCCT	GGGAACTCCC	TCTGAGCCTC	GGTGACCTCA	ACGTGAACAA
	130561	CCGGAAGGTG	GGCCTGGCCT	ACCTGATTTC	TTCACCGCTT	CCCAGGGAAG	CACTCAGTCA
	130621	ATCCCTAGAA	CTATGAGTTC	ACCGCAGGTG	ACTAAACACC	TTCGGAAGC	ACCCCGAACA
	130681	GCCGTGGGCC	AGAACCTCCT	CTCAGCAGAG	CCCTGCCTCC	CCTCTCTCTA	TGCGCCCCCT
	130741	AGCTCCAGGC	CCCTCCAAGA	CCCACAGCCA	CAGGTGACCA	CCTGGCTGGA	AGGCTAGGGG
	130801	TGGTTTTACC	ACGAAGCCCT	GCCCTTCCT	TGTCCTTGCT	GGTCCACACC	CCAGGCAGCT
	130861	GCCCTCCCTC	CCCAAGCGTC	TGCTCTGGGG	CCCCTCCCCT	AAGCCTGGGG	AAGGTGCTTA
11	130921	CCACAATGAC	TTTGTGGTG	ATGAGGTTTT	CTGGGCGATC	TGACCTGGAA	AAATCAAACC
	130981	ATGCAAAAGC	AAACGCTTCC	TGCATTGACC	TCCCCTCCC	GCCCATGTGG	GGTCTCCCTG
	131041	GGTGATGCTG	AGGGCTCTGT	AAGTCTTGGA	TTCGGTTCAG	GGACAAGGGG	AGATGCTTTG
	131101	CAGGAGGAGA	GGGGAGGGGA	GAGAGCAGAG	AGAGAAAACA	TCGGCCTTCC	TCTCTGTGTT
	131161	TCCCTAAGAG	AAGCCCGGCC	CCACTCCCGA	CCTCTCTGGC	AGGGGTAGGA	GATGAGGGAC
	131221	AGTCAGGCTG	GACTGACCAC	AGGGAGATCT	GGGTGAGGGC	CATTTCTCTA	GAGCCTGAGT
	131281	CCTGAGACAC	TCCCTGGGCT	TGCTGACATT	ACTCACGTTT	CTTCCATCAG	GAACCTCACA
10	131341	TCCAGTCTCT	CCTCTCCCTG	AAATCAGAGA	GCATGAGCAA	GTCATCAGCA	CATACCTGTG
	131401	TCTTCAGAGT	TGACACAGTT	CTTTTACAGG	TGTCATCTCA	TCCTCTCAAG	CGGTCTCCAA
	131461	GGGAGGCACG	TTAGTACAAT	TATTTCCAAT	TTATGGATGA	GGAACCTGAG	GTATGGGGCC
	131521	AGATGGGCTA	CCTGCCCACA	TTGCCACAG	CTAGTAAGTC	TGCAGAGAAC	AGGTCTCCCT
	131581	CCCTTGGAGT	TCCCACAAC	CTGGGCAAGA	CTGAAGGGGA	AGACAAGGAG	CACAGCCAAA
	131641	GGAGGCATGA	TGTGCAGGAC	CCATATCTTT	TTCTTGAGGA	GGGCCCAGAC	CCCCGAACCTC
	131701	TGCCGTCCCC	ACCCCCAGGC	TGCAGCAGGA	AGGGACGTGA	GGCCTTCACA	AAGCAGGCCCT
	131761	TGGCACCACC	AGCAGGAGTG	ATGGATGCTG	AGCTGTGTTT	TGAGGGGTGG	CCTTCTTGCC
	131821	ACAGGTGGCT	AAAATTCCCT	CACCTGGGTGA	TTCAGCCTGT	AACTGTAAAG	TGAAGCCTTT
	131881	TCCTTGCTTA	TTAAAATGCA	AATCTAAAAA	GGGGTTTAGA	TTTTTCTGTG	AATGAGTTCA
	131941	CCCCAGCAGC	CTCCCAACAG	GTGGGAGAAA	TAGTACTTTC	CTCAGCTTGG	AGTTGGAAGA
	132001	AGTGGGGCCC	ACGGGGACTT	CCCCACCCAC	CCACCCACT	GGGGACTCTG	GGAGAAGGTT
9	132061	TTCGGAGACA	GCTTGCCAGG	GAGGACTTCT	GAGATGTCAT	AGAGAACCCT	GAAGCAGATG
	132121	TCATCCTCCG	TGACTGAGTC	CTCATCATAG	ATGCTAAGCT	CCAGAACATT	CTAGGAACCA
	132181	GGAACAGCGA	CTTAGCATTT	TACCTGCTCA	TAGTGAGAGT	AGCTCACCTC	TGTTGAGCCC
	132241	TTACTTGATG	TCTAGAGTCA	GAGAAGATCA	TTATGCCACC	ACACGTCTGC	GCTTCAAAAT
	132301	AAAAAATAAA	ACTGTAAAAC	TTAAATATAA	GGAAATCTTA	TTAAATTTCA	AATTTCTCTC
	132361	TTCTAATTTT	TGTTTTTATC	AAACATAAAA	TTCTGAGCTG	GGCATGGTGG	CTCATGCCCTG
	132421	TTATCCAGCA	CTTTGGGAGG	CTGAGGTGGG	TGGGTACCT	GATGTCGAGA	GTCGAGACC
	132481	AGCCTGGCCA	ACATGGTGAA	ACCCCATCTC	TACTAAAAAT	ACAACAAATT	ATCTGGACAT
	132541	GGTGGTGGGT	GCCTGTAATC	CCAGCTACTC	GGGAGGCTGA	GGCAGGCGAA	TCGCTTGAAC
	132601	CTGGGAGGCG	GAGGTTGCAG	TGAGCCCACT	TCGCGCCATT	CAACTCCAAC	CTGGGCAATA
	132661	AGAGCAAGAC	TCTGTCTCAA	AAACAAAACA	AAACACATAA	AAGTCTGTCA	AATAATTCCC
	132721	TCTTCTCTCT	TCCCTGTCTG	TGTGTCTCTC	TCTCCCTGGC	TTTTCTGGCT	CCCTGAGGTG
	132781	CACCCTAAGG	TATGTGCTGG	TGATCTTTCA	CGGTCTTTCA	CGCATTATCT	CATCTCACCC
	132841	CACAGTGACC	TGATCAGGAG	GGGGCTCTTA	CTGTCTCTGT	TTCATAGCTG	GGCTTCGAGA
	132901	GGAACCCAGC	TTGTGGGAGT	CAGTCCCACA	GCCAGTTGGT	GAAAGGACCA	TGTTCAATTCT
	132961	TTCATACATT	CCCTTGCACA	CAGCACAGAC	AGGCCAGGGC	ACAGAGACCC	CCTACTGCTG
	133021	AGAGTTTCTC	TGACAGCCCC	AGAGGAAGCA	TGCTCTACCC	CAGGAACCCT	TGCATGGGGG
	133081	TCCAGGCCGT	TCACAAGGGA	GTGGCCAGGA	TCCCGGGGCC	TTACCTTGA	CCTGACTTTG
	133141	GATAAGGAAA	CGGAAGGCCT	CATTCCACAC	AGGATGACTG	GTGTCGGTGA	GCGTCTTGGT
8	133201	CTTAAACTTC	ATTCCAGGTG	CGGTGACAG	CTGTAGGATC	ACGTAAGGGT	CGGCCTCACT
	133261	CACTGCCAAG	GTTAAGGAAG	TCTTCGTGAG	GGAGTACTCC	CCACACCCT	CTCCATGAGC
	133321	CATGCCAGTT	TCCTACCAAG	ACCCTGGCGG	TCACCCAGG	TGACCCAGCC	CAGCCCAGG
	133381	GCACGTGCT	TGGCCCTTCC	TATGGATCT	GGTCTCCCT	GCTCCCAGCA	GAGCGCACAG
	133441	GGCGGTTACT	CACACAGGTC	AGCCCAGCGC	AGGTTCCGCG	CCTCCAGGAC	CCTCACTGTG
7	133501	AGCTGCCAGC	AGGTAGAGGC	CTCCCCCTGA	AGAGAAAATC	AAACTCCAGA	GTCCTTCCTG
	133561	CATTCCCTCC	CTTATGCCTG	GAGCTGCAGC	CCGGGCTGCC	GACACCCTC	ACCACTCTCC
	133621	TGGTACTTGG	TGGTGCTGAC	ACCCCTGCCA	GGCGTTCCGG	GGACACCCTC	TGGGGACCAA
	133681	AGAGCCTTTG	TCCACTTCTG	TTTTGGGGGT	CATGGTTCTG	GGCTGCAGCA	GTCAGGGAAT
	133741	AGCAGCACAC	CCAGATGTGG	AATTTGTGGA	ACCCCTTCCC	CTTCGGGAAC	AGAGCAACAT
	133801	GGGTAGTGGA	CAGAGCAAGG	AAGTGAGGGC	CAAGGAGTGG	CAACCACAGA	GGTCAGTTAG
	133861	ATGTCAGCCA	TTCCCAGAGC	TCCCAGGGAG	GGCAGCCCTT	TGGCAAAAACA	AGGGTTACCT
	133921	TTGGTGACCC	TGCTGCCAGC	ACTGCCCACC	CCTCAGCCCT	TCCGGACGGG	AAGCAGGCAT
	133981	ATGACTTAAG	GGAGGGTGAG	CCTCAGCAGC	CCAGTCCACA	GGCGCTACCC	TGGATCTCAG
	134041	CTATCTGCAC	TGTAAGGGGT	ATGAATGAGG	AAACTCCAAG	AGCATCCAGC	CATCACAGGG
	134101	AGCAGTGGGT	GCTGTGGGTG	AAAGGCCACT	GTGTATTAA	ATGCTGATGA	GGCCACATCA
	134161	GCCACAGACA	GCCCAGGTGT	GAGGGAGAAG	CAGCTGAGTG	GGCAGAAGCT	GGAGGAGAAG
	134221	ATGAGCCGAG	AGTAGGTGGC	AGGAGGGAAA	GCAGATGGCC	AGAGGGGCGT	GGCTGAGGCC
	134281	TGCTGGGGTG	AGAGCACTGG	GGGTGTCCCC	ACACCCTACT	GCTGAGGTCC	CCGGCACTGC
	134341	CACGATGGTC	AGGGCCCTGG	AGTTTCTGTT	CTTCTGAGC	CCTGCAGCAC	GGCTTTTCT
	134401	GGGCATCTGG	CACAGGTTTC	TTATAAGCCC	CTTTGATGTG	AGGTTACTCA	AATCGGTTTG
	134461	GCAGGGCAAA	TGGCATTTTT	TCCTGTGGGC	TGACCCCAAG	GAAGGCTGTT	GGCCTTGCAT
	134521	CCTCTTGGA	GGGAGCCTTC	CCAGGAGGAC	CACCTATGAG	GCACGTGGCC	CCAGCAGAAA
	134581	AGGCCCATGT	CCAGAGGGTC	TGTTAGAACA	GGGGCTGTCA	GGAAGGAACA	CTGCCCTGTG
	134641	CGGGGAGTGG	GGCAGTGGGA	GCCCCAGCCA	GGAGGGGATC	CTGCAGGGCC	ACAGCATGGC

134701	AGGACCCCCA	GGAGGCACCA	AGCACATGAA	GGCCTATGAA	ATGCTAACCC	CAGGAGCCAT
134761	GCAGTGTGGC	GGCCCTGCTG	AACAACCCAG	AGACCACCTT	CATTCCAGAA	AAAAACCCGT
134821	GCTGGGGCTT	CTGCCATTCA	GCGGCTCTGC	CACCAGGTTA	GA CTCCCCAT	TAAGTAAGTA
134881	CGATCTTCTC	AGACCCATATG	TCCCTTCACC	CGCTCCTCCC	CGACCCCTAG	GGAGCCAAAA
134941	GCCAGTGCAG	TGTTTGCACA	GTGAACATTT	GATATCGCAT	AGCTGGGAGA	CTGTTGTGAT
135001	AAAACCCAGAG	GTGTTTCTCTG	CTTTGCGCTC	TCCCGTGACC	AGACCCACTT	CCATGCATGT
135061	GGTTGGAAGT	TCTCGGTGCC	CACAAGGACT	CGTACCCGCT	CTGCACAGAT	GCTGCTTGCT
135121	AGGGGCTGGC	TCAGCACCCA	CGGGGACCTG	AATCTCCACT	CCCCACCAA	AACTGTAAAG
135181	GACCATGGGC	CACAGCAAAAG	CCATATGCCT	GGTGGCCCTG	CCCCTGATC	TTTCTCTCT
135241	GTGGTTCCAT	GCCAACGTCC	TATCATTTCCA	GGCAGCTTCC	CAGCAGAGCA	CATGTGACCT
135301	CCCCGAGCAA	GGTAGATTTG	TTTCTCTCTT	CTTGGGGATT	AACTCCAGTG	GCACCAAGCAA
135361	CATGTCAACT	CTATAAACCT	TAAGCCTTAT	AGAGGTTGTC	AGCAGCAGAG	ATTTTCATGAA
135421	CTTTGGGACA	TCTAGCTCCA	TTTCTCTCCC	TCATCTCAGAT	TCCACCACTG	CCGGCCACAG
135481	GGAGGTGATG	GCTGGGATGC	TGGGGCCCTG	CAGGGTATTG	AGTTTGAATG	TCCCGCTCCA
135541	TGATCAAGCA	GGTGAAGGCC	AGAGGCAGCC	CAGCCTCGCT	CCCTGCTCC	CCCGGCCAGG
135601	CTGATACTCC	ACAGGTCCAC	AGCTTGTGGG	CTGCTCTAAA	GTCACCTGGG	CTACTGAGAA
135661	GTGGCCCGAG	TTACTCACTG	TCTAGCCAGT	CACCATCCCG	GGTTAAGCA	CTGTGGTTCA
135721	TTTCTCCCTC	CCATCACTGA	GAGCTTCTCA	ACCTCAGAAA	GCCCACCATT	CCCCGGCAGG
135781	TTATTTTCCC	TAGAGAAAAA	TCCCATTCTT	CTTCACTCT	AAGATCATCG	CTAACATGGA
135841	GTGCCAGAA	CATGGTATTA	TTGGGACAGT	GATTACAATG	ACTGGTCTTT	ACGGGGGCTT
135901	AAAGGCAGAC	TCTGGAGCCA	GGTTGTCTGG	ATTTAAACCC	TGGCCCTGCC	ACTTACTTGC
135961	TGTGTGACCT	TGGACAAGTC	ACTAAACTTC	TCTGTGCTC	TGTTTCTCTA	TGTGTAGTGG
136021	AAATAATAGT	TTTGACCTCA	CATAGTATT	TGAGGCTTAA	AATGTTAATG	CCTGCAGACT
