

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 A₂ 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 A₂ (PLA₂) 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 A₂ cytosoliques (cPLA₂) 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 cPLA₂ sont spécifiques à la rétine et, par conséquent, que leur malfonctionnement pourrait être associé à diverses dysfonctions rétiennes.

L'objectif de ce projet de maîtrise consistait donc à isoler l'ADN complémentaire complet d'au moins une de ces cPLA₂ 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 cPLA₂ déjà connues au niveau de l'EPR et de la rétine neurale a permis d'établir que la cPLA₂ gamma est présente au niveau de ces deux tissus, alors que la cPLA₂ alpha en est absente. Cette approche ne s'est toutefois pas avérée concluante pour la cPLA₂ 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 cPLA₂ 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 (réteine neurale). Par conséquent, plusieurs expériences ont ensuite été réalisées afin de cloner l'ADNc complet de cette nouvelle cPLA₂, mais se

sont avérées vaines. Enfin, la présence de l'ADNc de cette nouvelle cPLA₂ dans plusieurs autres tissus permet d'exclure la possibilité d'un gène spécifique à la rétine.

TABLE DES MATIÈRES

1.2.3.6 Résumé comparatif des PLA ₂ des groupes VI, VII et VIII.....	23
1.2.4 Les phospholipases A2 cytosoliques (groupe IV).....	24
1.2.4.1 Isoformes des cPLA ₂	24
1.2.4.2 Domaine CaLB.....	26
1.2.4.3 La cPLA ₂ alpha (groupe IVA)	27
1.2.4.4 La cPLA ₂ bêta (groupe IVB).....	27
1.2.4.5 La cPLA ₂ gamma (groupe IVC)	28
1.2.4.6 Résumé récapitulatif des trois cPLA ₂	28
1.2.4.7 Régulation transcriptionnelle de la cPLA ₂ alpha	29
1.2.4.8 Localisation cellulaire, intracellulaire et tissulaire.....	30
1.2.4.8.1 Localisation stratégique pour la biosynthèse des eicosanoïdes.....	31
1.2.4.9 Activité spécifique envers les phospholipides et leurs acides gras	32
1.2.4.10 Mécanisme catalytique de la cPLA ₂ alpha.....	33
1.2.4.11 La phosphorylation de la cPLA ₂ alpha.....	34
1.2.4.12 Activation de la cPLA ₂ alpha et son interaction avec certains éléments.....	36
1.2.4.13 Inhibiteurs des cPLA ₂	38
1.2.4.14 Implication des PLA ₂ dans diverses pathologies	39
1.2.4.15 L'implication des cPLA ₂ pour l'intégrité des photorécepteurs.....	41
1.2.4.16 Implication des cPLA ₂ dans l'inflammation oculaire	42
1.2.4.17 Implication des cPLA ₂ de l'EPR dans la digestion lysosomiale.....	42
1.2.5 Nouvelles phospholipases A2	43
CHAPITRE 2 OBJECTIFS DE RECHERCHE	45
2.1 Objectif global	45
2.1.1 Premier objectif spécifique	45
2.1.1.1 RT-PCR à l'aide d'amorces spécifiques	45
2.1.2 Deuxième objectif spécifique.....	45
2.1.2.1 Approche expérimentale	46
2.1.2.1.1 RT-PCR à l'aide d'amorces dégénérées	46
2.1.2.1.2 Criblage de banques à l'aide d'une sonde d'ADNc	46
2.1.2.2 Approche bioinformatique	46

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
CHAPITRE 3 MATÉRIEL ET MÉTHODES	50
3.1 Préparation de l'ADNc de l'EPR et de la rétine neurale.....	50
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
3.2 Réaction de polymérase en chaîne (PCR)	55
3.3 Clonage des bandes PCR.....	60
3.3.1 Purification des bandes	60
3.3.2 Clonage des bandes	60
3.3.2.1 Préparation des bactéries compétentes.....	62
3.4 Purification de plasmides	62
3.5 Digestion des plasmides	64
3.6 Criblage de banques d'ADNc.....	64
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
3.7 Criblage d'une banque génomique.....	80
3.7.1 Titrage de la banque génomique	80
3.7.1.1 Culture des cellules hôtes <i>E. coli K802</i>	80
3.7.2 Isolation de l'ADN à partir des lysats de phages	81
3.8 Amplification rapide des extrémités de l'ADNc (RACE).....	82
3.9 PCR sur différents ADNc de tissus humains	85
3.10 Criblage bioinformatique	85
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
CHAPITRE 4 RÉSULTATS ET DISCUSSION.....	90
4.1 Recherche des cPLA₂ connues dans l'EPR et la rétine neurale.....	90
4.1.1 RT-PCR à l'aide d'amorces spécifiques	90
4.2 Recherche de nouvelles cPLA₂ dans l'EPR et la rétine neurale	92
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
<i>CHAPITRE 5 CONCLUSIONS</i>	112
5.1 Récapitulation des résultats	112
5.2 Conclusions.....	114
5.2.1 Atteinte des objectifs fixés	115
5.3 Perspectives de recherche.....	116
<i>RÉFÉRENCES</i>	118
<i>ANNEXE A ARTICLE SOUMIS</i>	133
<i>ANNEXE B SÉQUENCES DES NOUVEAUX GÈNES 1 ET 2</i>	141

LISTE DES TABLEAUX

1.1	Compilation des différentes PLA ₂ à histidine catalytique	17
1.2	Résumé des caractéristiques des PLA ₂ des groupes VI, VII et VIII à sérine catalytique.....	23
1.3	Résumé des caractéristiques des cPLA ₂	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 ₂ connues dans la littérature jusqu'à présent.....	25
1.14	Schéma de la cPLA ₂ alpha montrant les acides aminés les plus importants pour son activité	34
1.15	Modèle d'activation de la cPLA ₂ 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 λ 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 ₂ alpha dans le poumon	91
4.2	Obtention de la cPLA ₂ gamma dans l'EPR.....	92
4.3	Schéma représentant les HTGS trouvés.....	95
4.4	Séquences codantes des nouvelles cPLA ₂ 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 ₂ connues (alpha, bêta et gamma)	103
4.8	Représentation schématique des trois isoformes connues (cPLA ₂ 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 ₂ (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 messager
BLAST	<i>basic local alignment search tool</i>
pBPB	bromure de para-bromophénacyle
CaLB	domaine de liaison au calcium (<i>Ca²⁺-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- γ	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 oléïque
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 ₂	phospholipase A ₂ sécrétée
cPLA ₂	phospholipase A ₂ cytosolique
PLC	phospholipase C
PKC	protéine kinase C
PVP	polyvinylpyrrolidone
RE	réticulum endoplasmique
SCL	12-épi-scala-radial
SDS	dodécylsulfate de sodium
TNF- α	facteur de nécrose tumorale alpha
TPC	thioéthéramide phosphatidylcholine

CHAPITRE 1

INTRODUCTION

1.1 Composantes de l'oeil

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).

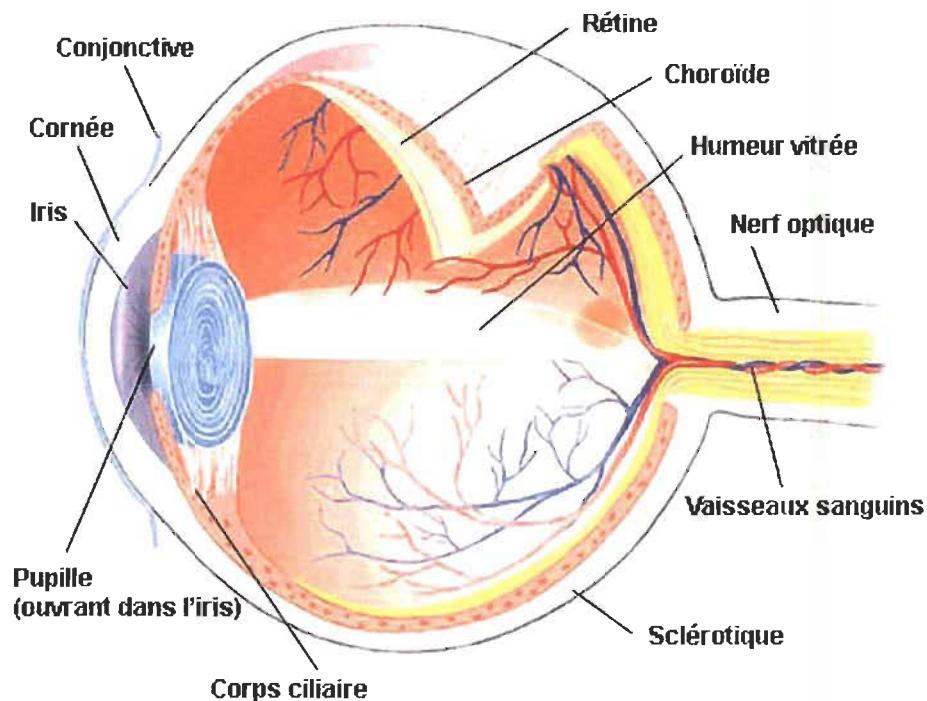


Figure 1.1 – Coupe transversale d'un œil humain. Image tirée du site <http://www.driesen.com>.

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 radiales 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 rétiennes (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 rétiniens, 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).

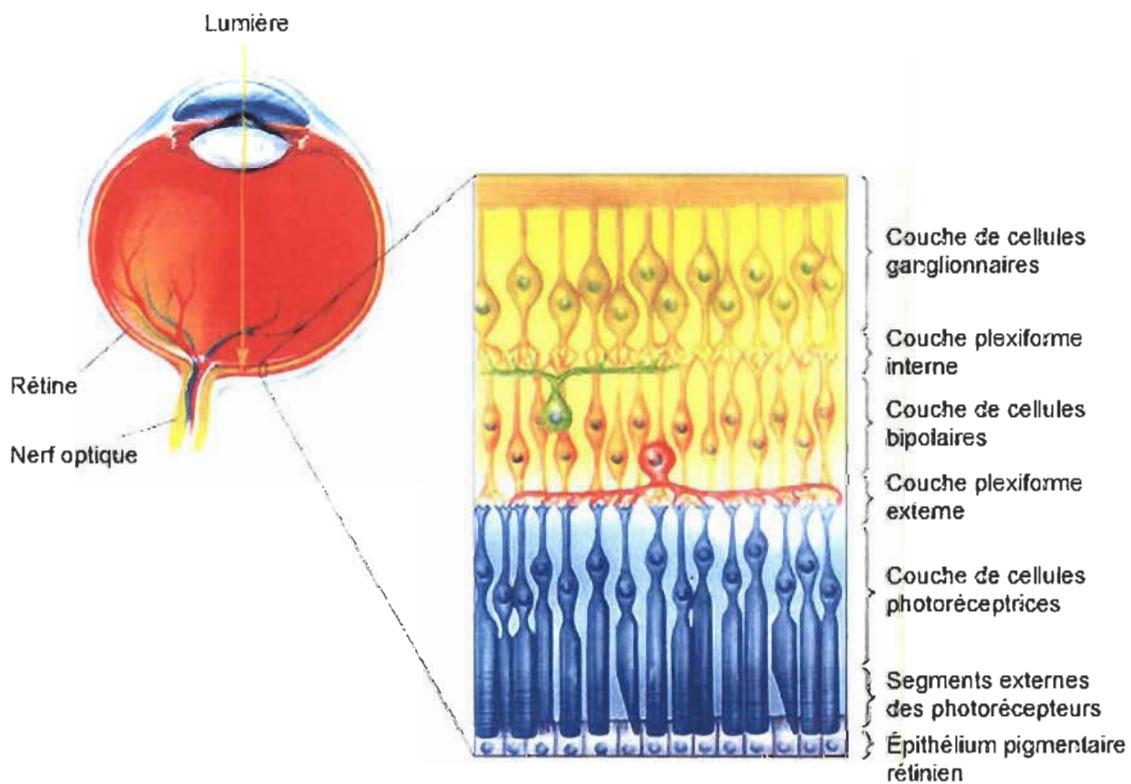


Figure 1.2 – Organisation laminaire de la rétine. Image tirée du site Internet <http://www-old.physiol.usyd.edu.au>.

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).

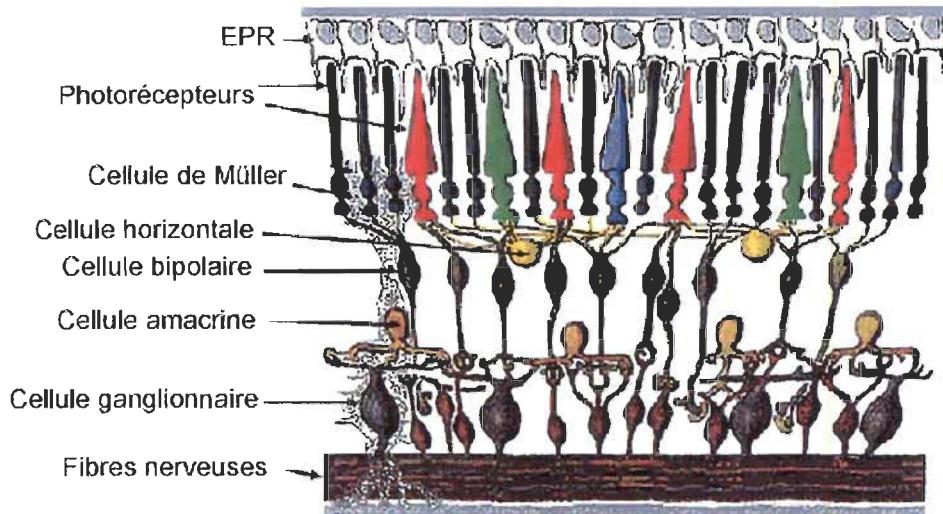


Figure 1.3 - Organisation laminaire de la rétine présentant différents types cellulaires. Image tirée du site Internet de WebVision.

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.

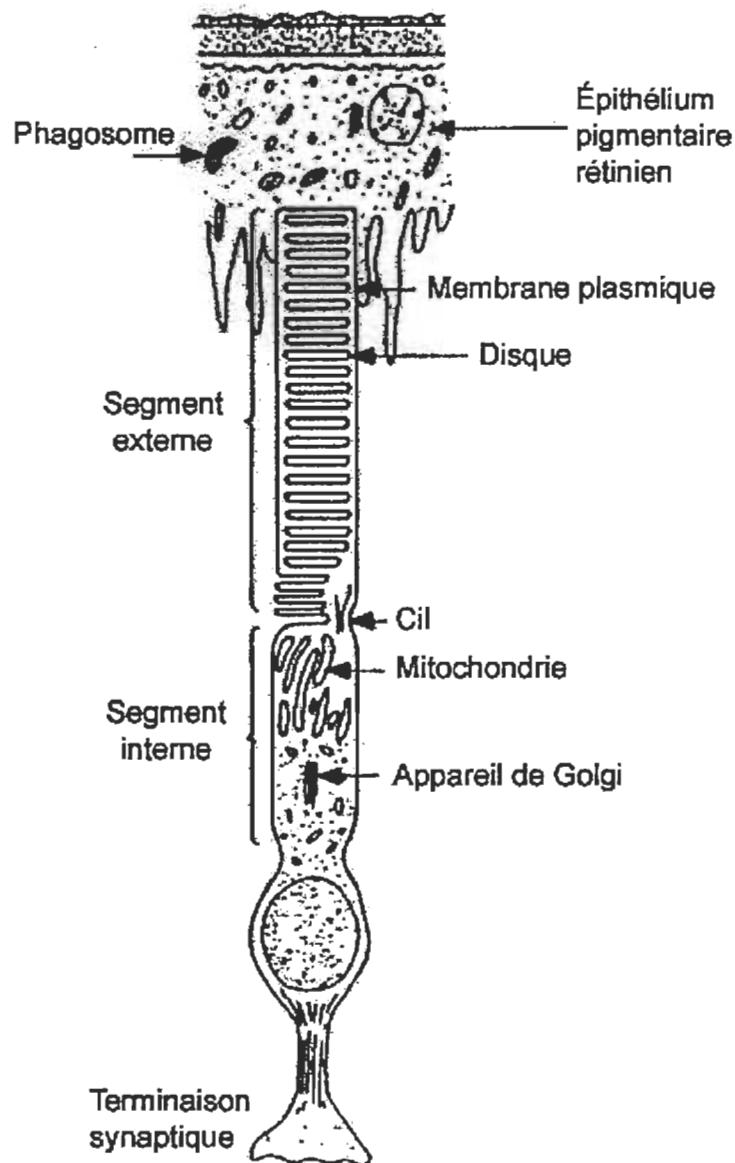


Figure 1.4 – Schéma d'un bâtonnet et de ses différentes composantes. Image tirée de Giusto (1990).

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).

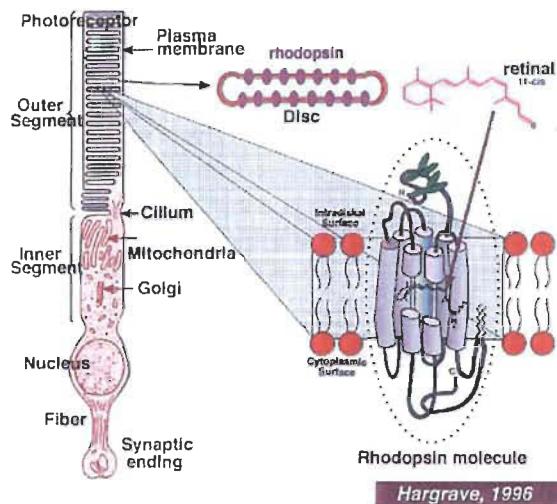


Figure 1.5 – Localisation de la molécule de rhodopsine. La rhodopsine se situe à l'intérieur du segment externe du bâtonnet, plus précisément dans la membrane du disque (Bernstein, 1999).

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 ω 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 cPLA₂, 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 rétinal, l'aldéhyde de la vitamine A. Un isomère du rétinal, le 11-cis rétinal, combiné avec l'opsine, une glycoprotéine, forme la rhodopsine. La lumière provoque l'isomérisation du 11-cis-rétinal en tout-trans-rétinal (métarhodopsine II), entraîne la dissociation de l'opsine et du rétinal et conduit à la conversion du rétinal en rétinol. Cette isomérisation du 11-cis-rétinal en tout-trans rétinal entraîne un changement conformationnel de l'opsine, et par conséquent l'activation de la transducine. Cette protéine G activera la *phosphodiestérase* (*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 Na+/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 rétinal, 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).

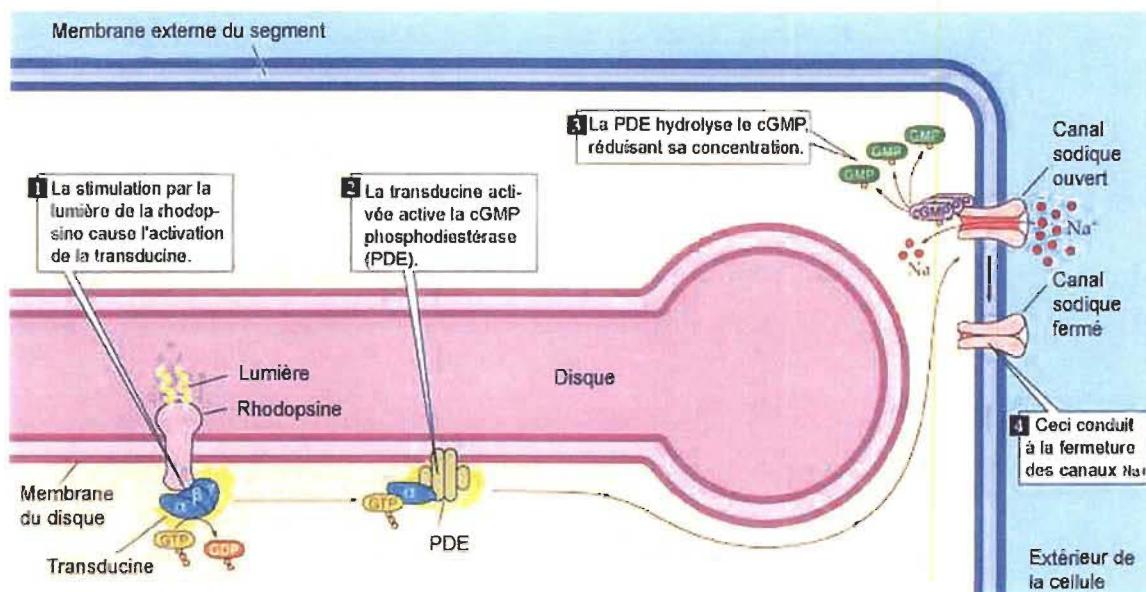


Figure 1.7 - La cascade de la phototransduction. La lumière entraîne la photoisomérisation de la rhodopsine, activant la transducine, une protéine G hétérotrimérique (1). La sous-unité alpha active la phosphodiésterase (PDE) (2), laquelle dégrade le GMPc en 5'GMP (3). La diminution de la concentration 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.

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.

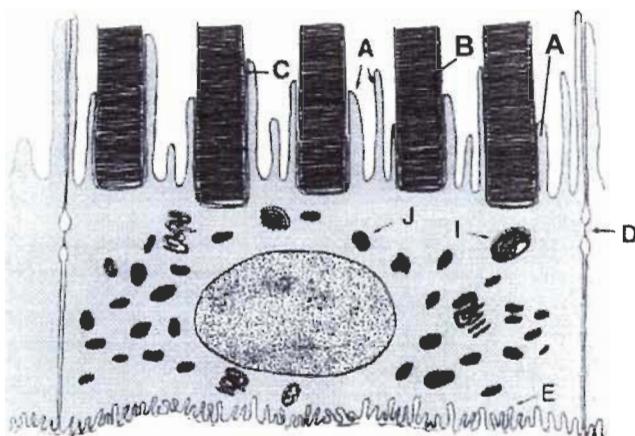


Figure 1.8 – Cellule de l'EPR. Ce schéma montre les microvillis courts et longs (A), le segment externe des photorécepteurs (B), la matrice interphotoréceptrice (C) et les jonctions étanches intercellulaires (D). Image tirée de (Clark, 1986).

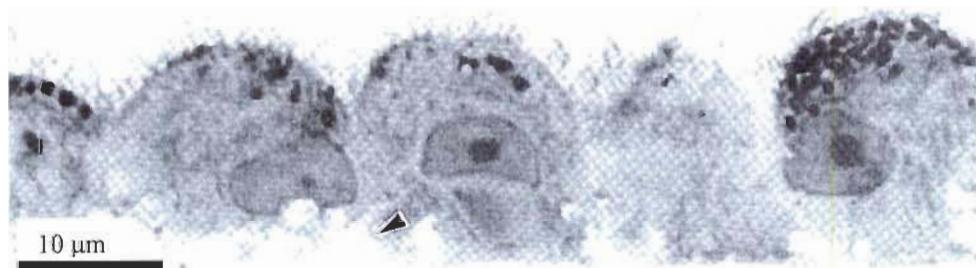


Figure 1.9 – Coupe transversale d'une monocouche d'EPR de singe. Les processus apicaux sont évidents de même que les citernes basolatérales (pointe de flèche). Quelques-unes des cellules contiennent des mélanosomes ovoïdes dans la partie apicale (haut de la photo) (Pfeffer, 1991).

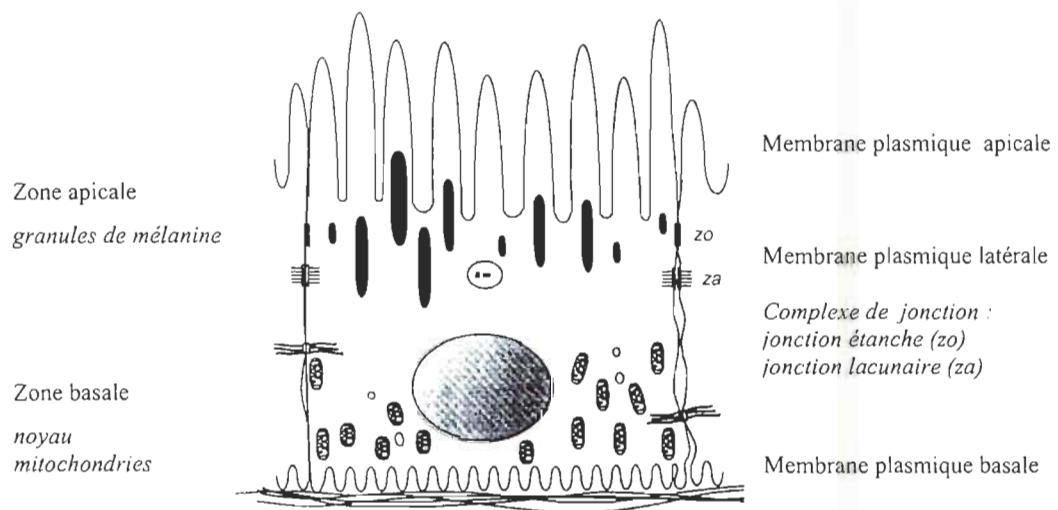


Figure 1.10 - Structure d'une cellule différenciée de l'EPR (Marmor et Wolfensberger, 1998).

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

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 rétinal qui a été photoblanchi lors du processus visuel est estérifiée 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 phagocytter (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.

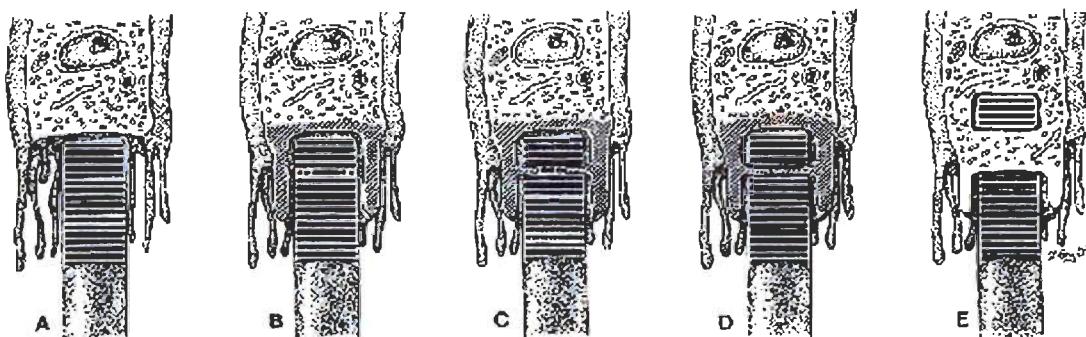


Figure 1.11 – Schéma illustrant le mécanisme de phagocytose des segments externes des photorécepteurs par les cellules de l'épithélium pigmentaire rétinien. L'extrémité distale des segments externes des photorécepteurs est enveloppée par des pseudopodes de l'EPR et il y a formation progressive du phagosome 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 sPLA₂ 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 parasites. 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).

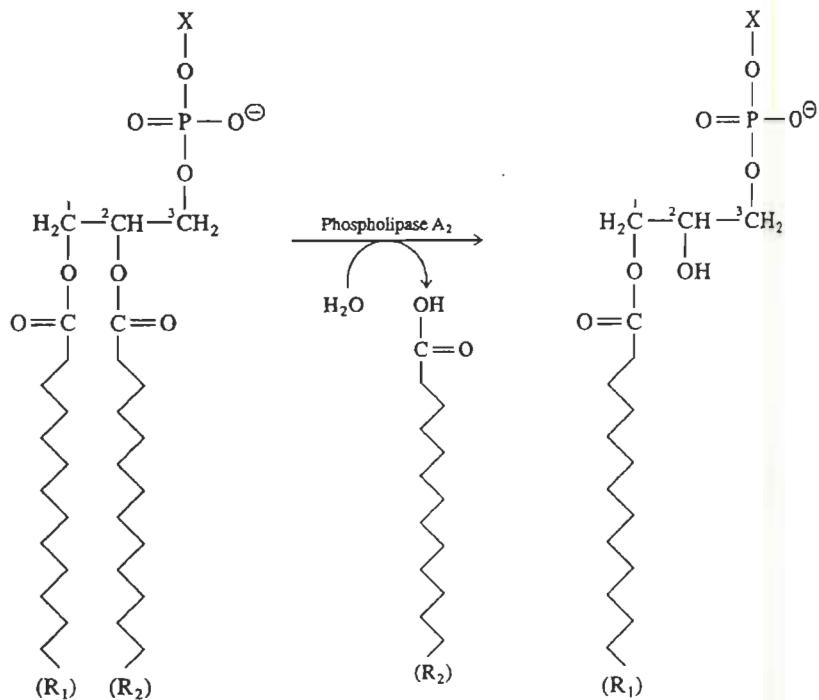


Figure 1.12 - Hydrolyse d'un phospholipide par la phospholipase A₂. La chaîne en position *sn*-2 du phospholipide (à gauche) est hydrolysée par la phospholipase A₂ et conduit à la production d'un lysophospholipide (à droite) et d'un acide gras (au centre). Image tirée de (Horton *et al.*, 1994).

1.2.1 La classification des phospholipases A₂

L'activité phospholipasique a été pour la première fois étudiée dans le venin de cobra à la fin du 19^e siècle (Stephens *et al.*, 1898). Toutefois, la première PLA₂ 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 PLA₂ a évolué au cours des années dans la littérature. Par la suite, plusieurs PLA₂ 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 PLA₂ sécrétées (sPLA₂) 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 PLA₂ s'est étendu aux PLA₂ non

pancréatiques, parfois appelées sPLA₂, 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 classifiée dans le groupe IIA. Ensuite, l'unique PLA₂ du venin d'abeille fut classifiée dans le groupe III. Plusieurs formes additionnelles de PLA₂ sécrétées utilisant une histidine catalytique ont été découvertes ces dernières années. Ces PLA₂ sont clairement reliées aux groupes I, II et III des PLA₂, 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 PLA₂ en tant que petites enzymes sécrétées devint désuète en 1991 avec le séquençage et le clonage d'une PLA₂ (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 PLA₂ 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 PLA₂ leur a valu l'abréviation de cPLA₂, mais elles sont actuellement classifiées dans le groupe IV (IVA puisque actuellement deux autres cPLA₂ ont été identifiées). Enfin, trois autres différents groupes de PLA₂ ont été découverts et classifiés, les groupes VI, VII et VIII, menant ainsi le nombre total de groupes à 12.

Actuellement, la classification des PLA₂ 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 PLA₂ doit inclure toutes les enzymes qui présentent une homologie de séquence importante. Si plus d'un gène de

PLA₂ 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 PLA₂.

1.2.2 Les phospholipases A2 sécrétées

Les phospholipases dites sécrétées (sPLA₂) 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 sPLA₂ 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 sPLA₂ 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 sPLA₂ 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 PLA₂ sécrétée du groupe IB est munie d'un propeptide et doit être clivée par la trypsine pour devenir une enzyme active mature. Les PLA₂ de ces quatre groupes possèdent toutes un poids moléculaire approximatif variant de 13 à 20 kDa.

Puisque les sPLA₂ 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 PLA₂ à histidine catalytique^a

Phospholipases A2 utilisant une histidine catalytique (sPLA₂)					
Groupe	Sous-groupe	Sources	Taille (kDa)	Ponts disulfure	Source de l'archétype
I	A	Cobra, venin de "krait"	13-15	7	Cobra
	B ^b	Panréas de mammifères	13-15	7	Humain
II	A	Liquide synovial, plaquettes de serpent à sonnette, venin de vipère	13-15	7	Humain
	B ^c	Venin de vipère gaboon	13-15	6	Vipère
	C ^d	Testicules de souris/rat	15	8	Souris
	D	Panréas, rate de souris/humain	14-15	7	Humain
	E	Cerveau, cœur, utérus de souris/humain	14-15	7	Humain
	F ^e	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 ^f		Abeille, lézard, scorpion, humain	15-18	5	Abeille à miel

Phospholipases A2 utilisant une histidine catalytique (sPLA₂)

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

Inspiré de (Six et Dennis, 2000)

^a Ce sont typiquement de petites PLA₂ 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 PLA₂ mature.

^b Le groupe IB présente une insertion de 5 acides aminés connus sous le nom de boucle pancréatique.

^c Le groupe IIB ne présente que 5 des 6 ponts disulfure hautement conservés.

^d Le groupe IIC chez l'homme n'est qu'un pseudogène

^e Le groupe IIF présente une cystéine additionnelle dans son extension C-terminale

^f 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 sPLA₂

Les inhibiteurs présentés ici concernent uniquement les sPLA₂ les plus étudiées, c'est-à-dire les sPLA₂ 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 sPLA₂ 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 sPLA₂ 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 sPLA₂ peuvent être soumises à un traitement au H₂SO₄ 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 sPLA₂ et cPLA₂ calcium-dépendantes (Tischfield, 1997).

Il existe plusieurs autres inhibiteurs de sPLA₂. 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'oléyloxyéthyl phosphorylcholine (OPC) et un produit analogue à la PLA₂, 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 cPLA₂ (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 PLA₂ du groupe IV qui seront détaillées à la sous-section 1.2.4.

1.2.3.1 Le groupe VIA

La première PLA₂ clonée et caractérisée de ce groupe est souvent référée en tant que iPLA₂, pour son indépendance au calcium. Elle est actuellement classifiée dans le groupe VIA-1. Cette PLA₂ 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 PLA₂ du groupe VIA présente une activité PLA₂ 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 PLA₂ du groupe VIA-2 présentent une forte identité avec les PLA₂ 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 PLA₂ 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 PLA₂ 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 PLA₂ VIA-2 lorsqu'elles étaient cotransférées. 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 PLA₂, bien que plusieurs les nomment respectivement groupes VIA-ankyrine-1 et VIA-ankyrine-2.

Plusieurs expériences ont démontré que les PLA₂ 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 PLA₂ 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 PLA₂ du groupe VI compte 437 acides aminés et présente 25% d'identité avec la PLA₂ 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 PLA₂

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 PLA₂ (*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 PLA₂ du groupe VIIA est en fait l'acétylhodrolase 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 PLA₂ 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 VIIIB

La PLA₂ du groupe VIIIB 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 PLA₂ 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 PLA₂ 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 PLA₂ 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-AH_{Ib} (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-Deicker, 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 PLA₂ 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 neurologiques génétiques.

1.2.3.6 Résumé comparatif des PLA₂ des groupes VI, VII et VIII

Le tableau 1.2 permet de reconnaître les principales différences entre les caractéristiques distinctes de chacune des PLA₂ des groupes VI, VII et VIII.

TABLEAU 1.2

Résumé des caractéristiques des PLA₂ des groupes VI, VII et VIII à sérine catalytique^a

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 ₂ , iPLA ₂ -A	84-85	Forme courte épissée, 8 répétitions ankyrine	22q13.1 (Pickard <i>et al.</i> , 1999)
	A-2	Lymphocytes-B humains, testicules	iPLA ₂ -B	88-90	Forme longue épissée, 7 répétitions ankyrine	22q13.1 (Pickard <i>et al.</i> , 1999)
	B	Cœur et muscle squelettique humains	iPLA ₂ γ, iPLA ₂ -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-AH _b α ₁	26	Intracellulaire, arrangement protéine G, triade ser/his/asp, dimère	ND
	B	Cerveau humain	PAF-AH _b α ₂	26	Intracellulaire, prot.G, triade ser/his/asp, active en hétero/ homodimère	11q23 (Lecointe <i>et al.</i> , 1999)

Adapté de Six et Dennis (2000)

^a 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 PLA₂ 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) séquençait entièrement cette enzyme qui fut nommée PLA₂ cytosolique, selon la nomenclature de l'époque. Ces PLA₂ ne présentent aucune homologie de séquence avec les PLA₂ connues et constituent actuellement le groupe IVA de l'ensemble des PLA₂ connues dans la nomenclature. Ce groupe IVA s'explique par l'existence de deux autres PLA₂ cytosoliques (IVB et IVC) qui s'ajoutent dorénavant au groupe IV.