136081	GCTTAGGAAT	AAAGTAAGCA	GTTTTTAGGT	AACTGCTATT	GCTACTACTG	TTTATTAAGT
136141	ATTTGCCATG	TTTTCAGGCA	CTGTGAATGT	TGCAGATACC	AGGATGAAAT	CACTTTTGTG
136201	AGACCCCGAC	AAATTAGGGC	TGGGAACGCA	TGAAAGAGCA	GAGCTCGTAT	TTACATGTCT
136261	GAGATGAGAG	TTGTTTCCAA	GGACTTCCTA	CAAAACCCAC	<u>AGAAAAACAGC</u>	<u>TTACATCTCT</u>
6 136321	<u>TGATGCCAAA</u>	<u>GTCCCTCTTT</u>	<u>GACAGGAAGA</u>	<u>GGGAAATGAG</u>	<u>CAGACCCGCC</u>	<u>CTGTTCCCTG</u>
136381	<u>CTAGCACCGG</u>	<u>TGCCCTGTGG</u>	<u>TGGGGTAATG</u>	<u>CAGGAGGTCA</u>	<u>GAGAGAAGAG</u>	<u>GAGTGGAACA</u>
136441	<u>CCTGGGCACA</u>	<u>GCCCTTGGTC</u>	<u>CTGGGGACTC</u>	<u>AGTCCTGTGA</u>	<u>ATGCCGGCTG</u>	<u>GAGGTGAGGC</u>
136501	AGTTTCCAGT	TGTGGGGCAG	ACTCAGGCTC	CCAAATCGCT	GACACTTAGG	CTTTTGTTTT
136561	ACCTGGGAGC	TAGACCTAGA	GAGAGCTGAG	GGAAAAGGGG	GAAGGGGTAC	AGCCTTCGCC
136621	AGGGAGATGA	GTTGACAAGG	TTCATCACTA	GACATCCATT	AGGACTGCAA	TAATTCAGAT
136681	AAGATGCTCT	CAGAAAACAC	GTGCCCAGTA	ACAGCATCTC	CACCAATGAA	CTGACAACCTC
136741	TGGCTTTGAG	CTCTGGAACC	AATGAACCTC	GTTTCCAAGC	AGTTTCTGTG	AGCCTCCCCC
136801	TTTTTTCCAG	TAAAAGCTTC	CCTTTCCCTT	TCCCTCACAA	TTGCCCTGGC	GGCCTGTCTAT
136861	TCGACACATT	CCGAGCCACA	GTCATTTCTC	ATTCCAGAGT	AAACACAACA	TAGAGATCAT
136921	TTTCTCTCGT	GTCTTCTCTG	AGGGTGACAA	TACTGTACTA	ACCATATTAC	CTACCGTGTG
136981	GGTGGCAAGC	CACCCAGGCG	CCGAGGCAAG	AGACCGAGGA	TACAAGCTGT	TCCAGTATAA
137041	TAAAATATAA	AACAAGAATA	GTTATACCAG	ATATAGATCT	TAGATACGAT	TATATATGAA
137101	TCTCATTTAAT	CATTAGTTTG	TAGCAATTAC	TCTTCAGTCC	AATATTATAA	TAATCTCTGC
137161	TCTATAATCA	TAACCTAGGA	AAAACCAGGC	CATACAGAGA	TAGGAGCTGA	GGGGACACAG
137221	TGAGAAGTGA	CCAGAAGACA	AGAGTGCGAG	CCTTCTGTTA	TGCCCAGACA	GGGCCACCAG
137281	AAGGGCTCCT	TGGTCTAGCG	GTGACGCCAG	CGTCTGGGAA	GATGCCTGTT	GCCAGGCGGA
137341	CCATGGTCTA	GCGGTAGCGA	AAAGTGTCAA	GGAAAAACAT	CTGCTACTTA	GCAGACCGGG
137401	AAAGGGAGTC	TCCCTTTCCC	CGGGGGAGTT	AGAGAAGACT	CTACTCCTCC	ACCTCTCGTG
137461	GAGGGCCTGA	CGTTAGGCTC	GCCCGCAGTT	ACCTGGAGGC	CTAACCATCT	CCCTGTGATG
137521	CTGTGCTTCA	GTAGTCACGC	TCCTAGTCTG	CCTTCATGTT	CCATCCTGTA	CACCTGGCTC
137581	TGCTTCTAG	ATAGCAGTAG	TCAATTAGTG	AAAGTACTAA	TAGTCTCTGA	TATGCAGAAA
137641	TAATGGCGTA	AGCTGTCTTT	CTCTTTGTCT	CCTCTCTCTC	TCTGCCTCGG	CTACCAGGCA
137701	GGGAAGGGCC	CCCTGTCCAG	TGGACACGTG	ACCCACGTGA	CCTTACCTAT	CATTGAAGAT
137761	GACTCACACT	CTCTACCCTG	CCCCTTTTGC	TTGTATCCA	ATAAACAAAC	GCACAGCCAG
137821	ACATTTCAGG	CCACTACCGG	TCTCCGCGCA	TTGGTGGCAG	TGGTCCCCCG	GGCCAGCTG
137881	CCTTTTCTTT	TATCTCTTTG	TCTTGTGTCT	TTATTTCTAA	ACTCTCTCGT	CACCGCACAC
137941	GGGGAGAGAC	CCACCGACCC	TGTGGGGCTG	GTCCCTACAC	TACTGTAACT	TTATTTATCC
138001	TCCCACAGCT	TCAAAGAGCA	GTA CTCCCAT	TTCTCCCAT	TTTCTACATG	AGGGAATTGA
138061	GGCTCAGAGA	GGCTAACTTG	CCCAAGGTCA	CAGAGTTAGC	AGGGACAGAG	TTAGAATTCA
138121	AACTTCAGAG	CCAGAGCTCA	CAACCCCGT	CTATTCTACC	TTCCAGGAC	TTTTGCTTCT
138181	TTGGAAGAGC	ATGCAACTCT	TCCCTCATTG	CCCCACGGAT	ATCTAGGATA	TCTAGACACT
138241	AGATTTTCAT	AGATTTGTCT	AATTTGACTA	ACAGATATTC	ATAGATTTGT	CACAGATTTG
138301	ACTAATATGA	TTACTGAGAT	AACCGACACT	TTAACTGATA	TGAATCTTTT	CCCCTGAGGG
138361	CATAGGGGCG	TCACATTCCC	TTCTGAAAGC	TCCTCCCATC	TGAAATACCT	CCTTAATTGA
138421	TGACTAACCC	GGAGGCTTCT	GGAGGAAATT	CGTCAGCCCA	GTCACCTGTA	GGATCCCCCG
138481	AATGCCTCCC	TGCTGAGGGT	CTCAGCCACT	CTGAACCTTG	CGCAGGGACC	CCAGAGGCTC
138541	TGGGCCATCT	TTCTTCCGGT	GGAGAGGCCC	AGGGCACACG	TTTGCCAGGA	AGTTTGTGTA
138601	GAGGCTCTGT	GCTGCCCCGT	GCCCATCTTG	GGAGATGTGG	CCCAAGAGAG	CTTAAGGGGG
138661	TTGGGTCTCC	CCAGCTGGGG	TCCAGAGTAT	CCTTGCCCGT	GATCCTGGGT	GTTCCGGCAGG
5 138721	AAGAGAGAAA	TGAGCAGACC	CACCTGTCTG	<u>TCTGCTGGCA</u>	<u>CTGGCGCCCT</u>	<u>GTGGTGGGGC</u>
138781	AATGCAGGAG	GTCAGAGAGA	AGAGGAGTGG	AACCCCTGGG	CACAGCCCCA	GGCCCTGCAC
138841	CTCCCTCAGG	CTGGGCTGAG	TCTCTTGAGC	CACACGGCAG	GCACGTTCCA	GAACAGGCTC
138901	CGGATGGCCA	GCCTAGCCTG	CTGGCCACAG	TCCCGCTCTG	GGGCAGGTTT	CAGGCCCTGT

138961 TTCCAGCAGA AATGAGGCTA GGCTAAACCG CTGCTGAAGC TTGCCAACCA CAGGAACCGA  
139021 CTTGAAGGCT GGAGCTTGAG GACCTGGTCC TGGGGACTCA GTCCTGTGAA TGCCAGCTGG  
139081 AGGCGAGGTG CTTTCCAGGT GTGGGGCCGA CTGAGGCTCC CAAATTGCTG ACACCTTAGGC  
139141 TTTTGTTTTA CCCAGGAGCT AGACCTAGAG CAGGCTGAGG GAAAAGGGGA AAGGGGCACA  
139201 GCCTTCGCCA GGGAGATGAG TTTAAAGGGG GAAATATTCC AGAAGCCTCC ACAATGGGCC  
139261 GTTCTGGGCT CAGGGCTCCT GATGGGACGG GTGGTGGAGC CAGGGCCGCC AGCCCCCTCA  
139321 GTGGCCTGCA TCCGTGGACT CCCCCGCATC CAAATGTCCG CTGATGTCTG ACTCGTGGGT  
139381 GGATCCACCT GACCTCAGGC TGGGGCAACA GTGTGGTCTT TCACTGAGGG GTCCCTCTCA  
139441 TCTTCCCCCA GGGCTCCTTA GTGTCTGCCC TGCCCCAGGC TGAGCCTCTT CCAGGAGCAC  
139501 CAGTCTGGAG CTGCTTGCTT GGATCCAGCT TGGGCCTGGG TCTCAGGGCC ACAGCTCTAG  
139561 AAGCCAGCG AGGTCTGCAT ATGAGGCAAA GGACCTCAGG ACAGTGACCT TGACACACTT  
139621 AGCCTCATTC CTGGCCACCA CCTGGGGCAC CGGGGGGAGA CCTCAGGAGC CAGCTCTGCT  
139681 GGCTTGCTC CTGGCAGAG GATGAAACT CTGGGCTTTG GAGAAACACA TGACAGAGGG  
139741 TGGGGGGGTG GCGAGAAGTT TGCAGGCCTC CTTGGCCTGG ATGCAGCCCC ACAGCTCTTG  
139801 GTGGCTACTA ATACTTGGA GTGGATGGGG ATTTGGGCCC CCAGAGCTCT GGAGACACAC  
139861 AGGGTGATCC CAGGGCTCCA GCCAGTTCTG GGAAGTCTGC CGGCCTGTGG TGGTGGGGCT  
139921 GTGGGCTTGG TGGGTCTCCT TTTGTGACAC ACCCAGCCAC TATGTCTCTC GATGATGGTG  
139981 GGGGAGAGGT CCCTAAGGAC ACTGCTGTGG TCGCTGGTGG TGGGGAGAGG AAGCTCTGCT  
140041 GACCACTCT ACGTTTCCCA GCCCCGCAT CTACGGATGC TGCTACCCA CCCACCAC  
140101 TCTGCATCCC AAATCTCAGA CTCCCTCGCC CTCTGGCCCA GTCTGAAATC ACCCTGATTC  
140161 CAGCTATCAC ACTTCCAGA ATGCACCCTG GCTCCAGGCC TGCCCTGACT GGTGCCCACA  
140221 TGCTCTGCAG ACACTGTAC CTGGTAAGGG TGGCCAGGTG GTCCCCCAGG TGACAGGCTC  
140281 TCCATGCTAG CCCTTCCAG TTTACTCCAG GCCCAGCCCT TGCCAAGATG CTGGCTCTCT  
140341 GGCTGAGTGG CAGCGACGGT CCCAGTGGGA TAGGACAGC ATGAATCTGC CAGCCTTCAA  
140401 TGAGCCACGA TCCTATAAAC AGGTGGGGC TCCACCCCA GGTGCTCTCT TCAGCCCTGC  
140461 ACCTGCCCTT CTTCCCCAAG AGGCTGGGTC AACCAGCATC TCCAGAAACC TCTGGCAGGG  
4 140521 CTCCAGGACC TGCAGGCGGA GCACCTCTTC CAGACCTGGA GTTTCAGGAC AGAAGATGCA  
140581 GCCAGGGGGA GGCAGACGAG TCCCAGGAGT GGTGGCAGGA AGCTGGTTCC CCTGAGGCTG  
140641 AGGGCAGAGG ATGGGGTGGG GGTAGAGCTG AGGAAACCCC CAGCTCCACT GGAGCCAGGG  
140701 CCAGCCTCTA GAGGAGGAGG GCTGAGCACC TGAGACTAGG CAAGGAGCCA GCCCTCCTGG  
140761 GCAGCTCTGT GCAGCAGGGA GTTCAGGGGA TGAACAACAG AGCAGTCGTG GTGACAGGCC  
140821 CGGGCTAGTG TATCCTCCTC ACTGACCTG AGCAAGCTCC AGGGAGCAGC ACAGGGGTCT  
140881 GGACATTCCC CACCACAAGC CCCTCACCTC CATGCGGCGG GCCCTAAGAA TTTAGGGGAT  
140941 TTCAATTAGA CTTTGTAAATA AAAGTAAAAG AGACCAATCA GGAGATTCTT AGGTTATGAA  
141001 AATTGTAGAA GTAAACTTTG TGAAGCCACC TCACCCTTAA AAGCATTGAG TGTACACATT  
141061 AGCACTGAAC GCTCTCTTAA AATGACATGA ATTGAGTCAT TTTACCATCA ATTCTGTAAA  
141121 CAGTTTGAACA GCGATGATGT TGTTTTTCAA AGCAATGAA ACTGAGTTAA TCTTAAATA  
141181 CTACAAGTGG ATCTAAAATG TATTATTTCA GAGTTGGCCA AAACAGGAAT CTAATGGTCT  
141241 ATGAACAAGT AAACATTATT TGGGTGGAAA TATCTGTCAT TTGCATGATA TAAATTTTAA  
141301 TAAAAAATG AATGTAGTGG TTGAGCATGG CAGCTAACGC CTGTAATCCT GGCACTTTGG  
141361 GAGGTCAAGG AGGGAGAGTT GCTTGAGGCC AGGAAATGGA GACCAGCCTG GGCAACATGT  
141421 TGAGACCCCA TCTCTACCAA AAAAATTTTC TAAATTAATG GGGCATGGTG GCACGTGTCT  
141481 ATGGTCCCAG CTAATCAGGA GGCTGAGGTG GGAGAATTGT TTGAGCCCAG GAGTTCAGAG  
141541 CTGCACTGAG CTATAATTGC ACCTTTGCAC TTCAGCCTGG GTGACAGAGT GATACACTGT  
141601 CTCTTTAAAA AAAGGGTGG GGGGAAAGA AGTAAATGTA ATATGTAGAA AACTACAAAG  
141661 AACAAGTCAT TTCAAGTGGG TCACAGACAA AGCAATACAG AATCACCAGG CTGGTGACTG  
141721 GAAAGGCTGG GATTGTAAGT TCTGTCTTGA AGCCACCCCA CCAGAACTTC CCTGGTCTCT  
141781 GGTGGCCTCC ATTGTTTAGG CCCGTTCCAG AGGAATCCTG GGACCTTAGG AGTCTGATT  
141841 AACCTCCCTG CCCCTGTCAT CCCCATACGC CCGCCGGGAT CCGCTGGGGG AACCTAGGGA  
141901 GGCAGAGCCG CAGCAACACG TTCCTTGAGG TGACCCAGC CCCAGCGCCC CACTCCAGTA  
141961 CGGGAGGGCC AGACACCCCC TCACCCCCAG AAGACAGCTG GCTGTGCCTG GAGGAGGACG  
142021 TGGTGTCTAG TTTCCCCACC TGTGCAATGA TGGTGTGGTA GATGCTGGGC CTGGTGAAAA  
142081 CTACGAAATC CCTAATTCAG AGATTCTTAA TTTGGTCTA GGTGCCATGG ATGGGAATTA  
142141 GGATACTTTT AACCTCCCTT GAAATGTAT ACAAAATTAT GTGACTATGT ATGCATGTGT  
142201 GAATGTTTCC AGAGAAGGGG TCTAGAACAG CACAGTCCAG TAGGTGACTA TGTATGCATG  
142261 TGTGAATGTT TCCAGAGAAG GGGTCTAGAA CAGCACAGTC CAGTATGTGA CTATGTATGC  
142321 ATGTGTGAAT GTTTCTAGAG AAGGGGTCTA GAACAGCACT GTCCAGTAGG AATGTTCTGG  
142381 GAGCTCTGCA GATGTGTAT TAACTTTTCT AGTGGCCACA AAAAAAGAGG TAAATGATA  
142441 CAGGTGAAAT TAATATATAT TTAGCCCAAT AATCTAAGGT ATTGTCAATT CAACATGTAA  
142501 TCAACATATA AAATTATTAA GGAGATATTT TACTTTTTTT CATACTAAGA CTTCAAAATC  
142561 CACTGTGTGT TTTATGCTTG CAGCACATCT TAGTTCAGCC CAGCCGCATT TCCAGAGTCC  
142621 AGCAGTTGCA TGTGGCTAGA GCCTACCTTA CAGAGAGCT CAGGTCTATA GCTTTAATCA  
142681 GATTCTCAAA AAGAGTTTGA AAAAGATGGA AAACCATTGC TCGATGGGTC ACACGTGGAT  
3 142741 ACCCCCCACT TGAAGAGAGC CGCAGGGTTC TGCATGGAGC AGAGGAAATA TTTAGGATGC  
142801 TCCCATCACA GTTTCCATGA ATTATGTGCC AAGACCTAGG GTTTTGACAC ATGGTGAGGA  
142861 CCTAAGGTGG GAGGGGGGAA GGACACGGAG CAGGGAAGCA CAGTGGCTGG CCTCGGACTG  
142921 AGGTCTCTAAA GATCTCCCCA CTCCAACCCT CCAAGGGCTA TGAATAGGCA GCTGGGAATA  
142981 GGGATCAGAC CAAGACAAAG GTCAGGACAA GAAGACAGAA GACAGGTAGA GGTGACCTGG  
143041 AGGCAGTCAG CCTGGAACAG GGCTGGCAGG CACAGCGGGG CTGGGCTGAG GCCACATGCT  
143101 GCTCACCCAG GCCACGGGTG GAGGGTCTTT AAAGGAGGGG CATGCCTAGA GAGCCCATCA  
143161 GAGCAGACCT GACCTAAGCA GCCATGCTGG AGGAAGTCTT GGGCGCCATG TTTTCCACTC



143221 TCTCATCCCT CTCCCTTGCT GGGCATCCTT GCCACCACTG TCTCTTCTCA ACGTTTACTC  
 143281 TATACTCACA GAAATGAAC ATCTATTTTT TCATTGGCA CTCAGGCTC CTTAATTTT  
 143341 CCCAGATTTT CTCTGGAAAC CAGCCCTCCC CGACTCTCTT CCCCTTGGT TTGGGTGCTG  
 143401 TTGACTCCAC CTCCAGCTGC AAGAACACAG CATGTGGTCC AAAGCTGGCC AGTCAGTTCC  
 143461 TGGAAAGCTT CTAACCAAGG CATTGTGGTC ATTGGACTCA GGACTCTGGC TGGTGCAATC  
 143521 ACAATAAATG CTGGTGCATG AAGTTCTAAG AAAGACACTT TGTCTTTCTT GCTGGGCAGA  
 143581 ACTTGGGAGC ATATGAAGCT GGGGCTGCTA AATGTATCCA TCTTGCCACC ATGAGAGGAA  
 143641 AATTTGTCAG GAAATGGAG CCAGTAATAA TAAACACCTA TATAACTAAT ACTTCCTATG  
 143701 TGCCAGGCAG AGGCGGGTTT ATTGTGCAGC TAATGATACT TAAGCTGAGT CTCTCCTTG  
 143761 CACACACACA TCCTTCCAAG GCTCCAGGGG AGCCCTAGCA GTAGGTTCGT GTGATCATAT  
 143821 ATTTTAGTAA AATTCACAGT AGTAAACGAT TTTATCCACA ATTGGTAAAG ACTGCTGTCT  
 143881 TCCCATTCTT ACCCTCTGT TAGCATTGGT GTGGCTATAG ATATTTTTAG GATCTGGTTA  
 143941 AGGAGAGGTT GAGTTAGAGA TATATTTTAT TTCAGTTTAG TGAGATGTAT TTATGGGGTT  
 144001 TGCAGTCAGC TTCACATATA GTGAAGTCAG CTTCATGTAC AGTCAGCTTC ACATATAGTG  
 144061 AAGTTTCACT TGATGTGGT GTGTCTTTTG GTGCCAGAAG AATGGATAG AGGAAGAGAA  
 144121 ATGAGATCTG AAATATCCAG AGCCAGAAAC CAATCTGAAA AATTCTCTA ATCCAACACA  
 144181 CAACATTGTA AGTGGAGAC TTGGTTTCC TTGACACCTG ATCAAAACAG AAGTTCTCTC  
 144241 CTGAAAAGGG TGTTTCCCAT CATGTCTTGA TTGCATGTTT TCCTTTATTT GAGGGTAATT  
 144301 TTAAGAAATA TTAATAGAAT ATGTATCGTG ATTCCTGTGC TACAGCAGCT AGCACAAAC  
 144361 CTACAACAGC TTATACACCA AAGAAGCTTG ATGGAGGCTG CGGCAGATTC TGTTTCCAA  
 144421 TAATGGTTGC ACCGATATCT CCTGTCCATG CTCTCCAGC ACCTTGCCAC TCCCATCAA  
 144481 AAAATAGAGG CCAACTCCCC TTCTCTTAAA TCTGATGGGC TTGTGATTGG TAACTTACAG  
 144541 GATGTAGAAA AAATGACTGC TTGACTCCCA AGGTAGATC ATAAGGTGAT GCAGCTCTCT  
 144601 TTTTTAATGG CTAGAACACT CCCCCTTAGA GCCCTGAGCC GCCAGGCGAA ACACCAACT  
 144661 CCCCTGAGGC GGCCATACTG TGGGGAAGCC CAGCCATAAG GACAGGCCAC CAGTGGGTGC  
 144721 TGGAGTCAGC AGTCCTAATC TCCAACTGGC CCAGCCAGG CTCCAGACCC ACGGACGAAT  
 144781 GAGCCTTCCA ATAATTTTAG CCCAGTGGT TGGTTCATCC CTAGCCAGGG CCCTTTCTGA  
 144841 ATTTCTGACA CACAAATAT CTGTGAGCCT AATAGAGTGG TTGTGTGTT ATGCCATTAC  
 144901 ATTTTGGGTG ATTTGTTACA TAGCCATAGT AAGTAAAGCA GAGGTTTTCC TAAATTTGTA  
 144961 ATGATCCTGA AGTTTATGT AACATTACCA GTAAAGAATT ATGAAGCTGA AAGAAACAAT  
 145021 CAATAATGAA CTACTATTAA CTATCAATAA TTTTCTAAAA ATTCCAATCA ACCAGACTGG  
 145081 AGGAAAAGAT GCTACCTATT TGCACCAAAG AAAGTGATAT TACAAAACCTG GTATTGTTAT  
 145141 ACAGCAAAGA AGAAATGAAA AAAAAACAACA ACTCACAAAA TGGGTGTTGA AAAAAAGGC  
 145201 AATCAAGGTG TATTCAGGCA AGCAGTGGAG GGAAGAGCAT GGAGGTGTGT CTGGCAGTTG  
 145261 ATTCACACAA ATATCATGGT GTGTTTTTTT TGTTTTTTT TTGTTTTTGA TATGGAGTCT  
 145321 TGCTCTGTGG TCCAGGCTGG AGTGCAATGG CATGGTCTCA GCTCACTGCA ACTTCCACCT  
 145381 CCTCGGTTCC AGCAATTCTC ATGTCTCAGC CTCCCAAGTA GCTAGTATTA CAGGTGCGCA  
 145441 CCACCATGCC TGACTAATTT TTGTATTTT AGTAGAGATG GTGTTTCACC ATGTTGGCAA  
 145501 GGCTGTGCAA GAACTCCTGG TCTCAAGTGA TCCACCCGCC TCAGCCTCCC AAATTGTTGG  
 145561 GATTACATTT GTGAGCCACT GTGCTGAGCC CATATCAATG TATTTTTTTA AATTTTGTAA  
 145621 TATTTTGGAA TTTATCAACC TTTTAAAAAT TTAATTTTTG CCTGATTCT TTTTATCATT  
 145681 TACAAAGAAT ATTAACTTTT GTACCTAACT TTGTGTCAAC TTGTATAACC TTCAGGCTCC  
 145741 ATAATACACG CCCTGTGTGC ACACATTTCA AAGCACTTTA TATAGAGAAC ACCGTTGTTT  
 145801 TATCTACAAT TGGCAAAGCT GCCTGTCTTC CTGTTCTGC CTGTCTGCCG ACATTGGTGT  
 145861 GGCTATGGCT GTTTTTAGGA TCTGGCTAAG GAGATCCCCC AAGCCCTGTC AGGTGAGGAA  
 145921 CTTTTATTAT CCTCATTCTA CAAATGAGAA CACGGAGGCA GGAGTTGGTT GATGAACTGC  
 145981 CCCAGGTGCA CCCTGGGAGT GGAGCCCAGA GGCCACAGTA CAGTCTCCTT CCAGAGGAAAG  
 146041 TGCAGGCAAA AGGGAGCAGA GAGAGCCCTT GGATCGAGCT TTGCCTGAAG GAGTAGATCC  
 146101 AGTCTTGGGC TTTTCACTTG GGTCTCAGCC CCTCCCCACC TACTTTAACC CCAGGAAGA  
 146161 AGGGTTTGGT CACTCCCTTC CTTTATCCT TGCTTAGAAC ACCAATCGAT GACAATGTCA  
 146221 CAGAGAAGAG GAGGCAGCTG ACTGACAGGG GTGTAGGAAG TATCAGGCT GATGACCTGC  
 146281 TGAGTTTCCA TTCCAGGCAA AAGCCCTGGA GCAGGTTGTG GGTGAGATAG CAAGATATGT  
 146341 AGCATGAGG AGGCATGACA CATGCTTCGG TGTTCGCGC TGAGTACCAT GGAAGGCAC  
 146401 AGCGGCGCT GTACTCTTGC CTCATGAGGA AGTGGTTATT TACCCTGTTG ACATGAGGGA  
 146461 AAGGACCATT TCCTGTCCAG GAGTGCTCAG GACCAAGCAG GAAAGGGACA CCTCACCCCT  
 146521 GCCCTCATTC ATTTAGGGC TACTGGAAAG GACAAGCCTT GCTGTAAGCA AGGGACAGGG  
 146581 CACCAACCC CCAAGCCATT CATTCATCCC ACAAATTGTT ATTGAGTGT TACCCTGTCC  
 146641 CAGACAGAGA GCTGTGAGG TAAAAACAGG CCTGGTCCCC ACCTCCAGTG TACAACCCAC  
 146701 CTTCCACTCC ATGTCTCTCA GCCAACGGGG TAAGTCAGTG TTCCAGAACC TGGGACATGA  
 146761 CACTTCAGGC CCTGGAGTGC TGGGGCAGAT ATACAATGCA GAACAGAGTC TCTGGAATTT  
 146821 CAAGGTAAAG AGCCATGGAA GAGCTGTAAA CAAATATTAG TCAGCTAATT AACTTAGTTT  