Les trois cPLA₂ sont souvent appelées cPLA₂ 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 cPLA₂

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 cPLA₂ 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 cPLA₂ 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 séries 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 cPLA₂ alpha et bêta possèdent le domaine CaLB en N-terminal, alors que la cPLA₂ gamma en est dépourvue, expliquant ainsi son indépendance au calcium. La cPLA₂ 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 cPLA₂ bêta, ce domaine qui ne compte que 19 et 5 acides aminés chez les cPLA₂ alpha et gamma, respectivement (Song *et al.*, 1999).

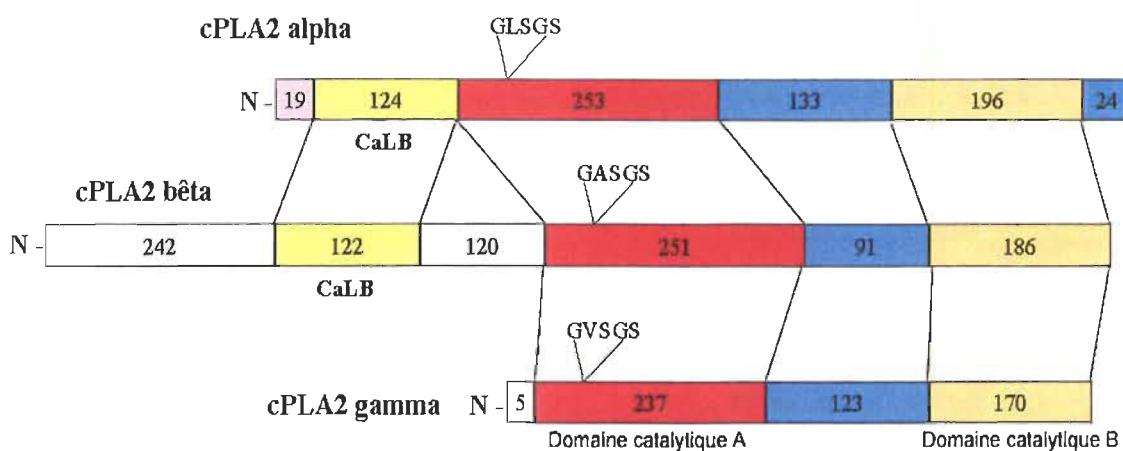


Figure 1.13 – Comparaison structurale entre les trois isoformes de cPLA₂ connues dans la littérature jusqu'à présent. La cPLA₂ bêta est l'isoforme le plus long, avec deux domaines supplémentaires, alors que la cPLA₂ gamma se reconnaît à sa plus petite taille, vu l'absence du domaine de liaison au calcium (CaLB) en N-terminal. Le domaine catalytique A apparaît en rouge, alors que le domaine B apparaît en orange.

La cPLA₂ 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 cPLA₂ bêta et gamma sont dépourvues de ce motif, suggérant que le mécanisme de phosphorylation connu chez la cPLA₂ 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 cPLA₂ 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 cPLA₂ 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

cPLA₂ 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 cPLA₂ alpha présente une homologie de séquence avec d'autres protéines possédant un mécanisme de translocation calcium-dépendant. Ce domaine est abrégé CaLB pour *Ca2+-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 cPLA₂ 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 cPLA₂ alpha vers les membranes intracellulaires et les vésicules lipidiques (Channon et Leslie, 1990; Clark *et al.*, 1991; Nalefski *et al.*, 1994). La cPLA₂ 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 cPLA₂ 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 cPLA₂ alpha et la vimentine dépend de trois éléments : du calcium, de la moitié N-terminale du domaine CaLB de la cPLA₂ 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 cPLA₂ alpha pourrait lier simultanément la vimentine et les phospholipides membranaires, d'une façon calcium-dépendante.

1.2.4.3 La cPLA₂ alpha (groupe IVA)

La cPLA₂ alpha est l'enzyme la plus étudiée de son groupe pour avoir été découverte la première. La cPLA₂ alpha est une protéine monomérique d'approximativement 85 kDa. L'ADNc de la cPLA₂ 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 cPLA₂ 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 cPLA₂ bêta (groupe IVB)

La cPLA₂ 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 cPLA₂ alpha (Pickard *et al.*, 1999). Elle a un poids moléculaire d'environ 114 kDa. Elle possède une grande identité avec la cPLA₂ alpha dans la région des acides aminés catalytiques Ser228 et Asp549, mais ne présente toutefois pas les quatre séries connues pour être impliquées dans la phosphorylation de la cPLA₂ 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 cPLA₂ 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 biphosphate phosphatase. Enfin, l'ARNm de la cPLA₂ 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 cPLA₂ gamma (groupe IVC)

La cPLA₂ 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 cPLA₂ alpha. Les acides aminés conservés de la triade catalytique chez cette cPLA₂ sont la sérine 82, l'aspartate 385 et l'arginine 54. Le gène de la cPLA₂ gamma est localisé sur le chromosome 19 près du gène de la calmoduline. Le transcrit a une longueur d'environ 3 kb.

La cPLA₂ gamma présente un motif d'isoprénylation (-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 cPLA₂ gamma à l'intérieur de la cellule (Underwood *et al.*, 1998).

1.2.4.6 Résumé récapitulatif des trois cPLA₂

Le tableau 1.3 résume les principales différentes caractéristiques des trois cPLA₂ connues jusqu'à présent.

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

la « cPLA₂ alpha » dans le titre indique simplement que les études n'ont été réalisées que sur cette cPLA₂ alpha et qu'il n'y a encore rien de publié dans la littérature concernant les cPLA₂ bêta et gamma.

TABLEAU 1.3
Résumé des caractéristiques des cPLA₂

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

Adapté de Six et Dennis (2000)

^a U937 est une lignée de cellules leucémiques monocytaire humaines.

^b RAW 264.7 est une lignée cellulaire de macrophages provenant de la souris.

1.2.4.7 Régulation transcriptionnelle de la cPLA₂ alpha

Des stimuli extracellulaires peuvent altérer les niveaux d'ARN messager (ARNm) et de la protéine cPLA₂ 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 cPLA₂ alpha dans divers modèles cellulaires. La stimulation de la synthèse de la cPLA₂ 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 cPLA₂ alpha (Lin *et al.*, 1992a; Schalkwijk *et al.*, 1993).

La région promotrice du gène de la cPLA₂ 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 cPLA₂ alpha a établie. Par exemple, de multiples séquences AUUUA dans la région 3' non codante de l'ARNm de la cPLA₂ 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 cPLA₂ alpha chez les cellules mésangiales glomérulaires du rat (Tay *et al.*, 1994a).

1.2.4.8 Localisation cellulaire, intracellulaire et tissulaire

La cPLA₂ 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 cPLA₂ 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 cPLA₂ bêta est exprimée ubiquitairement, mais plus fortement au niveau du pancréas, du cerveau, de cœur et du foie, alors que la cPLA₂ 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 cPLA₂ 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 cPLA₂ 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 cPLA₂ 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 cPLA₂ 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 cPLA₂ gamma a aussi été déterminée. Ainsi, le marquage de la cPLA₂ 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 cPLA₂ 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 cPLA₂ 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 cPLA₂ 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 cPLA₂ 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 cPLA₂ 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 cPLA₂ 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 cPLA₂ en conditions physiologiques. La cPLA₂ 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 cPLA₂ 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 cPLA₂ 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 cPLA₂ bêta et la cPLA₂ gamma indique qu'elles ont une régiosélectivité différente de la cPLA₂ alpha. En effet, la cPLA₂ 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 lysoPLA₂ plus importante que son activité PLA₂ (Song *et al.*, 1999). La cPLA₂ gamma montre une activité PLA₂ 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 cPLA₂ 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 cPLA₂ alpha

Les lipases possèdent une triade catalytique *sérine-aspartate/glutamate-histidine* bien conservée chez les séries protéases. Le mécanisme catalytique procède via un intermédiaire sérine-acyle (Hanel et Gelb, 1995). Le motif GLSGS présent chez toutes les PLA₂ du groupe IV (cytosoliques) rappelle le motif GXSXG 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 cPLA₂ alpha ont été identifiés par mutagénèse 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 cPLA₂ 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 cPLA₂ alpha peut soit se dissocier de l'interface de la bicouche ou lier un autre substrat phospholipidique, répétant ainsi le cycle (Dessen, 2000).

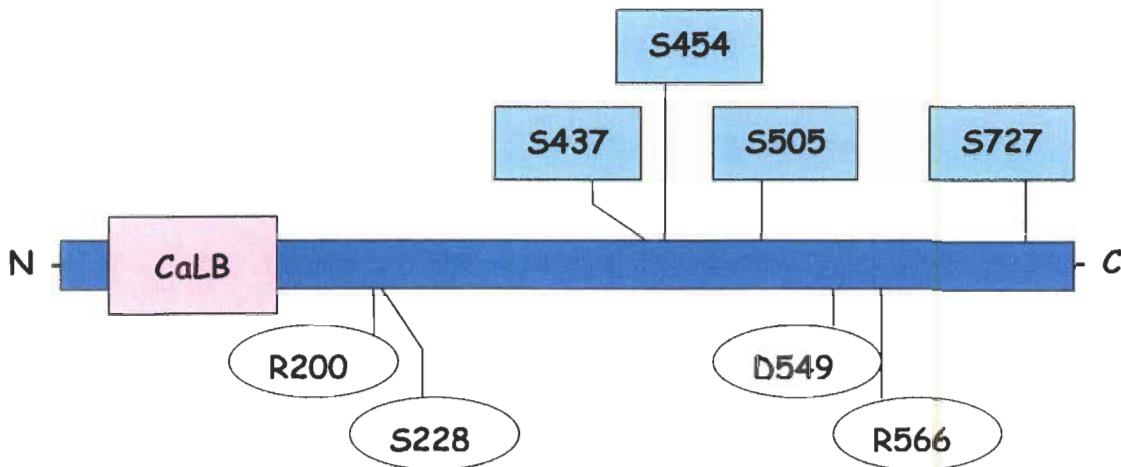


Figure 1.14 - Schéma de la cPLA₂ alpha montrant les acides aminés les plus importants pour son activité. Les différents acides aminés impliqués dans la phosphorylation apparaissent dans les carrés bleus et les acides aminés impliqués dans l'activité catalytique apparaissent dans les cercles blancs. Le 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 cPLA₂ 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 cPLA₂ 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 cPLA₂ alpha

La stimulation des cellules par divers agonistes qui induisent une libération d'acide arachidonique promouvoit aussi la phosphorylation de la cPLA₂ alpha (Lin *et al.*, 1992b; Clark *et al.*, 1995). Il a été démontré que la cPLA₂ 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 cPLA₂ 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 cPLA₂ (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 cPLA₂ 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 cPLA₂ 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 cPLA₂ alpha est un substrat pour la MAPK. L'activation maximale de la cPLA₂ 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 cPLA₂ 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 cPLA₂ 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 cPLA₂ 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 cPLA₂ menant à la libération d'acide arachidonique.

1.2.4.12 Activation de la cPLA₂ alpha et son interaction avec certains éléments

La cPLA₂ 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 cPLA₂ alpha (Qiu *et al.*, 1998; Hirabayashi *et al.*, 1999).

La cPLA₂ 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 cPLA₂ alpha, résultant en une augmentation de l'acide arachidonique libéré (Balsinde *et al.*, 2000). Il a été démontré que la cPLA₂ 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,

transduction intracellulaire, notamment, les protéines G et les kinases. En effet, comme il est mentionné un peu plus haut, la cPLA₂ alpha est activée en partie par la MAPK. De même, puisque ces kinases sont régulées par les protéines G (Xing et Mattera, 1992), la cPLA₂ alpha est indirectement modulée par ces dernières. La figure 1.15 présente le modèle d'activation de la cPLA₂ 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).

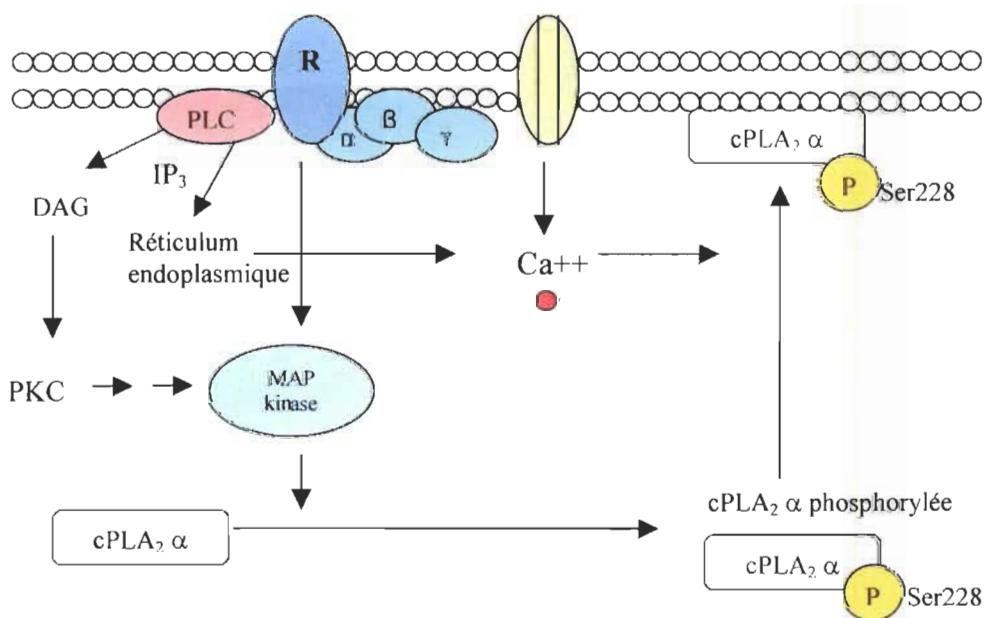


Figure 1.15 – Modèle d'activation de la cPLA₂ alpha. Dans ce modèle, l'activation rapide de la cPLA₂ alpha est causée par l'action synergique du calcium et de la phosphorylation de cette cPLA₂ par la MAP kinase. Lorsqu'un ligand se lie au récepteur approprié (R), ce dernier active la phospholipase C (PLC) via soit une voie dépendante ou indépendante d'une protéine G, conduisant à la formation de l'inositol triphosphate (IP₃) et du diacylglycérol (DAG). L'IP₃ entraîne la mobilisation du calcium intracellulaire. Alternativement, l'augmentation intracellulaire de la concentration calcique peut aussi résulter de l'activation de récepteurs couplés aux canaux calciques. L'augmentation de calcium entraîne la translocation de la cPLA₂ 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 cPLA₂ alpha au niveau de l'acide aminé Ser228. De concert, la phosphorylation et l'augmentation calcique entraînent l'activation optimale de la cPLA₂ alpha. Image tirée de (Lin *et al.*, 1993).

5, W-7, W-12 et W-13 s'avèrent être des inhibiteurs tant pour la cPLA₂ alpha que pour les sPLA₂ (Hope *et al.*, 1993).

1.2.4.14 Implication des PLA₂ dans diverses pathologies

L'action hydrolytique des PLA₂ 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 PLA₂, 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 PLA₂ dans un grand nombre de situations pathologiques. De ce fait, les PLA₂ constituent une cible très intéressante pour la découverte de nouveaux médicaments; des inhibiteurs spécifiques de PLA₂ pourraient intervenir dans une grande variété de processus physiopathologiques.

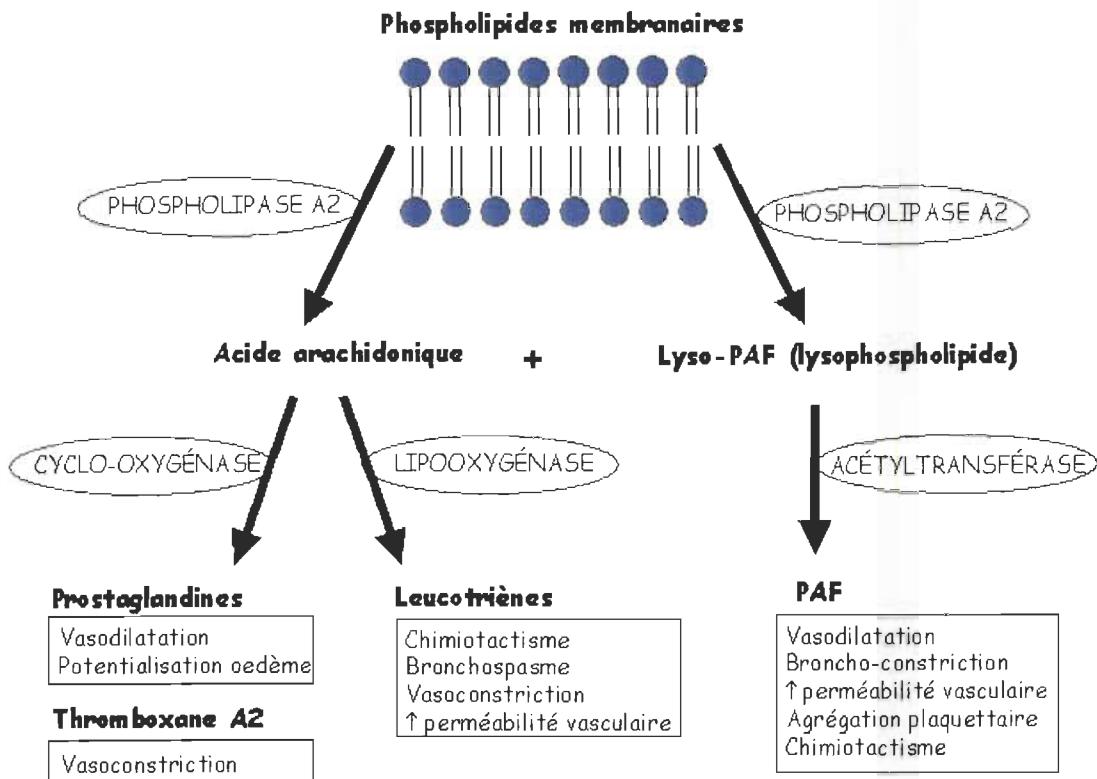


Figure 1.16 – La synthèse des différents dérivés de l’acide arachidonique et du lysophospholipide. Informations tirées du site Internet <http://www.mbi.ufl.edu/~hnick/cPLA2.htm> de la *McKnight Brain Institute* de l’Université de Floride.

À l'aide de souris *knock-out* cPLA₂-/-, différents rôles ont été attribués à la cPLA₂ 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 cPLA₂ est un important effecteur de l'hyperréactivité à ce niveau (Sapirstein et Bonventre, 2000).

Enfin, l'expression de la cPLA₂ 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 cPLA₂ dans un nombre important de situations pathologiques.

1.2.4.15 L'implication des cPLA₂ pour l'intégrité des photorécepteurs

Tel que décrit plus tôt dans la sous-section 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⁻) et superoxyde (O₂⁻) et le peroxyde d'hydrogène (H₂O₂) (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 PLA₂ 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 PLA₂ 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 PLA₂ hydrolysent préférentiellement les lipides peroxydés (Sevanian et Kim, 1985; Sevanian et McLeod, 1987).

En conséquence, puisque les cPLA₂, 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 cPLA₂ 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 cPLA₂ 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 PLA₂.

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 cPLA₂ sont susceptibles d'être en partie responsables du processus inflammatoire oculaire.

1.2.4.17 Implication des cPLA₂ 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 PLA₂ (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 cPLA₂ 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 cPLA₂ alpha a déjà été localisée ponctuellement dans des corps lipidiques cytoplasmiques de certaines cellules (Yu *et al.*, 1998) et que la cPLA₂ 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 PLA₂ 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 PLA₂ ont été obtenues. La sensibilité de ces fractions envers le pBPB, l'EGTA, le DTT,

la chaleur et le H₂SO₄ indique que les cellules de l'épithélium pigmentaire rétinien humain en culture contiendraient deux PLA₂ intracellulaires différentes. De plus, des expériences d'immunobuvardage de type Western ne montrent aucune réactivité entre les PLA₂ cytosoliques des cellules de l'épithélium pigmentaire rétinien humain et plusieurs antisérumis dirigés contre les PLA₂ sécrétées et la cPLA₂ alpha (Van Themsche *et al.*, 2001), ce qui suggère que les PLA₂ des fractions identifiées sont des PLA₂ différentes des sPLA₂ et cPLA₂ 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 PLA₂ spécifiques à l'EPR pourraient être impliquées dans ces fonctions-clé de l'EPR. De même, l'implication possible des cPLA₂ à divers niveaux de la fonction visuelle permet de croire qu'une ou plusieurs PLA₂ spécifiques à l'EPR pourraient sans aucun doute être liées à différentes dégénérescences rétiennes 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 PLA₂ cytosoliques présentes au niveau de la rétine et/ou de l'EPR. Ces cPLA₂ comprennent non seulement les nouvelles cPLA₂, 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 cPLA₂ 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 cPLA₂ connues.

2.1.2 Deuxième objectif spécifique

Le deuxième objectif spécifique consiste à identifier de nouvelles cPLA₂ 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 cPLA₂ 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 cPLA₂ 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 cPLA₂ déjà connue en tant que sonde marquée pourrait permettre d'obtenir l'une ou l'autre des cPLA₂ 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 cPLA₂ 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 cPLA₂ 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 cPLA₂ connues. Ce sont précisément ces séquences qui pourraient s'avérer être de nouvelles cPLA₂.

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 » cPLA₂.

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 cPLA₂ connues, le PCR est le premier outil utilisé afin d'obtenir l'ADNc de ces « nouvelles » cPLA₂. Ainsi, en se basant sur les séquences codantes déduites par les différentes analyses bioinformatiques, la séquence des amores 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 CaCl₂, 8,3 mM glucose, 15 mM HEPES, 3,5 mM KCl, 0,27 mM MgCl₂, 0,37 mM MgSO₄, 138,5 mM NaCl, 4,3 mM NaHCO₃, 1,0 mM NaH₂PO₄, 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 CaCl₂ 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 ReagentTM* (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- α -tocophérol, 50 mg/ml transferrine, et 0,3 mg/ml uridine dans du milieu de culture de base *Keratinocyte-SFM*, pH 7,2). Ces feuillets isolés d'EPR sont incubés à 37°C pendant 48 à 72 heures dans une atmosphère en CO₂ 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 microustique 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 microustique 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 microustique 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 microustique 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, ϵ , 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 $40 \text{ } \mu\text{g/ml}$. 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).

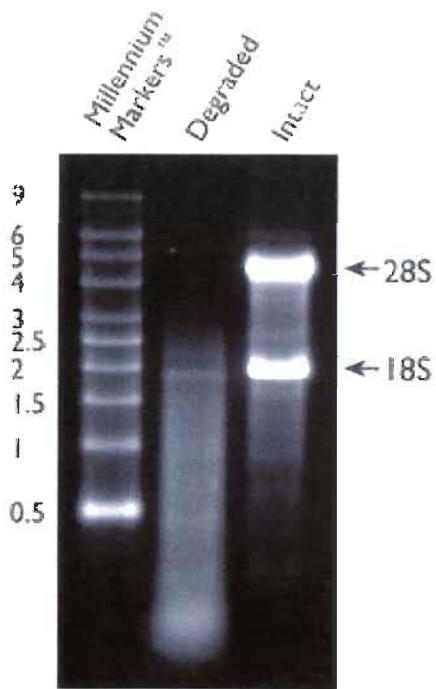


Figure 3.1 – Vérification de l'intégrité de l'ARN par migration sur gel d'agarose. Dans la deuxième piste, l'ARN apparaît dégradé, alors que dans la dernière piste l'ARN intact apparaît sous forme de deux bandes bien distinctes. Image tirée du site Internet de la compagnie *Ambion*.

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™ II*.

Pour un volume réactionnel de 20 µl, 5µg d'ARN total sont utilisés. Dans un microtube sans nucléase, 1 µl d'oligo (dT)₁₂₋₁₈ (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 $(\text{NH}_4)_2\text{SO}_4$	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_2 (25 mM)	2 – 2,5 mM
ADN plasmidique génomique/ADNc	1 ng / 20 μl 0,5 – 1 $\mu\text{g} / 20 \mu\text{l}$

TABLEAU 3.2
Compilation des différentes amorces utilisées

Nom de l'amorce	Séquence de l'amorce	Tm ($^{\circ}\text{C}$)	Produit attendu ¹ (en pb)
Amorces dégénérées pour les cPLA₂			
cPLA ₂ -1	5' GG(GT) (GC)(CT)(CT) TC(GT) GGC TCC ACC TGG 3'	70	1053
cPLA ₂ -2	5' (AG)TA GTC (CA)AA (GT)GA SA(GT) (GT)AT GAG 3'	47	1053
Amorces correspondant à la cPLA₂ alpha			
Position 5-5'	5' CAT TTA TAG ATC CTT ACC AG 3'	60	5 ^e
Position 51-5'	5' CCA CAA GTT (CT)AC GGT AGT GG 3'	66	51 ^e
Position 204-5'	5' CCT GT(CG) TGG AA(CT) GAG ACC TT 3'	66	204 ^e
Position 1798-3'	5' CTT (CT)GG AAA (GT)GG (CG)AG CTT GT 3'	65	1798 ^e
Position finale 3'	5' CTA TGC TTT GGG TTT ACT TA 3'	60	2070e
cPLA ₂ alpha courte (S)	5' ATT CTG CAC GTG ATG TGC CTG T 3'	67	1030

Nom de l'amorce	Séquence de l'amorce	Tm (°C)	Produit attendu ¹ (en pb)
cPLA ₂ alpha courte (AS)	5' ACA GCT GCA TCC AGT TCA TCA T 3'	65	1030
cPLA ₂ A1 (S)	5' GAG CTG AAA AAG GAT CCT GAC T 3'	65	440
cPLA ₂ A2 (AS)	5' TGT CCC TAG AGT TTC ATC CA 3'	63	440
Amorces correspondant à la cPLA₂ bêta			
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 ₂ bêta courte1 (S)	5' ATC AGC TTT GGA GGA TCT TGG A 3'	65	3565
cPLA ₂ bêta courte2 (S)	5' ATC AGG AGC TCT AGT ATA GGC T 3'	65	2149
cPLA ₂ bêta courte 3 (S)	5' TCT GAC TGC TAC GTG ACTTCTC T 3'	67	608
cPLA ₂ 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
Amorces correspondant à la cPLA₂ gamma			
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
Amorces correspondant au nouveau gène 1			
Gène 1a	5' CTT GGA CAG TGA GCC AGA CAG A 3'	68	1000 ^e
Gène 1b	5' CAG CAG AGC TGG AGT TTC TGC A 3'	68	1813 ^e
Gène 1c	5' CAT CAG GTA GCA TTC CTT GAG A 3'	65	3441 ^e
Gène 1d	5' CTT CTT CTT CEC CAC TGC GAG C 3'	68	3795 ^e
1A	5' ATG GGT GAG CCT GGA AAA GGT A 3'	67	1 ^{ère}
1B	5' TTC TTA GTT GTG TGG CAG AGC T 3'	65	523 ^e
1C	5' ATG GGA TGG TTT GCT GAG GGT A 3'	67	658 ^e
1D	5' AAG CTG AGG ACA AGG ACC ATC T 3'	67	1060 ^e
1E	5' ATG GAA GAG CTG GAG GTG GAG T 3'	68	1279 ^e
1F	5' ACT GCC TTT CCG ATG GTC AGG T 3'	68	1790 ^e

Nom de l'amorce	Séquence de l'amorce	Tm (°C)	Produit attendu ¹ (en pb)
1G (AS)	5' TGT AAA GGT GAC CCT GTA GCC T 3'	67	2298 ^e
1H (AS)	5' ACT TCT GCA GAC CCA CCT CGT A 3'	68	2656 ^e
1I (AS)	5' TTC CAC CTA GAG AAC TGG CCA T 3'	67	3062 ^e
1J (AS)	5' TGT CAA AGG TGG CCT CTG TGT A 3'	67	3709e
Amorces correspondant au nouveau gène 2			
Gène 2a	5' TTT CAG CAA CCT CAG GTC TGT G 3'	67	1 ^{re} 2
Gène 2b	5' GCT TCA ATC TCT GTG CAG AGG A 3'	67	1529 ^e
Gène 2c (AS)	5' TGG AGA AGT CTT TGT GGA TAC A 3'	65	2527 ^e
Gène 2d (AS)	5' TCT GAG CAA CCT CAG GTC TGT G 3'	68	3128 ^e
2END1	5' CAG GAC TAC TGT AGC CAC AAA G 3'	67	2518 ^e
2END2	5' TGT AGC CAC AAA GAC TTC TCC A 3'	65	2527 ^e
NEW2A	5' CTC ACG AAG ACT TCC TTA ACC T 3'	65	801 ^e 2
NEW2B	5' GAA TGT TCT GGA GCT TAG CAT C 3'	65	936 ^e
NEW2C	5' AGA TCA GGA CAA GCT GGA GCT G 3'	68	1215 ^e
NEW2D (AS)	5' TGA GAA GAG GTG GCA TTC CCT T 3'	67	2811 ^e
NEW2E (AS)	5' CTT CCT CCT TGT AGG TCA TGT T 3'	65	2979 ^e
2A	5' AAC AGG AAC TTC CAC TCA TTA A 3'	62	53 ^e
2B	5' AAC GAC TTA CAC AAT GTC TAA T 3'	60	261 ^e
2C	5' ATT AAA GCT ATA GAC CTG AGC T 3'	62	406 ^e
2D	5' TCC TGA AAC TCC AGG TCT GGA A 3'	67	501 ^e
2E	5' ATG AAG GAT GTG AAG CTG TTT T 3'	62	703 ^e
2F	5' TTC AAG GTT CTC TAT GAC ATC T 3'	62	994 ^e
2G	5' ATG TGG AGT TCC TGA TGG AAG A 3'	65	1085 ^e
2H	5' ACA TGG CAG CCC TAG AGA CAG A 3'	68	1313 ^e
2I (AS)	5' TCT CTT TGA CAT TGA GGA TCA A 3'	63	2068 ^e
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
Amorces spécifiques à différents vecteurs plasmidiques			
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	Tm (°C)	Produit attendu ¹ (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	-
Amorces utilisées pour le RACE PCR			
PolyT	5' TTT TTT TTT TTT TTT TTT TTT TTT TTT 3'	53	-
RACE S1	5' (CAU) ₄ TTC CTG AGC AGG AGG AAC CA 3'	80	1558 ^c
RACE S2	5' (CAU) ₄ ACC TGC AGG AGG ATG AGG TA 3'	80	1622 ^c
RACE S3	5' (CAU) ₄ GAC CTA CTT CAT TGG CAT CT 3'	77	1743 ^c
RACE AS1	5' (CAU) ₄ TAA GCT CCT GGT CTT GTC CT 3'	78	2342 ^c
RACE AS2	5' (CAU) ₄ TGT CAG GAA GCC TTT AAA TG 3'	76	2444 ^c
RACE AS3	5' (CAU) ₄ CAG TAG TCC TGG TGC AGC TG 3'	81	2509 ^c

¹ 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^{er} acide nucléique et T, le 2^e).

² 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'éthydium (concentration finale de 0,01 mg/100 ml) et migrant 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 *ΦXI74 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 MgCl₂·6H₂O et 24,65 g/100 ml MgSO₄·7H₂O) 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 sont ensemencés 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'abord ensemencées sur un pétri LB/MgSO₄ (pétris LB/MgSO₄ : 10 g/l bacto-tryptone, 5 g/l extrait de levure, 5 g/l NaCl, 10 mM MgSO₄ 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/MgSO₄ 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 MnCl₂, 100 mM KCl, 10 mM CaCl₂ 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 CaCl₂, 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, puis ensemencées 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 1x10⁸/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 un 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 MgCl₂, 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 cPLA₂ 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 cPLA₂ 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 » cPLA₂. Inutile de préciser que différents fragments (3) de cette dernière cPLA₂ 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 cPLA₂ 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 cPLA₂ 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).

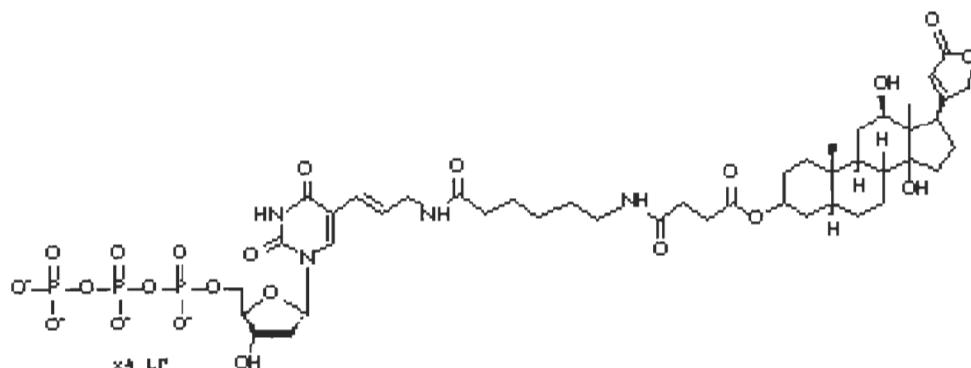


Figure 3.2 - Structure de la digoxigénine-11-dUTP. Cette molécule est utilisée pour la plupart des applications de marquage non radioactif d'ADN. Image tirée du site Internet de Boehringer Mannheim.

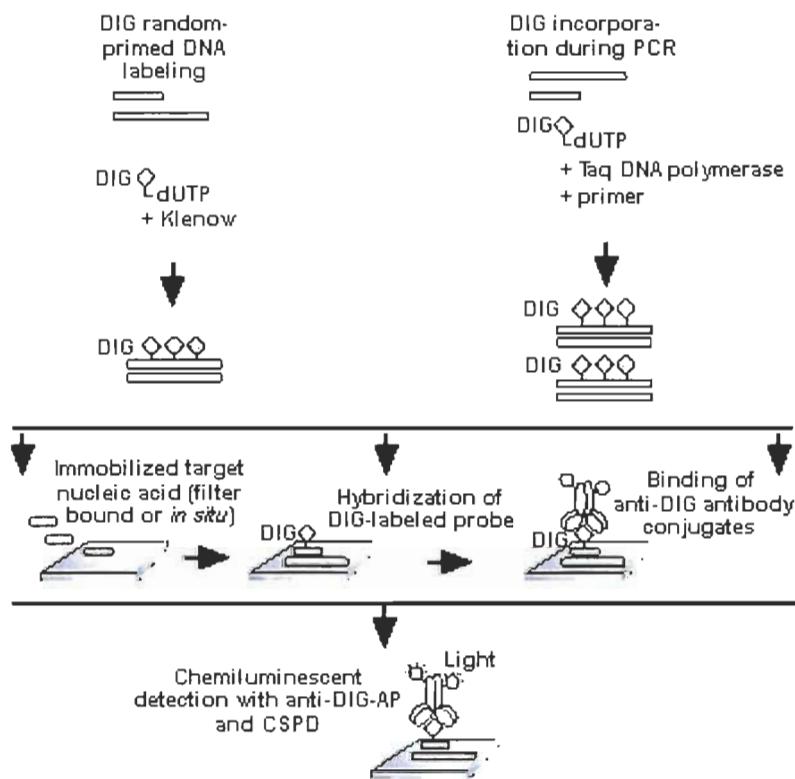


Figure 3.3 - Deux des différents types de marquages possibles et la détection au CSPD par le système DIG de Boehringer Mannheim. La première technique présentée (en haut à gauche : *DIG random-primed DNA labeling*) est celle qui est appelée marquage aléatoire dans le présent mémoire. La deuxième technique (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'elongation, concentration en MgCl₂, 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 MgCl₂ 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'ADN 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).