 146881 GGAGGCAGTC GTTGCAAGT CATGAGCTT GGGATAAGAA ATAGAAACAT GTTTTGTTTT  
 146941 GTTTTTTCTC TGAAATTTTT GTTTGTTTTA ATACAATTGA ACATCTCATT CCCCCCAAC  
 147001 AGTAATCAAA GCAAACATTT TATCATTTAT TTTGTATAT TATTCTTGGT ATATCAAACT  
 147061 AAATTCAGG AATGTAAGAA AATCTGACTT GGAAATATTA CATAGAAAAA GGCTGAATTA  
 147121 ACACATTAGC TTTTGTAGT TTTCTTTGTA AGTATACATT GACATCTACT TATCCGATAA  
 147181 CAGCACTGTG GTTTGAGTG AGTTTGGTTT GGTTTGGTTT GCTTTCCTTC CAGAGGGAAT  
 147241 GCAGTATGGC AGTATTATCT TTGGAACAAA TTAAGGAGGT ATTAGCAAAAT AACAGAGAG  
 147301 TTACTTCCAC ACATATATTT TCCATGGGCA ATAAGTAGAA AAGTATATTA TGAAGAAATA  
 147361 TGCTTCAACT ATGCTTTTTA AAATTTCTCG AGGTTTACTG GTTTGTTTTT CTTATTAATA  
 147421 TATAATGCAC TATCACTACT GTCTGTCCGG GGAAAAACAA ACCCTGAGGT TTGAGTTTCC



147481 TCCTTCAGCC CATAAAGTGA TATCAAGAGT AAACCTGGAT TGAGGTCACC CTGTTTGAGC  
 147541 CACCAGAGAG TTCTGTAGTA TATCTGGTTA TTTTACTAAT GTATACATAT TCTAGGTTAT  
 147601 AATCAAGATA CTCATCAAAT ATTTTCTATG AAAGTTTAT ATCCCAAACA GGTTCTTTGA  
 147661 TGATTATATCA TGCTACCATA GGAATTTTGA AGAAATTACA AATGTGTAA CAGAAATAGA  
 147721 CATGCAATAT AGAGGACAGA AAAGATACAG CTTGACCCCTG GGTTTTTCCC TTAGAATAAT  
 147781 TCCTTCCATA AATCAGACTT TATTTTCACA TATAGACTCA TCCTATAACT GGTCTTGGAT  
 147841 CATGAAGTTA ATGACATAGT TGGAAAAATC TAGGCCACAC ATTTTGTGG TCTTTTGTCA  
 147901 CCCTAGACCA CACACAAAAA AACATTCTCC AAGTGGATGC TCGGGTTAGA TTATGATGCT  
 147961 CTTACATTAA AAAATTGACT GAAGCTGAGA ATGTAAAGCA CACAACATT TGCAGCAGCA  
 148021 ATAGTCTTTC AAAAATGACT AAAACTTTTA AATACAAGGC ATACAAAAAC GACTATTAAG  
 148081 GAGAGTTTTT CAAACTATAT CTAGAAATAC AGGAAAAAT GTATTGCCTC ATAATAAAGT  
 148141 TAAATAAATA TCATTCTCTGA GTTCTATAAA CACTATAAGA GGTTATCGTT TTATCGGAGA  
 148201 ACCATGTTTT AAGACCCCTT GAGGTACCTG CTCCGACCTC TCCCAGAAC CCTCTGCCTC  
 148261 TCCTCACCAG GCCTTCTCCG TCTTGAGGGG ATCTCCCCGC TCCCCACTCC GCCCTGTTGG  
 148321 CCTTATCCTC TGATTCTGT TTTTATAAAG CTCCTCAAGCA ATGGAAATGA GTATGTGCTG  
 148381 TTTAAATGTA CTGTATTTGT CACTTTTATA TTCTAGTTCA TAATATTTTT TAAAAAATAA  
 148441 ATTGAGACTT TCATACATTT TCAAAACATA AACCAAAACA AAACCTGGGT AAGCGTGAAT  
 148501 AGGGCCCTCT AAAAGATTAT TTCTGTAGGC TGTTGTATCT GAATAGAGA CACATCTTTC  
 148561 TCTAAACTCC AGTAGTGACC TGCTGACCAC AGATTGGCA AGAGAATCCT CAGTCCAGAC  
 148621 GTATGTTTCA AGTTTCATCC TAATGCTCAA TTGCCAGGAG TTCTATGAAA TTTTGCCTC  
 148681 CTCTTCCAC ATGGCACTGA CTCATCTGAA AGGCAATAGG TTTCTGTAA ACACCCAAC  
 148741 CACCAACAGC TGCCAAGGGG GCGCACTGTG AGGTAATGAT CTGCCTGCCC TGGGTTCCAGG  
 148801 AGGGTGTGGG GTACACCTAT GCCACCTGCA AGGCCAGGG CCCCAGCCAG CCTGCTCCTT  
 148861 TCTGCCCTCG AGGAGGCAGC ACCTGAGTGC CCGCCAGTG CGGAGGAGAC TTCCTTCTG  
 148921 TGTGCCACCC GGGGTGGCCC AGAGGGATTT GAGGTGAAGT CCAGAGCTGA GGGCCAAGCC  
 148981 CAGCCCTCCC CATCACACTC CCTCAGCACC AGCTATAGTT GATGACTTCC CTCACATGAC  
 149041 AGGATTTGAA ATGCATATGT GAATGTGCTT TTAATGTTTT GCCCTTCTCCT CCTATTCTT  
 149101 TCTCTCGTCT TCAGCCAACTG GCACCAACTG TTAAGCTGG AGATGAGAGA AAGAAAGGGG  
 149161 CCCTCCTGGC CGGGTGCAGT GGTCTCACC TGTAATCCCA GCACTTTGGG AGGCTGAGGC  
 149221 TGGCGGATCA ACTGAGGTCA GGAGTTTGAG ACCAACCTGG CCAACATGGC AAAACCCCGT  
 149281 CTCTACTAAA AATACAAAAA AAATTAGCCA GGTGTGGTGG CGCACACCTG TAATCCCAGC  
 149341 TACTCGGGAG GCTGAGGCAG GAGAATCGCT TGAGCCATAG AGGTGGAGGT TGCCGTGAGC  
 149401 CAAGATTGTG CCATTGCACT CCAACCTGGG CAACCGAGTG AGACTCTGTC TCAAAAAAAA  
 149461 AAAAAAAGGG AAAAAAAGGG GAGCAGGGGG CCTCCTTCTG CCTCAAATCT CAGAGCCAAC  
 149521 GTTCCAGAAA CAAGGCTTAA AGGCTGACCA CTGTGAGGAA AATGAAGACT TTTCTAAATT  
 149581 CCCAGAGTTT CTTTACAG ACCATAAGCA GAGCTTTGAA ATGTGGGTCC TGTGAAGAGA  
 149641 GAGCTTTAGA GAAAGTGCCC ACCCAGCATA GGATGAAAG AGAAAGCCCC TGTACTTTCA  
 149701 GGAGGGAAGC AGGGAAGAG ACTCACATGT CTCCCAGCC CAGCCCAAGC CCTTGGGGGT  
 149761 ATTCTGCAGC AACAAATCAG GTGGGTGAAA TCCCTTAGGT GTGGGCTTAG CCCTGTCTCC  
 149821 ACAGGCACAT CCTGGGAAAA GTCACATCCC TGGACAGGAG CACTGTGTCC CGCATAGTAA  
 149881 TGCCTGCCCA TCGTGGGTGT GGGAAAGTAT TGGGGAAGGG CCTAGGTAGG ATGGGGCCTT  
 149941 GTGGTGGTGC CCTCTCATCT TCCTAGGCCC CTTGTGTGA CAGAAGGATC CAGAAGTGT  
 150001 CCCAAGCTCC CACTGGGGGT GTGGAGGTAG CTCAGCTGAC AGCAAGAGGC CTGCCTGAAG  
 150061 CTGCCATGGG TGCGGTGAGG TGAGGCAGAA GCTGGGGATG GTCAGCTTGA GCCAGGGCCT  
 150121 GAGCCGACCT GGCCAGCGTG GAACGCAACA GCGGTCCCGT CAGCTGTGGC TTTGGCTGTG  
 150181 CATAAAGCA GGAGTCCAAA GGCCTGGTCA GAACACCCCA CATAGATCCT GGACTACAAC  
 150241 CCTCCCAAGA CCACTGAGGG TCCCGCATGC TTAGCTGCCA TCTCGGCTCC AAGGGTGGGG  
 150301 TAGAGAGAGC TCCAAAATGA GCATTTCTTA AGAGAAGCTG AGTTTAAATCC AAAAGTGAAC  
 150361 CTGAATTTTT TTATGTTAAG TGGGAAATG TAATTCTCC CAGCCTATTT GTAACCTCT  
 150421 CTTACACAA CTGTCCAGAG GTCTGGGGAC TAGGAAATAA GATGTGCGAA GTTATGAAGC  
 150481 AGAAAAGGAA TCTAAATGG ATATCTGGAA AGAGATTAC AGAGGAGGCA TTTCTAAGA  
 150541 TGAGCCTTTA GAAAATGGGT TGGTGCTTAC AGATTATTGC AACACCAATT TACAGTATTC  
 150601 TGTGCTATTT AAAAGTGGAG TACCACTATT AGGTATAAGC TGTTACTTCA GAGTTCATC  
 150661 AGCAGCGGCA TTCAACCTTT TTGGCTATCT GTAAACCCCT TCCCCCATA TTCCCGCCTA  
 150721 CCCTCTTCTC AGAGCCTCTC ACCACACTCC ATCCTAAGCC CTCTTCTCTA CTCAGCTACT  
 150781 GCAAAACCCC ATCTCCTCCA TATCATATTG CCAGCCCTTC TCAAAAAAGC CTGAGTTGGG  
 150841 TCAGGTGCGG TGGCTCACGC CTGTAATCCC AGCATTTTGG GAGACTGAGG CAGGCAGATC  
 150901 ACCTGAGGTC AGGAGTTTGA GATCAGCTTG GCCAACATGG TGAACCCCTG TCTCTACTAA  
 150961 AAATACAAAA ATTAGCCGGG CGTGGTAGTG CACACCTGTA ATCCCAGCTG CTTGGGAGGC  
 151021 TGAGGCAGGA GAATCACTTG AACCCGGGAG GTGGAGTTTG CAGTGAGCCG AGATCATGCC  
 151081 ACTGCACTCC AGCCTGGGTG ACAGAGCAAG ACTCTGTCTC AAAAACGAAA ACAAATCAA  
 151141 AGGCTGTAGT AGAGGAGGTC TCCAAAGCAG ATGTTCCACC AGGCATGGGA GGGATTAGGC  
 151201 ACCAACCACA GCCCCTTAC CAGGGCAGAG CAGCAGCCCA GAGGCAGAAG TCCTGGGAGT  
 151261 TCCATTCTTA GAAGCCTCAA AGAGCAGGAA GGAAATCAGC CCCTTTCATC AAGGGCTGCC  
 151321 CCTCTGAGGA CACCCAGGGA GGATGGCACC AAGATCAGAG GTGTGAGAGC CAGGAGGACA  
 151381 GAGGACAGAG GCCTAGAAAT CCTTCTTTT GGCATTAAAT ATATGGAATC AGAATCATCT  
 151441 TGAAGCCAGT ATCAGTGACT CTATAATTCT ATGTAAGCTG AAAATTCTTA CTGAATACAA  
 151501 CTCTCCCTCG TGTAGACAGG CAAGCAGGGC CTGCAATAAA ATGTTAAACC AAACCTAAGGC  
 151561 ACATTTTGT TTTAAACCT CTCAAAGTTT TCCTTCATCT TTATACTTCT TTAAGCTTTA  
 151621 AGAAAGAAA GAAAATTATA ATCAACAAGG CAAAGCTCCT ACTGCAAAAA TAATGCATAC  
 151681 TTTAAAGAAT TTCTGGGCGG GCGCAGAGG CTCACGCTG TAATCCAGC ACTTTGGGAG

151741 GCCAAGGCAG GTGGATCACC TGAGGTGAGG AGTTGGTGAC AAGCCTGGCC AACATGGTTA  
 151801 AACCCCATCT CTAATAAAAA TACAAAAATT ACCTTGGCCT GGTGGTGTGT GCCTATAATC  
 151861 CCAGATACTT GGGAGGCTGA GGCAGGAGAA TTGCTTGAGC CCAAGAGGCG GAGATTGCAG  
 151921 AGAGCCGAGA TCACCCCACT GCACTCCAGC CTGGGCAACA AGAGTGAAAC TCCGTCTCAA  
 151981 AAAAGCGAGT ATCTGGACCA TAAATGTGGC AGTATTGAAA TTAGCAAGA TCATTCTGAA  
 152041 CAGAACATGC CATCTGTATA CAATGAAATC AAAGACCAAG ACAATGCTTT CTTTGATCTT  
 152101 AAAAGATTGC AGGCTGGATG CGGTGGCTCA TGAATGTAAT CCCAGCACTT TGGGGGGCTG  
 152161 AGGCGGGAGG ATTGCTTGAG CCCAGGAGTT TGAGACCAGC CTGAGCAACA TAGGAGACCT  
 152221 AGTCTCCACA AAAAATTTTT TAAAAATTAG CCAGGCATGG GAGTGTCTGC CTGTGGTCAC  
 152281 AGCTACTCAG GAAGCTGAAG TGGCAGGATC ACTCGAGCCC AGAAGGTCAA GTCTACAGTG  
 152341 AGCCGTGATT GCACCACTGC ACTCTAGCCT GAGTGACAGA GTGAGAAAAA AAAAAAATC  
 152401 ACAACTTCTA CAAAAAGATA TTAAGTAAAT TTTATTAAGA AGATGCCATC AGTTTTCAAA  
 152461 TAAGGGGTGG CTTTATTGGA AAATGAAGAA CTGAGAGAAA AGTTAGAAGT TTGAAAGTTT  
 152521 GAAGTTAGTT TCAAGTTAGC TCTTTACCAT TCCCACAGCT CATATGGATT CAGAAGAGCT  
 152581 TTTCTAGCAG GTAGAATTAT TGGCTCTCAA CTCCAAGATA ACAGCTTTGA CACACTCAGT  
 152641 TTTCTCAACA GCATACTTTT CAAAAATATC ACATATTACT GTCATTCTGA CCCAGCGGTG  
 152701 TCATGCCCTG CACCTGACC ACGCTGAGCT GCGGCGAGTA AGGCTGCTGT GTGGCTTGTG  
 152761 CTGCTGTTCC CACTGGTCTT TAAAAAGGGA TGTGGGTCC GGGCAGGCTG GCTCACGCCT  
 152821 GTAATCCCA CACTTTGGGA GGCTGAGGCA GGCAGATCAC GTGAGGTGAG GAGTTTGAGA  
 152881 CCAGCTGGC CAATATGGTG AAACCTGTG TCTACTAAAA ATACAAAAAA TTAACCGGAT  
 152941 GTGGTGGTGG GCACCTATAA TCCAGTTACT TGGGAGACTG AGGCACGAGA ATTGCTTGAA  
 153001 CCTGGGAGGC GGAGGTTGCA GTGAGCCAGG ATTGTGCCAT TGCATCCAG TCTGCGCAAT  
 153061 AAGAGTGAAA CTCTGTCTCA AAAATAAATA AATAAATAAA ATTTTAAAG GCGACGGGTG  
 153121 GCGGGGGGGG GGGTGTGCGG AGGAGACCAC TGGGTGAGAA TGAGTCTCCT CTTACAGGCC  
 153181 TAAAACTTAC TGACCTTAAA GAAACTCACA ATATTGGTG GATCCAAGGG GCTTGAAGGA  
 153241 ATGAGAATCA GAGGGACAGA GAAACAGACC TGCTCAGATC AAAAGATGCA TCGTGAAAGA  
 153301 GCCAAGAAAT CAGATCAGAT TTACATACCC CGCCACACAC ACTTCTCAC AGACACTGTT  
 153361 TAGCTAACG GAACAATGAA CGTATTTAAT GTATTGTCAT GGGGTGGCAG CTGTGTACCT  
 153421 GCGCCAAAGT CAAAGAGCTG ACTGTGGATA ATTATCACGG AAACGTGCA CACCACAGC  
 153481 CTGGCGCCAT GCGGGGATGG GTCCGTCCCT TCACAGCCAT CAGGAAGCTG ATCAGAGGAG  
 153541 ACCTGAGATT GTGAAGAGGG ACAGGGGAGG CCACAGAGTC AAGGGGTGGA GAAGAAAGGA  
 153601 TGGTGAATCA GAAAGCAAGA CTAGCAACTC AAAACAAAAT GCCTGTACAG ACATAATGTA  
 153661 CCGTGAAGA AGCCAGGCAT AGAGGAGTGT GATTGTGTGA TTCCGTGGAT GCCAAGCTCT  
 153721 AGAACGAACA CAGTTAATCA GTGGTGACAG AACTTAAAT TGTTGGTTACC TTTGGTTGAG  
 153781 GGTATTTTTC TGGTAAGAA CACAAGGGAA TCTTCATACA TGATGAAAT GTTCTGTATC  
 153841 TTGATCAAGG TGGTGGTTAC AGGGGTAGAG ACATGCAAGC ATCCATCCAG CTGCACACCT  
 153901 ACGGTCTTGT GCTTTTCTGT CTATGTTATA CTCAATAACA TTTAAAAAAA AAAAAACAAC  
 153961 ATTAAGAGTC TGCTGTTGGG GGTGGGAGCA AAGAGCTATA AGCAAACTGA TTGATGGTTC  
 154021 AACTTCTAGG AGAGCTCAGT TACCATGAAA GGTAAAGGTG CAGATTAACC TGAAAAATAGG  
 154081 ATTAGTTGGA AGTTGGCAGA AGTGTAAATG TCCCACAAAT CACCTGTCCA GTCTCAGCAG  
 154141 AAGGTTAAAG GTTTACTCTC TGGGGAGGAG CAAGCAGAAG ACATGGGAGT CCACATCTG  
 154201 AAATGGGAGT CATCACCTTC AGACCTCTTT CCCCACTCAG CCCTGAGAAT GCTGGCAGCC  
 154261 AGTCTTACAC CTCCCCTGCT GCCACCACCA CCAGTATCAT CAGGCAAGAG AACAGAGGAT  
 154321 TCCCTTCCAG GGAAACACAC CAACATAAAA GGAAAAACAA AAACAAACAA AAAACTACCA  
 154381 ATACTAATAG GTGATTGCTG TATTCCAGAG TTCCAGATCC TGCCTAATC AACCTACAAT  
 154441 GAAGTTTCTA TTCAGCCAGG TCAGACCATG TGCACAGAGC TTCCAGTGAT CACTTTAGTG  
 154501 CCTCACTCTT AAATGTGACC AGACAACCTG GGATCACCAG ACATTGGAAG AAATCCTCTA  
 154561 ACATGATTAA TGATTAAGAC TACAACACAA ATAAAGAGAC ACAGAAACAG AGACAGTGTG  
 154621 GGAAGTATGA GAAAACTTTT TACAAAGCTG TAATTAATGT TGTTAAATA AAAACTTCAG  
 154681 CCAAAATTA TGTAAAGGAG TTTAATTGAG CAATGAACAA TTTGCAAAAT GGGCAGCCCT  
 154741 CAGAATCACA GCAGATTGAG AGAGAGTCCA ACGCAACCAC GTGGTAGAAG AAGATTTATA  
 154801 GACAAAAATG GGAAGTGATG TGCAGAAATC AGAAGTGAGG TACAGAAATA GCTGAATTGG  
 154861 TTACAGCTCA GCATTTGCCA TATTGGAACA CAGTTTAAAC GCCTGACAGT ATGTGAGTGG  
 154921 TTGAGGTATG GCTGCTGGGA TTGGCCAAGA CCCAGCCACT GTTAAGGCAC AACTCCCAA  
 154981 GGTAGGTTCT CAGTCTGTGC TACCTATTAG TTAGGTTGCA GTTCTCCAC AAGGACTCAA  
 155041 ATATAGAAGT ACAGAACTCT TCTCAAGCCA TATTAGTTTC GCTTCAACAA TGTCTTTACA  
 155101 TGAAAAAGACA TAAAAAAGAC AATGAAGATA AAAGATGATG GATCTTTGCA AAAAAAGGG  
 155161 TGGGGATGCC ATGAAATGGA ATGTTTCAGAG TATAAAGAAC CCTCGTACAC TAAAAATATG  
 155221 ATGGCAGAAA TAAGTAAAT TAATGGAAAA ATTAGATAAT AGAGTTGAGA AAATCTCCCT  
 155281 AGAAAAAGCAA ATCAAAATTT CAAAGAAATG GGTCAAGGCG AGTGCTCAT ATCCATAATC  
 155341 CCAGCTCTTT GGGAGGCTGA GGCAGGCAGA TTGCTTGAGC CCAGGAGTTC TAGACCAGCC  
 155401 CGAACCAAT GGTGAACCC CATCTTACA AAAAAAATAC AAAAATTTGG TGGGCACGGA  
 155461 GGCATGTGCC TGTAGTCCCA GGTACTTGA AGGCTGAGGT GGGTGGACTG CTTGAGCCCT  
 155521 GGAAGCAGAG GTTGCTGTGA GCCATGATGG TGCCACTGCA CTTAGCCCTG GGTGACAGAG  
 155581 TGAGACCTTG CCTCAGAAAA AAAAAAAGAA AAAAGAAAAA AAAGAAAAAA AGAAATGAAA  
 155641 ATAGGAGATA AGACTATGTA TTTAAAAGAG CATGCCAGGA AGCCTAAAT ATTCTTTCTT  
 155701 CTTGCTCTCT CTTCTTCTCT CCCCTTCTCT CCCCTTCTCT CTCCTCGTCC TCCCTTCTCT  
 155761 TCTCTTCTCT CCCCTCTCTC TCTCTCTTTT TTCTTCTTCT TCTCTTTAGC GATGTGGTAT  
 155821 CACTTTGTTG CTCAGCTGGA GTGCAGTAAT ATCTTCTTCT TGGGAGTTCC AGGAATGGAT  
 155881 CAAAGAAATC ATTAATAAT TCAAGAAAT AATGCAAGAA AAGATTGCAG AACCAGAAGA  
 155941 CTCTCTAGAT AAAAGGGGTT GCTGAGTACA CAGAACACCC AGCAAAACAG TAAAAACAAAC

156001 AAACAAACAA AAAACCCCTCA TCAAAGGATA TTTATCATTAA AATTTTAGAA ATCTGGAAT  
 156061 AAAAAGGTCA CACACCAAGG ACTAAAAATC ATAATGGGAT TGAACCTATT AACAGCAACA  
 156121 TTGGAAGTTA GGAGACAATG GAAAAATGTC TTGAAATTTT TTACAGAAAA TTATTTACAA  
 156181 CTTAGAAGTT TTTTCCCAA TCAATCCGCT AGTCAAAAGT GAGAGTAGTG TATAGGCATA  
 156241 TACAGAAATG CAAGGACTCA AATTTTTTTT AGCCTTTCTC ATGAGCTACC AGAACAGTGC  
 156301 TTTCTAGATA TTAAAGTGCC ATGGATCACC TGGGAACCTG TTACAATGCA AATCTGGTT  
 156361 CAGCAGGTCT AGAGTGAGGC CTGAGCACTC TGACAAGCTC TCAAGTGACA TCAGTGTGC  
 156421 TGATCTGTAG ACCACACGGT GAGAGTGAGG TACTAGAGGA TACGTCTCAC CAGAATGAAA  
 156481 GAATAAACCA ACAAAAAGGA AGAAGTTGGA ACCAGGAAAT TGAGGATCCA ACATAAAGC  
 156541 AAGTTAATAT CCAAAATGAT GCTGAAGGGA GCTCCCAGGG CATTACCTAA AGGGAAGCA  
 156601 GTCCAAATTG GGACATATAA GGAGATATTA AAAGGAGGAG GAAAGTTGAG AGGAGAAAA  
 156661 AATGAACTG ATGATTACCT GATATATCAA GAGGAGATT ATACTTCTGG TGGAGAGTTT  
 156721 AAAGACAAAT TACTGATAGT AACATAGAAA ACTAAGCTAA TCAGGAAGTA AGGCAATTAT  
 156781 TAACCTCAAG GAAACTAGA ATGACAGTAA CCATATCATA CTACATAATC ACAATGTAAA  
 156841 AAATGAATAT TGACCAAAAA ATTATAACAT AACAAATTG AGGCACTGAG AGAGGATAGA  
 156901 GCAAAAAAAA AAAAAAATAA AAAAAAATAA ATTGAGCCCT CATCTCCAT AATAGGAAGT  
 156961 CAATAAAAAA CAAGAAATAA TAGTATAAGC ATGTTATTTA GAAATCTGGA GATAGGGCTG  
 157021 GGCACCGTGG CTCACAACCTG TAATCCCAGC ACTTTGGGAG GCTAAGGTGT GTGGATCTTG  
 157081 AGATCAGGAG CTCAAGACCA GCCTAGCCAA CAAGGTAAAA CCCATCTCT ACTAAAAATA  
 157141 CAAAAATTAG CTGGGCGTGG AGGTGCACGC CTGTAATTCC AGCTACTTGG GAGGCTGAGA  
 157201 CAGGAGAATC ACTTGAACCC AGGAAGCGGA GCTTGCAGTG AGCCAAGATT GCGTCACTGC  
 157261 ACTCCAGCCT GGGAGACAGA GCGAGAAAAA AAAAAAAG AAATCTGGAG ATAAATACCG  
 157321 AAAGACACAA CTAAGAGAGT CCTTAGCAGT TGCTGCTGGG TGAATAGGAG CGGGGGAGGA  
 157381 AGCTAATGTT TCTTCATTAT AAGATTGTC AAGCTGTTTG ACTTCTAAA AAATAATTAC  
 157441 TTGGGCGGGG CATGGTGGCT CATGCCTGTA ATCCCAGCAC TTTGGGAGGC CAAGGTGAGT  
 157501 GGATCACCTG AGGTGAGGAA TTCGAGACCA GCCTGATCAA CATGGAGAAA CCTCGTCTCT  
 157561 ACTAAAAACA CAAAAATTAG CCAGGTGTGG TGGCACATGC CTGTAACCCC AGCTACTTGG  
 157621 GAGCTTGAGG CAGGAGAAATC ACTTGAACCC AGGAGGCGGA GGTGTCAGTG AGCTGAGATC  
 157681 ATGCCATTGC ACTCCAGCCT GGGCAACAAG AATGAACTT CGCCTAAAAA AAAAAAATA  