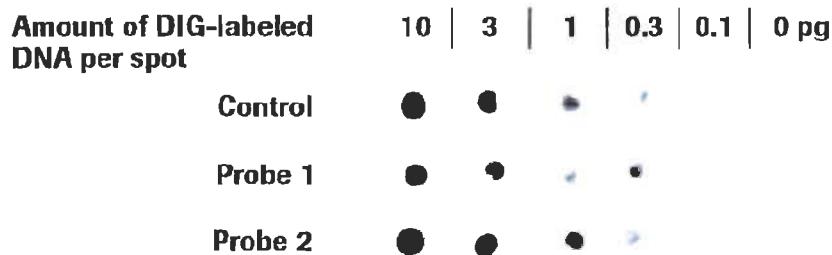


Figure 3.4 - Résultat d'une quantification. Exemple du résultat obtenu dans le cas d'une quantification parallèle de deux sondes différentes (deux rangées du bas). La première rangée (le contrôle) est utilisée en tant que référence pour estimer la concentration des sondes fabriquées. Image tirée du site 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 λ Trip l Ex2TM 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 (λ Trip l ExTM & λ Trip l Ex2TM 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 prélevera 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 ensemencer 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/MgSO₄ (pétris LB/tétracycline/MgSO₄ : 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 MgSO₄) est ensemencé. En utilisant une de ces dernières colonies, un bouillon LB/MgSO₄/maltose (bouillon LB/MgSO₄ : 10 g/l bacto-tryptone, 5 g/l extrait de levure, 5 g/l NaCl, 0,2% maltose et 10 mM MgSO₄) de 15 ml est ensemencé 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 MgSO₄ 10 mM.

De grands pétris (150 mm) LB/MgSO₄ 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 MgSO₄·7H₂O, 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 µl de tampon de dilution lambda 1X et 200 µl 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/MgSO₄ top agarose (45°C) sont alors ajoutés à chacun des tubes (LB/MgSO₄ top agarose : 10 g/l bacto-tryptone, 5 g/l extrait de levure, 5 g/l NaCl, 10 mM MgSO₄ 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 1x10⁶ pour être intéressant; un titre de 1x10⁹ 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 pétri – sur le pétri 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 duplicita. Elle est retirée du pétri 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 µJ/cm² 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).

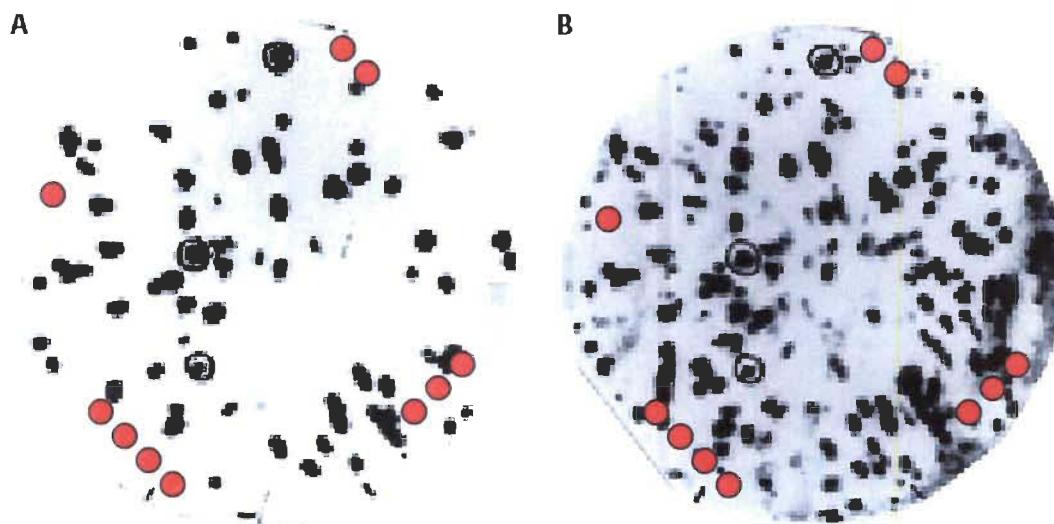


Figure 3.5 - Résultat typique d'un criblage de banque d'ADNc (banque de phage) où les clones positifs sont nombreux. La membrane de gauche serait le résultat du *duplicata* (A) alors que la membrane de droite serait la membrane initiale déposée sur le pétri (B). Remarquez les points rouges qui ont été 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 (encerclés). 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 pétris 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 µl 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 λ 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/MgSO₄ 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 MgCl₂ 1M sont ajoutés à la culture bactérienne.

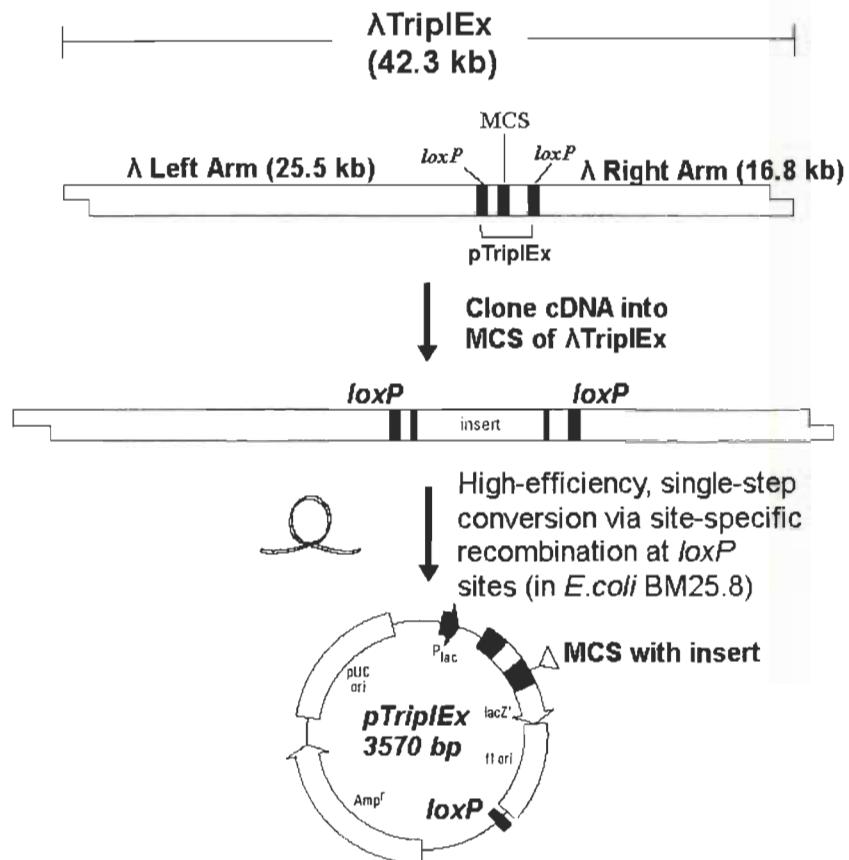


Figure 3.6 - Conversion du phage recombinant λTriplEx en pTriplEx. Le site de clonage multiple (MCS) est situé à l'intérieur de la région plasmidique, laquelle est flanquée par des sites *loxP* aux jonctions λ. La transduction de λTriplEx dans la souche *E.coli* BM25.8 promouvoit la libération de la *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 sont ensemencés 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/MgSO₄/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 MgSO₄ 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 µl seront utilisés pour la transformation des bactéries compétentes *E. coli DH5α*.

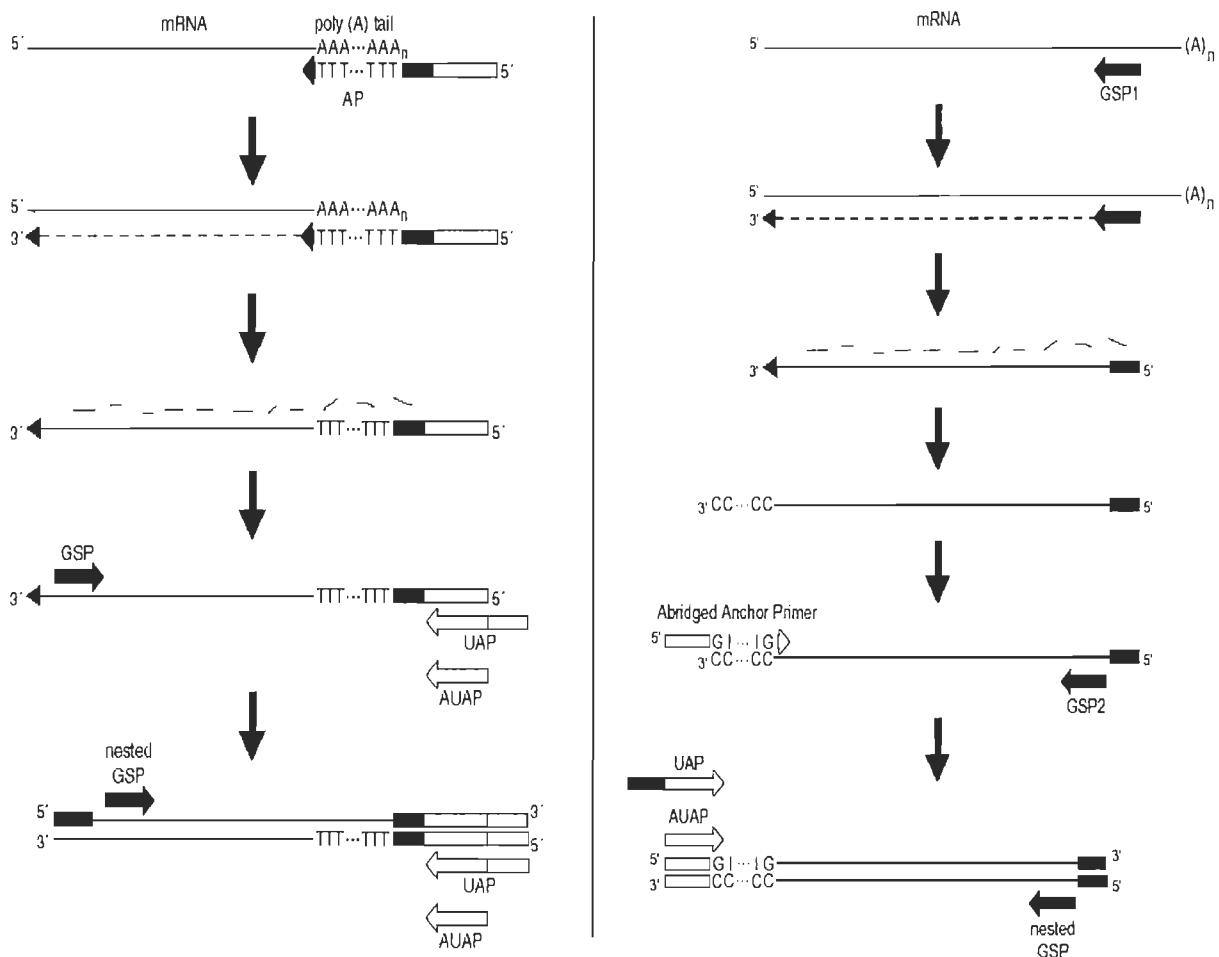


Figure 3.7 - Schématisation des procédures pour le système du RACE. Dans chacune des procédures, des amores spécifiques sont utilisées (GSP : *gene specific primer*). La procédure pour obtenir l'extrémité 3' est présentée à gauche alors que la procédure pour l'extrémité 5' est présentée à droite. Dans 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'amores 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).

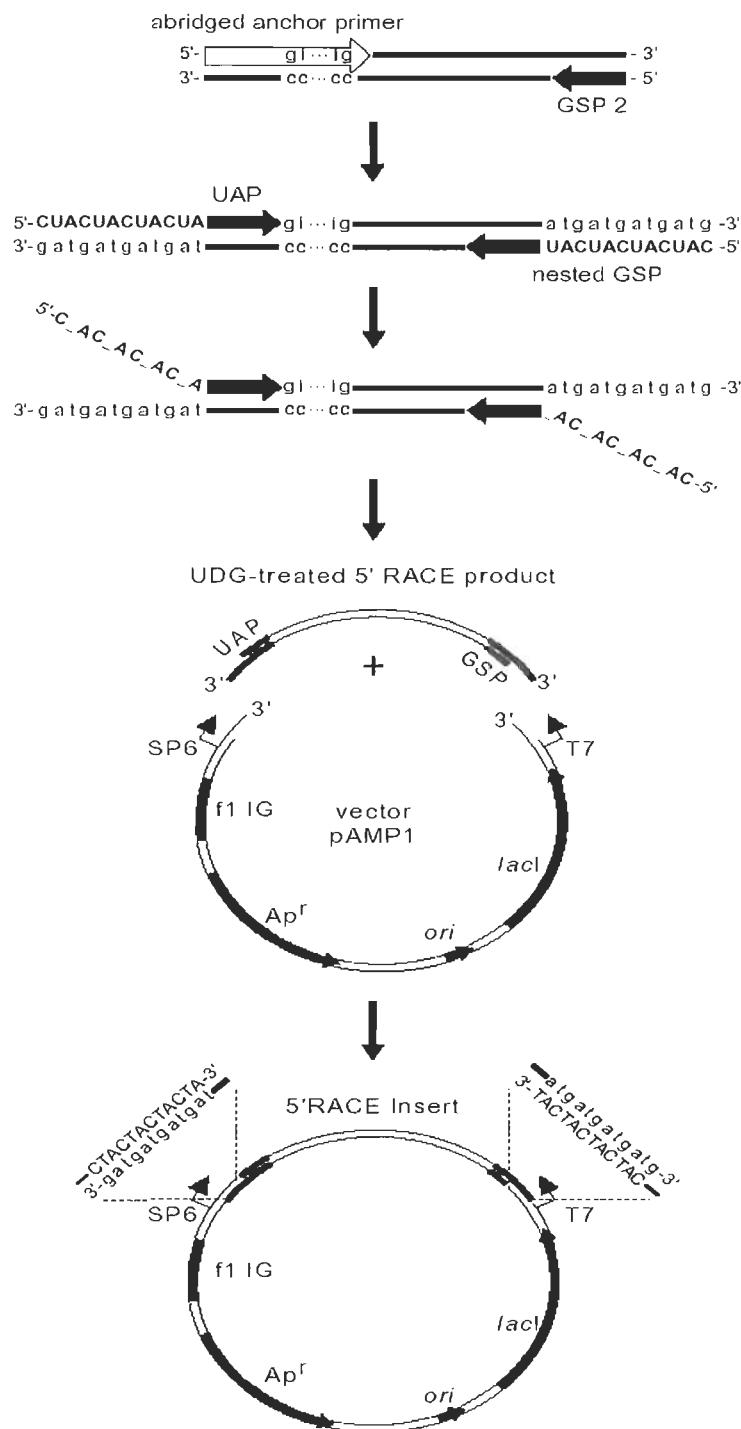


Figure 3.8 - Schéma explicatif pour le clonage des différents produits obtenus lors des amplifications des extrémités de l'ADNc par le système *CloneAMP* de *GIBCO BRL*. D'abord un premier PCR utilisant l'amorce AAP (*Abridged Anchor Primer*) et une amorce spécifique (GSP2), puis une amplification subséquente avec une amorce spécifique à laquelle est ajoutée la séquence de clonage (CAU)₄. 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 MTCTM 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 cPLA₂.

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

Dans un premier temps, c'est la séquence protéique de chacune des cPLA₂ 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 classées 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 cPLA₂ déjà connues et non pas une identité parfaite, ce qui indiquerait qu'il s'agit simplement de la séquence d'une PLA₂ 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 cPLA₂.

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 cPLA₂. 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 cPLA₂.

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 cPLA₂ 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 » cPLA₂ 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 » cPLA₂.

CHAPITRE 4

RÉSULTATS ET DISCUSSION

4.1 Recherche des cPLA₂ connues dans l'EPR et la rétine neurale

La détermination des cPLA₂ 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 cPLA₂.

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

Afin de trouver quelles cPLA₂ connues étaient présentes dans l'EPR et la rétine neurale, des amorces spécifiques aux trois différentes isoformes de cPLA₂ connues dans la littérature jusqu'à présent ont été utilisés en RT-PCR. Ainsi, différentes combinaisons d'amorces pour les cPLA₂ 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 cPLA₂ 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 cPLA₂ alpha.

La cPLA₂ 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 cPLA₂ dans d'autres tissus. D'autres amorces devront donc être utilisées, plus courtes encore, afin d'obtenir cette cPLA₂ dans les tissus où elle est exprimée en grande quantité, comme c'est le cas dans le pancréas (Pickard *et al.*, 1999).

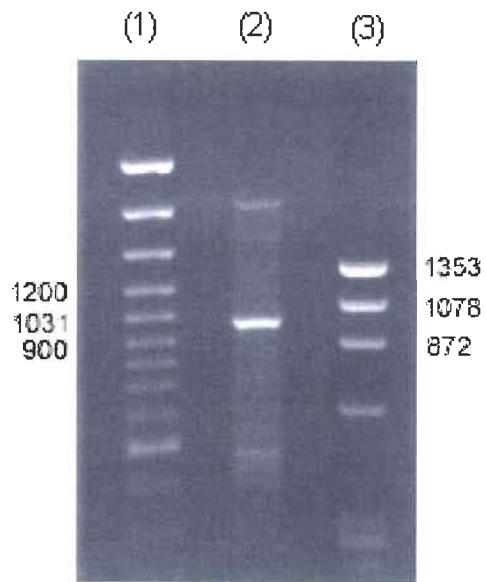


Figure 4.1 - Obtention de la cPLA₂ alpha dans le poumon. La piste (1) contient le marqueur de poids moléculaire GeneRuler 100bp DNA Ladder Plus. La piste (2) présente le fragment de la cPLA₂ alpha obtenu et la piste (3) contient le marqueur de poids moléculaire φX174 Hae III.

La cPLA₂ 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 cPLA₂. 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 cPLA₂ gamma.

Ainsi, une seule des cPLA₂ connues a été clonée au niveau de l'EPR et de la rétine neurale, la cPLA₂ gamma. La cPLA₂ alpha n'est définitivement pas présente dans ces tissus, mais d'autres amorces doivent être essayées en ce qui concerne la cPLA₂ bêta. Un des problèmes possibles avec l'amplification de cette cPLA₂ 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 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 cibler l'ensemble des cPLA₂, c'est-à-dire qu'elles sont dessinées de telle sorte qu'elles peuvent s'hybrider sur n'importe laquelle des cPLA₂ connues, et par conséquent, de toute autre nouvelle cPLA₂ 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 cPLA₂.

Puisque même la cPLA₂ 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 cPLA₂ 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 cPLA₂. 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 cPLA₂.

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 cPLA₂ 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 » cPLA₂. 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 cPLA₂. 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 B.

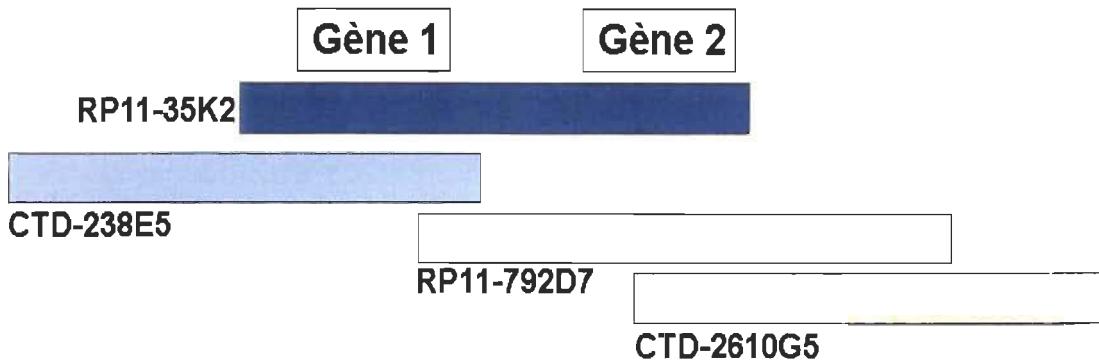


Figure 4.3 - Schéma représentant les HTGS trouvés. Il montre la position des différents HTGS trouvés en superposition avec le clone RP11-35K2 au niveau de chromosome 15, plus précisément dans la région s'étendant de 15q11 à 15q15. Les clones sont présentés dans le sens 5' vers 3', de la gauche vers la 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 cPLA₂ présentant une identité d'environ 40% avec la cPLA₂ 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 cPLA₂. 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.

cPLA₂ nouveau gène 1

```

atgggtgagcctggaaaaggtaaccacaagtccattctgatctacaggactggcagcct
M G E P G K G T T S S H S D L Q D W Q P 20
ggcctaaggggcggtccagggtggatggcaggcccggccaggcttggagtg
G L K G R S R V V W M A G P G P A F G S
gtcaggctaggcggtcacccagatgcagggtggacctgggacgcggccctgagcagcgc
V R L G G H P D A G W T W G R G P E Q R
cgagcagacccctccaccggcaggacaggaaccggcatcctcgccatggcagcaatggc
R A E P P P E A Q E P G I L G D G S N G
tgccccgtggctggtcccaaaggccacccgtcccaacttcagaccggccggccagaa
C P A G W S P K R H R S H F Q T R A P E 100
gcgcatcccagttccctctgggctctgcccagtgcacagacggcaccgctct
A H P Q F P L L G L L L P V Q E Q H R S
gggcccagttccacttcaggccctgtctcccaactgcgcgtctcccaactgcgcgt
G P S S T S G P C S P Q R C S P A A L
ctccctcgccaggccgacccgtggcccaactcgccgcggcccccaggccgcagaca
L L C Q R A T W P S P I A A A A P R A Q T
gttgctgagggtctacgaggcagataactacagggggcagccactttagttgtggca
V A E G Y E A D T T G G Q P V L S C V A
gagctctgtaaaccaggggccccaacccctggccacagaccgtaccgtccgtggcc
E L C K P G A P N P L A T D Q Y Q S V A 200
tgtttggaaaccaggccacacaggcggatgtggtaacactggcggagactcccaacttcatg
C L E P G H T A G E Q W A S S P D F M
ggatggttgtctgggttatgtctggcccaactcgccatggcgtccaccacccatcgaa
G W F A E G M S G P A M G I G S T T S E
atgtggttgtctgggttatgtctggcccaactccaccaggaggccctgaccctgggagc
M W L S D P S P D P T P T Q E P D P G S
ccttggaggggctgtccatgccacctgttgcactgtggcatccggatgaaaaat
P L E G L S P C H L L T V R V I R M K N
gtccggcaggctgtatgcaccaggtaggtatagactggcaccctgcctgcaggctccc
V R Q A D M Q P V G I E L A P C L Q A P 300
agcgtaccggagacacactgaaagggtgtggccaggccgggtggggggccactgttt
S V P E T D L K G V V Q A R G G G A S V
ctggaaaaggccaaagggttcaaggggctgagcaggccctgtgagccagacacagac
L E K P R E G F K R A E Q V P V S Q T D
tgttttgtgagccctgtggctgcccaccgcctctcagaagaagctgaggacaaggaccatc
C F V S L W L P T A S Q K K L R T R T I
tccaactggccaaatccacaggatggaaaggctcaactccacatccggccggactg
S N C P N P E W N E S F N F Q I Q S R V
aagaacgtcttaggttagtgcgtatggcactggacacacatggcaccatctc
K N V L E L S V C D E D T V T P D D H L 400
ctgacagttcttatgacccatggcactgtttccgaaagaaaaccacgtgaagttt
L T V L Y D L T K L C F R K K T H V K F
ccactcaaccgcaggcatggaaaggactggaggtggagttctgtggaggagactccc
P L N P Q G M E E L E V E F L L E E S P
tctccacccgtggatggccctcgatccatggcgtctgggttaattatcttccgggttcc
S P P E T L V T N G V L V V I I F L G S
tgtagctccagaggccacggctggctgtctcaggaaacaggaccaaggagaaaaaa
C S S R G H G W L L L S G E Q D Q G R K
cagtggcccaactggctgtgtccatctgcacccatggcaggatggactaaacgag
Q W A Q L G L C P I L T S A G V R L N E 500
gccagccaaatggggcaccaggcactggggcactggggctgtacagaggaga
A S Q M G H R Q H W G T S W G F C T E G
ggagtgaaggacccctggatggtaacgaaatcttggatggactggccgg
G V K D L L V M V N E S F E N T Q R V R
ccctgcttggaaaccctgtggccaaacctctgcctgttccaaaccggctgcctgttccac
P C L E P C C P T S A C F Q T A A C F H
taccccaacttccactccagggtgcacgtggaaactggccaaagacttggagctgt
Y P K Y F Q S Q V H V E V P K S H W S C
gggctttgtggccctcgatggcccaactcgccatggcaggccctgcactggcttcc 600
G L C C R S R K K G P I S Q P L D C L S

```

gatggtcaggtatgaccctgcctgtgggtgagagttatgaattacatgaaatcata
 D G Q V M T L P V G E S Y E L H M K S T
 ccctgcctgagacactggacgtgcggctgggcttcagccctgtgcccagcagagctggag
 P C P E T L D V R L G F S L C P A E L E
 tttctgcagaagcgaaaggctgtgggccaaggccctgaagcaggtgtcagactggag
 F L Q K R K V V V A K A L K Q V L Q L E
 gaagacctgcaggaggacgagggtggcgctgatagccatcatggccactgggggtggaaaca
 E D L Q E D E V P L I A I M A T G G G T
 agatccatgacccatgtatggccacctgtctggggctgcagaagctgaacctcctggac
 R S M T S M Y G H L L G L Q K L N L L D 700
 tggccagactacatcaactggtctatcggggcacctgggaccatggctacccgtaccgt
 C A S Y I T G L S G A T W T M A T L Y R
 gaccctgactggcctccaaaacttggagccctgtatcttgaggctggagacatgtg
 D P D W S S K N L E P A I F E A R R H V
 gtaaggacaagactacccctccctgtccagaccatccgcaaattccaggaggagctc
 V K D K L P S L F P D Q L R K F Q E E L
 cggcagcgcagccaggaaaggctacagggtcacctttacagacttctggggctgtgata
 R Q R S Q E G Y R V T F T D F W G L L I
 gagacctgcctggggacgaggtagaccaagaccctggactctgagagactgtcctaaa
 E T C L G D E V D Q D P G T L R D C P K 800
 ctgcggatccaggccctgtctgtggggcagcccgagggggtgtgcctatg
 L R I Q S R P L S G G Q P E G V V P M
 ccattggcccttcagagtgttggctgttcagctggggccgatacagcagcccg
 P L V P F Q S D L A G V S W G R Y S S P
 agaaatgaatgaaactgtcagatcagctgtctttgagctgcggccagaacccctg
 R N E C K L S D Q R A A L S C G Q N P L
 cccatctacccatcaatgtcaaggatgtatgaaagcaaccaggactcagagactgg
 P I Y L T I N V K D D V S N Q D F R E W
 ttcgagttccccctacgggtgggtcagaaatggggccttacccctccggag
 F E F S P Y E V G L Q K Y G A F I P S E 900
 ctcttcgctccaggatcttcatggggcgtgtggtaagaggatccggagtcgaatc
 L F G S E F F M G R L V K R I P E S R I
 tgctacatgttaggcctgtggagcagcatcttccctgaacctgtggatgcctggaaac
 C Y M L G L W S S I F S L N L L D A W N
 ctgtcacacacccctcgaggagttttccacaggtggacaaggagaaatgcaggacatc
 L S H T S E E F F H R W T R E K V Q D I
 gaagacgaggccatctgcctgaaatccccaatgtgtatgtaacatcctggagaccacg
 E D P I L P E I P K C D A N I L E T T
 gtagtgcattccagggtcatggctgtccattttccgagaaatccattccatcggtcc
 V V I P G S W L S N S F R E I L T H R S
 ttctgtgtcgtttccacacttcgtctgggtcagtcgcacaccaactacccctcag
 F V S E F H N F L S G L Q L H T N Y L Q
 aatggccatctctaggtggaaatgtgtccattggctgtcaggccaccagggt
 N G Q F S R W K V P L S I G C Q G T R C
 gttggtaacggggcgtttccagacacagtcgttagatgtttccaaaccaggctg
 V G N G K A W L S D T V L D G F P N Q L
 acccgatcccgcaaccacccatgtgtccatggactgcgttcttgcactccagctac
 T E S A N H L C L L D T A F F V N S S Y
 ccggcccttcaggccaggcgaaacccatcatccaccaactactgtgct
 P P L L R P E R K A D L I I H L N Y C A 1100
 gggcccacaaaggcaacctgggtccagtcgcctccaggatccaccatccac
 G S Q T K A T W F Q S P F Q A T H H I H
 ccaagtgtgttgtggatgcgaacaccaacgaaagggttttgcattttactgggg
 P S L C L D A N T N E R A F V H F T W G
 agcagcttgcactggagccctgaaacaaacctgtgtgactactgcactgtcagaacatc
 S S L Q L E P L K Q T C E Y C T V Q N I
 cccttcccaatacgagctgcagatgagaatgaaatctcaaggaatgtcacctgat
 P F P K Y E L P D E N E N L K E C Y L M
 gagaaccccccaggaaaccatgtccccatgtgactttctccactcatcaatgacact
 E N P Q E P D A P I V T F F P L I N D T 1200
 ttccgaaaaataacaaggcaccagggttagagcgaagccctgaggagctggagcaggccag
 F R K Y K A P G V E R S P E E L E Q G Q
 gtggacatttatgttccaaaactccctatgcaccaaggagctgacatacacagaggcc
 V D I Y G P K T P Y A T K E L T Y T E A
 acctttgacaagctgtgtgaaactctcagactataacatcctgaataataaggacactctc
 T F D K L V K L S E Y N I L N N K D T L
 ctccaggctctgcggctcgactggagaagaagaagcgcctgaagggccagtgtccctcc
 L Q A L R L A V E K K K R L K G Q C P S 1280
 tag
 -

cPLA₂, nouveau gène 2

```

R G Q N P L P L Y L S L N V K E N N L E    700
acactggacttcaaggagtgggttagttctccctatgaggtcggttcctgaagtac
T L D F K E W V E F S P Y E V G F L K Y
ggggccttcgtccctcctgagcttcggctccgagttcttcatggacggctgtgagg
G A F V P P E L F G S E F F M G R L M R
aggatcccgagccccgatctgcttctgaaagccatctggagcaacatttctccctg
R I P E P R I C F L E A I W S N I F S L
aacctgctggatgccttgtatgacccatcggacttctggggatgcctggaaacagcacatc
N L L D A W Y D L T S S G E S W K Q H I
aaggacaagaccaggagcttagagaaggagccctgaccacccctcgccgg
K D K T R S L E K E P L T T S G T S S R     800
ctggaggccctgtggctgcagccaggcacggcgtggccaggcattaaaggcttcctg
L E A S W L Q P G T A L A Q A F K G F L
acaggcaggcccctccaccagcgcagccccacttcctccaggccctcagtcaccag
T G R P L H Q R S P N F L Q G L Q L H Q
gactactgttagccacaagacttccacccctggcagactaccatgtgactccatgccc
D Y C S H K D F S T W A D Y Q L D S M P
agccagctgaccccaaggagcccccgtctgcgtggacccgcctacttcatcaac
S Q L T P K E P R L C L V D A A Y F I N
accagctccctccatgttccggcaggccgaggctggacccatctcccttcgac
T S S P S M F R P G R R L D L I L S F D     900
tactccctatctgcgccttcgaggcactgcagcagacggactgtactgcggccgg
Y S L S A P F E A L Q Q T E L Y C R A R
ggcgtccctccccccggatggaaaccgcctcaggaccagcaccaggccaaggaaatgc
G L P F P R V E P S P Q D Q H Q P R E C
cacctcttcagacccgcctgccccgaggccctacaccctgtcaacttccgctggc
H L F S D P A C P E A P I L L H F P L V
aatgccttcataaggaccactcagccccgggtccagcgcagccgcagagctccag
N A S F K D H S A P G V Q R S P A E L Q
ggtggcaagtggatctcaccggggcacctgcctacaccctgtccaacatgacctac
G G Q V D L T G A T C P Y T L S N M T Y     1000
aaggaggaagacttcgagcgcctgtcggtcaactacaacgtgcagaccaggccag
K E E D F E R L L R L S D Y N V Q T S Q
ggtgccatctgcagggccctgaggaccgcgtgaaggacccgactctagaggcgaggcc
G A I L Q A L R T A L K H R T L E A R P     1040
ccaaggccacagacactga
P R A Q T -

```

Figure 4.4 - Séquences codantes des nouvelles cPLA₂ 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 cPLA₂ 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 cPLA₂.

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 cPLA₂ 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.

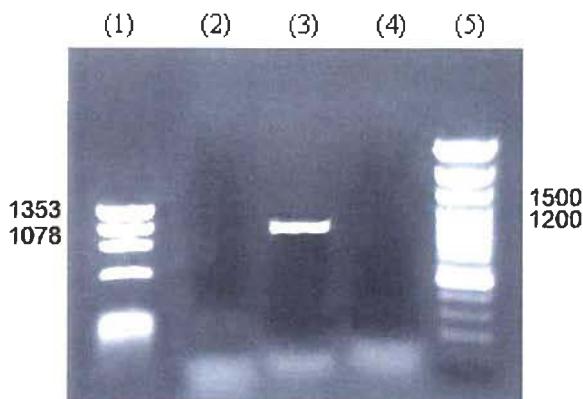


Figure 4.5 - Obtention d'un fragment d'environ 1100 pb correspondant à l'ADNc du nouveau gène 2 trouvé par des analyses bioinformatiques. Les pistes (1) et (5) contiennent respectivement les marqueurs Φ X174 Hae III et le GeneRuler 100bp DNA Ladder Plus. La piste (3) présente le fragment obtenu par RT-PCR 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 Analyse du fragment obtenu par RT-PCR

Le fragment obtenu a été cloné, séquencé et analysé afin de déterminer son identité avec les cPLA₂ connues. De même, il fallait vérifier l'exactitude de la séquence prédictive 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 cPLA₂. La figure 4.6 présente la séquence de fragment obtenu pour ce nouveau gène 2.