 157741 CTTGATCAAA ATATACATT CAAATTAAT GAAAATTAT TGAATATGT GATGTTTGT  
 157801 CATTAATAAT AAATGACTT ACAACACGTA TCACAACAGT TTGCCACTCC CCAGTGTGTTG  
 157861 GGGCTGCAAG ACCACAGCTC AGAGGCAGTT CATTCCTTT AGATGCTCAA TTCTAGGAGA  
 157921 AGCCATATGG GCAGGAAACA CGTGGAGTAC AGCAACGGCA AGTTCAAGAG ATGCTTGTGA  
 157981 AACTTCTGTA CATTATCAG GTTCTAGAT GAGACAAGAT CATGCTGTAT CTACTTTTTT  
 158041 CAAAAGAAGA ACTTTTCCCT CCCAGCTATA TCCACCCAGT TAGCCAATGG ACTAGTCAGG  
 158101 ATTTTTCAG TTGCAAGTAC CAGAAACCAA ATTGGAATG GCTTAAGCAG AAAAGCAATT  
 158161 CATAGAGTCC CATAAAGCCC ATGGGAAGCC CTGGGAAAAA CCTAGGTTCA AATCCTACCT  
 158221 TCAAATCCTA CTTTCAGTAT TTACCACTGT GTGACTTAAT TGCCTATAAA ATGAATACGA  
 158281 TAATAGTATC ATTTGGGATA TGAATAGAGT TAACATTTAC TTAGTGCTG GTGCATAGTG  
 158341 TATTAGTTCT CCCACTGCTA TAAGGAAATA CCCGAGCCTG GGTAAATTTAT AAAGAAAAGA  
 158401 AATTTAATTG GCTCAAGATT CCACAGACTG TACAAGAAGC ATGGCTGGGA AGGCATCAGG  
 158461 AAACCTACAA TAATGGCAGA AGGCAAAGAG GAAACAGGCA CACCTTCACA TGGCCAGACC  
 158521 AGGAGGAAGA GAGAGAAGGG AGACGTGCTA CACATTTTAA AATAGCCAGA TCTCCTGAGA  
 158581 ACTCACTTAT TATCATGAGA ACAACAAGGA GGAATCTGCT CCCCATGATC CAATCAACTT  
 158641 TCACAAGGCC CCTTTCCAA CACTGGGATT ATAAGTGGAC ATGAGGTTTG AGTGAGGACA  
 158701 CAAATCTAGA CCATATCACA GAGCTAACAC TTAACACATA TTATAACATT ATTATTGCCG  
 158761 GGCAGCGTGG CTGATGCTG TAGTCCCAGC ACTTTGGGAG GCCGAAGCGG GCGGATCATG  
 158821 AGGTGAGGAG ATCAAGACCA TCCTGGCTAA CGCGGTGAAA CCCCATCTCT ACTAAAAATA  
 158881 CAAAAAATAA CAGCTGGGCG TGGTGGCAGG CGCTGTGCT CCCAGCTACT TGGGAGGCTG  
 158941 AGGCAAGGAG ATGACATGAA CCCGGGAAGC AGAGCTTGCA GTGAGCAGAG ATCATGTCAC  
 159001 TGCATCCAG CCTGGGCGTC AGAGCGAGAC TCCATCTCAA AAAAAACAAA AAAAAAAG  
 159061 AACATTATTA TTATAAGCTA CTATTCAGGA ATCAGAGCAA ATCTGGAGAA TAAATTAATA  
 159121 AGTTATCCTT GGACACATTG TAATGTAAT GCAGAATACC AAAGATAAAG AAAAAATACT  
 159181 AAAAGTCCTC AGAAAAAAG ACATAAGTTG CCTATGAAG GATAGACAGG CTGGAGCAGG  
 159241 CTTCTCAACA GCAGTAATAA AAGCCAGAAG AAAGTAAAT ATCTTCAAAA TGCTGGGAGA  
 159301 AAAGAACCTG CCAACTGAGA ATTTTATACT CAGCCAAGCT ATTATTAAAG AATGAGCATT  
 159361 ACATGAATAA TTTATTGGAT CAACAAAAAT GGAAGGTTTA CCACTAAGAA ACCTCACCAG  
 159421 GCTCTCTAT AGATTTACTC CAGGAAGAAG GAAATAGAAT AGGTTATTAA ACATATGGTA  
 159481 AACCAAAATA TTGACTATAC AGATAATAAT CATAATGATG ACTAATATGG GGAATAAAAA  
 159541 ATGAAGGAGG GAAGTAGAAC ACCAAACAGT GACAAATGAA ATGGGAAGAT ATGATAGAGC  
 159601 GAAGATTTTA AAGCACCTGG CCTTGTCTG GAATAGGGTA AAGAAGTGGG TTAACTTTAA  
 159661 ACTTTGCTAA TTCAAGTAAA CAGGTTAAAA TTTAGAGATA CTTACAAAAA GAATAAAATT  
 159721 AGAACATGTA AATTCCAAGA TAATAACAG GTTAAATGT AGAGGTACTT AAAAAAGAA  
 159781 TAAATATTAGA GTATGTAAAT TCCAAGGTAA TAAAGAAAGG GGGAAAAGAA AAGGAATCCA  
 159841 AAAGTTTTCA GGGGAAAAA AAAGAAAAAC TCAGAAAAAA AGAGTGAATA GAAAGCACAA  
 159901 AATAAGATAG CAGGCCAGGC GCAGTGGCTC ACACCTGTAA TCCAGCACT TGGGAGCCCC  
 159961 AAGGCAGGCG GATCACCTGA GGTGCGTAGT TCAAGACCAG CCTGGCCAAC GTGGTAAAC  
 160021 CCCATCTATA CTGAAACTAC AAAAAATTAG CAGCCATCGT GGCAGGCACC TGTAATCCCA  
 160081 GCTATTCAGG AGGCTGAGGC AGGCAAAATT CTTGAACTCA GGAGGCAGAG GCTGCTGTGA  
 160141 GCTGAGATCG CACCACTGCA CTCAGCCTG GGTGACAGAG TGAGACTCCA TCTCCAAAAA  
 160201 AAAAAAATAA GATAACAGTA ATAAATCCAA ATATATCAAT AATCGTAATA AATGCTAATG

160261 GACTAAACTA ATCCATTAATA AACTATAAAT TTAATTTTCT GTATGCTTGA ATTTCTCCTA  
 160321 ATAAAAAAGAG TTTTAAAGAG AGATTGTCAA TCTAGTGTTT TTTTAAACT CCACCTGTGG  
 160381 CTGTTTGTAAG GAGCACACTG AAAACATAAG AACACACAAA AAAGTTGAAA ATAAGAGCCT  
 160441 GATAAAGATA CACCAGGGAA ACACATGATCA TGAGAAAAAT GGTGTACATA TAGATAAAAG  
 160501 AGACTCTAAT GTAAACATCA TTAGGAATAA GGAGGATTGC TGTGTAAAGA TCAATGGAAAT  
 160561 AATTTCGCCAG GAAGATATAA CAGTTATCTA CTTGCATGTA TCTAAATAAC ATAGCCTTAA  
 160621 AATGCATATGC TCTAATTAATA ACTTGACAAA ATTAATAATGT CCTAAGCAAA ACTTGACAAA  
 160681 TCTACAATTC TACCTCTCTC ATCAATGGAT AAATAAAACA TATAAAATAT TATTGAGAAAT  
 160741 ATAGAAGATT TGAACAACAC AGCGAACAAG CTTAATCTAC AGGACATATA CAGACCCCTA  
 160801 CAAATACATA TACAGGACAT ATACAGCCAC AACATTTAGG GAATACACCT TCTTCTCAAA  
 160861 CACAATGGAA CATTTTGTAC AAACATTGGT CATGAACAGG GTTTTATGAT AAGTCTCAAA  
 160921 AATCTGAAAA AGGTACCATA CAGATAACAT TTCTTTTACC AAGGTGCAAA TAAGGTAGAA  
 160981 ATCAATCACA AAAATATGCC TTACAAAATC CATGATTTGA GAATTATGAT ACACATTACC  
 161041 AAATGGAAGAA AGAATAAACT TACTTTTTTA AAAAAATTAG CCTTTTTTTT CAGAACAGAA  
 161101 AAACCTGTGAA GATAGTACAG TGTTCCCAAG GACCTGTGG CCAGTTTTCC TATTATTAACT  
 161161 ATCAAAACATG AGTATATTTG CTACAATTAA CCAATATTGA TATGTTATTA TTAATAAAAT  
 161221 CCATCCCTTA TTCAGATTTT CTTAGTTTTT ACCTAATGTC CTTTCTTTT CCAAGATTCC  
 161281 ATCTAGGATA TGACATTTAG TAGTCACGTC TTAGGCTCCT CTTGGCTGTG ATGGTTTCTC  
 161341 AGACTGCCCT TGTTTTGAT GACTTTGACA GTCTTGAGAA GTACTGGTCA GGTATTTTGT  
 161401 AAAATGTCCT TTGATTGAGA TGTGTCTGAT GTTTTCTCA TTATTAGACT AGAATCATGG  
 161461 TTTTGAGGGA GGAAGACCAC AGAGGTAAAG GGCCATTTTC CTCACATCAC ATCAAGGGTA  
 161521 CATGCTATCA ATGTGCCCTA TCAATGTTGC TGTTGACCTG ATCATCTGGA TGAGGTAATA  
 161581 TTTGCCAATT TTCTCTACTG TAAAGTTACT TCCTCCTTC CATCTGTAC TCTTTGGAAG  
 161641 CAAAGTCACTA AGTGGAGCTC ATACTTAAGG GGTGGGAGT CATATTCTAC CTCCTTTATT  
 161701 GAGGAGTATC TACATAAATT ATTTGGAATT ATTCTGCATG AAAGATTGT CTCTTCTTTC  
 161761 CCATTGATTT ATTTTATTAG TATTATTATA TATTAGTTTA TTTATATTAG TATGGACTCA  
 161821 TGGATATTTT ATGCTTTAGG TTGTAATCCA ATGCTACTTT ATTTCTTTTA ATTTTAAAAA  
 161881 ATTTTTTATT TTAATTAATT ATTTAATTAA AAAAAATTTT AAACATATAT TAATTTTTTT  
 161941 TTTGAGACAG GGGTCTTGCT CTGTCACCCA GGCTGGAGTG CAGTGGCATG ATCTCGGCTC  
 162001 ACTGCAACCT CCGCCTCCCA GGTCAAGTG ATTCTCCCGC CTCAGCCTCC CAAATAGCTG  
 162061 GGATTACAGG TGGGTGCCAC CACACCCAGC TAATTTTTTT TTTGTATTTT TAGTAGAGAT  
 162121 GGGGTTTTAC CATGTTGGCC AGGCTGGTCT TGAACCTG ACCTCAAGTG ATCCACCCGC  
 162181 CTCGACCTCC CAAAGTGCTG GGATTACAGG CATGAGCCAT AGTGCCAGT CAACCTTGTT  
 162241 TTCTTTTGTT GCTCAAATTG TTCCACCATT GGCCATTGGG GAGAAATTGT TGTGCTCTCC  
 162301 TGTGTCTCTT TTGACACGTA TGTATTGTTT TGCTTTTGA GCACCTCCTT GCTTTTTGAT  
 162361 CAATCTCTTA CTTTCTGGCA CTACAAGATG CTTTAGGATC ATCTTGATA CTGACTGTTT  
 162421 CCACTCTAGA ATGAACCATT TCTCCAAGAA ACCCTGGTTC TTTCTTGTG AAGAATGCTA  
 162481 TTAGAAGCCA AGATCTGGGC ACTGTGTGTG CTATTGCTA CTGGGGTGTG ATTACTCTA  
 162541 GGTCTCTCTA GCAAACAGAG CTAGGCTATT GTGTGTGTGT GTGTGTAACA CATATTTAAT  
 162601 ATATTATATA TGTTTATATA TATGATGTA TATGACATAT ACTGATCTGT GTATACACAC  
 162661 ATATATATAA TTATTTTAT ATCTATCCAT CAATATCTAA ACTAAACATG GTATCTCTGA  
 162721 CTCTAATCTA GCACCACATT GTTCATTCTA GTAATCCAAA AGCCTTTTGA AAAACTGAAT  
 162781 TACTGGCTGG GTGTGATGGC TCATGCCTGT AATCCCAGCA CTTTGGGAGG CCGAGGTGGG  
 162841 TGGATTGCCT GAGGTGAGGA GTTCGAGACC AGCCTGACTA ACATCGTGAA ACCCTGTATC  
 162901 TCTACATAAG ACACACACAA AAAAAACAGC CACGCATGGT GGCACACGCT TGTAGTCCCA  
 162961 TCTACATAGG AGGCTGAGAC AGGAGAATTG CTTGAACCCA GGAGGTGGAG GTTGCACTGA  
 163021 GCTGAGATTG TGCCACTGCA CTCAGCCTG GGTGATAGAG CAAGACTCCA TCTCAAAAAA  
 163081 AAAAAAAGAA AAAAAAAGAA AAAGAAAAAC TGAATTAGTA ATTCAAAAAT TAACAACAAT  