En comparant la séquence prédictive 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 prédictive par les analyses bioinformatiques. En effet, l'alignement de la séquence prédictive 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 prédictive, 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.

```

gcttcaatctctgtgcagaggagcaggccctcgtggacagggaaagcagggtggccaaag
F N L C A E E Q A F L S R R K Q V V A K
gccctgaagcaggccctgcagctggacagagacctgcaggaggatggatccgttgcgtg
A L K Q A L Q L D R D L Q E D E V P V V
ggcatcatggccacaggagggtgcccggccatgcactctacggccacctattg
G I M A T G G G A R A M T S L Y G H L L
gccttcagaagctggccctccatggactgtgtggacttcatggatctggctct
A L Q K L G L D C V T Y F I G I S G G S
acgtggacaatggccaccctgtacggggaccctgagtggcagaggacactggaggaa
T W T M A H L Y G D P E W S Q R D L E G
cctatcagatacggccggagcacctggccaagagcaagctggagggtctttccccagag
P I R Y A R E H L A K S K L E V F S P E
cgccctggcagactaccgcgggagctggagctgcgggctgagcaggccacccacgacc
R L A S Y R R E L E L R A E Q G H P T T
tttggacactgtggcgcttagtgcgtggacttcatgcgtcggccaggatggatcag
F V D L W A L V E S M L H G Q V M D Q
aagctgtcaggacagagccgcctggaaacgggtcagaaccctctgcggctacttg
K L S G Q R A A L E R G Q N P L P L Y L
agcctcaatgtcaaagagaacaatctggagacactggacttcaaggagtgggtgagttc
S L N V K E N N L E T L D F K E W V E F
tcccccattatgcgggttccatggactacggggcctcgccctctggactcttcggc
S P Y E V G F L K Y G A F V P P E L F G
tccgagttcttcatggacggctgtggatggggatccggagccccggatctggcttctg
S E F F M G R L M R R I P E P R I C F L
gaagccatctggagcaacattttctccatggacttgcggatgcctggatgacccacc
E A I W S N I F S L N L D A W Y D L T
agttctggggagtctggaaacacgcacatcaaggacaagaccaggagcttagagaaggag
S S G E S W K Q H I K D K T R S L E K E
ccccctgaccacccctggggacctccgcggctggaggcctcgccggcagccaggcacc
P L T T S G T S S R L E A S W L Q P G T
gcgctggccaggcatttaaaggcttcgtacaggcaggccctccaccagcgcagcccc
A L A Q A F K G F L T G R P L H Q R S P
aactctccaggccctccagctgcaccaggactactgttagccacaaaagacttctcca
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 cPLA₂ 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 cPLA₂ connues dans la littérature. Ce fragment du nouveau gène 2 présente une identité d'environ 53% avec le segment correspondant de la cPLA₂ bêta.

Des analyses plus poussées de ce fragment ont permis de vérifier la présence de la séquence consensus lipasique GXSGS propre aux cPLA₂. 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 cPLA₂ connues, les domaines caractéristiques aux cPLA₂ 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 cPLA₂ puisqu'il présente, à première vue, toutes les caractéristiques de cette famille.

nouveau gène 2	FNLCAEEQAFLSRRKQVVAKALKQALQLDRD---LQEDEVPVVGIMATGGGARAMTSLYG	57
beta	FGPCAEQQAFLSRRKQVVAALRQALQLDGD---LQEDEVPVVAIMATGGGIRAMTSLYG	57
alpha	MALCDQEKTFRQQRKEHIRESMKKLLGPKNSEGLHSARDVPVVAILGSGGGFRAMVGFS	60
gamma	PGLQKEEKAVERRLHLVALKKLR-----IEADEAPVVAVLGSGGGLRAHIACLG	52
nouveau gène 2	HLLALQKLGLLDCVTYFIGISGSTWMAHLYGDPEWSQRDLEGPIRYAREHLASKLEV	117
beta	QLAGLKELGLLDCVSYTIGASGSTWALANLYEDPEWSQKDLAGPTELLKTQVTKNKLGV	117
alpha	VMKALYESGILDATVYAGLGSSTWYMSTLYSHPDFPEKGPEEINEELMKNVSHNPLL	120
gamma	VLSMKEQGLDAVTYLAGVSGSTWAISLYTN---DGDMEALEDLKHRFTRQEWDLA	108
nouveau gène 2	SPERLASYRRELELRAEQGHPTTFVDLWALVLESMLHGQVMDQKLSGQRAALERGNPLP	177
beta	APSQLQRYRQELAERARLGYPSCFTNLWALINEALLDEPHDHKLSDQREALSHGQNPLP	177
alpha	TPQKVKRYVESLWKKSSGQPVTFTDIFGMLIGETLIHNRMNTTLSSLKEKVNTAQCP	180
gamma	KSLQKT-----IQAARSENYSLTDFWAYMVISKQTRELPESHLSNMKKPVEEGTLPP	161
nouveau gène 2	LYLSLN--VKENNLETIDFKEWVEFSPYEVGFLKYGAFVPPPELGSEFFMGRLMRRIPEP	235
beta	IYCALN--TKGQSLTTFEFGEWCFSPEVGFPKYGAFIPSELFGSEFFMGQLMKRIPES	235
alpha	LFTCLH--VKP-DVSELMFADWVEFSPYEIGMAKYGTFMAPDLFGSKFFMGTVVKKYEEN	237
gamma	IFAAIDNDLQPSWQEARAPEATWFEEFTPHHAGFSALGAFVSITHFGSKFKKGRLVRTHPER	221
nouveau gène 2	RICFLEASSGESWK--QHIKDKT-RSLEK-----EIWSNIFSLNLDAW-----YDL	279
beta	RICFLEGIWSNLYA--ANLQDSLYWASEP-----SQFWDRWVRNQANLD-----KEQ	280
alpha	PLHFLMGVWGSAFS--ILFNRLGVSGSQSRGSTMEELENITKHIVSNDDSDDESH	295
gamma	DLTFLRGLWGSALGNTEVIREYIFDQLRN-----LTLKGLWRRAVANAKSIGHLIFAR	274
nouveau gène 2	TPLTTSGTSSRLEASWLQPGTALAQAFKGFLTGRPLHQR-----SPNFLQG-LQLHQD	331
beta	VPLLKIEEPPSTAG-----RIAEFFTDLLTWRPLAQA-----THNFLRG-LHFHKD	325
alpha	EPKGTCENEDAGSDYQSDNQASWIHRMIMALVSDSALFNTREGRAGKVHNFMIG-LNLNTS	354
gamma	LLRLQESSQGEHPPPDEGGEPHTWLTEMLENWRTSLEKQEOPHEDPERKGSLSNLMD	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 cPLA₂ connues (alpha, bêta et gamma). La protéine associée au fragment du nouveau gène 2 présente une identité de 53% avec la cPLA₂ bêta.

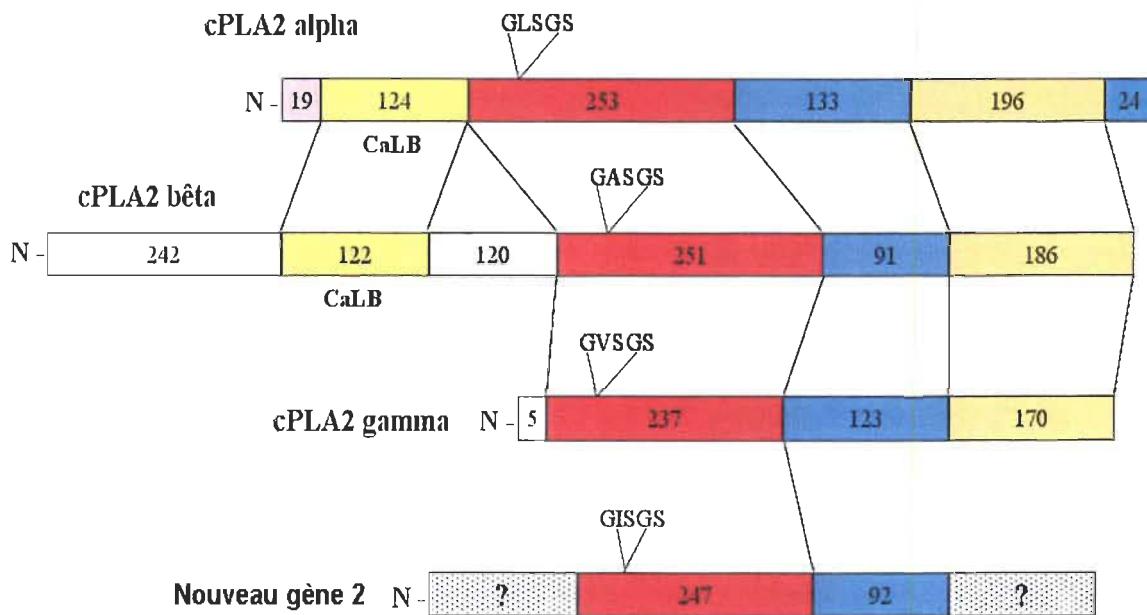


Figure 4.8 - Représentation schématique des trois isoformes connues (cPLA₂ alpha, bêta et gamma) avec le fragment du nouveau gène 2. La nouvelle cPLA₂ (nouveau gène 2) semble être organisée de la même façon que les trois autres isoformes de la famille dont elle fait probablement partie.

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 cPLA₂ sont beaucoup plus spécifiques à l'ADNc recherché que ne l'est la sonde fabriquée à partir de l'ADNc de la cPLA₂ 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 cPLA₂.

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) - *sondeB* (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 amores utilisées pour l'amplification) qui ont servi en tant que sondes (figures 4.9).

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

```
ACCTGGAGGGACCTATCAGATACTGCCCGGGAGCACCTGGCCAAGAGCAAGCTGGAGGTC
TTTCCCCAGAGCGCCTGGCAGCTACCGCCGGAGCTGGAGCTGCGGGCTGAGCAGGG
CCACCCCACGACCTTGTGGACCTGTGGCGCTAGTGCTGGAGTCCATGCTGCA
```

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

```
GAGAAGGAGCCCTGACCACCTCGGGGACCTCCTCGCGGCTGGAGGCCTCGTGGCTGCA
GCCAGGCACGGCGCTGGCCAGGCATTAAAGGCTTCCAGGGCCCTCCAGCTGACAGGCAGGCCCTCCACC
AGCGCAGCCCCAACCTCCTCCAGGGCCTCCAGCTGACCCAGGACTACTGTAGCCACAAAG
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.

```

gcttcaatctctgtcagaggagcaggccttcctgagcaggagaagcaggtggtgccaaag
F N L C A E E Q A F L S R R K Q V V A K
gccctgaagcaggccctgcagctggacagagacactgcaggaggatgaggtacccgttgt
A L K Q A L Q L D R D L Q E D E ♦ V P V V
ggcatcatggccacaggaggaggtgcccggccatgacactacgtacggccacctattg
G I M A T G G G A R A M T S L Y G H L L
gccttcagaagctggccctcttagactgtgtgacacttcattggcatctctggctct
A L Q K L G L L D C V T Y F I G I S G S
Acgtggacaatggcccacacgtacggggaccctgagtggtcgcagagg[acactggaggga] Amorce SondeH
T W ♦ T M A H L Y G D P E W S Q R D L E G
Cctatcagatacgcggggagcacctggccaagaga[acagctggaggctttccccagag] Amorce SondeH (AS)
P I R Y A R E H L A K S K L E V F S P E
cgccctggcgagctaccgcggagctggagctgcggctgagcagggccacccacgacc
R L A S Y R R E L E L R A E Q G H P T T
tttggacactgtggcg[taagtgcgtggagtcac]cgccaggtgatggatcag Amorce SondeH (AS)
F V D L W A L V L E S M L H G ♦ Q V M D Q
aagctgtcaggacagagagccgcctggAACGGGGTcagaaccctctgcccctacttg
K L S G Q R A A L E R G Q N P L P L Y L
agcctcaatgtcaaagagaacaatctggagacactggacttcaaggagtgggtgagttc
S L N V K E N N L E T L D F K ♦ E W V E F
tccccctatgaagtgcgttctgaagtgacgggccttcgtccctctgagctttcgcc
S P Y E V G F L K Y G A F V P P E L F G
tccgagttctcatggacgctgatgaggaggatcccggagcccccggatctgcttctg
S E F F M G R L M R R I P E P R I C F L
Gaagccatctggagcaacatcccctgaacactgctggatgcctggatgacctcacc
E ♦ A I W S N I F S L N L L D A W Y D L T
agttctggggagtcctggaaacagcacatcaaggacaagaccaggagctt[gagaaggag] Amorce SondeB
S S G E S W K Q H I K D K T R S L ♦ E K E
ccccctgaccaccggggaccccttcgcggctggaggcctcgtggctgcagccaggcact Amorce gène 2c (AS)
P L T T S G T S S R L E A S W L Q P G T
gcgctggcccaggcat[taaggcttc]tgcacaggcaggcccctccaccagcgcagcccc
A L A Q A F K G F L T G R P L H Q R S P
Aacttcctccaggcctccagctgcaccaggactactgt[tagccacaaagacttctcca] Amorce gène 2c (AS)
N F L Q G L Q L H Q D Y C S H K D F S

```

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.

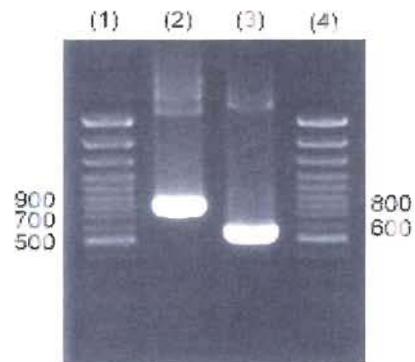


Figure 4.11 - Résultat obtenu pour l'amplification par PCR des deux EST utilisés pour la fabrication d'une sonde. Les pistes (1) et (3) contiennent le marqueur GeneRuler 100bp DNA Ladder Plus, alors que les pistes (2) et (3) contiennent respectivement l'insert amplifié des EST 2068898 et 2739188.

TABLEAU 4.1
Compilation de différents EST

Numéro de l'EST (IMAGE) ¹	Provenance	Longueur lisible approximative ²	Localisation probable ³
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

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

² 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).

³ 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 cherché 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 cPLA₂ (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 cPLA₂ (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.

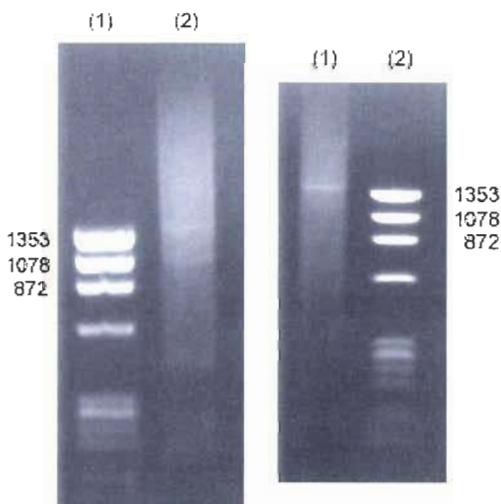


Figure 4.13 - Résultats obtenus avec la technique du RACE. Résultats obtenus pour la première ronde d'amplification pour le RACE 3' (à gauche) et pour la troisième ronde d'amplification (à droite). Le gel de gauche présente le marqueur Φ X174 Hae III dans la piste (1) et le produit de PCR dans la piste (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.

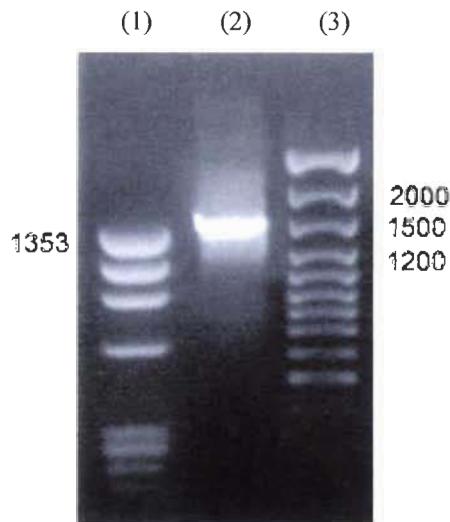


Figure 4.2 - Obtention de la cPLA₂ gamma dans l'EPR. Les marqueurs habituels sont encore utilisés ici (Φ X174 Hae III et GeneRuler 100bp DNA Plus) dans les pistes (1) et (3). La bande d'environ 1500 pb en piste (2) est l'ADNc de la cPLA₂ gamma.

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 cPLA₂ dans l'EPR et la rétine neurale

Suite à la vérification de la présence de cPLA₂ connues dans la rétine neurale et l'EPR, la recherche de nouvelles cPLA₂ 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 cPLA₂ connues sont présentes au niveau de l'EPR et de la rétine neurale. La cPLA₂ gamma s'est avérée présente dans ces deux tissus, alors que la cPLA₂ alpha en est absente. En ce qui a trait à la cPLA₂ bêta, aucune des amores 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 amores 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 cPLA₂ 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 cPLA₂ bêta est absente ou présente au niveau de l'EPR et/ou de la rétine neurale.

Par la suite, la recherche de nouvelles cPLA₂ a été entreprise par diverses analyses bioinformatiques réalisées sur les banques de données de la NCBI. Dans un premier temps, les séquences de cPLA₂ connues ont été utilisées pour repérer des séquences présentant une identité avec ces cPLA₂, 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 cPLA₂. 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'amores spécifiques ont été déterminées et celles-ci ont été utilisées en RT-PCR. Une seule combinaison – parmi une soixantaine de combinaisons - de ces 29 amores 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 cPLA₂. En effet, non seulement l'alignement avec les autres cPLA₂ connues est presque parfait, mais ce fragment présente les domaines reconnus chez les autres cPLA₂, 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 cPLA₂ 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 cPLA₂. Ainsi, l'ADNc complet de ce nouveau gène 2 devait être cloné afin de permettre des analyses plus poussées de cette probable nouvelle cPLA₂.

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 cPLA₂ 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 10⁹, 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 cibler 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 PLA₂, ou cPLA₂. 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 cPLA₂ déjà connues. En effet, tous les algorithmes utilisés indiquent que l'homologie de ce fragment cloné avec la cPLA₂ bêta se situe autour de 50%, ce qui est très considérable puisque l'identité présente entre les différentes cPLA₂ connues est d'approximativement 30%. Par conséquent, il est fort probable que cette nouvelle séquence soit une partie de la séquence d'une nouvelle cPLA₂.

La recherche des extrémités du clone obtenu permettra non seulement de vérifier l'alignement de cette potentielle cPLA₂ en entier avec les autres cPLA₂ connues, mais aussi de déterminer l'activité phospholipasique de cette nouvelle protéine pour ainsi confirmer sa classification au sein de la famille des PLA₂.

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 cPLA₂ bêta, qui est la seule des cPLA₂ qui n'a pu être obtenue par RT-PCR. En effet, la cPLA₂ 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 cPLA₂ bêta, pourrait aussi expliquer la difficulté à amplifier les séquences codantes de ces nouveaux gènes par PCR. En fait, puisque les séquences des amores 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 amores 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 cPLA₂ 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 cPLA₂ soient spécifiques à l'EPR. Ainsi, puisque la cPLA₂ gamma n'était pas connue au moment de leur travaux, il est possible qu'une 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 cPLA₂ 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 cPLA₂ présentes dans l'EPR et/ou la rétine neurale a été partiellement atteint. Il a été déterminé que la cPLA₂ alpha est absente de l'EPR et de la rétine neurale, alors que la cPLA₂ 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 cPLA₂ bêta dans ces deux tissus.

Finalement, l'objectif consistant à rechercher des cPLA₂ spécifiques à la rétine neurale n'a pas été atteint puisque la présence de l'ADNc de la nouvelle cPLA₂ 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 cPLA₂ 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 cPLA₂ 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. Tramposch (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 PLA₂ 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 Debora 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 Ca²⁺ transients in rat cardiac myocytes via inhibition of distinct K⁺ 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., Grünwald-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. Ilincheta 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 A2-alpha from the venom of *Crotalus adamanteus*. A new classification of phospholipases A2 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 Ca²⁺ response required for continuous translocation and activation of cytosolic phospholipase A2. *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 A2 inhibitors and calmodulin antagonists as inhibitors of cytosolic phospholipase A2. *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 A2 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 A2. *Biochem J* **327** (Pt 2), 455-60.
- Jacob, M., P. K. Weech et C. Salesse (1998). Phospholipases A2 of rod outer segment-free bovine retinae are different from well-known phospholipases A2. *Biochim Biophys Acta* **1391**, 169-80.

- Kapus, A., R. Romanek et S. Grinstein (1994). Arachidonic acid stimulates the plasma membrane H⁺ conductance of macrophages. *J Biol Chem* **269**, 4736-45.
- Kol, S., K. Ruutiainen-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 Ca²⁺-dependent human platelet phospholipase A2. *Biochim Biophys Acta* **878**, 394-403.
- Kramer, R. M., C. Hession, B. Johansen, G. Hayes, P. McGraw, 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 Ca²⁺(+)-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 (cPLA₂) in thrombin-stimulated platelets. Evidence that proline-directed phosphorylation is not required for mobilization of arachidonic acid by cPLA₂. *J Biol Chem* **271**, 27723-9.
- Kramer, R. M. et J. D. Sharp (1997a). Structure, function and regulation of Ca²⁺-sensitive cytosolic phospholipase A2 (cPLA₂). *FEBS Lett* **410**, 49-53.
- Kramer, R. M. et J. D. Sharp (1997b). Structure, function and regulation of Ca²⁺-sensitive cytosolic phospholipase A2 (cPLA₂). *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). cPLA₂ 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 Ca²⁺-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 USA* **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. Strifler, 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. Strifler, 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 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 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-Honigmann, 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 (cPLA₂). *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 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, cPLA₂-gamma, that is prenylated and contains homology to cPLA₂. *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 Ca²⁺ 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 Ca²⁺ 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).

Abstract

Purpose. Clone and sequence the phospholipases A2 (PLA₂) expressed by human retinal pigmented epithelial (RPE) cells. **Methods.** Primers of different PLA₂ 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 PLA₂s of group IIA and group V as well as the cytosolic PLA₂ gamma which is a member of the group IV PLA₂. **Conclusions.** RPE cells express three known phospholipases A2, i.e. two different secreted PLA₂ (sPLA₂) and one cytosolic PLA₂ (cPLA₂). Indeed, the sequence of both group V sPLA₂ and cPLA₂ gamma corresponds to the published sequences of these enzymes. However, given that we could not obtain the sequence of group IIA sPLA₂ 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 sPLA₂.

Introduction

Retinal pigment epithelial (RPE) cells are responsible for the shedding and phagocytosis of the outer segments of retinal photoreceptors (for a review, see ¹). The ingested phagosomes will then be hydrolyzed by enzymes of RPE cells such as phospholipases A2 (PLA₂). These ubiquitous enzymes play crucial roles in normal cellular functions by participating in the metabolism and turnover of phospholipids ^{2,3}. PLA₂s constitute a diverse family of enzymes that hydrolyze the *sn*-2 fatty acyl ester bond of phosphoglycerides, producing free fatty acids and lysophospholipids.

PLA₂s 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) ⁴ which are located at position *sn*-2 of these phospholipids ⁵. It was shown that the rat retina strives to maintain a constant proportion of docosahexaenoic acid ⁶. PLA₂s 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 PLA₂. They have been isolated from different sources and classified into three major groups: secretory PLA₂ (sPLA₂), cytosolic PLA₂ (cPLA₂), and Ca²⁺-independent PLA₂ (iPLA₂). There are six subgroups of sPLA₂ in mammals: groups IB, II (A, C, D, E and F), III, V, VIIA and X. Except for group VIIA, sPLA₂s are typically small proteins (\approx 14 kD) which contain multiple disulfide bridges. In addition, they all use a catalytic histidine and need μ M Ca²⁺ concentrations for hydrolysis ^{7,8}. cPLA₂s, which are also called group IV PLA₂, are subdivided into three subgroups: cPLA₂ α , cPLA₂ β and cPLA₂ γ . They are much larger than the sPLA₂s (61-114 kD) and use a catalytic serine instead of a histidine ^{9,10}. These three cPLA₂ isoforms contain two homologous domains named A and B which are interspaced with a sequence specific to each isoform. cPLA₂ α and β contain an amino-terminal domain called C2 that has been shown to be involved in their Ca²⁺-dependent binding to membranes. In contrast, cPLA₂ γ lacks this C2 domain and, accordingly, it is calcium-independent ^{11,12}. cPLA₂ β has additional domains next to the C2 domain, making this protein the largest one in this group IV. iPLA₂, or group VI PLA₂, like the other cPLA₂s, uses the same catalytic nucleophilic serine but are not dependent on Ca²⁺ for membrane binding.

We have previously shown that PLA₂ activity can be measured in bovine RPE cells ¹³ and that these PLA₂s are different from known PLA₂s ¹⁴. More recently, we have evaluated the sensitivity to different treatments of partially purified PLA₂s from human RPE cells and performed immunoblot experiments ¹⁵. These results suggested that human RPE cells probably contain two novel types of intracellular PLA₂ enzymes, one of them being similar to a bovine intracellular RPE-PLA₂ recently identified ¹⁴, and that cultured human RPE cells secrete a sPLA₂ enzyme. In the present paper, we present our results on the cloning of the types of PLA₂ expressed by human RPE cells.

Methods

EXTRACTION OF RNA FROM RPE CELLS AND AMPLIFICATION OF DIFFERENT PLA₂S
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.¹⁵. RPE cells were either cultured as described¹⁵ 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^{16,17} (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 sPLA₂ and cPLA₂, respectively. The reaction mixture contained 2 µL of cDNA template, 2.5 U of *Taq* DNA polymerase, and 10 % DMSO, 1.5 mM MgCl₂, 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 PLA₂s 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 PLA₂S

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 positive 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 sPLA₂ as well as cPLA₂ α and cPLA₂ 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) sPLA₂s. 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 sPLA₂s migrated to ~ 300 and 550 bp, respectively. Sequencing of these bands revealed the presence of both gIIA and gV sPLA₂s. These sequences possess 100% homology with the known sequences for these sPLA₂. However, only the sequence upstream of the catalytic site (H48 and D49) of gIIA sPLA₂ could be amplified. Other primers designed downstream the catalytic site failed to amplify either part or the full sequence of gIIA sPLA₂. These results suggest that RPE cells express an isoform of the gIIA sPLA₂. In contrast, the entire cDNA of gV PLA₂ was successfully amplified with a polymorphism substitution on the allele 140 (T140C)¹⁸. Figure 1B

shows the electrophoresis of the PCR products prepared using primers for cPLA₂ γ . 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 cPLA₂ γ .

Discussion

The results of the specific amplification of PLA₂s by RT-PCR show the presence of transcripts of gIIA and gV sPLA₂s as well as cPLA₂ γ in RPE cells. We have obtained the complete coding sequence for gV sPLA₂ as well as cPLA₂ γ , 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 results¹⁵ that an isoform of gIIA sPLA₂ could be present in RPE cells.

The results presented here partly contrast with our previous conclusions on the PLA₂ content of human RPE cells¹⁵. Indeed, on the basis of the sensitivity of different PLA₂-active fractions eluted from cation-exchange chromatography to para-bromophenacylbromide (pBPB), Ca²⁺/EGTA, DTT, heat and H₂SO₄, we had concluded that human cultured RPE cells probably contain two different intracellular PLA₂ enzymes. In addition, we had performed immunoblot experiments where no cross-reactivity was observed between the RPE intracellular PLA₂ enzymes and several antisera directed against sPLA₂ and cPLA₂, and thus concluded that intracellular RPE-PLA₂s are different from well-known secretory, cytosolic and Ca²⁺-independent PLA₂s. However, only antibodies against cPLA₂ α were assayed as antibodies against cPLA₂ β and cPLA₂ γ are not commercially available. In fact, the most striking difference between these data and the present results is that cPLA₂ γ is calcium-independent^{11,12} whereas we have previously demonstrated that intracellular PLA₂ from RPE cells are calcium dependent. It can thus be concluded that other types of intracellular PLA₂ should be expressed by human RPE cells and that these enzymes could be novel types of intracellular PLA₂ enzymes.

We also previously reported an additional RPE-PLA₂ enzyme that was secreted and which exhibited sensitivity to pBPB, Ca²⁺/EGTA, DTT, heat and H₂SO₄ characteristic of sPLA₂ enzymes¹⁵. This PLA₂, with an apparent molecular weight of approximately 22 kDa, cross-reacted weakly with an antiserum directed against porcine pancreatic group I sPLA₂ but strongly with an antiserum directed against N-terminal residues 1-14 of human synovial group II sPLA₂ (which is now called group IIA sPLA₂), which suggested that this extracellular enzyme was a member of the sPLA₂ class of enzymes¹⁵. These data are in good agreement with the present finding of the expression of a group IIA sPLA₂ in RPE cells. In addition, our previous observation of a molecular weight of 22 kD for this enzyme¹⁵ 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 sPLA₂ and supports our conclusion that this sPLA₂ is an

isoform of the known gIIA sPLA₂. Moreover, the present observation of the expression of an additional sPLA₂, namely a gV sPLA₂, by RPE cells could not be documented in our previous report ¹⁵ as antibodies against this enzyme became only very recently available.

It is very interesting to note that RPE cells express sPLA₂ 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 receptors²⁰, is involved in the phagocytosis process in macrophages and is also present in the plasma membrane of RPE cells^{21,22}. Moreover, it has been demonstrated that binding of different types of sPLA₂ to M-type receptor leads to the activation of phagocytosis^{23,24}. It can thus be proposed that gIIA or gV sPLA₂s 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 sPLA₂ increases the rate of disk membrane/plasma membrane fusion by five-fold, in a dose-dependent manner²⁵. 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 sPLA₂s induce activation of cPLA₂ γ ^{26,27} which could also take place in RPE cells and be responsible for POS phagocytosis. The exact role of gV sPLA₂ is not fully understood but it might be related and complementary to gIIA PLA₂²⁸.

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 A2, *J Biol Chem*, 1994, 269, 13057-60.
8. Six DA and Dennis EA, The expanding superfamily of phospholipase A(2) 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 PLA₂ 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 A2, 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 A2, cPLA₂-gamma, that is prenylated and contains homology to cPLA₂, *J Biol Chem*, 1998, 273, 21926-32.
12. Pickard RT, Strifler BA, Kramer RM and Sharp JD, Molecular cloning of two new human paralogs of 85-kDa cytosolic phospholipase A2, *J Biol Chem*, 1999, 274, 8823-31.
13. Jacob M, Weech PK and Salesse C, Presence of a light-independent phospholipase A2 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 A2, *Biochem J*, 1997, 327, 455-60.
15. Van Themsche C, Jacob M and Salesse C, Human retinal pigment epithelium secretes a phospholipase A2 and contains two novel intracellular phospholipases A2, *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 A2 of rod outer segment-free bovine retinae are different from well-known phospholipases A2, *Biochim Biophys Acta*, 1998, 1391, 169-80.
20. Nicolas JP, Lambeau G and Lazdunski M, Identification of the binding domain for secretory phospholipases A2 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, Greaton 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 A2, *J Biol Chem*, 1996, 271, 250-7.
24. Valentin E and Lambeau G, Increasing molecular diversity of secreted phospholipases A(2) 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 A2 activates the cascade of mitogen-activated protein kinases and cytosolic phospholipase A2 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 A2 induces phospholipase C γ -1 activation and Ca $^{2+}$ 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 A2S in regulating arachidonic acid release. Type IIa and type V secretory phospholipase A2S are functionally redundant and act in concert with cytosolic phospholipase A2, *J Biol Chem*, 1998, 273, 14411-23.
29. Seilhamer JJ, Randall TL, Yamanaka M and Johnson LK, Pancreatic phospholipase A2: 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 A2 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 A2 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(2)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(2)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(2). 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(2) 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 A2 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 A2, *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 PLA₂s present in RPE. (A) PCR products of sPLA₂s. Lane 1 is the Gene Ruler100 bp ladder Plus (Fermentas, Burlington, On). Lanes 2 and 3 are, respectively, PCR products of gIIA and gV sPLA₂s. (B) PCR products of cPLA₂. Lane 1 is the λ DNA- Hind III digest and lane 2 is the φX174-Hae III digest. Lane 3 shows the cPLA₂ γ PCR product.

Table 1. Primers used to amplify different types of PLA₂ expressed by RPE cells

Types of PLA ₂	Forward primers	Reverse primers	Length of PCR products (bp)
IB ²⁹	TGGTCATCTCAGTTCTTTC	TCAACTCTGACAATACTTCT	482
IIA ³⁰	AGCCACCAAGGAGGAGCAGG	GTCATGAGTGACACAGCAGC	309
IIC ³¹	CTGGACAACCTCCACCCCTCA	TGGAGTTTGTCCCTGCCACA	406
IID ³²	TGCCAGCATCTGCCTCCACT	AGGAACAGGGTAGAGGGTGT	480
IIIE ³³	TGCTCCTTGTGACCTCCCT	CAGCTTGTGGGATAATGGG	437
IIIF ³⁴	ATGAAGAAGTTCTTCACCGT	TTGGTACTGCAGGTGACCT	480
III ³⁵	GAGGATGGACCATGCCTGGCACA	TTACGGATGTCTTAGTCGTGCTGA	N/A
V	GTGGATACCAATGTTCCGAC	CCTAGGAGCAGAGGATGTTG	549
VIIIA ³⁶	AGATGGTGCCACCCAAATTG	TCCTAATTGTATTCTCTAT	1325
X ³⁷	GGAATTGCCTATATGAAATATGGT	GGAATTCAAGGTAGTCAGTCACACTTG	327
XII ³⁸	TTTGC GGCGCATATGGAGCTGGCTGCTGCCAAGT	TTTAAGCTTCTAGAACATCTGCACTAGCTGTCGGCAT	717
cPLA ₂ α ³⁹	CATTATAGATCCTTACCAAG	CTATGTTGGTTTACTTA	2987
cPLA ₂ β ⁴⁰	CACTTCTACCGGGACTGGGTCT	TCCTGGTTGTTGCAGACATTGT	2855
cPLA ₂ γ ¹¹	AGTTCCATAATTCTGGGCTC	GCCACGTTCATCAACTCTCAA	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 TACAATTTT CCATCTTAA AAGAATGAGA AATTAAGCCA
61 CCATTTAGAC CTATGTTTC TATGTAAATA AAACTGAAAA AAAATTTTT TTTTGAGAC
121 TGGGCTCTGC TCTGGGCC AGGCTGGCTG GAGTCAGTG GTGCAATCTC GGCTCACTGC
181 TGCCTCGACC TCCTGGGCTC CAACAATCT CCTGCCTAG CCTCCCAAGT AGCTGGGACT
241 ACAGACACAT GCCACCATGC CTAGCTAAC TTTTGAGAA ACAGGGTTTC ACCATGTTGC
301 CCAGGCTGGT CTTGAATGCC TGGGCTCAAG CTATCCCCCT ACCTTGGCCT CCCAAAGTAC
361 TGGGATTACA GGCGATGAGCC ACCGCAACTG GCCTAAACT GAAATTGTTT AGCAACCTTA
421 ACGGCTATA TTCTAAACCT TTAATTGATG TCATTGCTAG TTTTCACTTC ATCCAAAGTT
481 TCTACGTTTC TGTGAAGTTG TGGCACCCCA GAATAGATAC ACCATTAAAA TATGTCCTC
541 AAATTTACAG TAAATCTGAT GATATCTTT ATCTCTAAAT TTAGAGAAAA AAATCTGCTA
601 ATAGCACACG GCATTACTG AGTGCTTGCT ATATTCTGA TATGGCGCTG AGCCTTCTAT
661 ATTGTTTCAC CTCATGCTCT GAGGGTGGATG ATGACATTAC CCTCGTTTA TTTTTATTTT
721 TTTTTTTTTT TTTTTTTTTT TTTTTTTGAG ACGGAGTCTC GCTCTGTCGC CCAGGCTGGA
781 GTGCAGTGGC GGGATCTCGG CTCACTGCAA GCTCCGCCTC CGGGGTTAC GCCATTCTCC
841 TGCCCTCAGCC TCCCAGTAGT CTGGGACTAC AGGCGCCCGC CACTACGCC AGCTAATTTT
901 TTTGTATTT TAGTAGAGAC GGGGTTTCAC CGTTTTAGCC GGGATGGTCT CGATCTCTG
961 ACCTCGTGTAC CCGCCCGCCT CGGCCTCCCA AAGTGTGGG ATTACAGGCG TGAGGCCACCG
1021 CGCCCGGGCC TACCCCTCGTT TTATTTACAG GTGGGAAAC TGCGGCATCA GAGAGGCCAA
1081 GCGGCTTGCC CAAGGTACA CAGCGGATGT TCGAGTGGAA ATGGAATGCA AGCATTCA
1141 CTCCAGAACT TGCAGTGTCT TCAGAAATGG CCTCAAGTTA GTGGTTGCT CAGGGGTGAA
1201 GACCAAAGCA AAGTTCAAGGG CCTCATCCCA GGGTGTCTCA CTTGGCATGA GGGACGAGGA
1261 CCCCCATTC CTCTCAGCTG AGGGGAAGAG CTCTCCACAA TGTCCTCCCTG CACGGTCCCTC
1321 TGGCTACCC TACAACAAGG GCCAGCTCTC CCTACTCTC CTGGAGTAAA GCTGGGCTCA
1381 GGAGGTGCTA CCGTTCAAA ATGTGTTCAAG CTTTAATTT TAAAATATCT TCAAGAGTTT
1441 CTCCATCAAA ACCTCTGTAT CCTAATAGC CTAGAAATGT TCGTGTGAT CAATTCTTAC
1501 ACATTCCACA CAAATCTGAG AAATGCTGGT AAAAGTTACT GTGAGGGATG CCCACTCTGT
1561 GAAACTCCAG CCTCTATGGA GTGGATGCGAG AGCGTGTGGA CTCCATATAC CCATGGGAAT
1621 GTGGAATTTA GTGTTTACAG CCCCTCTGGG ATCTCAAGGG TGAAGACATT CAGCAGCGAG
1681 GCTGGAATTC AAAGCCAGAT CACACTCTG CCTCCACTCA CGTCCAGTTT GATTTTACG
1741 CCGCTGGGCC TGGATGGGGC TATGCTGGA GGCTGTATG ACTGAGGGAG GAGAGACTGC
1801 TGGGGAGGGG GGGCTTGGG AGAACAGCAG CCAGTGTGCG CCAGCTCCCC GCAGCACACA
1861 CAGTTCGCAA TTCACGAAAA TGTGAGGTGC GGGAGTGGGT TTGAAGCTTG ACAGGCACCTT
1921 TGCTCCAGAA TCAGTGTCAA TAGCAACGGT GGGCTGGCCA AGACGGCTA CAAAAGGTCA
1981 CACACACTGA TTAAAGGGAT GAACAGACG AGAACTGCG AGTGTGCTT TTGGGAAAGA
2041 GCTCGGTGTG GACCTCCACC TTGGGTGGCT CGCGCAACA CGGCCCTTC TCGGCACCTT
2101 CACCAAGCCT GCCCAGGAGG AGGCCCTCA GATTCCGGGC CAGGGAGGTG AGCATACGGA
2161 TATCGTAGTT AAGCAGTGTAG TCCCCCAGCT GGACAATCAC TCACACAGCA CAGGTGTTGA
2221 AGTCAGACAT CACTGTCGAG GAGTCCTGA ATTACTTCT GATGTTTCAT TCCTGTACAA
2281 GTACAGCAAA TCCCACACT TTTGAAATAA CAGGGGTTTC AAGGAAAAAT ATAATGAAA
2341 AAATTGTTTA AAGAAGAGCA GATGTTTCTC ACATTCAAGT TTATATTGCA ATCGAAGCCA
2401 GAGTTTCTA AGTAGCGACC CTACTGGACG GACCTTAAAC AATAAAATCAT TTTTAAAGTG
2461 TGCTTCAAGT GTGAAATAAT TTGAAGTCAA AGGTGTCCCC CTTTGACTGT GACCTTGTA
2521 AATCAATCTC TGGATCTCG ACACAGAGGA AGAACTAGAA TTAGGAAGAA TTACTGTATT
2581 ATAAAGCTTG CTTGCATTAT CAATTGTTT GGTGATACTT TCACAGATAA TACAAAGTTA

2641 ACATTCAGTG TTCAATTCA CCATCAAAGA AATACCAGTA AAGACAGAGC ACATTAACAA
 2701 ATTTTGAAA ACTCGGAATC ATCAATGCTG GCAAGGGTCC TGTGAGATAA ACACCCCTAT
 2761 ATATGTTGGT GGGGAATAA GTTGGTCTAA CTTCACTAAA AAGCAATTG ACACCATGTG
 2821 TTAGAACCT TTAAAATGGG GGTCCTGCTG GGAGCGGTGG CTCGTGCCTG TAATCCAGC
 2881 ATTTGGGAG GCCCGCAGTGG GAGGATCGCT TGGCCTAGGA GTTCGAGACT AGCCTGGCA
 2941 ACATGGAGAA ACCCCGTC TACAAAACAT AAAAAAATTA GCCAGGTGTG GTGGCGTGCA
 3001 TCTGTAGTCC CAGCTACTTA AGAGGCTGGG GTGGGATGGT CGCTTGAGCC CAGGGAGGTT
 3061 GAAGCTCGAG TGAGCTGTGA CCATGTCAGT GCACCTCCAGC CTAGGTGACA GAGTGAGACC
 3121 CTGTCTCAA ATAATAAAAG AAAAATAAA CTCAAATGGG GATTTGCGG AGTTGACCTC
 3181 TAACAGCTCG TGCTGGT ATTGAGGGC CTGTTTCTC CTCTCTCTTG GGCTGTCATT
 3241 TCCTGCAGG CATGGTCTAT ATACTCACTA CTTTAGAAC CTGGAGTCTG GTAGGAATGA
 3301 TTTGACTCG CCTCTTATCT TTACAACCGA CAAGACTGAG GTTGTGCA GGAACATGGC
 3361 TTGTTGCTGG TCAAACTCAC AGGGTGGCCA TGCAAGGATCC AGAACCTGCA GCCCCTTCT
 3421 CCTGGCCCCA GACTCTGCT CTACACCAAG GCACAGCACC ATCACCCCTGG GTCCTACAC
 3481 TGCCCCCAGC TCCCTGGTC TCTAAGCAGG ACCTGGCAGA TCCCTGACTT TACATATGAC
 3541 ACTCTAAGCT GATTGCCCTG AAGAAGATT TTCACTGAGG GACTGCAGAA ACTCAGATTC
 3601 ATCTGTTACA CTGCTGGG AACAAGTACA TGTCAGTGC TGAGGGAAAC TCGGAACAAAC
 3661 CAAGACACCA AGTATCTGAC ATGGAAAAAG GGAGAAATGAA GTATGGGATC CACACACAGC
 3721 AAGCTGAAAG GGAAGAGTCT GAGCTGATGC TGAGGTGCAA TGTGCATAAA TGGGTCAATC
 3781 ACACAGCCAC ACGGGATAGA CAGCCTGATC AACTGTGTTA GACATTCTC ACGGTGGGAG
 3841 TGCCCTCAGAG GCTTCCCAA AGCTACGAGT TTCAGGGACT CTAGTATTAT CTGCAATAAC
 3901 AACAACTCAA AGGAACAAGA GGACACAGGA AGACCAACCGT GATTCAAGAGC TGATTAGCCC
 3961 CTCCACAGGC ATCCCTGCT TTGGAGACTT GATTTTGTCT CTTGGCTCCA CTTGGCTCCC
 4021 CCTTCCCTGT GTTCTCGGAG AGGCCTGCCA ATTCCAGCAA TGATTGCTCC AAGGTTCTCG
 4081 AGCCAAGCCA GCCACATAAA AAGGTATGGA CATTGGATT CCAAGCAGCT GGGCGGAAGG
 4141 GATGGTGGGG TGGATGGGAG ACCTGACATC CCCTTGTCT GGTCCGATGA GGTCCCCAGT
 4201 TCAGACTCAT CCAACTGGCT CTGGGTACCC CGGGTTGCCA GCATTCAAGCC CAAGCTCAC
 4261 GAGCTACTTT GATGCAAGAG TATATATGGA TGGTGTCTC TTTTTTCAT TGCTTAGGCT
 4321 AGAATTTTT TTTTCACTTC ATAGAAGTGA ATGTCAATT TGACAAGGGG AAGAAAAAAA
 4381 ACTCTCAAGA TATTGAAAAC AGAGATACTC TCAATATGTC CAAGCACCAT CTTCTCAACG
 4441 ATGAGCTAA GAATGAATAA AAACCAAACA AATCTAAGTG GCACATTATT CAAACAAATG
 4501 CTGGAGTAGT GGGCAAGCA TTACACTTTA AATAAATATT CCTAACCTT GAGGACTCGT
 4561 TATGTGCTAA AGTCTTACA TGAATAGTCT CATTTCATCC TAATCACCTT ATGGGCAGGT
 4621 ACTATTCTC TCCCTATTTC ACAGATGAGG AAACATGGC TCAAATGGG TAAGCGGTTT
 4681 ATTAAGATGTA AGGATTCTT ACTACTCTTA TGAGGACACC ACCCACCTA AATCAGAGCC
 4741 TGAGAGGCCG ACGGGTTCCA CTGGGAATGTT TACAGGAAT GTGTCCTGTG AATCATAACA
 4801 AATACACAAA TGTAAGATT TAAAATAAT ACGGCACATA TATTTTGAAT GTTGTATGAG
 4861 CATTATTTC AAATTATCAA GCCAAATCTA AAGAAATGGT TTGTGTGGC TGGGCGTGGT
 4921 GGCTCACGCC TGTAACTCCC GAACCTTGGG AGGCCAGGT GGGCGATCA CGAGGTCAAGG
 4981 AGATCGAGAC CATCTGGT AACACGGTGA AACCCCGTCT CTACTAAAAA TACAAAATA
 5041 AAATTAGGCC GGCATGGTGG CGGGCGCTG TAGTCCCAGC TACTCGGGAG GCTGAGGCAG
 5101 GAGAAATGGCG TGAAACCGGG AGGCTGAGCT TGCACTGAGC CGAGATCGT CCACTGCACT
 5161 CCAGCCTGGG TGACAGAGCA AGACTTGTG TCAAAAAAAA AAAAAAAA AAAAAAAA
 5221 GAAATGGTTT GTGTGATGAT AAACGAGATC TAAATATGTT TTCGTGCATA CCTTGGCATC
 5281 TGTGCATGGA GTTTGGGGT TCAATTCTT GCTATTGCTC GAAACATTAA TAAAATAAGT
 5341 TTTGAAAGT TTCACTGTAAT ATATTTTCTC TCAGAAACT TACTTGATCC ACTATTTCAC
 5401 TAGGAGAGGG ATGAAAAGGA AGAGTTACTA CTTCTACTCT GATTGATTT GAGCCTGGAA
 5461 AGATGTTTA AGAGCACTT CTGGCCGGG ACGGTGGCTC ACACCTGAA TCCCAGCACT
 5521 TTGAGAGGCT GAGGTGGATG GATCATGTGA GGTCAGGAGT TCAAGGCCAC CCTGTGGCAA
 5581 ACACGGTGAAC ACCCGCTCTC TACTAAAAGC ACAAATTA GGCGGGCGTG GTGGTGGCA
 5641 CCTGTAATCC CAGCTACTCG GGAGGATGAG GCAGGAGAAAT TGCTGAACC TGGGAGGCGG
 5701 AGGTTGCAGT GAGCCAAAT TGACCAACTG CATTCCAGAC TGGGTGAGAC TCCATCTCAA
 5761 AAAAAAAA AAAAAAGGAG TTTCTATTGT TAAATCTACT CTTGTGTTT ACACCAGGAA
 5821 AAGTAAGTGG GAGAATCAAG CATTATTAC TTCCAAAGTC TTACAAGATA GGCCAAGATA
 5881 TATAAATATT TCCTGACAAA ACAGTATGAC CTGCACTACT ACTGCTCTA GATACTAAC
 5941 AATGACAGGT TACTCTGAG CGCTACTGTA CAGTTATAA TGCAATTAC ACATTAGCT
 6001 TATTTCATCC TCACTGCCAT GGAGCAGGT ATTTTTATTC CTAGTTATA GAGGAGAAAA
 6061 CTGCTTATGC CAAGTAAAT AGCTTACAGG AGACCACATT AATACTGGGT AATGACTACT
 6121 TTTAACATTT TGATGTTATT CCATCCCCTC ATTTTTCACT GATTTGAAAT TTCCCTTATA
 6181 AAATTGAGAT AAAAGTTAAA CAATTTGCA TCTTGATTT TCACCAATTG TCTTCCCTCA
 6241 TACCATTAAT AATTCTTGTG AACATGATT TTAATGGCT GCACACTATT CTAAGTGTGAA
 6301 TATGTATCAA TTATTTTGT CATTCCTGT TAGGCATTAA AATAGCTTTC TATTTTTG
 6361 TGTATAAAT AATGCTGAGA TAAACATCCC TGAACAAACA TCTCATCTAC TTAAAAAAGT
 6421 CTCCCTTGAA TTACTGAAAA ACTAAGTTAG GCTGATCTTC ACTTACATT TTTTTTTT
 6481 TTTTTTGAA GATGGAGTCT TGCTTGTCA CCCAGGCTGG AGTCAGTGG CATGATCTCA
 6541 GCTCACTGCA ACCTCTGCCT CTTGGGTCA AGTGAATCTC CTGCTCAGC CTCCCAAGTA
 6601 GCTGGGATTA CAGGGTCCG CCACACACTG GGCTAATTGTTA TGTATTTTA GTAGAGACGG
 6661 GTTTTACCA TGTTGTCCAG GCTGGTCTCA AACTCCTGAC CTCAGGTGAT CTGCCGCCT
 6721 CAGCCTTCCA AAATGCTGGG ATTAGAGGTG TGAGCCACCG CGCAGCCTC ACTTACATT
 6781 TAAATGTAAGAA AACTAAGTT TATTTCTT CTTCTTTAG TTAAATCATT GACATGTTAT
 6841 CACCAGGACT GTTTGTTCAAG ATGATGTCAC CTTTGAAACT GCTGTGCTA CTAAGGAAAT

6901 GGACATTAGG TCATGAAACC AATAAACTGG AATTGCTGG GTAAAACCTT TCGTATCTCA
 6961 GGCTATTTC CCAGCTATGG ATGGGATTTA TACTCCTCAG GGCTGAGATG TGCTGAAAG
 7021 CACACTGCTG TCTACCCACT GCCATTACCT GTAAGTTGCT CCACTCCACC TTCCATGCTG
 7081 TTTTGGAGG AGGGTAAAG AAAGGGAGGG GAATTCACAT AGAATCAAAT CAGTTCTTA
 7141 CAGTTCTATT TGTTCTCCA GTCTTCAGT CATGCAATGA CTATTATTG AGTGTGTTTG
 7201 ATGTTGTAAC TTGTAATTAA GGAGTATGAC AACAGACTTT AGGTTGAAA CTCTTGATC
 7261 TGGGAGTACT GTAAATCGTC ATGAGGAGGT TATGGAACAC AGCATATCCC AAAGTTACTT
 7321 GACTATGGAA CCCTTTTATT TTCTTGAAC ACCTACGAAT GATCTCCAC AACTGCATTC
 7381 CACAGAACAT AGGTTAAAGG AAAAGGCTAA AAGGCATCTA ACAAACTCCA TATCCATT
 7441 GTTTTTCT CAAATAGAA TCATTTAAAT GAAATCCTT TAAAACTAGT AATAGATAAT
 7501 TTCTTAATGC AATTAATTAA TAGATAAGAA GCCACCAGCT GACATAATCT CTTAGGGTGA
 7561 AATAACATAG GCAATACCAT TAAGATTAGA AGCTTGATAT GAATGCCTAC TTCCGTTAT
 7621 TATGAAATAT GGAACAGAAG GGGAGACAAA GTAATCATGA AGGTAACCTGA TACAGTTATC
 7681 TATCTAGAAA ACCTAATCTA CTGACAAGCT GTCAGAATAA CTTCTTGAGT CAAATGACCA
 7741 GATATAAAAC AAAATGTAC AACTCAGTAG CTTCCCTTACA TATCAGCAAT ACCCTGCCA
 7801 AAAAAAAATAG AGAAAATAAA AGCCCCATTCA TAAGAGCAAG AGGGCATAAA ATACCTAATA
 7861 CATTATAAG AGCTCCATAG GACCTGTATG AGTAAATCAA TCACTTTCC AAGAAGGTCT
 7921 GCGTTAAAGC AATGGTGAGC TGCTGCTTGG GAAAGAGCAA TACAAAATG TAACCTAGAG
 7981 ACAGAAATGCA GTTCCAGGAC AAATGGATCC CAGTGGGATT TTTTTTGTG TGTTTGAAG
 8041 AACCTAATCA AGAAAAGTCT TCTCAAAGAA TGACAGAAAA TGAGAATTGA TAGCATTG
 8101 AAAATAAAGA GTAATGCACG CAGAAGAGTC TTGAGCTAAA ATTATTAAAA AGTATTCTAA
 8161 AGTGACAATA TTTAAGATGG TATGGTGTG GCATAAGAAAT TTACAAACAG AATAACACGA
 8221 GACAAAGCTC CTAACAGACT CAATATAGTA TTTAATAACT TAGCGTATGA GGAAGCCTCA
 8281 TGTGACAGCA GAAAAAGGTA GATTTAATAA TAGTGTAGA AAATGACAAA AGACAAAAAA
 8341 AAAAAAAAC CCAACAAAT AATATTGAGA CAGATTTAC TTCCAGGGTGG GTTAAAAAGT
 8401 TAAATATTAA AAAAGTAATC CTATTTTAAAGA AAGAGCGCCC GGCGCCCGT CTGGGAGGTG GGGGGCGCCT
 8461 AGGGAGGCGG GGGCAGCCC CAACCCGGCC GCCGCCCGT CTGGGAGGTG GGGGGCGCCT
 8521 CTGCCAGCC ACCCTGTCTG GGAAGTGAAGG AGCCCTCTG CCCGGCCGCC ACCCTGTCTG
 8581 GGAGGTGTAC CCAACAGCTC ATTGAGAACG GGCCATGATG AGGATGGCGG TTTGTCGAA
 8641 TAGAAAAGGG GGAATGTTGG GGGAAAGAAA GAGAGATCAG ATTGTTACTG TGCTGTTG
 8701 GAAAGAGTA GACATAGGAG ACTGCATTTT GTTCTGTACT AAGAAAAATT CTTCTGCCTT
 8761 GGGATGCTGT TAATCTATGA CTTACCCCCC AACCCCGTGC TCTCTGAAAC ATGTGCTGTG
 8821 TCCACTCAGG GTTAAATGGGA ATAGGGCAG TGCAAGATG GCTTTGTTAA ACAGATGCCT
 8881 GAAGGCAGCA TACTCCTTAA GAGTCATCAC CACTCCCTAA TCTCAAGTAC CCAGGGACAC
 8941 AAACACTGCG GAAGGCCGCA GGGTCTCTG CCTAGGAAAA CCAGAGACCC TTGTCACAT
 9001 GTTTATCTGC TGACCTTCCC TCCACTATTG TCCTATGACC CTGCCAAATC CCCCCCTCGG
 9061 AGAAACACCC AAGAATGATC AATAAAATACT AAAAAAAA AAAAAAAAGA GGCAGGGCAC
 9121 GGTGGCTCAC GCCTGTAAATC CCAGCCCTTT GGGAGGCGGA GACAGGAGA TCACTTGAGG
 9181 TCAGGTGTTG GAGCCAGGCC TGTTAAACAT GGCAAAACT TGTCTCTACT AAAAAATACA
 9241 AAAATTAGCC AGGCATGGTG GCCCATGGCT GTAAATCCCAG CTACTTGGGA GGTTGAGGCA
 9301 CAAGAACCCG TTGAATGGGA AGGTGGAGGT TGCAAGTGC CTACTGCACT
 9361 CCAGTGGGG TGAGCAGAGT GAGAGTCTGT TTCAAAAAAA AAAAAAAA GAATATGAGT
 9421 GAACATTCC TGAGTCTTAA GATAAAACAG GATTTCTAA CCAAAAAATA AAAGATATTA
 9481 TAAAAAAAGGA AAAATTGTTA GATTGATCA CAAATATTAA ATATATCTGC ATGGAAAAAC
 9541 TATTACAAAAA ATTACAAAGGC AAATTTCAAA CTGGAAAAAC ACTTTTACAA CTGTGACAAA
 9601 GGATTGATAT CTGTTACTGA GAAAACATT TACATTTTATT TATAAGAAC AGACCAGGGC
 9661 CAGGCTCGGT GGCTCATGCC TATAATCCCA GCATTTGGG AGGCCGAGGC AGGTGGATCA
 9721 TGAAGTCAGG AGATCGAGAC CATCTGGCT AATACAGTC AACCCCGTCT CTTCTAAAAA
 9781 TGCAAAACATT TAGCCGGAG TGTTGGTGGG CGCCTGTAGT CCCAGCTACT CCAGAGGCTG
 9841 AGGCAGGAGA ATGGCGTGA CCTGGGAGGC AGAGCTTGCA GTGAGCTGAG ATCGCGCAC
 9901 TGCACCTCCAG CCTGGGTGAC AGAGTGAAGAC TCCATCTCAA AAAAAAAA AAAAAAAA
 9961 AAGGACAAGA CCAACACTTC AAATGAAAAT TGGACAAAGT TCACAGACAA TTCAGACAA
 10021 TCACAAAGTA ATATGGAAGA CAACTATATT TTATAAGTTC AACGTCACTA GTAAGAAAATA
 10081 AACAAATTA AATAAGTCT TTCTGGCTGA GTGCCCATG GCTCACCCCT ATAATCCAG
 10141 CACTTTGGAC TTTGGGAGGC CAAGCGTGA GGATCGCTT AGCCCGAGGAG TTTGAGACCA
 10201 GCCTAGACAG CAACAGGAGA TCCCAACTCT ACAAATTTG AATAAAATTGG CTGGGTG
 10261 TGTTGCAATG TACCTGTAGT CCCAGCTACT CCAGAGGCTG AGGTGGAGG ATTGTTGAG
 10321 CCCAGGAGTT TGAGGCTGCA GTGAGCTATG ATGGCATCAC TGTACTCCAG CCTGGGCAAC
 10381 AGAGCAAGAC CCTATCTAAA ACACTTTTT TTCCCTATCAA ATTAATTAA ATTAGTTTA
 10441 CTGGCAGTTT GGTGAATTAC TGATCCTCAT ATTTCTGGTG GGAGTCTAAT TTGGAATAAC
 10501 TGTCTAAAAA AAGAAAGAT AATACATAGA TTTAGGAATA TATATAAGCC TGAAAAAAA
 10561 AATTATATAA TATATATTAA GAGACAAGGC CTCACACTG TGCCCCAGACT GGAGTTCACT
 10621 GGTGCAATCT CAGCTTACTG CAGCCTCGAC CACCTGGCT CAAGTGATCC TCCCACCTCA
 10681 GCCTCCAGG TAACGGGAGC TACAGGCATG CACCAACATG TCCGGCTAAT TTTGTATT
 10741 TTTTTAGAG ATGGGGCTCA CCATGTTGCC CAGGCTGATC TTGAACCTT GGAGTCAGT
 10801 GATTGTCCTA CCTTGGCCTC CCAAAGTGC GGGATTACAG GTGTGAGCTG CTGGCCCG
 10861 CCTATTATCT ATATTTTTG AGGCACTAAT TTATGCTTAG GCTATTCTAA GAAAAAAACT
 10921 AGAAATACAA AGATTATGG CAAAAAATGT TCAGTTCTATT GCAACATTAT GTATAACAGC
 10981 ATAAAAAAACAG AAACAAACCTA AATGTCCAAC CTGTAATCTA GATTATGATG TAGTCATTGA
 11041 AGTTATTTC GAAGACTATA TAATGACATA AGAAAATGCT CATCGTATCA TATTAAGCAA
 11101 AAAAAAACCC CATTATTTAA ATTATATTAA CACAACTCTC AGTTTGTAAC GCATGAATTC

11161 TATGTCTTC A TCTAAGTCAC TGAGAGAAAT GTGAACAGAT GGGGTTGGA ACGGATGGAA
 11221 GGTCAGCAA AGTCAGCCTC CCTTTCTGG GTTGAGGCTA GCCCTCTCC ATTCTCCAT
 11281 TAACTGAACC TTCATCTGGC AATGGTCCC AAATGTTTC TAGGGGAGGG CACATAAGAA
 11341 TTACTGTGAG AGATGGTTA GAAAAACAAA AACAGATTG GGAGCCTGAC CCACAGAGAT
 11401 TTGCACTCAG GAGGCTGAG GTGGGGCCCT GGAATCCGTG TCAAAAGGCC TCCAAGTGT
 11461 TCTGATGCAC AGCCACGTGC AGAAGGAACC ACGGCCTTCT GGTCGAGGT TTTGAAAAGG
 11521 ACATTGAACA CCTTTGTAA GTCACTTCT GAATGTGAGT TAGGAAAGAT AATGTCAAGGA
 11581 TGGAAAAGC CTTCCTCC TCACATTTCT CCAAATACTC TGTTGCTGCT TGTCAGGTGG
 11641 ATGATGTGAC AACTGGCCA GCTGCTGGCC TGAGCCTGCT CCCCTGCTCG ACTGAGCTG
 11701 CCAGCGCTC CCTGCCACC CCCCCCCCCA CCTACTCAC TCCCTGCTAC CCCAACACC
 11761 CCACCCAGCC CCGCTACCC TGCACTCCT TGCCCTGCTC TTCAGGACTT TTGCTCCAAG
 11821 TCTTGCTCC TCTTTGTCT CTGAGATGCC AGCTGCTCCC TCTCCATGGC TTCCCTCTG
 11881 TCCACTTTTC ATTCTGCTGA GTTGTGCTT GTCTTACCC AAACCCCTGT GATGCTTGTG
 11941 GCCAGGCTC CTTCCCCCAG GCCTAGCTGC TGAGAATTG ACGCCTCAC TCCCACTCAT
 12001 TCCTCAACCC GCGCACATCTG GCTTCCTTA CAACCACTCA CAGGCACACC TCTTGGCAAG
 12061 GTCACCTGGC TCCCAAAGGC TAATCTACTG GTTCTGTTTC AGTCTTCTTC CTGGAAGCAT
 12121 CTGACACTC TGACCAACAC CTCCTCTCC GGCTTGTG GGGGTCACTG TTGCTTACT
 12181 TGCTCGTAT CCCCTTTCT TGTTAGCAGC ATCCTCAGGA GCTGCTTCC CCAGCTCTCA
 12241 GCCCACGTGG TTGGTGGGGC TGACCTCCCC ACTCCCCACA TTTCATGCC CTGCCCCCTAA
 12301 GGATTGGTTC AGGGATGGGC ATATGGTCAG GGTCAAACCA ATGAGACTCA AGCCTGAGAC
 12361 TTTTACTGGG GCTGCTGAGC CAGCAGAATT CGAGGCTGGA AATGGAGGTG GCCGCTGCC
 12421 CACCCCTGAG AAGGAACACA ACATGAGGAG GGCCACAGAG CCAACAGAGT CCTGGTGTGT
 12481 CTCCCTGAGT CCTGGGCCA GCTGAAGCT TGCACTCCT GGTCTGATGG GTTTCAGGAA
 12541 AACGTTCTCA TCTTTTATC TAAGCTAGCT TTGGTTGGAT TTCTGTCACT TGCAACCTAA
 12601 AGAGATCTGA GACAGCACTG AGCCACTACC CTCCAGGCA ATTACCCACC TCTTTGGTAC
 12661 TTTCCCGGA GGGCTCCAC AACGCTCTGCT TCCCTCTACT TAAGACTCCT GTGCAAGGAC
 12721 TTCTACTTCC CAGATGCTGG CTACTAAATG CTGACACTCC CAAGTCACAA GACCAGCCC
 12781 GATATCTCCC TGAGCTCCAG ACCCATGTCT TACCCCTCAT GGACATGCC ACCTTGGTGT
 12841 CCACTAACAC ATGACAGGTC TAAATCTAAA CTTATCATTA TGCCCCATCC TTCTCTTCT
 12901 TGGATCCCA TCTCAGTGA CAGCTGACT CCCAGTCTAC CCCAGAAACC GAGGCTCAT
 12961 GGGTGACACC TCCCCACCC CACTCCCTGG CCCTCTTTGG GCAGCTAATT CTCTGCTCAT
 13021 TCCTCAGGGC TCAGCTTTA GGCTGCATTC TCAGGGAGGC AGCCTTTCC AACCCGCAAG
 13081 CCTGAATGAG GGTTCCCTC CTATGCCAGG AGCCACCTGT TGTTCTAGT GTTACAGCCC
 13141 TTGTCACTCA GAGCTGTGGC CTGTTGACTC TCTCTGTGA TGCTGGCC CAACCCGTAA
 13201 ATCAAGATCA TTTGAATAA TGAACACTAG CATCTGTAGG TGTGATCTTA TATTGTATAT
 13261 TGAAAATATA GTCTAATTT CTGTAAAAAT TAAATATCTC AATATTTAG AGAGATCTGT
 13321 AAAGGGTATA ATATTGGAA CAGAAGATTG ATTTTCAGGA GCCCCAGTAT CTGGATGAGT
 13381 GTGAGAACCA CTAAATAGGA CCCATAGTT TTCTGTGTGA GATTTTATTT TTTTAAAAAT
 13441 AGATGAACCA GGCTGCTTAT TAGTGAATGC CTTCACTTCC CTTTCCCTTA ATTACTGAGT
 13501 CTAGAAACTC ATCAAGGGAA TGAAGCAGGC CTTTGTCTT ATTCAAGAATC CATTCCCTG
 13561 CTTTTTGAA AATGAGAAC TCTGCTGCT TCCAGTGTTC CAATACTTT CATATTCTC
 13621 ACAATTCTT AAAAGCCACC AAATACAGTT CTTTGAGCAG AAATGCAAGT TCTTCATCC
 13681 ATGTGTCCGG CCAGAGATT CCGAACCTC GCCTGTGGCA GGCCCTGGGA GAGCTCTCA
 13741 GCCTGTGGTT TTAATCCCT CTTGAACACA GCTGGGCTCC CTTCTGCCAT CTCTGGGCA
 13801 TTTTCTTGA GGAAAAGGAG GAAGCTGAGT AGTTCTGCTT TCTGTTAATA TTGAGCAAG
 13861 AGCCCATGAG TGGGCGGATC ACTTCCTGAT TCTTGTCTT GTTCTAATCA TACCTAAAAA
 13921 GGCTTTTTA ATGTCCAAGG CACTTTTACCT TGCACTGCCC ACTCCAGCTT CTTGATCCCT
 13981 CATAACGCCCTTGAACAC CTTGCACCC TGACATGCA CCCCTGCCCTT TTTTTCCCC
 14041 TATTAAATAC TATAGTTAA TGTTAATATC TAAAGTTTTT AAACACAATT ACGGTATCAA
 14101 GCATTTTATT AAAATTTTT ATCATACTTT GATTTTACCA ACTTGAGTGT TATGAACAAT
 14161 GAACATGTTT ATACATTGAC AAGAATACTC ATACAGGGAT TGTGGGATGG CTACTTCAC
 14221 TCCCTTTCC CTCATTTTA CTCTTTTTT TTTTTTTTT GTAGAGATGG GGTCTCACTA
 14281 TGTTGCCAG GGTGGCTTG AACTCCGTAG CTCAAACGAT CCACCTACCT CAGTCCTCCA
 14341 AAGTGTAGG ATTACAGGTG TGAGCCACTG CACCTGGGCC CTCACTCTTA CTCTTAGTT
 14401 CCTTTGCCCT TGGGTCTTC TAGTTTTT CTACTTTGT TTCTTTCTGG CATCCATTTC
 14461 TATTAAATAT TATTACACAT AGATGAGAA ATTACTTGGC TGCGCTAGT TGCAGAATCA
 14521 TGAAAATTAA ATCACATGAT GCTTAGTGT GAAAGACTGA ATTGTTTGG TTTTTTATTG
 14581 AGATATAATT CATATACCAAG AAAATTATT CTTTACAGG GTGCACTCCT TTAGTATATT
 14641 CACAAAGTT TGCAACCAAT ATCATTATCC AATTTCAGAA CATTTCATC ACATCAAAGA
 14701 GAAACTCCAC GCCCACTAAC AGTCATTCTC TATTTCTGTT GCATGCGGCT ATCCAGTTGT
 14761 CCTAGCACCA TTGTTGAACT TATAGTTCC CCCATTGTTG TAATAGTTT TAAGTGGTTT
 14821 CTTAGGATTG TCTACATAGA AGATCATGTC ATCTGCAAAG AGTTTACTT TTTCTCTCC
 14881 AGTCTGGAG TTTATTCTT TTCTTGCCTT ATTCCCTGCT TCTCTACCC TTTAAAGCCT
 14941 ATATAGTTTC TTGAGTGTACT CCTCTCTTTT TTCTCTATTA GAATAAGTT TCTGGGGAGA
 15001 GGGAACAGAG CATATGAAGG GAGAAAGTGG AGCTCTAAA AATGTGTAGT TTGGGACGGG
 15061 GCGCGGTGGC TCACGCCGT AATCCCAGCA CTTTGGGAGG CCAAGGCCGG CAAATCACCT
 15121 GAGGTCGGGA GTTGAGACC AGCTGACCA ACATGGAAAA ACCCCATCTC TACTAAAAT
 15181 ACAAAATTAG CCAGGGCTGG TGGCACATGC CTGTAATCCC AGCTACTCGG GAGGCTGAGA
 15241 CAGAAGAATT GCTTGAACCC GGGAGGGGG GGTGCGGGTG AGCCGAGATC ATGCCATGGC
 15301 ACTCCAGCTT GGGCAACAC AGCGAAACTG TCTCAAAAAA AAAAAAATTG TGTAGTTGC
 15361 ATTCAATTCA AATGGACTTC ATACTTTTC AGTGGAAATG TCTTCTCTCA CAGCACCTGG