 163141 AAAAAAATAA CTCAGACCC AGATTACTTT ATAGGCAAGT GTTACCAGGT AAAGACTAGA  
 163201 CACCATTTTA AATGTGCATT TGA AAAACAT CAAGACCTCA ACAAATGGAA AAAATATGCC  
 163261 ATGTTCCCTA GATAGAATCT TTATCTAAAG ATATCAATTC TCTCAAAAAA CAATCTCTAT  
 163321 ATTTAATGCA ACTTTAAATA GAATCCCAAA AGGGTTTTTC ATGAAACTTG ATGAGCTGAA  
 163381 TCAAAAAATT ATATGGAAGA TCAAAAAAAT AGTAATAGCC AAGACAATAC TAAAAAGAAC  
 163441 GGAAGGAGGC TTGCCTTACT GTAAATTAAT TACAGCAGCT GAGTCTTGGA GCAAGAACAG  
 163501 ACACACAAAT ATAACAGAAT ATAGAGCCCA GAAACAATAC CTACACACAG CTGGAAACTT  
 163561 GTTATATGAT AAAAGTGGCA GTGTCGATCA GGGCAGAAAA CATACTAGCT ATTGAATGGT  
 163621 GCTAGGGAAG ACATAAAATC AATTCAGGCC CTATATTTTA TTTTGTGAGA CTGGGTCTGG  
 163681 CTCTGTCATA AGGCTGGAGT ACAGTGGTGC AATCTTAGCT CACTGCAACC TCCACCTCCC  
 163741 AGGCTCAAGT GATCCTTCCA CTTAGCCTC CTGAGTAGCT GGGACTACAG GTGCAAGCCA  
 163801 CCACACCTGG CTAATTTTTA TATGTTTTGT AGATACGAAG TTTTGCCACA TTTCCAGGC  
 163861 TAGGCTTGAA CTCTGAGCT CAAGCAATCC TCCCGCCTCA GTCTCCCAA GTGCTGGGAT  
 163921 TGCAGGCGTG AGCCACCGTG CCGGCGCAGG CCTATATTT AAAAAAGTAA GATTAAAACT  
 163981 CTTAGAATAA AATATATGAT ATTTAGAACA TCAAGATAAG GAAGGATTTT TTA AAAAAGA  
 164041 AACAAAACCC ATAAATCACA AAGGATAACT ATTGTTTATT CAATTATATT AAAATATAAA  
 164101 ACTTTAAAAA AAAAGACACC TTCCAGTGTG GAAAAATAA TTCACAGACT GAGGGAAGAT  
 164161 ATTTGCAATT TATTTAACCA AAAATGTATT AGAATAAAAA AACATATTTA TATGTTTACA  
 164221 AATCAGAAAA AGACAAAGAA CTGAATAGAA AAATGGGCAG AATATACAAA CCAGCATATC  
 164281 ACAGATGAGC AAACCAGCAC AGACAAAAAC ACACCAAAAA TATGCTGCAC CTCACATAA  
 164341 TATGAGGAAA TGCAAAATTA ATCCACAGTG GGATACTATC TCGAACACAT CAGATTGCA  
 164401 AAAACTCAAA AGCCTTACAA GCTCCCTGAA CTCAGCATTT TCCCTTCTCT TAAGTCATTC  
 164461 AAAGGCAACA GGAGAAAAA ACAACAGAAA CACAACTCT ACGTTTGATG AAATTTAAAA

164521 AAACATTTGA ATCCCAATCC ACAAATATAT GTGGACTACT AAAAGCAATG AAGATCAAAC  
 164581 AGGGGTGTGG AGGGAAAGGG ATCAAGAATG CCTACAACCT CTGAGCTCAC AGACACCTAC  
 164641 AGGGTGCCCA TGCCAAGGGG AGGAACACAC CTTAAACTGC AAAGCCAAAA TAATGGATGG  
 164701 GGTGCATGGA CTGTGTGGCAG GAACTTCTCT GGAATGGATT TAGTTATGGA AAATATTGGG  
 164761 GGGCGGGCTG ACTGACAAAT CCCAGAAGCA AGACTCTTTA AGGAGAAGCT GTGTATTCCA  
 164821 ACAGCTGCCT ATCTCATTGC GCTGAGCAAC GATAAAGCTA AAGGAACTCA TGCACCTGA  
 164881 AAGGACAGTT CCTGCCGCAG CACAAGCAGC CTATGCCCCC CCCTGAGACT CTTATTCTGA  
 164941 AATAGCTGGT GTCTGTAGCA GGGTTCTTCA CAGAAACAAA TAGGATGTGT ACATAGAGAG  
 165001 AAAGAGATTT ATCATAAGGA AGCAGCTCAC AATTATGGAG GCTGGCAAAAT CCAGAATCTG  
 165061 CAGCGTGGCG CAGCAGGCTG AAGACGCAGG AGAGCTAATG GTGCAGGTGA GGTCTAAAGG  
 165121 CAGAGTGTGC TGCAGAACGG GCTCTTGCTC GGGGAGGTCA GTCTTTTTCG TCAGTTCAGG  
 165181 CCTTAACTG ATTGGATGAG GCGCACCTAC ATTAGGCAGA GCAGTCTGCT TTAAGTACAG  
 165241 TTCAACCAAT TAAATGTAA TCTCATCCCA AAACACCCTC CAAATTGACA CATAACATTA  
 165301 CCCATTACAG TGTCTTAGGT CTGGTTCCCTT AGAAGCAGAG CCTGAGGCAG GGACTGGGGA  
 165361 CAGTTGTGTG ACAGGGAGAA TACTCCAGG AGACAGGGAG AGAGGGAAGC AGGGAGGAAA  
 165421 AATAGCTAAA CAAACACGTG GTCTCAGCTG CAGCCTCGCT TCAGTCTGGT CCCACGGGGA  
 165481 CCTCTGCAGC ACAAATTGCA AGAGGTGGTC CCACTTTGGA GCAAGGGGCC TGGCCTTCTG  
 165541 AGCCCCCTCG TCTGTGGGG ATTGACTGCT GACGGCTCA GGGTTACGGG TATGAGTGAG  
 165601 CAGTTAACTG CTCAAGTGCG ATGGTTTCAAT GCATGTGCTT CCTTGGCTAG GCTATGATG  
 165661 CCAGCAGTTC AGTCAAACAC TAGCATAGAT ATTGCTATGA AGGTATTTTG TAGATGTGAC  
 165721 TAATTTACAA TAAGTTGACT TTAAGTAAAG GGCATTACCC TCAGTAATGT GGGTCGGCCT  
 165781 CGTTTAACTA GTTGAAAAGC ATTAGGAGCA AAAACTGAGG TTTCTCAGAA AAGGAGGAAT  
 165841 TCTGCTTAA AGACTGCAAC TCTTGCCCTG GTTCCAGCG TGCCAGCCTG CCCTACAAAT  
 165901 TGCAAGCTTG TCAGAGCCCA CAGTTGCATG AGCCAATTCT CTCCCTCTCT CATTGATATT  
 165961 GTCTCTCTGA AGAACCCCTG GTGATAAAGC AGGTGTGGGA CTCTCATTCA GCCAAGCACA  
 166021 GATCTTTTGA GAAGGGGTTA GCTGAGGGCC ATTAGCAACC AACCTCTCTA GCAGCCAGGC  
 166081 TGGTCAAAGG AAGCTGTGCA GGCCACCAAC AGCATATATC ACAGTCCACC TCCCAAACCA  
 166141 CTGAGATCCA ATTGCTTCCC ATATTAACTT CATTCAATCT AGGCACAGCT TTTCCAGGAT  
 166201 TCTGGCTGGT CACAATTGTG GGGAACATTT ACAGGAGAAG AGTTAGTGGA ATGAAGTACA  
 166261 GCCCACACCA CCACAGTTTG TCTCAATCTG TCACTGATAC TCATCATCTT CCTCCTCCAC  
 166321 CACCCACTCT AGGTTCTCTT GCCTCTCAGC GAGCACTCCT GCTGCTCTCT GTGGCTCTCA  
 166381 GAGGGGTGCT TCCAGACCTT TATCTCAGAG GAGTCTGAAG CTCAAGTTAT GCTTTACTCA  
 166441 GACCATGGTT GCTATATTTA CCCATTTTGA CATGCAAAAT TTCACATAGA AAAAAGCAAA  
 166501 CCCAGGATCA CCTGAGTGCC AAACATATTC CTCCCTGCCC CCATCATCCC ACTGTATAGC  
 166561 AGCAGCTCTG TGTCTTCTG ATGATCAGAT CAATTACCCC TGCTGGAGCT GTAACCCCTT  
 166621 CCTTTGCTG CTGATCTACC AGCATGAGGA GGCCAAAGTG GCCAGGCTGC AGCCATAGCT  
 166681 GTAAGATTAT TGGGACTCCT ATGGCGTCCT CCAATAGAAG CATTTCCTGC TCCTGCCCCC  
 166741 AGGAACACAG AGCTCTAGAC CCACAGAGTC TAAAGTTTCT GTGAAGAGAA GCATAAATTC  
 166801 CCCAGGTGGG TCACTTAGAA TAGTGATGAG CAGAGCCACT CCTACTTCTA CCGCTTGGAC  
 166861 CTTGGACCAT CCATTCTACC TAGTGGGGAC AAATGCACAT GCAATGGCCA TTGGCTTCAG  
 166921 ATGTGTATTG CATCCTGAAG ATAGCACACT ATCCCCAAAC TAGCTCCTCC ATGCATCTTC  
 166981 AGGAGGCCAT CTCGCCACTC TATTGGGCAA TAGTTGGTGG TTGATGTGAT ATGCACCACA  
 167041 CTGCGCTCCT TACTCTAAGT TGCTGTTATG CACAGTCATA TGATAGATTAG ACATTGTGTA  
 167101 AGTCGTTGGA TTGTCGTGCT TGCTGAGGTG CAGGAAGCCG GAAAGGTAAT CCACATACAG  
 2 167161 AATGTGTGTC CATTTCAGTA AGTATGAATT GTCCTTTTCA GGACAGAAGG GGTCTGTAT  
 167221 AATCAGCTTG CCATCAGCGG TCAGTTGAGC TCCTGGGTGC CTTATTGCAG ACTCGGTGTC  
 167281 AGTCTTTGCA GCTGGCAGGT TCGACAATCA GCAGTGTGAG TAGCTAGGTC AGTCTTGGAA  
 167341 TGGGGATCCT ATGCTATTGG ACCCTGTATC TGCTTCCATC TCTGCTCCAT TTGGGAGCAC  
 167401 AAACATTAAG ATGGCTGAGG ACAGGGCTGA TTGACAACCA CTGGACGATC CATCCACTC  
 167461 ATCTGGTTGT TCAGTCTTTC TGCGAGTGGT CAGCACTGTC ATTAGCATAA AGATGCTCAC  
 167521 CCTTTGTGCC TACTCCGAA AGGGCTCAGC TGCACTGCAG ACTGGGCAGA ATAGGCCAAG  
 167581 TGACAGGCCA GCCCTGGTAC TGGTGACTAA CAGCCTGCTG GACCTCCCTG CCCTAGTCTT  
 167641 GAGCTTGAGC TCTGCCGCTT GGCTCTGGGC CTGGCCCTGG AGGGGTCCAC AGGTTGCAAG  
 167701 AAGTACCTTG TGCAATCCAC ACTGGAAGCC TGCTCTCACC CAGATCAGGG TGGAGCTGAG  
 167761 GGGGCCCTCA GGGTTTCAGG GCCTGGGTGT AACAGCAGCC AGTCAGATCA GGGTAGATAC  
 167821 ACACAGATAC ACACACACAC ACACACACAC ACACACACAT ATTTACCTTG GGGCTCAGCC  
 167881 CAGCTCACTC TGGAGGCTGA AGGCCTCCTT TCTAGCCCTG TATCTGGCTT CATTGGAAGT  
 167941 TTCCCAATAG CAAAGACTTT TTGTTGTTGT TGTGAACTC AGGCCCTTA CTTTCTCCAC  
 168001 ATTCATGCCT AAAACACAAG ATGCGAAAAA AAAAAGCCAA GTGCTCAAAG GCATACGGTA  
 168061 GTATGGTGGC TAATTCTTTG GGTGTTGGAA TTAAAAAAAA AAAAAAGCC CTGAGTTTGA  
 168121 GTTCAAGGTA TGGCTTTGCT ACTTTTGACT TTAATGTTCT TGATAAATA ACAATTCTCT  
 168181 GAGCCTCAGT CTCTACATCT ATAAAACATG AATAATATTA CCTCCCTCAC AGAATTGTGA  
 168241 GGATAAAGGA AAACACATAT GTAAAATGTT TGACACAAAA CAAGTATTCA ACAGAGGCTA  
 168301 TATAACAATA TTATTAATAA CAAGAGTTTG TGATAAATA GGGAAACACAG AGAGAAAACC  
 168361 CAAAATAAAT GCAGAAGACA GAAAGAAATG GGAACACCCG AAGACCCTGT GGCAGCATG  
 168421 AAAATCAGAC CTCCTTCAAA AGAACCTGCT GCAGGCTTAT AGTGACCACC AGTCCCTGCG  