15421 GGCCAGGGAT TTCTGGAGG AATCTTGAA ATCTGTTTC CTCAAGTTT GGACACAAGC
 15481 CTGGGTGAGC TCCATATCCA TCCTCCTCG GCTAGTACAG AGTCTAAGAT GACATGGTCA
 15541 CTTTCCCCCA TGCCCATCTC CCCTCCTTCT CTGGCCAATT CTTTATTGCC AGTCAAAGGC
 15601 AGAAGAGCCA ACAGCAATAA GAGAAGATTG CCACCATGTC AGCACAGGTT CTGTGTCATC
 15661 AGTGACTCTCT TGAATGTGC AGGAAGCCAG GGAGATAAGA ATCTATGGCT ATGTAGATTA
 15721 GGTACACCCC ACAGCCAAGT ATGAGCTTCT AGGGCCCAAT CGCTAGAACG AAACAACCCA
 15781 GACCCTAAC CAGGCTCTAA CTAAGGCAGG ATCCAGGATG CTCAGCACCC TGGGCAACC
 15841 CTTCCGCTCA AGATGAATCC ATGTTTCACT TACAGTGTCT GGGGTGATAT TTAACTCCAA
 15901 GTTCACAAAGT CTCCAGGGT CTGCCCCCTT TGTTGGGGA TTCCCTCCCG GGTGCTAAAT
 15961 CACTCTCTCC TAGTCTCAAT CTGAAAAGCA ATTCAAGGTT CCTGTGGCCT CACCAAGCTT
 16021 CCTGCCACCA CTGATTTCTC TCTAACCTCA CTGGTCTACC TTGTGTACAT CACCTAACAA
 16081 CAAAGCCACT TAGCTCAAGG AAAGCCTCAA GAACAAATGG CTTCTAATA AGGTGATTCC
 16141 TCCCTCACCC TAGGGTGGCC CACAGCTCCA GCACCTGCTG AGAGGTGGCC AATTGGATG
 16201 AGGTAGCTGC CAGGAATCAA ATCTCAGAAC CTTCTATAAA ACTGCGCTTA GCTATTCCAC
 16261 AAATTTCTCC AAATAATGAA ACAGAGAGAA GAGTGAAGAT ATAATGAGAG AGAAATCTGG
 16321 AAAATACAA AAGATAAATT CATGTTCTG TGAGTTGTAT TCATCTCCT GGGAGTTGAT
 16381 AACAGGCTG GCTAGCTATC CAGTCTATGG GAGCAGAGAG AGGGCTGCAT TATCTCCAGG
 16441 TTGGCACAGC ATCTAGCAGG GGCTCCCTGA TGACACCTGG GCACAAACAAG GACCTCACAG
 16501 CAAAGACAAA CAGGAGGAAA TGCCCTAGGGT GTGGGACAT GCCACCACTA AATTACAAAA
 16561 ATATATGCCT TTCTCTTTG TATGTTCAA CAGTACCGCT TCCCACCTT GTACCACTTG
 16621 GCCATAATAC TTTATTATCC CTAGTACCTA AAAACTAGAC AGTATTACTG AGACCTAAA
 16681 TGTGACATTG GTCATCTTTC AAGGTGCCAG TTACTGTAAT ACAGCTATGG TTCATTGTA
 16741 GTGTCGGGT GACGGTACA TATTTAATT TATTTTGTC TTGTCATAA GCACCAAATA
 16801 AAGCAATTTC GAAATTACTT TCTTCACACT CAATTATTA TCTAGAACGA CATTCAATA
 16861 CACAGTATAT TTGTTTGTG GAAATTCTTC GCAAGTGTAC GCAGTGTATC TGGAGGCCG GATTTAACAT
 16921 ATTTCTGTG TGATAGCTA CTCTATTGTT AGGAACATTAG TAGTCACTG GAGTGGAGAA
 16981 AAAGAGGTG CTAACAAACA GAAATGGCAT GTATCCAAAT ATAGATGCAG TTTTCTCAAC
 17041 ACCTGTTGCA CATCAAATGA TGAAGAGATT GTTACAATT CCAAAACCTT GGACAAAAC
 17101 CTTAAATGAT CAGCACTTTG TTCAACTG TACACGGCTT AAGTCAGAAA AATAAGTTCT
 17161 TCTTACAAA AACAGTGAGC CTAATATGC TGGCTGTTT AGAAATAAAA CAGGTGATCA
 17221 ACAAAAGCAAT CTGTTTCTGT CTAATAATT TCCCAGCAA AGTCACCTCT TTTGAAACTA
 17281 TATGATTGAA ACCACAGGGT ATGCACAGTG ATTGTGTGCT TTCAGAAACA CGGTAAAGCAT
 17341 GACAATATT TTCAAGTGTG AAGGGTAAGC GAGGAGGGAA TGACCCACCT AAGAAATGGA
 17401 AACTAACAA TAGTAAATATG ACACTCCCCC TTAGTTCCAC AGGCTGGTGC GGCTCTACA
 17461 CGCAGCTCAC TTCTACGGG CGGCTAAAGG CGGTGCATGA ACTGAACGTA GGCGCCCTCT
 17521 TTATCCCCCTC CCCACAACAC TGTGGGGTGG CACGTAAGA CCTCAAGGGG AGGGGGCCTG
 17581 GCGGATACGC TCTCTGCTG CCATTCAATT GTTCAAGAAA TAGTGAACAA AATGAATAGT
 17641 TAAACAAACT CTCCAACCTAG GACTCTGTCC CATTGGATA CCTACGGCCT CTCCCCCAT
 17701 TTCCATCCCT AGGCATCACT CCTCTGTCAA GAATATGCTG ATTATGGCTC GGCCCCCCTC
 17761 TCTCCCTGG ACCCTCCCGC CAGGGCGACC GGACGAGGGC AGCAGGGAGCA ACGCCCTCCG
 17821 AACCGCAAGG AAGGAGGGCT CGGAGTGC GAGGAGGGAA TGACCCACCT AAGAAATGGA
 17881 GCGGGGCTGG CCAGCCGGGG CGGGGCCGAG GGACCGGGC GCGGAGCTCC GGGAGGGGTC
 17941 CCCGAGGTG CGCCGGCCGG GAGGGGCAGC GGCGGGAGGC GGCGGACTGC GGCGCACACG
 18001 TGGGGAGGGC GGAGCGGGGC GGGGGCCGGG CGAGGGCGGG TGACGGTACC TGATGAAGGT
 18061 GGTCTTGCCG GTGCTGTACT GGGCCACCAAG CAGGATCATG GGCTTGTCT CGAAGTCGGC
 18121 GTCCCTCAGC GCAGGCGAGT GGAACCTGTG GAAGCGGTAC GCCTCCCTCCA GCGGCAGCAC
 18181 CTTGCGCAGG TAGAGCGAGC CAGCGCCGCC CGTCAACCGTC TGCAACCGGT CGCGGCCGCC
 18241 AGCGCGTTCG CGCCCCCCCCC CCTGCCGCC CATCCAGCTG AACATCCTGC CGCCAGTCCA
 18301 CGCTCGATG GACCCCTGCT CCGGGTTCGA CTCTCCCCGG GCCTTGTGCT GCCACAGCCA GTCTCGGCC
 18361 CCGCGCTGGC CGCCGCCCGA GGAGGGCCCG GCCTTGTGCT GCCACAGCCA GTCTCGGCC
 18421 GCTACGTAC GCGCGCTCGC CCACTATTCA GACGAGCTCT CTCAGGTCTT TCCTGATTGG
 18481 ATCTCTGCCA GAGAGGGCTG AAGGCGTCAG CCAATTGAGA GGAGGTTCTT AATGAATTAT
 18541 TGATGACGGA GTCATTCTAG AGGACTGGC GGGGAGATGG TGTCTTCCCG GCCGGTGGC
 18601 ACTGACCGGT CATGCTCAGG GCCCCCTTCT GTGTCGGT GCCTCACCTC TGGCCAGGTG
 18661 GCATCGGCTT CCGGGTAAAT CTGAGCTGCC CAAAGCGAGC CTGAGGAGGC TGCTCCATCT
 18721 CCTGGAACTG TTACTCATTC AGGCCTCTGG ACCTGATAC TGGGGATGAA AAGTCAGAA
 18781 AGTGGACAGG CAGGGTCTGT CTTCCTGGGG CGGGGTGGT ATGACTGCAT TTACAGGCC
 18841 AAGAGACCA GGCTCACGGG CATAATTGAC ACAATTGAGA GGGGAAGTGG ACTGGAATCT
 18901 ATTTGAGTAG CCGTTTCTCC TACTTCTCCA CATTGCCCTC ACTAAATTCT GGATCAGCTA
 18961 AAGGCTTTT CCCCTTAGTT TTCTTTATGT GTTTTCTGT AAAGCCGAAA ATTAGTCAT
 19021 TACTCTCATT ATTACATTAT CCTCATTATT ACAACTTTAT TTCCACATAT TGACATGTGA
 19081 GTCAATTGCA CCTCTAGCT GTTTTTGTT GTTGTGTTG TGTTGTGTT TTTCTGAGAC
 19141 GGAGTCTCGC TCTGCGCCC AGGCTGGAGT GCAGTGGCAC GATCTCGACT CACTGCAACC
 19201 TCCGCTTCCC GGGTCAAGC AATTTCCTG TCTCAGCCTC CTGAGTAGCT GGGACTACAG
 19261 GCGCTCACCA CCATGCCAG CTAATTGGT TATTTTATGT AGAGATGGGG TTTCACCCATA
 19321 TTGGTAAGGC TGGCTTGA CTCCTGACCT CAGATGATCC GCCTGCCCTG GCCTCCCAA
 19381 GTATTGGGAT TATAGGCGTG AGCCACCGCC CCTGGCCTGT TTCTTTCTA GTTAAAAGG
 19441 ATACTCACT ATCTAAGTTG AGATTGGGG AAATAATCAC CTTTTGTGTC CTTGATTGTA
 19501 GCTTAGGAAT ATTTCTTAGT TTGTTTAGAG CATTCTCAA GTGTGATCCA TGGGGCCACC
 19561 TGCATATATG CATTATGAG CTGGGATGCT GTTAAAATG CAGATTCCCTG GTCTTCAGTC
 19621 CAGACCTACT GAGTAAAT ATTGAGGAGA GATGGAGCTC AACAGATGCC CTAGGTGACT

19681 CTGATGTGGC CTAAGTCTGG GGAAACAGTA CACCCCTGCAG TACTTCTCCA ACCTAACAGA
 19741 AATGCGGATT CTGATTCGGT AGGTATGGG TGGGGCCCAT GACTCTACAT TTCTAACAGG
 19801 CTCTCGGGTA TTGCCAATGC TACCCATCCC TGGACCATAc TTGGAGTATG AGAGAACAGG
 19861 ACTTTGCATC AAAAGATCTG GGTTTGAACC TTGACTATGT AAATTAAATGG CTGTGTGATT
 19921 TTAAATGTGA TGAGTTACCT CTCTGCCCTA GTTCCCTTAA GAGCAAATG GAAACTCCTT
 19981 TGGAGTTGTT GTGTGGTCA CATAGTTGGG GGGATTGTG TGTAACCTAGG GAAGTGAAA
 20041 GAAAGCCAAG GGAAAATAAA GTTAGACATC TGCCCTGTG CAAAATAAA CAAAGATTGA
 20101 AATGTAAAAA GAAAATGAAA CCACTGCACT ACTAGAAGAA GGTACAAGAG ATTTATACAA
 20161 TTTGGCGGTG GAAAAGCTC TCTAACCAAC ATTAAGGACA GAAATTATGA AAGCAAAGGC
 20221 TGACATATT GACAACAAA GTAAAAGCTT CGATGTATT TTTTAGAATA CATAAAAACA
 20281 AACGGATATA AATACAATCT TACCATATGA TGCGAAATC AGCTCCTTAG TATTACCCA
 20341 AATGAATTGTA AAACCTTATAT CCACAGAAAA GCCTGTACAT GGATGTTCAT AGTGGCTTTA
 20401 TTCATAAGTG CCAAAAATTG GAAGCAACTA ACATGTCCTT CAGTAAGTGA AAGGATAATC
 20461 TGTGGTACAT CCAGACAATG GAATCTTATT CAGCACTAA AAGAAATGAG CTATCAAGCC
 20521 ATAAAAAGAC ACAGAGGAAA CTAAATGCC TATTACTAAG TGAAAGAAGC CAATCCGAAA
 20581 ATGCTATGTA CTGACGATT TCATCTTAT GACATTCTGC AGAAGGCAA ATTATGGAGG
 20641 AAGTAAAAAG GTCAGTAGT TCTAGGTGTT AGGGAGAAGG AAGGAATGAA CAGGCAGAGA
 20701 ACAGAGGATT TTATGGCG TGGCTCTACT TTTTATGGT CTATAATGAT GGATACATGT
 20761 CATTATATAT TTGTCAAAAC CCACAGAAATG TACAACACCA AGAGTGAATG TAAGTGATAA
 20821 TGACTCGTGA ATTTCAAGATC AGTTTAGGT TCACAGAAA ATGGAGCATA AAGTACCGAG
 20881 AGTTCCATA TACCCCCACC CCCCCGACAC CACAGCCTCC CCCACTATCA ACATCCCTGAA
 20941 CCAATGGAAC ATTGCGTTACA ATCAATAAAC CTAACCTAAA TAGATACATC ACTAAATTAA
 21001 CTTAAATAGG AACTAGATAT AAAACTAATA TAAATGTAGA ACTAAATGTA AAACCTTAAGA
 21061 TTATAAAACT TCTAGTAGAA AACACGGGAG AAAATTGTT TGAACCTGGG TTAGGGAAAG
 21121 ATTTCTTAGG TGACACAAA AGCACAAATCC TAAAGGAAA AAAAGTTGG TAAATTGGAT
 21181 TTCAATTAAAC TTGAGAACTT CTGACCTTG AAAGTACTGT GAAGAGAATG AAAAGACAAG
 21241 ACAAGGCACA GAATATACCT GCAGATCACA TATCTGATAA AGGACTTGT TCCAGAATAT
 21301 TTAAAAACTT CTAAAAACTT AATAGTAAAA ACCCAATT TAAAATTCC ATTTAAAAAA
 21361 TGGAAAAAG ATCTAAACAG ATATTTACA AAAGAAGATG TATGGATGGC AAATAAGCAC
 21421 ATGAAGACAT GTTCAACATT GTTGTGTTGTT AGGAGAATGC GAATTAAAAC CACAGCAAGA
 21481 TACTACACCA TACCTATTAG AATGTCTAAA ATTAGGAAAG ACTGATCATG CCAATCATTG
 21541 ATGAGGATGT GGAGCAACTG GAACCTTAC ACATTGTTG TCAGAATGTA AAATGGCACA
 21601 ACCACTTGG AAAACAGTGT GGCGATTTCT TAAAGGTTG AACGTACACT TACACATGA
 21661 CCCAGCAGGG GAGGAGCAGA GGGAGGATT GCAAAAGGC CTGAAAGCT TTTGGGCTG
 21721 AAGGGTTTAT TCATTACCTT GATTGTGGTT ATGGTTTAC CATTGTATAC ATATGAAAAA
 21781 ACTCATGAAA TTATGTACGT TGGATCCTCC ATATCTCTGT GTTCTGTATC CATGGATTCA
 21841 ATCAACCACA GATGGAAAAT ATTTGGGAA AAAAATGGAT GGTTGCATCT GTACTGAACA
 21901 CATAACGACC TTTTTCTTCTT TTCAAGTATT CCTAAGCAAT ACAGTATAAC AACTATTTTC
 21961 ATATCATCTA CATTATAAT AAGTAATCTA GGTAGGATT AAAGTATATG GGTTCTGT
 22021 GCATGGCGTT TAAAGTTAA AGAAAAGTAT GTGGAGGATG TGCTTAGGCT ATATGCAAAT
 22081 ACAATACCTT TTTTTTTTT TTTTACAATA CCATTTATA TAAGGGAGTT GAGCATCCAT
 22141 GGTTTGGTA TTCCCAAGGGG TCCTGGACTC ACGGATAAT GTACTTTAA TATGCACAGT
 22201 TGATCATATG TTTTACAC CTTGATTAGT GTTAAATAT GGTCAACTT TTTGACTCG
 22261 GAATTTTAT CTTGAGGAAA TTATCCTGA GGGTATTAT ACTGAGGAAA TGATTAATGT
 22321 AAATTTTAT CTAAGACGAT GGGTGTCAAGC ATGGAGCTGT CTATCATCAT GGAGGATTAG
 22381 TTAAGCAAT TATGGTTCC TATGTGGTTA CTGAAAATCTA TGTCATTGTA CAAGAATATT
 22441 AATGATGTGC AAAATGCTC AGAATATTATT TTTAAGTGG AAAAGGCTAG ATTACAAAAG
 22501 TGAATGTACA TGCACTGCT AGGGAAAACA ATATGGAGT TCTGCACAAA TTAAAAAATA
 22561 GAACTACTGG CCAGGTGCAAG TGCGTCACGC CTGTAATCCC AGCATTGGG AAGGCCAAGG
 22621 CAGGCAGATC CCTTGAGGTA AGTAAGGAAT TCAAGACCG CCTGGCAAC ATGCCAAAGC
 22681 CCATCTCTAC TAAAAATAAC AAAAATTAGC CAGGCATCAT GGCGGGCACC TGTAACTCCA
 22741 GCTACTCGGG AGTCTGAGGT AGGAGAATCA CCTCAACCCG GGAGGTAGAT AGAGGTCGCA
 22801 GTAAGCTGAG ATTGTCTAC TGCACTCCAG CCTGGGTGAC AGAGCAAGAC TGTCTAAC
 22861 AAAACAAAAC AAAAACTACC ATATGATTAA CCAATCCCCA ATTGGGTAT ATATTTAAAA
 22921 GAATTTAAAG CAAGCTCTCA AAGAGGTATT TGTCTACCCA TGTCACAGC AGCATCATTC
 22981 ATAATAGCCA AGAGGTGGAA GCAAGGCTAGT GTCCCATGAT GGATGAATGG ATAAACGAAA
 23041 TGTAAACACAC ACACACACAC GCACACAGAG GAATATTATT CAGCCTTAA ATGGAAGAAT
 23101 ATTCTGGTAC ATGCTACACC CTGGAACACG GATGAACCTT GAGGACATTA TGCTGAGTGA
 23161 AATAAGACAG TCACAAACAG CCAAATGCTG TATGATTCCA CTCATATGAG GTTCTAAAG
 23221 TAGTCAAATT CACAGAAGCA GAAAGTAGAA TGGTGGTTAT CAGGGCTGG CGGGAGGG
 23281 ATGGGAGGT GTTCAATGGG TATAGAGTT TAGTTTGCA AAATGAAAGA GTTCTGGAGA
 23341 TCTGTTGCAC AACAGTGTGA ATATGCCCTGA AACTGCTGAA CTTTACACTT CAAATTATTA
 23401 AAACGGTAAA TGGTGTGCGT GTTCTTACCC ACACGCTTT TAAAGTGGAC GTACAGGATT
 23461 TTGTTACCTT TGTTTAAAC AGTTTAAAGT ATCTAAGAAA AAGATGGAA GGAAAAACCC
 23521 CCAAATAATA ACGATGTCTG AGTTTGGAC TGGTAAATGA TCTTAGTTT CTCTTTATAA
 23581 GCTGTAGCAT TTTCAAATG TTACACAGTA ATCACGTTCC TGTTAAAAA GCAGAAAAAA
 23641 TTATCAGGAA GAAAAAAATG TACTGTCTTT AATCTAAAAA AAAAAAAGA AAAGAACAGA
 23701 AAGTCCTAAA AGCCCTGGC CCAGAGGAGG CACCACTCCC ATCTTGCCTC GCTCCCTGCT
 23761 TGCCTTCATG CAGTGCAGC TTTCTTATT ATTATTTTG TAGGCACATA GTAGGTGTGTT
 23821 ATATTCTATGG GGTAGCAGCA GCTTAATTAA TTTTGTTGGG CACATAATAG GTGTGCAATAT
 23881 TCATGGGGCA GTGACAGCTT TTTTCTTAT TTATTTGTT TTTGTGGCA CATAGTAGGT

23941 GTGCATATTG ATGGGGCAGT GACAGCTTA TTTTATTGTT TTTGTGGCA TATAGTAGGT
 24001 GTGCATATTG ATGGGGCAGC GGCAGCTTC TTATTTACTT TATTGTTTT GTGGGCACAT
 24061 AGTAGGTGT CATATTG GGGCAGCGC AGCTTCTTG TGCTTGTCT ACCCTCAATG
 24121 AAGCCAGACA GTGGCAGATC CTCCCCGCCA TAGTCATGTT GCATCTGTT GCAAATCATT
 24181 CTGCTCTTG GTTTCTTC TTCAGACTAA ACAGTCCCAA TTGCAGTCAC TCTTCATCT
 24241 GAATCCAGT GTTTTCAGG TCAAATCTA TCACAGGAG TTCTACAAGG GAATGCCATT
 24301 AGTTGTCGG AAATCTATTG TATTGAATGC ACGCTTGAAA TTAGGAAGCT GGTAATAATG
 24361 GTCAGAAAGT TTGCTCTAGA CCATGTTAG TGGTTGGCT TAAATGGTGA TCAAGTCCAG
 24421 TGTTTACTT CGAAAAGTGG ATTGAGAGT GGTTAAAGGC CGTGTGAGA
 24481 CCAGAGCGAG TGCTGTGGCC TGGTGTCTC ACACGCTGAT GGCTTCTC TGTGCCACT
 24541 CACCCAAGTC ACTGCCACCA CCCAGAGGAG GGGCTTATAG GACACACAGG TGGAGATGGC
 24601 TAGGAGGTAC TTGCAAGAAG TCTGCTAGCT GCGTATAAAT TGTCTTCTT TCGCAGCAGA
 24661 ACTCTGATGT TTAGGTGGC ATATTGCCAC CCATGCAAAA AACAAAAACT TCCAGCCTCC
 24721 CTGCACTTAA GTGTGATGAT GAGACCAAGT TCTGGTTGGT GATATGTTAG CTGCAGGATG
 24781 TGTGGACTT CCAGGAAGTC CCTTTAAGAA GGAGTTCAAT CTTTTTCTC CTTCTTCTG
 24841 GTGGCTTGAATATGGTGT GATAGCTGGA GCTCCAGGGC TTACCTGGA CCATGAAGCC
 24901 AAGGACCAACA TAAGGGTGA TAAAGCAAG AGCCAGAAGA AGCAGGGTCC CTGACGAATC
 24961 GTAGGTGTC CTTGAACATC CTTCTGTGG ATGTTTTAAT CATGAGTGTG TTTAAGCCAC
 25021 CATTTCAG AAATAAACAT GCAATTCTA ACCAATAATC AATTGTTAC CAGAAATGGA
 25081 GTCTGCCACA ACGGAACCTA AAATGTGACA TGTGCTCAGG GGACAAGGGG GAAGGTCAAT
 25141 ATTAAGGATT CAAATATTGT AGCTGGAATT TAAAATATGG CTGCACAGTT CTTGACATGC
 25201 CCTCCATCCA GAGATGGTT CATTGCTCCC TCCCCTTGA TCTGGACTGG CTTGTGACTA
 25261 CTTCTGCCA TAGATTCTA TGGAAGTGG AGTACGTAAC TCTTGTGAACT CTGAGTGCCTA TTTAAGAAGA
 25321 GGGCCATGCT GCTGCCACTG TAACAGACAC TCTTGTGAACT CTGAGTGCCTA TTTAAGAAGA
 25381 CTGACATCC TGAGTCCATC ATGCTGGGGT CACCAAGGTG CTCTGATTGA CACTCCTAGC
 25441 TGGGATTGCC TTTCATCTT CACAGACTTC CCAGGCCAGG CACCAGACAC AGGAGTGA
 25501 GAGTCGTGG ACGAGGGCCC TTCAAGCCAG GCTGTTCCA GATCCCAACC ACTGGAGTCC
 25561 TCACCAGCAG TTCAGATCAT CTGACCTGTG GCCCCAAACA TGGAGCAGTT ATGCCCTTT
 25621 CAAATTCTG ACCCACAGAA TCTGTGAGCA TAATAAAATG GTGGTTAATT TATGCCATTA
 25681 TACATTGAGG GCTTTGATG CTCAGCAGTA GATAACTGGA ACAGCTGGT ATTCTTACTA
 25741 TGTTAAGATA AAAGTTTAC TCTTTTCA AGAGAAGAGA TCCAGACATA GGCAGTCCAG
 25801 GACAAGTTG GTCACTTTG AGTCAGTAGG AACGCAGATT CCTTTATCTT CCCTGCTGCC
 25861 TGGAAATGTTG ACATGATGGG TGCGGCTCCA GCAGAAATAT TGGACAGAT GGTAATCTT
 25921 AGGATGACAG AAGAGAGCAG AAAGGAGCTT TGGTCTCTGA TAATAGAGAA ATAACCTCT
 25981 CTCTTGTGTTA AGCCACTGTA ATTTAGATC TTCAGGTTAT ACGGCAAAAG CATTGCTAAT
 26041 AGGCACAGAA TTGGAATTCA TGGGTTAGA ACCAGGGAGAT TGGCTTGGCT AAAGCTAAA
 26101 GGTTAAGAAT TTTCAAATTG GAAGCAGTAG TTGAAGTCTT GGGTATAAT GAGGTTGCTT
 26161 AAGGAGAGTA CAGACTGAGA AGAGGAGAAGA GCTAAGGAGC CTCGAGGAAT GGCAGCAGTT
 26221 AAAAGAGAGG CCAGAGGAGG GGCTCCCACCT TTCCAGGGG AAGAAAACCC AGACAGAAGG
 26281 TGTCAGTGA ACACGACTTC TTCCCTCTG TGTCTGGTTT GGGAGCCCTT CCTATGCCTC
 26341 TTCTGGACAC AGCTAACATG CAGCTTCTC TTCAATCATC AAGGGCTCTG TTTTACATGA
 26401 GGGCTCCCA CATGGGACCA GAAGGAACCT CCTTAAAGTT TGAATCAAT TGTAAATTCT
 26461 TCTTGTGCC CCAGTGCCTA CAACAGTGC TGGCATGTAA TGGGCAGGCA GTAGTTATT
 26521 GCAAGCTGAA CTGAATTGAA CTGCTTTGG CAAGGAAGTC CTGGGAAGGG AAGGTTTCCA
 26581 GAAGGAGGGA GTGCTCTGCG GTGCAACAC TGAAACAAAT CAAGTCAGAT AAGGCCCTTA
 26641 CAAGACTGGT GGGTTGGCT CAGTCATCT GCAGAAATGG AGGCCATTGG TGACCTTGAG
 26701 GAGAAAGTGTG TTCTGGGGG GGCAGGGGG ACAGGGTAAC TGGTGTGTTG GAGCGGGTGG
 26761 GGTCTGGTT GCTCTTTGG GAAGAAAGGG TGGAGAACAG GAGGGACGGG TTTCATT
 26821 TTTGTCAATT TCATCATCTT CATCATAGGC AGAGAGCAGG GAACTGGAA GGGAGCCTCC
 26881 CCTGACTTCT GCAGGATACC ATGTGCTTGC CCTGTCAGAG TTGATGACAC TTTTGTCTT
 26941 TCCTTAGCCA ACTCACCATT TAACCTCTA TGTATGCATA ACAATCTTT CATGGTTTTC
 27001 TCCCGAACCA AAATGTAAGT TCCGTGAAAG CAGAGACTGT TTACAGCTTG CTGTGCACAA
 27061 ACTAGGTGTT CGTCTGGTGT TAATGCTCTC TGTTCTCTTG CAAGCAATGT TTCTGAAAT
 27121 ATACCTGCGC TTGGAACACG ATGGATCTGT ACGAGCTGGT GCCCCACTAG TTGATGACAT
 27181 GGCTTCATGA AAACAGAGGA GCACCTCACC ACTAGAAAGG ACCCTTGAG TAGGAAACAA
 27241 AGCTGCATGA GAAAAGGTGG GATGGGATGG AAACGGCTCA CTCACCCCC CAGGAGTGG
 27301 GAGGAGGAAG CCTGAAGGGG TGAGGGCAGG GTTATGCATA TTTGCAGCA TCAGATAATG
 27361 AAAGCAGATG AGTCAGCAGC ATTCACCCAC ATGGCATT TATTCCATGC ACAAATAATA
 27421 TGATGATTTT ATAAACCTCG TTTCAAAAC TTTTGATGAA TCCAAGGCAT GGCTGCCTTC
 27481 GCCTTACGAT AAACAGTC TGTCTGTG TGCTCTACT CTTTGATACT GGAAGGAGAG
 27541 TCTGTTGCAG TGAGAAAGAA AAGATGTCAA AGTGGAGGCA GCCAGTTGCT CTGGTGGAGG
 27601 GGAGGGCTGG TGCTCCAGGT TAGCATTAC TGGATGGGG CAAGTTTGC AGGTCAGAG
 27661 CTCAGTTTAT CAGTGGCTCA AGGGAGCACT GATTCTACAG TTCAAGCTC ATGGGAGGG
 27721 GTGGCGTAAAGT TCCAGGCCTC TCAAACATTCA ATATGCAAG GAATCACGGG GGGCCCTGG
 27781 GGATGGGAAT CTCCCTGAAA ATGCAGCTTC TAATTTAGTA CAGCTGGCT AAGGCCCTGAG
 27841 ATTCCGCATT TCCAACCAGC TCCCCACTGC ACTTTGAGTA GCAAGGATGT ACATCGGAGG
 27901 CCAATTTC TAAACTTTG GTACCTGTCA CCACGGTTT CCTCACAAACC CAAAGACTGC
 27961 TTAGGGAGCA TGGTGGCCCA GAAAGGGGAC TCCGGCTTC AGGGTTGCTG GGTTTCTG
 28021 AGACCTCATG TCTCTGTGCC TTTCTGTCTT CACCTTCTAT CAGCCTGGGC CTGTCAGTGG
 28081 CTCCGGTTAT ACATTCCTT GTTGTGTTCC AGCAGCTGAA GGGAGACCTG CTGACTTGGG
 28141 GTTTCTCTGT GTCCCCAGGGT TTTCAGAG ATTCAATTAGG AAGCAATCTT AAAGGGGG

28201 CTTCTTCCTC TGACGGCTAA CCTAGTCCTT TTTCACTAAT TCTTCCAGT TTCAGTCTCT
 28261 GTGGGTGACT CTCCAAGAGT TGAATGCAC TGGTCACGGT GTTGACAACA CAGATAAGAG
 28321 GAATCACATT TAGAAATAAG TAATCCCATG AGGCAGCCAG AGAGATGGC AGTAAGAGGC
 28381 TGAGAGATGC TCTTGGAGCA GGGTCGTCAG TGAGCACCGA CAGCTGGAGT GGGCAGCGCT
 28441 GGATGCATCG CCTGCCTGGC CCCAACCTC ATTACCCAC AGCCTGCCTG ACCTTGCTGA
 28501 ACTCCATCCC AGATCCTCAT TTGGTGAACC TGGGCTATAG ACTTAGGCTT TGCCTGGCAT
 28561 TTGTTGGAG CTGGATGCTT GGGCATGGCT GTCTGGTCAT GATGACATGG ATCACCAAGCC
 28621 CTGGAGCTCT GGGTACGTGG AGAGGCCAG TGAAAACCGC CACCTGTTG GCTTGTGTT
 28681 GGCAGAGACT GGCTTATTTC TACGGCTCCA TCTAGATCAT CTAGCTTCAT GCTTCCTATT
 28741 AATACTCCGA TTCTAAAAAA TAAATCAT GAAATATAAT TGCCTATAAA AATAATGAGA
 28801 CTAGGCCGG GTGCGTGGC TCATGCTGT AATCCAGCA CTTTGGAGG CCGAGGAGG
 28861 TGGATCACCT GAGGTAGGA GTTCCAGACC AGCCTGTACT AAAAATCAA AAAATTAGCT
 28921 GGGCATGGTG CGGGGTGCCT CTAATCCCGG CCACTTGGGA GGCTGAGGCA AGAGAATTGC
 28981 TTGAACCTGG GAGGAGGAGG TTGCACTGAG CCGAGATCGC GCCATTGCAC TCCAGCCTGG
 29041 GCAACAGAGT GAGACTCCGT CTAAAAAAA CCAAAATCA CAAGACTAGA ATTGCCCTAA
 29101 ATTTTGAAGT GTTTTCCAG TGAACGCCG AGTGCCTCTC AGGTGCCTCT GTGCATTTC
 29161 CTAGGCCCGC CACACGAGGT AGAGGACAGA GAGAGGAGGG CTGCCAGCCC TGGGCTCTGC
 29221 CTCTGAGTCG AGGTTCTGG GGGCCCGCCT TCTTCCCGAG CTTTTCCCTC TCCACAGATG
 29281 TGTGGAGGC TGTGGCTATG AGACCACTT CTCTATAGGG AAGCACTTGA AGAAACTCTA
 29341 GTGGGTATAA AACCCCAACA GAGCTCCTCT CTCTCTCTT CTGACTTTCT CAACCTTTTA
 29401 AAATTTTTA CTTTCGTTA GAAAGCCAG GCCAACCTGC TCACACCCAT TGCCGGAGAG
 29461 CAGAGCTGGC TGCAGTAGAA CCTGGTCAGC CCTGAGGGGT CCTGCACCTC TGATGGTAGA
 29521 AGCTGACACA GAACAGGGGA GGTCCCTGG GGCCTAGGG AGGACACTGGC CCTTCAGGCG
 29581 CTTCTTCCTC TCCACTGCGA GCCGAGAGC CTGGAGGAGA GTGCTCTTAT TATTCAAGGAT
 26 29641 GTTATACTCT GAGAGTTCA CACAGCTTGC AAAGGTGGC TCTGTGATAG TCAGCTCCTT
 29701 GGTGGCATAG GGAGTTGG GACCATAAAT GTCCACCTGG CCCTGCTCCA GCTCCCTCAGG
 29761 GCTTCGCTCT ACACCTGCGG GCAGGAGGAT GACTTTGTAT AGACTCTGTC ATGTTAGTT
 29821 GGAATGACTT GTTCCACAT CTCTCCCAA GGCCCCACCAA CCTGCACTCA CGGACACACAA
 29881 CACCTGCATG AAAACCTGTA CACACGCACA CCCTCACACA CGCACCTGCA CACCTGCAAG
 29941 CTCCCTAGGG CGTGGTGCAG ACTTGCCTAT TGGAGGATTG CCCTAGCACC AAGCACAGTA
 30001 CTGGGCACAT GGCTGAATGT CAGAAAATGT AGAAATGACTG AAAGAATGGG GGATGACACT
 30061 ACCAGCCCAT GGGGGTGGAGG GAATGGGAGG CGCCCTTCTT TGTGGAAAGCC AAGGTTGGCT
 30121 CAGGCTGGAG CTGAGCCAGG CTTCCCCGT AGTGCCTTA GAGACTCTAC AGATCTAAAG
 30181 GCATCCCCGG CCAAGGGCA GCAGGTGCTG GGAAGTGGAC ACAGAGGCG CTCACCTGGT
 30241 GCCTTGATT TTGAGAAAGT GTCAATTGATG AGTGGAGAGA AAGTCACGAT GGGGGCATCG
 25 30301 GTTCCCTGGG GTTCTCCAT CAGGTAGCAT TCCTTGAGAT TTTCATTCTC ATCTGGCAGC
 30361 TCGTATTG GGAAGGGGAT GTTCTGCACA GTGCACTGACT CACAGTTTG TTTCAGGGG
 30421 TGGAACAGCA CAGAGGGCGT GTTGTAGCAT TAGGAGGAGT GGGAGTGGAG GAGAACACAA
 30481 GAGTTCTTAG CCCCAGCTTC CTGATTCCC CAGCTAACAA CTCAGCTCCC CAACTGCTT
 30541 GCCAGTGTAT GTTACCTTG GACTAAAGCC AGGTCAAGGCA AAAAGGCCCG AGGCAGTAGG
 30601 TGGATGCACC TCCAGTTGCA AGTGTGCTT CCAAGAAAAA TGGACAAAAG CCCTTCGTT
 24 30661 GGTGTTGCGA TCCAAACACA GACTTGGGTG GATGTTGGT GTAGCTTGG AAGGGCAGTG
 30721 GAACCAAGCT ATACACAGCC AGAGCTGCC AGGCCCCCTGT GGGATGTGCA TGGGCTTGG
 30781 AGCTGACAGT AGTACAGGGC TAGAAAAGAC TCTGCAAGGG CATCTCTAA AGCATGATGT
 30841 CGCAGGAAGA GCACTGCCTT AGAGTGTCA GACCCCTCA GTTCTGCCCTC TCGTTCTCT
 30901 GCTCACTAGC TGTGTGATCT TAAGCAAGTC TCTCTGTGTC TTGGTTCTCT CATCCATAAC
 30961 ATGAGAAATG ATAGCACTG TGCCGTCTCC CTCCAAGAGC TGTGTGTAAGG GCCAAACAAG
 31021 ATCACACGTA GGAAAGTGCA TTGTTACCCA TTACAAACT TAGATTGTT TCATTGTTGC
 31081 CCTGTTGGTC TGTGTGACT TCACTCGCC TGGCTTCAAG GTAGAACTGT ATACACATAG
 31141 AACTTCCAA AGTAAATTAA AAAATAAAAA ACTTCTGGTT CTGACAAGTG CCCTCTCTC
 31201 TCGCCAGGCT TCCCTCCAC CTGAGAGCAC TTGAGTGTAA TTCAAGGAGG GTCAATAC
 31261 CTGGGGAAAG GGAGAGTGTG ATCAATTGAA AAATAGGATA TTACAAGCTC TTGAATGGAA
 31321 TGGGCTCTAC TAAGTTATT TTTAAAAACT CATTTCCTGT AGCTGAAGAA TTGCCAATA
 31381 TGGCAAATGC TCCAAGGCAA GAGTAAAATT CGAATTTCGT AAGTCACAT GGGCACCAGT
 31441 GTGTGAGAGC CAGCTGGCTC CCCGTGAGGT CTGGGTGTC GACTTCCCAT CATGACTGCT
 31501 GCTTGGCTTC CTCTGATGCT CTCTGGGGC TGTCCAGAGC ATGCCTCTTC CCAGTTAC
 31561 GCGAGTGGAG CCACAGTACT CGGGCCCCAT TTAGGAGAG GCTGCAGCCC ATGCTCAGGA
 31621 GGGAGGCTGC GCACTTGCTT TTGCTGGGA CCCAGCACAG TAGTTGAGGT GGATGATGAG
 31681 GTCGGCTTT CGCTCTGGCC TGAGGAGGGG CGGGTAGCTG GAGTTGACAA AGAACCGAGT
 23 31741 GTCCAGCAGG CACAGGTGGT TCGCGGACTC GGTCACTGTT TTTGGAAAC CATCTAGCAC
 31801 TGTGTGCTGAA AGCCAAGCCT TCCCGTTACC AACACACCTG GTGCCCTGAC AGCCAATGG
 31861 CAGAGGCACT TGGGGGGACG GAGCGCCCTG TCTGGAGCAT CTTCTGCCTC AGCCGACCA
 31921 GGACGCCAC TGAGTCCAG TGCCCTCTGG GTGGCTGAGG AAGAACAGCG TTATCTGG
 31981 CTTTTAGCT GTCTTATTGAT GGATATTGAT ACTTGTGGGG TATAAGGGGC TTAAGACCAC
 32041 CCCATAAAATA GTGTAAACCC TGCCCTCTGA GTCACTGTC CCATTGGCC CAGATGGTGA
 32101 GGATGAGGCT TAACAGGCCA GGCTGAAGCA GCCATTGCA GGCCACGGCC AATTAAAATA
 32161 TTAAGAAGGC GAAAGTCAT GCCTTAGGGC CTCAGCAGCA GAGGAGGGGG CACTGAGCTG
 32221 ACGCTGCAGG ACCACAGTTT AAGGCCTGGC TCTGCTGCTT ACCACTGGGG GGTCTCTGAG
 32281 CTGGGCATC CCTGCCAGT TCCAAGCGGG CACCACTCAC ATTTCCTTTTT TTTTGAGAC
 32341 AGGGTCTTC TCTATTGCAAGGCTGGAGT GCAGTGGCGT GATCATGGCT CATTGCAAGC
 32401 TTGACCTCCTC GGGCTCAAAT GATCCTCTA TCTCAGCCCC CAAGTAGCTG GGACTACAGG

32461 TGCTCATCAC CATGCCAGC TAATTAAAAA CAATTTTTT TTAAGAGATG GGGTCTCAC
 32521 ATATTGCCCA GGCTGGTCTC AAACCTCTGG ACTCAAGAGA TTCTCCAGCC TCGGCCTCCC
 32581 AAAGTGTGG TGTTACAGGT ATGAGCCACT GTGCCTGGCA CCACTCACAT TAAAGGGCAT
 32641 CCCCAGAACAA CAACTGCAGA AAGTAAGACC CCCTGTTTCC CCATCTGTAA CATGGGGCCA
 32701 ACTTCTACCT GCTAGGAAGTG TTGTACAGAT TCTGTATAGA ATGACTGTAC AGATGCAGTG
 32761 AGTTAAGTAC CCACACCACT CCCTGCTGCC AGACACATAAC AGTGTCAAG ACGACACCAC
 32821 CAAATGTGGA AAACACCATG GGGAGAGAGG TGCCCTGGCCC AGTGAACATA GCCAGAGAAG
 32881 TTTTCTTATT CTTATCCAT GGAGCCTTG ACATCCACAT CTTCCTCCCT CATATGGAGG
 32941 CAGGCCCTCAA GGAGTAGGCCA GGTAGGGTTA CCTTCTCCACC TAGAGAACTG GCCATTCTGG
 33001 AGGTAGTGG TGTGCACTG CAGGCCAGAC AGGAAGTTG GAAACTCAGA CACGAAGGAC
 22 33061 CGATGGTAA GGATTCTCG GAAAGAATTG GACAGCCATG ACCCTGGGAT CACTACCGTG
 33121 GTCTCCAGGA TGTTAGCATC ACATTTGGG ATTTCAAGGCA GGATCGGCTC GTCTCTGCA
 33181 GGGGAGGGAG GGATGAAGGG CAGGTATGA GAGGTGGAGT CCCCAGATT CCTATGCGAA
 33241 GCCCCACATG GTTGCCTGGC ACCCAGGGT GAGCACTGTA AAAGCAGCTC ATACCTTCCT
 33301 TGCTGGCTC TCAAGGTCTT GGCACGGGGG TTAGGAGGAT TCAGGTGAGA GACATGGAGT
 33361 AAGCAGAGAG GTAGTAGAGA GGGAGGGGG GCGAGACAGC TAAGAAGCGT CCTGGAGAGG
 33421 ACAGGGAGTT TCAAAGAGT ACCTGTGACT GCTATAGACC CTACTCTCT CACTCTCCT
 33481 GGGTGTGGC AGCCTTGCTT CCAGGCCCTTG GCGCTGGTGT CACTCACCTA CAGTCACCCCT
 33541 CCCCTCAGG CAGGGAGAC ACTGTAGGCC CAGCTGCCCT CCCACGTCTT AGGCCATAGG
 33601 TGACGATGCG GAAACATGTG GCAGGGGGCC TCATCCTCGG TTTGTGGTCC ACACCCGCC
 33661 CCCCATCACC CAGCCATGAG GAAAAGCCGG GGGCCGCC CCATCACCCCA GCCATGAGGG
 33721 AAAGCCGGGG GCCGCCCCCCC ATCACCCAGC CATGAGGGAA AGCCGGGGC CGCCCCCCAT
 33781 CACCCAGCCA TGAGGGAAAG CCGGTGTCA CCATCACCCCA GCCATGAGGG AAAGCCGGGG
 21 33841 TCACCCACCG ATGTCCTGCA CTTCTCCCT TGTCCACCTG TGGAAAAACT CCTCCGAGGT
 33901 GTGTGACAGG TTCCAGGCA CAGCGAGGT CAGGGAGAAG ATGCTGCTCC ACAGGGCTGG
 33961 GAGCCAGGGC CAGCGTGAAGC AGCTCCACCC TGCAAGCCCCA GGCCCCACAC TGACATTTGA
 34021 TTCCCCACTC CCCCCACCCC CCCACCCAC GCAAGCCAGG CTGCTCCACA GGCGGGACAG
 34081 ACTTCAGACT TCCATGTCTT TCTCCCTTC CCACTGCAAGC TCAACACAGT GACTGAGGGG
 34141 CTCAGCTCTC AGCTCTTCTC TTCCCTCCCC TCAGTGCAGG GAGACGTAAA GGTGTATAAG
 34201 GCACAGCCCC GGCAGGGATT TAATGCTGCC AAGTGCAGGC TTTGAGTTA AGTCCAGCTG
 34261 CTGTCAGTGA ATACATCTAA AGCATTAGG GCAGCACCTG GCACATGGCG AGCACTCAAT
 34321 AAATGCAAT GCTATTGTT ATTATGATT TAAATGTAAG CATTAGACA TTGACCCATT
 34381 TGGAAATCTCT ATAAGGTGGG TTTACTCTT AAGCCTAATT AACTAAGGAG AAGATGGAGG
 34441 CTTCGACAGG TGAGGAATC TGGCCAATCT TACACAGCTA GTGAGTGAGA AGCCAGAGTT
 34501 CAGAGTTCAA GCCCACGTGT TTTGCCACA GTTTCTAAA CCACATGCTC ATCAGCATAG
 34561 GTGTTCAAGGG GAAGGAGAGG CCACCGTGGG CTGGGGCCTT CTGGGGGAAG AAGTGTAGGG
 34621 ACTGAGTTGG ATCTTCAGG ATGGGTTGGT TTGGGATAGA CAGAAGGGGA AGGCACAGAC
 34681 AACTAATTAA AATAAGGGCG TGTGGGGCT GTGAATGTGG GTCTGTTCT GAAGAAGGCC
 34741 TTGACGGGC GCACGTAAGA CTCTGAAGGG AAGAACCTGAG GTGGGTCTCTG CATTGGTGG
 34801 GCCCAGCTCA TGAAGAACCG AGGCCTTAG GGTGGGAGGC TGGGCTAGGC TTCTCTGGTG
 34861 ACCCTGCCCT TTACCCGTG TGCTCCCTGG ACCCTCGGGT CCCAGGTCTG GGTTGGCCCC
 34921 TGGAAGCAGG CACACTGCTG GCTGAAGCTA GACGGGAGTG GGTGGGAGTT ACAGAAGCTT
 34981 CCGTCTGAGT CTGCTGTTCC CAGTGTGGT CTGAGCAGGC TGCAAGGGAGA CTGCTGGAGA
 35041 CCAGGAAATT GGGAGAGCCC CCAGACAAAC TCTGGTCCC AGTGAATTGG GACCACTCCT
 35101 AGGATGAGAC AATGCTGGCA AGTGCCTCGA GTGTCACATA ATCAGGGGCC AACCCAGGGT
 35161 GGCCCAGCAC TGGCAGCAG CCCCCCTGTCA TTCCGCCAGT TGCAAGGCAG TCACCTAGCA
 35221 TGTAGCAGAT TCGAGACTCC GGGATCCTCT TCACCAAGCCG CCCCATGAAG AACTCGGAGC
 20 35281 CGAACAGCTC GGAGGGATG AAGGCCCAT ACTTCTGCAG ACCCCACCTCG TAGGGGGAGA
 35341 ACTCGAACCA CTCTGCACAT GAAGCAGCAG GACGGGGTC AGTCTCAGGC CAGGGAGCCG
 35401 TCCACATAGC CGGGCCCCCCC TCCTGCCTGC AGCCACTGGC TCAGGCCCTG GAAAGCCTGG
 35461 GACAGGGAGG CCGCAGTCC CCCAAAGGAC ACATGGGGGC ACCCTGGCC AGCTCTGCT
 35521 GGTGTAGCAA GGGTCAGCTT TGCCATGCTT TGCAAGGAGG ATGGCTCATT TGCACCTCAG
 35581 AGCCTGTGA ATGGGACAAA AACTGACCCCT GAAGCAGGGC AGCTCTCTCG AGTCTCTCTC
 35641 TCTCTGGCA CGCTAGAAAT ATCTTCCTT CACAACTTGT GTCTTCCCCA GTGTCATTAA
 35701 ACACATGCCCG AGGGAAAGCTG CAGACAGTGC CTGCGGGCTT CACCCATGCA GCAGGGCAGA
 35761 TTGGGGTAGC AGGAAGTGGATT TCTCAGCTG TGTCAGTGTAG AGGTTGACAA
 35821 AGATGAGTGG GATTGGGGGA CCTGGAGGTG CTGGGTGCTT TTGCCACCA AGAAGTCCCC
 35881 CCCTCCACCC CACTTGTAAA TCACCAAGCCA CTGGTGTAC CTCTGAAGTC CTGGGTGCTT
 35941 ACATCATCCT TGACATTGAT GGTGAGGTAG ATGGGAGGG GGTCTGGCC GCAGCTAA
 19 36001 GCAGCACGCT GATCTGACAG TTTGATTCA TTTCTCTGTG GGGAAACAAA ATGGTTAAAG
 36061 AGAACAGCA GCCCAGAGGG TGAGGGAGGC AGTGCAGAAAT GAGAACACAA TCTGGACCCC
 36121 CGCAGCCACT TCCTGGCTCC TGAGCTCACC GGCGCTGCTGT ATCGGCCCTCA GCTGACACCA
 18 36181 GCCAAATCAC TCTGAAGGGG CACCAATGGC ATAGGCACCA CCCCTCGGG CTGCCACCA
 36241 CCAGACAGGG GCCGGGACTG GATCCGAGT TTAGGACAGT CTCTCAGAGT CCCAGGGCT
 36301 TGGTCTTAGG GGTCAAGCAGA TAAAGAGGAA GGGGTGGAG AGGCTAGCAC TGCCCTCTT
 36361 CTGCCCTCATC GTCTCCCTC TGCTAGAACT GGAGCAGCTC CATGTTAGC ATCTCCATT
 36421 TTTAAAAAAA ATCAGCTTGG CATTTGAGC TCCCTCAACT GAGGGTAAAG AAGTTGGGG
 36481 AAAAGAAGTT TCCCTCCCTC CGCTCCCTCG CCTCCCTCCC CCTCCCTTCC
 36541 CCTCTTCTT TTTTTTTT TTTTTAACAA AAGACTTAGG ACTTGAACAC TAGATGGTTG
 36601 CTGAGGTCACTCA TCAATCCCC AGAGTGGGAGC GCCTTCAGA TCAGCCGGGG CCACAGTCCA
 36661 CCTCCCAAAG TGAATCAGCC ACTCAGAACG AGGCAAAATC AAGGTCTAG AGGCCAGGC

36721 TCTGACCCCA GGCCAAGCAG CTGACTCCTG GGTGACCAAC CCTGGTCACC CCAGTGTGCA
 36781 AAGGCTACAG AACCGTGTG GAATGGCAAAC CACATGTATA TGCCCAATGA CATGTGGGTC
 36841 TCCCTTGAAA GAAAACAGGG TCATAACTGG GCCCAGAGAT CTGGGCATCA CAGAGCCAC
 36901 TGTCTGGAA GTAAAGCTGC AGGCTCAGGG AAGAGCCGGT CCACATTCA TGTCAAGGTG
 36961 GCCCAAGGGAA AAGGGGCTCC TGCTCTGGCC AGAGCGAAG AGCAGAGGCC TCTCAACAAAC
 37021 AACCTGCC TTGCTTCTCCT GGGAACCCAG CTCCCTCCCA CCAAACCCCC AGAATGCTGA
 37081 GACTCTGAT GTCGGAAGAC TCGGTGCTTC CTCCCTCATCA TATGCGGCAC TCATGATTGT
 37141 TACAATATTG TTGAAAATGT TAACTGAC TTATTGATGT AATTCCCACT TAAGGAATAG
 37201 TAAAGAGCAG GGCCCTGAATT TCAAACCCAT CCTGACTGAA TCCCAGGGCC GATGCCCTTG
 37261 GCCTCTGCTC GCTCTGCTTC ATAGGGAGGA CAGGTTAAGT GGTGGGTTAT CACAGGGGCA
 37321 GGGTGGTGG GATCCATCG AAGTGGCGAG GGCCTGGGT TGBAAGGCAT GTCCGATTCC
 37381 TTTCTGAATC CCCATTGCG AGCGCAGTTC CTGGCATGCA GGAGGTACTC AAGTAATTAA
 37441 ATTGAAGGAA TGAGGAAAGA TTGAAGATGA GGGTCCGGTC ATTCTCAGGT GCTCAAACGT
 37501 TCTGGCTTC TGGAAGATGG AACCTACTAT TATAAAAAT AATTGCTGCA ATGACTTTCT
 37561 GGCAAGTCT GATTGTAATT CAGAACCTCC CGGACACCTT TGGAGTGGAG CAGGGAACCG
 37621 TGGCCTGGA CAGAGTGGG AAAGCTCTCT TTGGGCTGCA GCAACACAA AAGTTGTTAA
 37681 TCTCGGCTTG TGCAACCTAA AGGGCACAGA TGTCAGCTGA GCCCACCTG ATGTCAGGA
 37741 TCGAACACCC CCTTGGAGAT CCTGTGCTG TGCCATGTTG CCAGGAAAG GACCAACCTG
 37801 TGCTGGTCG CAGCTCTCAC AGCAGCAGCC ACCCAACACA GGCTGGTCT CTTCTGGC
 37861 ACGTATTAGT CAGATCGCTG GATCTCAGAA TAGGAAGGAA CCTCATTGGG CCCACTGGT
 37921 GCCATGAGAA CGGGCCCTGC AGACCTCAA CTGCAGAGAG TGTATTGAC TGAGGACCCC
 37981 AGCTCTGATG CTCTGAATC CACAGCTGAG TCTGTGCTGA GGCTGCACTT CCTGTGACTG
 38041 AGTGGGGCAG GGATACTGAG GCTGGCCCGT TCTGGGGAGAC GCGGGACTCC TGTGACAGCC
 38101 AGCTTGGCT GGAGGACCCCT CCAAGGGCCT TGCAGACCCAC ACGGTGGTCT GGGATGCTGC
 38161 CGCTGACCT TTCTCCCTC CCCACCCCTT AAGGGTACAG TCGCTTCAAG CGCTCACTGC
 38221 TCTCCAGTCT TCCCCGGCTT CCTCCCCGTT TTTCCCTCAC AGGTGCTCTG GCGTGGCTCT
 38281 CAGAGGACGT GCTGCCCTAG GCAGGTTTGC CTCCAACACG AGGGTCTCAG CCATCCTCTA
 38341 AAGGGTACAC GGGAGACAAT CCCACCACCT TCCTCGTCA TAGAGTCCGA CACCTTATCG
 38401 GTCTTCCAGC CCAAGGATAC TCCCGCAGGT CTAACCTAAT CCCCTGTGTG CAATGGAAGT
 38461 CCTGACTCCC CTGAGGAAGA AAGAGAGCCA GGCATCAGCC AGGGCAAGGT GGGGCCAGG
 38521 AGAACGGCAG CACCTACCT GTCCCCCCAGG CAGGTCTCTA TCAGCAGGCC CCAGAAGTCT
 38581 GTAAAGGTGA CCTCTGTAGCC TTCTGGTG CGCTGCCGGA GCTCTCTG GAATTGCGG
 38641 AGTGGTCTG GGAACAGGAA GGGTAGCTTG TCCTTACCA CATGTCCTG AGCCTAAAG
 38701 ATAGCAGGCT CCAAGTTTT GGAGGACAG TCAGGTCAC GGTACAAGGT AGCCATGTC
 38761 CTGAAAGAG ACAGGCACAG CTGATAGGGC TGACCCGGCC CCTGCTGAT GAGTCTGGAA
 38821 CCCGGGCAGG AGGGAAAGGT GGCAGGCCAT TGAGAACCT AACCTGCCAG GAGGGAAACC
 38881 TGAGAGTAG GGGTAGGGCA GAGGCTGGC AGACAGTGGAG GAGCAGGGGA GGTGAACATT
 38941 ATTACTCATC ATCACAAAGTG CTACAGTTTC TCAAATGCC ACTGTGTGCC AGGCATAACA
 39001 GTGGTTAAGA TGAAGATTCTT GGATCCAGAA CTTTCTGGAT CAAGCCAGGC TCTGTGCTC
 39061 ACTAGCTGTG TGATCTTGGG CAAGCTACTT AATCACCGTT CCTCTCAGTT TTCTGTCTG
 39121 TGGAAATGGGG AGAATTACAA TAATTTGACC TGCTTCAGAG TGGTCTGGT AAGATTACAT
 39181 GAGTTAAGGT GATGTGGGGT AGCTAGAACAGA GTGCATGGCA CAGAACAAAT GCTGGAGCC
 39241 ACATTTAAC CCAGGCTCGA CTGATTACAG GTCTGGCTG TTTCTACTCT GCCATGTTAA
 39301 TTCTTATAAT TCTGATGTTG ACTGAGGATT TATTTCCCTT TCTCTCTTA ACTTACATGT
 39361 TGGTTGTCGG GGGGAGCAA ACAAAAGAGGT GATCCCACTA GCCCTGGTC TGGCTCAGT
 39421 CTGCCTCTT TTGGCTGGCA TCCAGCTTTT CTGCGTCAGG ACAGGAAAGG GCCCTTTCA
 39481 TTCTGCGCCT GAGAATGACC TCCCTGGCAG TGCACTAGTG CTGGGGTGC ACATTTAAGG
 39541 AAACAGGGC CTCTCCCTC ACAGGACTTG CTTTTGGGA GGTGGGAGGT GGTCAAGGAGA
 39601 TTGTTCAACA AAAAGCACCC TGCTCTCCAT AAGCTTACCC CATTCTTTT ATACTTTAAA
 39661 AAAGCTTCCA GAAGCCATTG CAAATGAGTT TTAAAGAAAGA TACGACTGTG GCTTCCCA
 39721 GGCCAGGGGA CCCACCCCAT CCCACCCCGG CTGGGCTGA GTTCTCTG CTATCTGTCT
 39781 ATCGCATAGT AGTCCGGGAG GACGCAAGGTG CCCCTCAACC TGCTGGAAGC AAGTTACCTT
 39841 CTTAACCTAC TACCGGGGGC AGCCATTAGT CATGGAGGTT AAGGATGTGG GCCCCGGACT
 39901 AGGTCAAT TCTACTCCAC TGTCATTTA TGAGTGTCT CTGGCAAAATT ACTCTGTTTC
 39961 TCCAGCTGAA AAATAATGAC AATGCTCCCTC TCATAGTTTT TTGTGAGGAA GAAACAAGAT
 40021 GAGCCATAGG AAGCAGTGC TCCAGTGTCTT AGCAAGAGGC AAATGCTCTC TTCTGTCTCA
 40081 CCTGGGCTGG CGCCTCTGCC CCCGGGGGGG GGTGTGTGGC ACAGGAGGCC TTGGCTTCC
 40141 AAGGCCATT CACATCCATT TGCTCACACG GGGGCTACCC AGAGCTCCTC CAGTATACAA
 40201 GGCACAGCAA GGGACCTCAG GGCTCACAGG CCCACCAAGTC ACAGACAGGA CACGGCCCTG
 40261 CTTGATAGGC CACCCCTGAAC CAGGGCCACG CAGGGTTTG ATGGGCCCA CAAAAAATGT
 40321 TTATATGTTT TTATTTTTTG TTATTTTTAA AGACAAGTCTC TCCCTATGTG GCCGGCCTG
 40381 GTCTTGAGCT CTCGGGCTCA AGCAAGGCTT CTGCTCTGC CTCCCAAAGT GCTGGGATTA
 40441 CAGATGTGAG CCACCATACC TGCCACATA AAAATATTA AAATTTATAT TTCTATGCTGG
 40501 CCACATAAAA AATATTTAA TTATATTTTC ATGACTGTGTT TATAAGACAA AATGATTCA
 40561 TATTATTTAT TAAACATTT TCTTCACCT ATAAGTTGT TAATGAAATG ATCCTGATTT
 40621 TAAAGAAAT TAACCTATTT TTGTGGGCCCT CAAAAGTAC AGTGAAGCCGC AGGCACTGTG
 40681 CCTTCTGTGG GTAAGCTGAC CTCACCCCTGG CCTCCCTGTG GGGGAACCCA GTGCTCAGCC
 40741 CCTGGCCCTG CTCACCCAGCA GAGGGTGAGA AGGAGGGAGG TCAGCTGCAG GGCCTGCCC
 40801 GTTGCATGTC CCCAGCAGTA TGGGGTCTAG TTCTGGGTCATACAG CCCCAGGGAA
 40861 ACAGTAAACAC CAGGAAAGGC CAGGGCATAC AGTGTCTCTG GCATACTTAG TAGGGGAGA
 40921 CGCTCTTACA CCCACCGCAT GAGGTCAAGAG CTACTGTTAT CGCAGTACACA

40981 AGCCGAGCCT CCAAGGAGTC AAGCTGCATG CCCAAGGCCA CACCACGTCA TGTGGGGCAC
 41041 TTCCCGAGCT TCAGGAGTCA CTTTGTCTT GGGTGGCGG GAGGAGCTTC ACCTGAGACC
 41101 CCAGGCCAGC CTGTTCTGG CTTGCCCTGG GCTCCACAGA CAGTGGCTCA TGTCCTCTTA
 41161 CCAGGTGGCC CCTGATAGAC CAGTGATGTA GCTGGCACAG TCCAGGAGGT TCAGCTCTG
 16 41221 CAGCCCCAGC AGGTGGCCAT ACATGGAGGT CATGGATCTT GTTCCACCCC CAGTGGCCAT
 41281 GATGGCTATC AGCGGCACCT GGGGTTACAC AGAGGCAGGC GGTTGGGGAA GGCTCCAGAA
 41341 GCAAAGCTTG GGAGAAGACT TAGAATGACG GCAACTTTGA AAGAACATC GAGGAGGCAG
 41401 AGCTGGCTAC AGAGGGCAGG AGTGCAGCG TTGAGCCACC TGACACCTGC CTTTCCACAG
 41461 AGCACTGGGT CTAGAAATAG ACCTCAGTC CACACCCAGA AGTTGGTGT AGTGAAGGAC
 41521 TCTGAAGCCT AACACAGGCT TGTTAAATCA CAGAAAGCTG TGGCTAGCAG AGTTCTTAGC
 41581 ACTGTTAGTC TAGCCCTCGG TTAACAGGTG ATAAAAGCTGG GGTCCGAGGG TGGTGGGCAG
 41641 CTTGTTCAAG AACCTACTA ATGTGCATC CCTCAATGGG ACAAAACAAAC ATCTTCTGTA
 41701 GCTGGATCTG AGGGAAAGGCA CCAGGCCCTGG CACATGGTGC CTGCTCAGCA GGTACTGTAC
 41761 TGGCAAGGCC AAGAGTCAAT GGACTGTTGA TGTGGGAAGG AGCAGGGTTG CCAGCTCTGG
 41821 GTCAGCAGAG GTTAAAGAGG CCCCTGGAAAT GTTCCCTAAAG AGATAGGGCC AGGTACAATG
 41881 GCTCACACTT GTAATCCCAG CACTTTGGGG GGGCGAGGG GGCAGATTGC TTAAGCCCAG
 41941 GAGTTCAAGA CCAGGCTGGG CAACATGTAG AAACCCCATC TCTATAAAAAA AATACAAAAAA
 42001 TTAGCCAGGC ATTGTGGCTC ATTCCCTGGAG TCCCAGCTAC TTGGGAGGGT GAAGTAAGAG
 42061 GATTGCTGTA GCCTGGGAGG TGAGGCTGC AGTGAGCTGT GATGGTGCCTA CTGCACTCCA
 42121 GCCTGGGTGA CAGAGCAAGA CCCTGTCCTA AGAAAAAAAAA AAAAAAAAAA AAAAAAAAAATA
 42181 GGAGGGAGGA AGGGGCTTGT TGTCAGACT CATTCTGTC GTTTCCCTT GAGTGGGGC
 42241 TGTAAGAACCC TTCCATTGGC TAAAGCGCTA TTACAAGTCC TCCTCGGGCA CCCGCCAGGT
 42301 CTCTGTCCTC ACCTAGGGTC AGCCTCAGAA AGCAGGCCGT TGGAGGGCTT GCCTGCTGCC
 42361 CAGGCAGGTG TGGGGCTTGG TCCACACAAAG GGAGGCAACA GAGGGCTGTG ACAGAGGCC
 42421 ACCGCTGCT CCCGTGGAAAGAAGGGGGGG TCTGGCCTA GGAGTGCAG TTGCTTCTCT
 42481 CCAGCTTCCCT AGCGCTGTTA TTTCTGGGAG TGTGTTTACCG TGTGCGGATT TGAGAGGAGA
 42541 GACAATTGAG GAAAGAAGGA CCTAAGGCCTT TTGGGTGGA GGTGTGGGT GCAGAAAGGTG
 42601 GTACAGCTTA AAATGTTTC CTGGCTGGGG TGAGGAGGGC ATTTTATCCC TGAGTAAGGC
 42661 CCCTTAGCC ACAGCAGGAA CTCTTAGTGC TTGGGAGGAG GCGCTGGGGC TCTGTCCTT
 42721 CTGCTTCAT TTTGCTCTCC CCTCTCTCTG TGAATATAAA ACAGAACATC TTCCAGCAC
 42781 GCTCTCACCA GGGCCCTGT ACCTCAGCAG TAAAGACATC CATGGAGGTG GCGCCTGGGG
 42841 TAGGGCAGGG CCTGGGTGG CTGGGAGGGC ATGGTGTGG AAGGACCAAG TGGGGCCAGG
 42901 CTGGCACCAT CCCTGATTAC CCACCTGTC CTCCCTGAGG TCTCCCTCCA GCTGCAGCAC
 15 42961 CTGCTTCAGG GCCTTGCCCA CCACGACCTT CCGCTCTGC AGAAACTCCA GCTCTGCTGG
 43021 GCACAGGCTG AAGCCCAGCC GCACGTCAG TGTCCTCAGGG CTGGAGGGAG AGAGGACCC
 43081 GTATTGGGCC TGCAGCCTCT ACCTCCCTGG AGTAGACACA GCTTTAGGGT CCATTGGACC
 43141 TAGGCTCAAA GCCTACTCTC GCCAGCGACC ACCTCTGTGA CCTTGGGTCT CCACCTTCCC
 43201 ATCTGTAAG TGGGGCAAT CGTGTGAAC TCTCAGCACT GTCTTCAGCA CTTAAAGAGA
 43261 TAACAAATGC AAAGCCCTCA AAAGCTTAGC TGGCACAAATCG AGAGCTGTTA CTCAACAAAAA
 43321 GATAACTGTTT GGCACTGTGT TAGGGACGCT TTGAGTCCTC TGTCTCTCG GCCAGATCTG
 43381 GAAAGTCAGG AAAGCTTCC CCCAGGTGGC TTCTCTAGCC CAGGGCCAGG CATGGGTGAC
 43441 AAAAGGCCTG GAGATCATT TGCCAAATCC CTCTGAACAAAG GGCCCCACGG TATGGACTCT
 43501 TCTGTGCTAG GTAAACAAAC TCCCTGTTCC TATGGAGCTC ACGTTCTCA ATGAGCTCTT
 43561 CCTGCTCCTA GAATAACAAAT TATTCTGTGT GTGTGTCATA TTGACAAAAA CCATCCCATT
 43621 CCTGATCTCA TTGTGTCACC TCATAGTCTG CAGGCAAGGTG GGGCAAGGCAC CTCTGTGACA
 43681 TGCACAATGC TTGTCTCTT TTTTTTTTTT AATCACATC CTGCCCCAAA CAAGTCCAGG
 43741 CTCCCTGGCCT GGCACTTCAAG ACCCTGCTG ATCTTCCAA GCTCACCTCA AACTACCCGC
 43801 GAAACTCTG GGTCAAGGTCC CAGAGCTCCC AGAACAGTC CTGTTCCGG GTCTGGCAC
 43861 AGGCCCTGCTC TTTCACCCAGA AGTGTCTTCC CTTTGCCAC CTCCAGCTG TGCACTCT
 43921 GAGGCCAGC TCCCCAGTGA CTCCAGATGA ATGTACCTTG ACCATTCCA ACTGAGCCTT
 43981 GGCTCACTGA CAAAGCCACA TCTCTTGAA GGCTGGGCCT AGGAGATTTC CATCATGTCT
 44041 CCTCTTGTC CAGCACAGGG CTGCAGACAT GGCTGTTAGA AAATAATTCT TTCTTGGCC
 44101 TCAACAGGGAA ATGAATAATA ATAATACAAAC AACAATAATA ACAATAACAT TAGCTGACGT
 44161 TTGCTGAGCA TCCCACCTG CAGATGGATT ACTGAGTACT TCGCACGTGC TGTCTAGCTG
 44221 GATCCTCAGA GCAACCCAGT GATGCAAGG ATCCTTGTAA TTCTCCCAAT GTACAGATGA
 44281 GGAAACACAGA GACTCAGTAA GGGCAAAAG CTTTCCAAG GGCAATAGC TAATCAAAC
 44341 GTCAATAGAA ATCATGGCCA CTTGATTTCT GACAATCAAG ATTTCTGAAA ACGATGTTGT
 44401 TTTCAGCAGA TGGTGTGAG ACATCTGGAT ATCCACATGC CAAAGAACGA GGTTGAGCCC
 44461 CTCTTTTACA CCATGTAGAT TGATTACCTA AAAATTGATC ATAGACCTAA ATGTGAGAGC
 44521 TAAAACCTAA CAAGCTCTAG AAAAACACAT GGAATATAT TTTCTAGGACT TTGGTGGGC
 44581 GCGGTGGCGC ACGTGTGTAAC TCCAGCACT TTGGTGGAGCC GAGGAGGGTG GATCACTTGA
 44641 GGTCAAGGAGT TCGAGACCAAG CTGGCCAATA CGGTGAAACC CGGTCTCTAC TGAAAATACA
 44701 AAAATTAGGC AGGTGTAGAG GTGTGCTATC GTAAATCCAG CTACTGGGGA GGCTGAGGCA
 44761 GGAGAATCGC TTAACCCAGG AGGCAGAGGT TGCACTGAGC TGAGATCACA CCACCTGACT
 44821 CCAGCCTGGG CGACAGTGTAG ACTCCATTTC AAAATAAAAT AAAATAAAAT AAAATAAAAT
 44881 AAAATAAAAT TTTCAGGACA TTGGATTAGG CAATGATTTC TGTTCTTA GATATGACAC
 44941 CAAAGCACA AGTGACAAGA TAAAAAAACA AATTTCCTCA AAAGTAAAGT CCTTTGTGCT
 45001 TCAAAGGACA CCAACAAAGA TGTGAAAAGA TACGCCATAG AATAAGAGGA AATATTGCA
 45061 AATCATATAT CTGACAAGGG ATTGTATCT GGCATATGTA AAGCACTT ACAACTTAAC
 45121 AATAAAACAAACACAAATAAC CCACATAAAAA AGTGGGCAAAGGATCTAATA GAAATGTCTC
 45181 CAAAGAAGAT GATATGCCAG CCAGCCCCATG GAAAGATGTT CAACATCATT AGCCATAGGG