 168481 CACCCTCTGG AGCTATCACT GCGTTCACAC CAAGGCCACG CTTTCTTGAG TTACTCTTAG  
 168541 CCAAAGACAG AACGTGGTGG GGCCTAGTGT TCTGCCCAT CTTGTGGGAT ATGGGACCCC  
 168601 TCTAATGGGC AACTATGGCT CCAGGACTCC CATCAGCCGG GCTGGAATTT TCTTAGAACT  
 168661 GAGCTCTGAG CCCTTCTTTC CATGCCCTTG TCTTCCCTG TCTCCTTTCA TAGGTGTCCA  
 168721 AATGGTATTA CAATCTGAAG GCTCTCCTGC CCCTTTCTCT CCTCATTAC AGGTGTTTAC



168781 TTTAACAAGT CTCTGGCACA AATAATCCCA TCTTGGCCAC TTTAGAGCAC CTGAACTAAC  
 168841 ACAGGCCCCCT TCCCTAATGC TGGGCTTTCT TAGAGGGAGG GTCCCTGAGAG GATCTGTCCC  
 168901 AAATCTGTGT TGTGCAACCAC CTGGGTGAGG GCAGCCGCAG TTCCAGGTTG ACATGAAATT  
 168961 CTTAGATCAT GGAGCTTGAA TGGGAAACAG GAAGTGAGTG TCCTTGGGAC ATGCTATGAT  
 169021 GCCTGATGGT TCATGTTCCT CTCTGTACCT GAAGTCGCCC CTCAGATTAA GCAATCCCTG  
 169081 CCTCTTCCTG CCCCCACTCC CTAGGCTGGC TCCTCTGGCC TGCACCTCAA ATTATCCTG  
 169141 CTTGTGCATA AACATAATTA GTGTGAATGC TTTGTTTGGT TGAACCAAGT TTCTATATAG  
 169201 GTCTGGGGAG GAGTTTATTG TTGGCTTCTC TGAAGCCATA AATGTAAAGG AGAGCTTTAG  
 169261 GTGGGATGTG GCTCCCTGGG CAGGGAGACA AGCTGCAGAC GTGGCCCAAG CCTAGGAGAG  
 169321 CTGGGAAACT CCAGCCAGG GAGGGCTGTG GCTATGCCCA CTTGCAGGTT GGGTTAAATC  
 169381 CCCAGCCTTT GCAGGGACAG AGTGTAGCCC TCAGGGTCCT GGGGGTTCTG GGTGAGATAA  
 169441 GGGTCTCTGG AGACTTCCCT CCACCCCTCA TCCCAGGACC CAGTGCATCA GGGAAAGTCTT  
 169501 CAAGGTCTGC AGCATGGCGC TGTGGCCGCA CAGGTGCAAC AAGTGGTCGT AAAAGTCTTT  
 169561 CACCTTGAAC ATCATGTAGT ACAGAGAATA GGGTGAGGAA AAGTCACAAA GGTGTAATTG  
 169621 CTTTCCCTGC CTCCAAGGAG AGAGGAGATT AAGAGCAGGT CACAGAAGGC TTCCACCCAT  
 169681 CTAGTAAGCT GAGGCTCTCA GGTGGGAAC AAGCCAGGCT TCATGCTTGG AAACATAGAA  
 169741 ATGTTGCATT TTCTGGATTG AGACACATGT AACCACATGT AACCAGCACT ACAGTGTGAT  
 169801 ATAGAGTAGA TTCACTACCC TCAAAAGCTT CTGTGCTCTG CCTCTTCATC CCTCCCTCCC  
 169861 CCTTAACGCT TGCAATCAC TGATTTTTTT TTTTACTGTC GTTTAGTTTT TCTTTTTTCA  
 169921 GAATGTCATA CAATTGAAAT CATACTCTGT ACTCTTTTCA GACTGGCTTC TTTCACGTAG  
 169981 TAATACACAT TAAGTTTCCT CCATGTCTTT TCAAGGCTTG AATGCTTATT TCTTTTTAGT  
 170041 GCTGAATCAT ATTCCATTGT CTGGATGTAC CCCAGCTTGC TCATCAATTC ACTCACTGAA  
 170101 GGACAGCTTG GTGTCTTCTA AGGTTTAGCA ATTATAGCTA AAGCTGCTAT AGACATTGAT  
 170161 GTGCAGTCTT TGTAAGAGA TAAGTTTTCA ACTCTTTTGG GTAAATATTC CAAAATGAAA  
 170221 TGGAGAATAT GTTTCGTTTT TTAAGAAACC TTCAGATTG TATCCAAAGT GACTGAATCA  
 170281 TTTTGCATTG CCACCATTAA TGAGTGGAAG TTCTGTGTC TCCATGTCTT CGCCAGCATT  
 1 170341 TGGTGTGCTT AGTGTCTGAG ATCTAGTCAT TCTCATAGAT GTGTAGTACT ATTTCAATTG  
 170401 TCCCTTAATT TGCATTTCCC TGATGATGTA TGATACGGAG CATCTTTTCA TATGCGTACC  
 170461 TGTCTCTCTG ATTTCTTCTT TGGTGAGGTG TCTGTAAAGG TCTTTGGCCC ATTTTTTAAT  
 170521 AGGGTGTGTT TCTTTTATC AAGTGTTAAG AGTACTTTGT ATTTTGGATA ACAGTCCTTT  
 170581 ATTAGATATG TCTTTTGCAA ATATTTTCTT CCAGTCTGTG GCTTCTCTTC TCATTCTCTT  
 170641 GGCAGTGTCA TTTGAGGAGC AGAAATGTTT TTTTGTGTTT TTGTTTTTTG AGACAAAGTC  
 170701 TCACCTTATT GCCCAGGCTG GAGTGCAGTG GCACAATCTC GGCCCACTGC AACCTCTGCC  
 170761 TCCCGGGTTC AAGCAATTCT CCTGCCTCAG CCTCCCAAGT AGCTGGGATC AGAGGCATGC  
 170821 GCCACCATGC CCAGCTAATT TTGTATTTTT AGTAGAGATG CGGTTTCACC ATGTTGGCCA  
 170881 GGCTGGTCTC AAACCTCTGA CCTCAGGTGA TCCACCCGCC TCAGCCTCCC AAAGTGCTAG  
 170941 AATTACAGGT GTGCCACCAC ACGCCTGGCC AAGAAATGTT TTATTTTTAT AAAATCCAGC  
 171001 TTATCAATTA TTTCTTTCAT GGATGTTGCC TTTGGTATTG TATCCAAAAA GTCATTGCAA  
 171061 AACCTTAAGT CATCTAGATT TTCTCCCGTT ATCTTCTAAG AGTCTATAG TTTTACATT  
 171121 TACATTTAAG TCTCTTATCC ATTTTGAGAT AATTTTGTG AAGAGTATAA GGTCTGTGTC  
 171181 TAGCCATTAT TTTTGTGATG TTTTGTGATG TGATGTCCA GCTGCCCGAG CATCATTTGT  
 171241 TGAAGAGACT CGGCTTTATT GTATTGCCTT TGCTCCATTG TCAAAGATCG TATATTTATG  
 171301 TAAGTCTATT TCTGGGCTGT CATCTGTGTC CATTGATCTA TTGTCTCTG TGCTGTGTCT  
 171361 TCATCCCTTC CTCCCCCTTA ACTTCTGTCA CCAATATCAC ATTGTCTTGA TTATCATAGC  
 171421 TTTATAGTTA GTCTTTAAGT TAGATAGTGT CAGTCTACCA ACTTTGTCT TCTTTAATGT  
 171481 AGTTTGGCTT ACTCGGGGTT TCTTGCCCTC CCATACAAAC TTTCGAATCA GTTTGTGAGT  
 171541 ATCCACAAAA TAGCTTGCTG GGATTTTGAT TGGAAGTGCA TTAAGTATAT AGATCAAGTT  
 171601 GTGGAGAAGT GACATCTTGA CAATATTGCA TCTTCTTAAC TATGAATATG GAATATCTCT  
 171661 CCATTTACTT AGTTCTTTAA ATTTCTTTCA TTGTGGTTTT ATAGTTTTCC TCATATAGAT  
 171721 GTGTGTCATA TTTTGTAGCT TTATACCTAA GTATTCAATT GTGGGGATGC TAAATAGTAT  
 171781 TGTGTTTTTT AATTTCAAAT TCCACTCATT CATTGCTGGT ATATAGGAAA GTGATAGACT  
 171841 TTTAGATATT AAACCTGTGT CCTGCAACCT TGCCATAATT GCTTGTTCAC TTGGAGTTTT  
 171901 TCTAGTACAT GGTATGGAGC ATGTGAATCC AATTTTATTT TATTTTACAA ATAACATCT  
 171961 GGTATATACA ATGTCCCTAT TTAAGAGTCA TTTTCTTCTT CAGTGAATTA TCATGTATAG  
 172021 TTTACACTAA ATGTTTCATAT GTATTGCCAT CTATTTTAT ACTGTGTATT CTGTTCTGTT  
 172081 GGTATATGTT CAGCAGCATA TAATTTTAAAT TATAAGGCTT TTATCATGTT TTAATGTTTG  
 172141 ATAAGGCTCA TCCCCTGCCC CCCATTACCA CCATAATTAT TGCTTTTTCA AGAGATGGGG  
 172201 TCTCACTATA TTGCCCAGGC TGATCTCAAA CTCTGGCCTT CAAGCGATCC TCTGCCCCA  
 172261 GCCTCCCAAA GTACTAGGAT TATAGGTGTG AGCCACCGTG CCTGGCTCCT GCAAGTTTAT  
 172321 CTGTTTCATAT AAACCTTTACA ATTAACCTCC CTAGCTTAAA AAAATAGTTA TTTTTAATTG  
 172381 AATTACATTG GGGGAAATTG ATGACTTTAT GACATCAATT TCCCCCAATG TAATAAATAT  
 172441 ATACTTATTT TTATTATTAT TATTTTGTGA GACAGAGTTT TGCTTTGTCA CCCAGGCTAG  
 172501 AGTGCAGTGG CACAATCTCG GCTCACTGAA ACCTCTGCCT CCCCAGGTTT AAGCAATTCT  
 172561 TGTGCCTCAG CCTCCTGAGT AGCTGGCCAT CAGGCACCTG CGACCACACC CAGCTGATTT  
 172621 TTTGTATTTT AGTAGAGACG GGGTTTCGCC ATGGTGGCCA GCCTGGTCTT GAACTCCTAA  
 172681 GCTCAGGCGA TCAGCCAGC TCAGCCTCCC AAAGTTCTGG GATTACAGG CAGAGCCACC  
 172741 ACACCCGGCC TCTACTTACT TATTTTATTA TATTTATTTA TTTATTTATT TTGAGATGGA  
 172801 GTCTCGCTCT GTCAACCAGG CTGGAGTGCA GTGGTGCAAT CTGACTCAC TGCAACTTCT  
 172861 GCCTCCTGGG TTCAAGTTAT TCTCCTGCCT CAGCCTCCTG AGTAACTGGG ATTACAAGCA  
 172921 CTTGCCACCA TGCCAGCTA ATTTTGTGTA TTTTGTAGTA GACCGGGTTT CACCATGTTG  
 172981 GCCAGGCTGG TCTTGAACCT CTAACCTGCG CCTCCCGGGT TCAAGCAATT CTTGTGTGTC

```

173041 AGCCTCCCTA AGTGCTGGGA TTACAGGCAT GAACCACTGC ATCCAGCCCA TTTATTTATT
173101 TTAGATTCAG GAGGTATACG TATAGGTTTG TTACATGGGT ATATTGCATG ATGCTGAGGT
173161 TTGGGCTTCT AACCCAAGTA CTTCGCCCAA ATGATATGGG ATGGGACAGG GAAGTTCTGG
173221 TTAGAGAAGG GCGGGGTCCC TGGCGAGGGC TCCATCCCA TTTTGTGCTC AAATGTTGTG
173281 TTTCCCAAGA CCACCTGGC CTGCCACACT CCCATCCTGT GCCTATAAAT ACCCTGAGAC
173341 ACTAGCGGGC CAAGCAGCTG GATGTTGAGA GGAACACAGC CATGGAAGAA GACACAAGCC
173401 ACCTAAAGAT GGCAAACTG AAAGAGCACA CCATAACACA CGCCCACTGG GGCTTCAGCT
173461 GTAAACATTC ACCCCCTAGA CACTGCCATG TGGTCGGAAC CCCACAACCT GTCTGTATGC
173521 TCCCCTAGAG ATTTGAGCAG CGAGCCACAC CCCCATCGCA CGCCCTGCAG GGGGATAAGG
173581 GAACTTTTCC CATTTCAAA GTATTGAATA GGTAGTTGTT CAACCCTTGC CCCCTTCCTT
173641 TCTCCCTGCT TTTGGAATC CTAGTGTTT ATTGTTCTG TCTTTGTGT

```

## LÉGENDE

XXX : LES NUCLÉOTIDES SOULIGNÉS FONT PARTIE DE LA SÉQUENCE CODANTE

XXX : LES NUCLÉOTIDES ENCADRÉS DÉSIGNENT LES CODONS DE DÉPART OU LES CODONS STOP

XXX : LES NUCLÉOTIDES OMBRÉS INDIQUENT QU'IL Y A UNE INCERTITUDE SUR LEUR IMPLICATION AU NIVEAU DE LA SÉQUENCE CODANTE