45241 GAAATGAAA TCAAAACCAC AGTGAGATA CACTGCACAC CCACTGCTAT GAATAAGTGA
 45301 GAGGCACGGA CAGAATGCGG GTACTATACC AGGGTGTAGA CTTCATGTGT AATTCTAAC
 14 45361 TCTCACCC GGGAGGGAAA GATGAAAAGC TTGTCAGAGG TGACAATTCA GAGCCAAGTG
 45421 ATCCCAGATC CACACTGTCC ACCCAGACCC CACTCAGCAC ACACGTGCAC ACACAGGC
 45481 TCCATCCCTG CCTCTCTCTT CTGGACACAA TCTCCCCACC ATCTTCAACT CAGCCTCCAT
 45541 ACACTCTTG GACCACACTG TAGCTCATCA GTCTCCTGCC CCACAAGACT CCTCCATCCT
 45601 GCCTGAGAGA GACCCCTCACG AGTCACATTG TCTCTTGAC CTCCCCAGGC TCTCCCCACC
 45661 TCCCTTACAC AGGAGCTCCA GGGGAAGTAA GTCAAGGGGGC CTTCCCAGAC TCAGGC
 45721 ACCACAGGCA GGGTCATCAC CTGACCATCG AAGGGCAGT CGAGGGGCTG GCTGATGGGG
 13 45781 CCCTTCTTC GAGAGCGGCA GCAAAAGCTGG AGGGATGGGT GGGTTCTGTG AGAGGGGCTG
 45841 GGGAGCTCCT CTGGCACCCC CCACCTGTCC CAGTCCCCCA CACCTCCCC AGTGTCA
 45901 GCCTGTACCC CCACAGCTCC AGTGACTCTT GGGCACTTCC ACGTGCACCT GGGACTGGAA
 12 45961 GTACTTGGG TAGTGAAGC AGGCAGCGGT TTGGAAGCAG GCAGAGGTTG GGCAGCAGGG
 46021 TTCCAACGAG GGCGGGACAC GCTGGGTGTT CTCAAAGGAT TCGTTCACCA TCACCCAGGAG
 46081 GTCTTCACT GGGAGAGAGG ACAAGAGGGT GAGACCTCTG GAGCCTCTCA CTACCCACCC
 46141 GCAACTTGGT CCAGCAAAA ACTTATTGG AGGTATCCAG GAGGAGTTA GAATCACCTG
 46201 GGGCACCAGT TCAAATTCCC CAGAGATTCT CAATCTATAG GGTAGGGCAG GGCCAGGAA
 46261 CCTGATTTC AATGAGGTTT ATGAGGATGC AGACAAACAT AGTCCCCAGA AAACCTTTG
 46321 AGATGCTGAG AAGGTCTCA GGACCACTCC TTGGCTCAGC TTTGCCAGAA GATGCCCAA
 46381 AGTAACCTGC CCACTGTAGA GGAGGCAGAC CCTCCGCACA CTGTCGCTT TGATGTAGAC
 46441 GCCTTCTCTG TCCTCCCTCC CTCCCTCATG TCTGGTAGCC TTGCATGGC AACATGGGGT
 46501 GGGGTGGGA GTAGGGTAGG GGTATGGTAT TCATGGATCT CTTAGGTCTC TGCCAGCTAT
 46561 AAAGTGTAG GACCCCTCTG GAGCTGTGGT CTGCGTAGAG GAAGGGGCAT AGTTAGAGGG
 46621 GCCCTGGGCT TCGGGTGCT ACCACCTCAC TGACCTAGGC AATCTGAGGT CCTGGCATGG
 46681 CACAGCACAG ACAAGCGTGG ATAGAGGGAG CAGCACTGGG ACCCCAGACC TGAGATACTC
 46741 CAAGGTTATG TGGGTGCTTG GGGCAACACG CCCTTTACAG AGCAGGATGG GAGGCCAGGG
 46801 GCTCAGGACC CCTGGGGGTG AGATGATTCT GAGGGGTTT AGCCGAGCCTT TTGGTCCCCC
 46861 TGAGCCCAGG GAAGCCCTGG ACTAGGTGGA GTCTGATGCT GATGGGATT GGAAGCAGCT
 46921 CCCCTTGTG CCCCAGTAGG CAGAAAAGGC AAAACAGCTC ACTTTCTCC CTCTCCCTCC
 46981 GCCTCTTGA TTGTCATGA ACTTCCAGGC AGGAGACTTG TCGAGACTGT AAAAACAAAA
 47041 GGAGGGTCAC AGAGTGTCTG AGCACCAAGCC ACCTCCCTTG TCTTCCAGGC TGAAGGGAGA
 47101 TGAGGGACCA GAGGATACAT TCAGCATGG TCTGACTGGG CTGATGCTGG GGAATGAGA
 47161 ATGTGTTGAC CACATTTC TGAACCAATC AAGGCATGCC ATCTCACTCA GAGATAGGAT
 47221 ACTTGACCAC GAAGCTGTCC CAGGATGTCT AGGACCCACG GTCACAGTT ACTAGGGTC
 47281 AGGTCTCAG CTGCAAGTCG GGAATCTGAC CTGAGGGAA GTTCAGCTCT GGGCTTTCAA
 47341 AGTGGGGCT TCCGACAAA TAGCCTTGTCA AAGATGCTGA GGCAGATCAT ACTCCTTTGC
 47401 CACACATCC CAAAGACTCT TCTTTTTGAG GGAGGGTCT CTTGAACCC TGAGGC
 47461 GATGCATTTA ATTAAAAATAG CTGTACCAAC AGGCCAATA TGAATTTTG TTTTGGC
 47521 AGGGTCTCAC TCTGTCACCC AGGCTGGAGT GCAGTGTCA AATCACAGCT CACTGCAGCC
 47581 TTGACCTCTT GGGCTCAAGC GATCCTCCCA CCTCAGCCTC CTGAGTAGCT GGGACTACAG
 47641 ACACATGCTA CCATGCTGG CTAATTTC TATTTTTTGAG TAGGGATGGA GTTCACTAT
 47701 GTTGGCCAGA CTTGTTTGA ACTCCGGGC TCAAGTGATC CCCAACCTT GGCCCTCCAA
 47761 AGTGTGGGA TTATAGGCAT GAGCCACCAT GACTGGCTAT GAAACAATT TTAATACAGA
 47821 AATTGTGCTT GTAGGCCAGG CACAGTAGCT CACACCTGTA ATCCCAGTAC TTTGGGAGGC
 47881 TGAGGGTGGC AGATGCCCTG AGCTCAGGAG TTTGAAACCA GCCTGGCAA CATGTTGAAA
 47941 CCCGCTCTC ACAAAATACA AAAATAATAGC TGGGGGTGGT GGCCTGCTTG AGCACGGGAG
 48001 GCAGAGGTG CAGTGTCCG AGATCGCAGC ACTGCACTCC AGCCTGGCG ACAGAGTGA
 48061 ACCTGCTCTC AAAAAAAA AAAAAAAA AAAGTAAAC TGTGCTTGA GAATTACTGG
 48121 GTTGTGGGG TACAGGCCA AGACTGAGTC TGAGAGGAGG CTGCACTTAG TGGGGTTGTT
 48181 TATTGGGGAG GCTCTTAAGG GGAAGGGTCA CTTTTGGAT ATGAAGTGT CGGCCAGGGT
 48241 GATGTCCCCAG CTACCTCCCTC CCTCTGTACA GAAGCCCCAG CTCGTCCCC AGTGTGC
 48301 GTGCCCATG TGGCTGGCT CGTTAGTCT AACTCTGCA GAGGTAGGA TAGGACAGAG
 11 48361 ACCAAGCTGG GCCCACTGTT TTCTCCCTG GTCCCTGTTCC CCTGAGAGCA GCAGCCAGCC
 48421 GTGGCCTCTG GAGCTCAGG ACCCAGGAA GATAATTCA ACCAGCACGC CATTGGTGAC
 48481 GAGGGTCTCA GGTGGAGAGG GACTGAAAAA CAAAAGGAGA ACTCCTGGTC AATTCTAGCA
 48541 TTCTGACTG GTTCTCTCA ATGGCAGCAA AGGGAAATAA TAGGTTCAAT AACAAAATAT
 48601 CAGAGCCTTC TTTGTTCTCA TAAAAAAATAC TAGCTCGGA AACCTTAGCC TTATAGTTG
 48661 TCAGGGAGGA AACTGGCTA CATTGAAAC AGAATGTCCA GGAAAAAAAT AGATTATAG
 48721 GTTAAAGAG TAGGAGAGTT CTATTTTTT GAAAGGATAA CTGTTTTTA TGTCTAACCA
 48781 GATGCAAATG CATGGGGTTA GAAGAATAAA TAGGTTGTAA GAAATGCTG CAACCTGCCA
 48841 GCTAGGAGCAG ACATAAAACAG TTTGGTAGT GAAAATGATC CATTCTCCAT TCTGTGCTT
 48901 ATCTAACTTA TGCAATTAAAC AGTCTGGCC AGCTGAGAGG GTGAATGTTA AATGCCCTG
 48961 TGTTATAAAG TTGAGAACAT AAGCCTCAA GGTTCAATT AGCCTGGCAG AGTCCACAGC
 49021 ATCCCTTGTCA AAAATAAAGC CCAAGGTGAC TGTGCTCCA GATCAACTT ATTGTTGTTG
 49081 TTGAGAAGGA GTCTCGCTCT GTGCCAGG CTGGAGTGCAT ATGGCACGGT CTGGGCTC
 49141 TGCAACCTCC ACCTCTGGG TTCAAGTGTAT TCTCTGCCT CAGCCTCTG AGTAGCTGG
 49201 ATGACAGGTG CGTGTGCA TGCCCAACTA ATTTTTGTA TTTTAGTGA GACGGGTTT
 49261 CACCAAGATCA ACTCTTATTA CCAAGCATGA CATGTTTCAT TTTGCAAGCC ACTTTCTAAT
 49321 TCAGGCCAGG AGTTTCCCCC GCCCTCTAAA AATTCTGTCA AATCAAGAAT TGGAGGGAG
 49381 AGCTTCGATT TTCTCAATG TTAAGTAAAT AACGTTGATT TTAATTGCC AATCAGCTGT
 49441 TGATGGGAAG TAAGTCCCT CACATAGTTT GAAGATCCCC TAGACAGAGG ATTCTCAAG

49501 TGGCAGTTGA CGAGCCCTCT TGTCCTGTCT CCCACCTGTC ATCCTCCTTC TGTTCCCGGG
 49561 CTCATCACAA GCACAGGGGG CTCAGCCCTG CCAGGTGAAG AAGAGGGAGG AGTGAGAAC
 49621 AAGCTGCCCA GTTACAGGGC ACTCTGCCCG GCCTGAGGCG GGGGTTGCCT TTCTTCTT
 49681 CTTCTTCTC TTCCCTTCCT TCCTTCCCT CTTCCCTTC CTTCTACCT TCCCTCTTG
 49741 TTCCCTCCTT CCAGACCTCC AGCTCTCTGA GTCCTCAGCA CAGTCATGTC TCCAGGATCC
 49801 TGGATCGTT TCCTAATTGA AAATGATTG GGCTGGCGC GGTGGCTCAT GCCTGTAATC
 49861 CCAGCACTT GGGAGCCGA GGCAGGTGGA TCACCTGAGG TCAGGAGTT GAGACCAACA
 49921 TGATGAAACC CCATCTCTAC TAAAAATACA AGAATTAGGC TAGCATGGT GTGTGTTCCCT
 49981 GTAATCAGG TTACTCGGGA GGCTGAGGGG GGAGAACATCGC TTGAATCCAG GAGGTGGAGG
 50041 TTGCACTGAG CTGAGATCGC ACCACTGTAC TCTAGCTGG ATGACAGAGT GAGACGAAG
 50101 AAAGAAGAAA AGGAAAGAAA AGGAAAGAAA GGAAAGAAA GGAAAGAAA AGAGAGGGAG
 50161 AGAGGGAGAA AGGGAGAAAA AGGGAGGAAG GGAGGGAGGG AGGGAGAAA GAAGGAAGGA
 50221 AGGCAGGAG GCAGGCAAGA AAAAAGATTG GAAAGAACCC TGCAAGGTGG GCCTGAGCCT
 50281 CCCATTGACG TGGTCCAGA GTTGGCTGCA GAGTTGGCTG CTTGCGGCCA GGGTTGGGGT
 50341 GTTGTCTCA ACAGTCTGG TCAAGGAGA GACCCCTGGG CTCTGACCTC TTTCTCAGCT
 50401 GGCGCTAAC AAGGGCCCCA TCAGGCTATGGT GGCCAGAGTA CTAGATCTT AAAGCAATGC
 50461 GGAATGTAGG GGGCCTGAGG CAGGCTATTG GGCCAGAGTA TCCTACACTA TGGAGAACG
 50521 TGAGGGCAGG AGAACTGGGT AAGTGCAGTG GGGCCAGGGC GCCAGGAGT AGACTGGAAG
 50581 CCTGGGTGAT TCTGATGCAC AGTGAAGAGT CAGGGTTAGT GCCAGGAGT AGACTGGAAG
 50641 AGGGTGGAGC AGCAAAAGG CGCACCTGCC TCTGAAATGC TCGTTCCCAG AAGCTACTTA
 50701 ATGGCCAGGA CCTTGGTGGG CCCCGGGGCC TGGGCTACT CACCTCTCTC CCAGCAGGAA
 10 50761 CTCCACCTCC AGCTCTTCCA TGCCCTGTA GGGGAGGGAG GGAAGGCAGC TGTGAGTGGC
 50821 GTCTGCACAT CTCTGCTGAC CAGAACCTT TGCCACCTG GGAGCCCTG TGGACCTGCCC
 50881 TCCGCTCTC CCCACAGCT GCCTGCTGCC ATGGAGTGTG CTTGTGGTCT CCTCTGCCAA
 50941 TCTCTGTC CAACTTTG ACATCAGGAG CATCACTGAC AGCCCCAGGG AGGTCAAGGGT
 51001 GTGAGCTGA CCCCCAGCCC CAGCATGTAC ATTCCCTGTC CCGAGAGCAT CTCCCTTCTG
 51061 TCACCTCCCT GGCTTCCAC CGCTCTCCTG CTTTCCAAG GGACTCAGTG CCTGTGGCAA
 51121 GGGAGCGCTG AGATAAGCCT TCAGAAAGCC CATTCTTCC TTCACTGAGT CCCTCACACC
 51181 TCTTGCTGTG TACTGGAGC AAGCCAGGGC CTCACTGGC CTTACTTTCC CCACCCAGAAT
 51241 GTAAAGCTAGG TGAGGGCAGG GATGCTAGTT CCCAAGGGGA CCCAGTGCT AGAACGCTGC
 51301 CGGCCCTCA CAGGCACTCA AATGCTGGCA GAATGATTGA ATTGACAGCA TCCTGGGGGC
 51361 TCCAGCTCTG ACTGTCTGAT ATCTGAGAT TCGCCACCT TTACCCCTCAT CTGTCCTG
 51421 TTCTGGGGC CATTCCAGCA ATGGGACTCT GAGGGAAAGA AAGGAGGGAA GGGAAACAAAG
 51481 ACTTACTACC TCTCCGGTGT GTTCTGTGAA CACTGCTAGT TGCCCTGAATT ATGGTTCTC
 51541 ATTTAAGTCC CACAACAACC TTGTTGGTA GGTATTATGA TCACCATTTT GTAGATGAGG
 51601 AAACAGAAAC TTAGGGAGTT ATTTTACGA CATCACATAG TGAGTGACAG TGCTGGGATT
 51661 TAACCAGAG GCACCATCTT TTCTCAACTT GACTGTGTAC ACAGAGAAAGA CCCTGGGGAA
 51721 TGGTTTGTG GGAATCAGCA GCAGCTATGA GTTCTATGA AAGCCGGAGT CTGAGGGCCT
 51781 GTACCTGTCA CCCACCTGGC GTTGTAGTGG AAACCTCACG TGGGTTTCTC TTCGGAAACAA
 9 51841 GAGCTTGGTG AGGTCTAGA GAATGTCAG GAGATGTC TCTGGGTGCA CTGTGTCTTC
 51901 ATCACAGACA CTCAACTCTA GCACGTTCTA GGGGAGAAGG AAGGATGCCA GTCTGGTTAC
 51961 CACGGTTGCA TTGACCCCTT CCCAGAACAA GGGGCTTGAC CTCTGCTCTA GCAAGGAGGT
 52021 AGGTCTGGGG GATGGGAGAC AGCCCAGAAC TCCCTGCTCA GATGTGACAG ACATTGTTAG
 52081 GGAAACGAAT TAAGGAAGGC AAGTGTCTAT TCCCTCTTTT CCAACTCCCT GAAGCCTCT
 52141 TGGCCTCTG TTCTCACTTT CCTGCTAGAG AATTTCGCTT CTTTTACAAAG CTTGGTAGAG
 52201 ACCTGTGAAG TATGGAAGGA TTAGTTTCAG CCTGAGGGGA AGGATGGATG GGAAAGAGCA
 52261 AGCAAGACCCC CTGCCAGGAT ATAGAATGTG GTCAGCACTG GGCCCTGGG CACTGAAGAC
 52321 AGCCCCTTG GTCTAGAAA ATCCAGATTA GGGAAATCGAA CATCTAACAC TCCAGGGAGT
 52381 GATGTGCAA AGCTTGTAGG GTTCTGTATA ATTGGGGGG TATATGAACCT CTATCTTGT
 52441 ACTCATCAGC CCCTAGTAAG TTCTCATGGA TGGTCCAAA GCAGCTACCT TGTCTCATCTGG
 52501 CTGAACCTCTG AGGTCACTGT TCTCAGGTGA CATGGACTAA TCTCTGAGGC CTCTTGTCC
 52561 AAGGATCTCTG GGGGATGTGG GTTGAAGCTG CCTCTTAAAT ACATCTAGTG GACATAAGGT
 52621 GTTCTGGAG GTTCCCAACG GGATGTGATT GTTGCACCTTG TAATTACTAG AAACACAGGT
 52681 ATGTCCTCAGA ACCAACCCAG CAGCTCCACA ATCCAAAGGA AGCTCAAATT CTGATTCTCA
 52741 GATGGCAGAC TGAACCCACA TCAAAATTCT CCTCCCTTAC CAAACTCACA TGTATGATAG
 52801 AAAAGGTATT TAAAAATAAA TAATGATAAA GCCCCCTTGG AAAAGATACG TCACGTAACA
 52861 AAAATAAAATA AGTATCCCTA AACATGGAG GCAGTAAGGT GGAGCTGTCA AGCCTCAAGC
 52921 ACCTCCAGAC TGGGGCTGAT GCCACTATGA TGAGGCCAAA TTTGTTGAA GGGATCTCTG
 52981 CTGAGGTGCTGT CACAACCTCG GATGCTCAAA GGATAATTAT CCCATGGGAG ATGTGAACCTC
 53041 AACTTAGAGT CATCATATAT CTGAAGAAGA CCAATCCCCT GGAACAACCTC AACAAATAAA
 53101 ATACACACCC AAAGAAATGG AAATAATAAA GCAATATGAA AAGAATTAA GAGGGAGAGA
 53161 ACCCTCTGGC TGGAATTAGA CTGGCCTCAG ACCTAGTGAA TAGAATGGTT CTTCACCTCT
 53221 GTGTATTCCTC CTGCTTGGAG AAGGGCTCTGG AAACACAAATC TGTTTGGAT TAAGAGACAC
 53281 CTAGTTCTAG CAACAAATTAG GACAAGAGAG ACCAGGAGAA AAAGCTAGAA AGAAAGATAG
 53341 ACTCAGAGAG GGGCAACAT GAGACAGGGCA GAAGGAAGTA GTGGGTGTGA GTGGGAAGAC
 53401 AGGTTGTAG GACAAGAAGT AGGCAGAAAG GAAAACAAAG AGGCATAAGA ACACAGAGGC
 53461 AAAGAGAAGG ACAAGTGGGC AAGAGGGGAG TAAATGCAG ACAGGGAGAAGA CAGGTGCAGT
 53521 CCTCCATGTC TCACCTTCAC TCGGCTCTGG ATCTGGAAGT TGAAGCTTTC ATTCCACCTCT
 8 53581 GGATTTGGGC AGTTGGAGAT GGTCTTGTGTC CTCAGCTTCTC TCTGAGAGGGC GGTGGGCAGC
 53641 CAGAGGTGCTCA CAAACAGTC TGTCTGGCTC ACTGTCAAGG AAAGATAAGT GGAACCAATC
 53701 CAGGTTCAAGA GAGAACATGT GGAGTCAGAA GGGAACTTCA GGCAGCCTC TTCCAGAGGCC

53761 AGGCCCATC TCAGCTCCAG TTCCCTCTCCT CATTCTATGGG TCTAACTTTC CAGAGACTCA
 53821 GATCCCCAAG AAGTCACAGA TTGCAGGCTG TGGGTCTCGC AGAGGGCGAC CCAATCCCAC
 53881 TGGCCATTCT CATAAGACACA GATGGCTCC GGCTTATT AGAAGCACCC TTGTCCTGAA
 53941 GCAAACACCT GCTGGTTGG GACAAGGTAT GGTCACACAG TTTCCCTCT GCCAAGGTAA
 54001 GATCTGGAAG ACCAGTCCCC AAGGGTCCGG ATTACTTCTT GTCTTAGCCC AGGGGACCAG
 54061 GAAAACAAA GTAGGGCAA CGCCTCCCT TGGGCTGAGG AGAAGTGAGG ATGAGAAGGA
 54121 ACACCTGCT GTGAGCTGGC CAAGAGGAAG GGCCCCAGTG GGCAAGGACT TGAGTGAAC
 54181 TAGCTATCA GCAAGATCTA GGAGCAGGAA CCCCATCTC CATATGGCTT AATGCCCTCA
 54241 CTGCACTCCAG CCAGGCTCTT CTTACCTATC CTGATCTGCA GCATCAGGTT ATGGCCCAAG
 54301 AAACCTCCC CTCTGGCC ACATACTTCTT TTTTCTTAAATGAAACAA CCCAAGTGTG
 54361 ATGTTGCTGG ATATGCTCTG GAGAAGGGAC AGAGGCTTG ACCCAACAA GCCCTCCCAA
 54421 AATGACAGAA GGTCTCTAA GTCTTCCACC TGGAGCAACA GAAGGGATAG GGTGAAAATA
 54481 GTAGCCCAGA CTGGCTCCCT GCTCTAGTCT ACCTGGGAGG AGGGAGGCAG AGGGGCCATG
 54541 ACTGCCCTG TCACTCTCCT GGCAGTTGGT GAGGGTGAAGG CCCAGGGTAC ACACAGATA
 54601 CATAGGAAAT GGTGAAAGCC AGCAGTTCC AGTGTAGGAG AGTCCTCTG ATGGTAAAGC
 54661 TGCGAGGAGG ATGGAGGTTACACTGGT GGACTGATTAAACTAGTTT TTAATTATTA
 54721 ATAGTCATG ATGAGAGTGG AGCCGGCTT TTCTTCGCCC TGCGTTGTCA CCACACTGAA
 54781 AGGTTATCA GAATCAACAG CCCTGATCTA AATCACAGGT GTTCTAGCT TCTGATTTC
 54841 TCCCTGAAGT CTGGTTAAC TTTGAGGCTT CATCATCCAG TGAATTGGAA TTCACAGCCC
 54901 TCCTCCTATT GAAGAAAGGT GAGATGACAC TCTCTTTAG ATGTCATGAT AATCACTTAA
 54961 TCCAGACAAA AAGAATGATT GAAACCAAC AGAATCAAA GAAAATGATG ACCTCATCAA
 55021 CCTACTCTTC CCTCTCATGC TCAACCTGCT TTCTCTTATC ATCACTCAGG TGTCAGGCC
 55081 AGAACATCAGT GTCATCCTTG ACTCTCACAT TCAGTTGATT TAACCCTTG TCTCTCTCAG
 55141 GACCCCAAC TTCTCATCTG CCCTGACTG GCCCTGGTGT AGACCCACAT AATCTCTGAG
 55201 GCCACTTCCC AGATAGGCTT CCCATCTCTT GTCTTGGCCCC ACTCTAATCT TTTCTCCAAA
 55261 GAGAAGCCAC AAGGAGCTTT TTAGATATGA CCATGCCATT CCCAGTTG AGATCCCTCA
 55321 GGGGCTCCCT TCTGCCCTCA GGAAGGAGTC CCAGGCACTC AAAGTGGCTG ACCCTGCTCA
 55381 CTTCCCCAGC CTCAGTCTCA TCTCTCCACC CTCCCTAATA CCCCACCCCT GCCCTGTGTG
 55441 TGCGAGCATG TGTATGTGTG CACACATGTA CATTACACACA CATAACGATT
 55501 CACTCAACT TACACACACG CACTCACCGT TTAGACTCCA GCCAAACACA GTCACTGTA
 55561 TTCAAAACTG CTTTGCACTA TGAGCTCCC TCTGCCAGA ACACTCTCC ACTTTCACCC
 55621 ACCCTCTCCC TTGCTAGCCC CTTGGGACTT CAGTAGGAAC CTCCTCTGGA AGCTTCCCT
 55681 GAGCACCCAA GTCTGGACCA GGTACCCAG AGCCAGAGT CCTCTGCTCC TCTCTGGGG
 55741 AGCACCCCTA TGCTGCTCTG CAACAAACAGG TGGACTCATC TGTCTCCATC ACACGTGGC
 55801 CACAGGATAC AGGTACAGAT ACATCACTGC AGATACAGGT CAGAGGTAG TGTGTCTCAT
 55861 TGCTCTGCTG ATAGGGAGAC TCACCAAGGAA CCTGCTCAGG CCTCTGAAAC CCTTCCCTG
 55921 GCTTTCCAG AACACTGGCC CCCCAACCCC GGGCTGGAC CACACCCCTC AGGTCTGTCT
 55981 CCGGTACGCT GGGAGCCTGC AGGCAGGGGT CCAGCTCTAT ACCTACTGGT TCTGTAATGC
 56041 ACTTGGCCCG CATGATTGA TGGACCCCTG AAGAGCCCCCT TCCCTCTAAA ATTGTTTTCC
 56101 TGTCATTCTG CTGCTCTTAA ATTCTCGCA TAAAGCCCAT TAAATAAAAAA CTCAGTATT
 56161 TTGGCCTAAG GCTTAAATGT CTAGTTAAGA TAAAGAAGTA TTCTGGCAG AGACAAGATT
 56221 GATCAGACAC CCACATTCCA AGCATTAGTT TTTCAGTTCT TTTGCAGAGA GCTGCTCCC
 56281 CTGTCACCC CCAGCAATGA ATGGTTAGGC AACGGTGACC TGGTTCTGAG AATTTCATGG
 56341 ACTATGGGAA TTTCCTAGGC ACAGATAAGT TCTAATCCTT GTTCTTAGTG GAGTGACAGA
 56401 GGGCTTGGTC ATCTGAATGA CATGTCCTCT TTCTCTGAAAT GTGGGTGGCA GAGATGACTC
 56461 AATCATCTCT ATAGATGTTC AGAACCTTGT CTCTCTAATA CCTCTGGAGA CACCTGTCTG
 56521 ATATCCCTTG GGGTCTGGT TTCACTTTA ATTCTCAAG AGAAACAAAAA ACTCTGACA
 56581 ACTGGGAGGT TTCTAGGCCT GATTGCTCT CAGTCTGGC CCAAGGCTCT TTCACTTATC
 56641 GAGGAGTCAG CCCTGAATCT TGAGTATTCA CTGCCATTG GCCTCTGGAA AACTTTAGGA
 56701 GACTCAGGCA CAAGTTACAA GACTAAATGG AAGTGTGTT TATTTGGAT TCAAACAGA
 56761 AGGAAAAAAAT TTGGGACAAG GGTAAACCCAG GACCTAAAAG ACATCGGGAT AGTCTTAATC
 56821 TAACCCAGGC AAGGTAGCAA CACAGCCCAG CCTGCAGCAA CTGTTGAATC CAGACAGAC
 56881 ATGACTGGG TTGGCACCCC GAGATGCCAC ATGTCCTTC CGCCTTGTTC TTGGGGAAA
 56941 TGAGGAAAAC GTAAGCAACA GGGGGTTTC CAAAGAGGGC CAGGTATCGC CAAAGGGCAG
 57001 CCTGTCCTGCC CAAGGGCAT TGAGGGAAAG GAGGGGCTT GTAGAGGCCAG ACTACAGCCT
 57061 GTTCTCTAG CTGCTGCTC TTCTCTCACAC CCCGTACCT GCTGAGTGAC AGGTGTGGCA
 57121 GGAAGAAGAC AGAAGAAAGG ACCCTGGACA GTGGGAAGTC TGAGCATGCA GGTGCAGGCA
 57181 TGCATGGGG CATGCAGCCA GGTGACTGGA GAGGCTGGGA CAGACACACC TCAAAAGGAG
 57241 TTGGAGGGCT GGGCATAGTG GTTCACGCCT GTAATCCCTG CATTGGGA GACTGAGACA
 57301 GAGGATCACT TGAGCCAGA AGTCGAAAC CAGCTTGGGC AACATGGTGA AACCTGTCT
 57361 CCACAAATAA TTTTTTTTT TTAAGGTAG CCAGGGTGTGG TGGCCCATGC CTGTTAGTCCC
 57421 AGCTACTTGG CAGGCTGAGA TGGGAGGATC ACCTGAGCCC AGGAGGTCCA GGCTGCAGTG
 57481 AGATGAGATT GAGCCACTGC ACTCCAGCCT GGGCAACAGA GTGAGACTCT GTCTCAAAAT
 57541 AAAATAAAAA TAAAAATAAA AATAAAAAGC AGAAGGGGTTT GGAGGAAGGT AACCACTGCT
 57601 GTGGATATGT GGGAGCTGGG TTGTCCTAACAG GATGAGGGTG AACTCTGCC AAGAGTGGGA
 57661 AAGTGTCTGA GTGACCCACCG TGGCATGGAC AAGAAAGGCA AGAGAGGTGG CCACGCTCTG
 57721 GACCCAGAGG ACACCTCCCTG GATCCCTAAA TCATTCACCG ATAGGCTGGG AAAGACCCCA
 57781 AGAAGCTGAGA GGAAAGTCC CTGGCCCTTT GTGGGGCCAC TTGGCAGACA CAGACCCAAA
 57841 CTGGAGGCTA CTGGAGGCCAC TCCCTGCTGGC GCAGGGCTCTG GACACATTCG CCTCCAGAGG
 57901 CTCTGTCTGC GGAATCCTGC TTGGTCCCTT TGCTGTCTCT GGGGCTCAGG GCCCTCCAC
 57961 ACTTATCCAT TCACACTCAG CCCTGCATTG GAGTGCACAGC TTCCCTACAGG TTCCCCAGACT

58021 CTCCGCACCT GCACTCTAGC ACAAGCTGCA TACAAGGCCT CAGGAAGTCA TGCAATCTCT
 58081 TCCTCAAGAC CCCAGGTACC CAAACACTTT GAAGGCATCC AATCCCTCCT CTGCCTCACT
 58141 TCCTGGTTCT CCATACTCTGG CTGTCTCCT CTCTTGCCAC ACTCTCCACA CCTGGGACCA
 58201 GACCCTAATT CCCTCTTAGG CAATCATTAA CCTCTTGTG CAAAATGGGC CTCCAGGGCC
 58261 TCTCCTCAAG TTGCTGGGAG CATCGGGTAG TGTGTGGATG TCTGGCTACA GGTCAGGCC
 58321 GCAGGTGCTC AGGGGGTGTG GCTCCCCCTCC CTCAGGACAG CCTCTCCCCA AGATCAGGGCT
 58381 CCCAAGGCCT CTGGATACTC AGGTCTGCTCC CCTCTCAGAC CCCAGGCTCT CTGTCAGGCC
 58441 TTTCTTGGCC ACTCCAGCT CTCATTGTT TCCTCTGCCT CTAAACTCCC TGTTGACAAG
 58501 ATGGGAACGT TGTATTCTA TAATGCTTT GGACATAGAT AGCTCCCTT GTAAAAATGT
 58561 ACACCTAAC TCAAGGGCTCT GTGCTCTTC TGCCCTCTC AGGCCCTCCC TAGCCTTCG
 58621 GGATCCCTGT GGCGCTTGGAG TCATTCTGT GTGGTTACCG GTGAGCAGGG ATGCTCCCCA
 58681 CCTGTGGGGC CTGGATATAC AGCTTTTCC CCCTTGGGGG CCAACTTCCC CTCTGGATGG
 58741 GGCCAGGGT GGGCAGTGTG GGAGGGGCTT TCAGTTCTCC TGATCTCTTA AAATGCCACC
 58801 CAAAGACAGG GGTGGGGAGG ATTCTGCCA CTGAGATGCC TCCATTGGCC CCCGTGGCCT
 58861 GGAAATCGAG CAGGGCTCTGG GATGAGGGGG GAGGTAGGG CTGCCAGGAA ACAGCACTGC
 58921 ACAAACAGTG CTGCTAAGTC AGGCCATAC CATGGCTGA ACTGTAGAAC TGGGTGCTGT
 58981 GGACTCTGGC ATGGTCCCCA GAGCCCAGTT TCAGATCTT GTCTTGAGAA GGGGCTCTGG
 59041 GAGTCACCTG GCAGCTGGCT CCACGATGCC TAAGTGCTC TCCTTCCACT TCACCCCAAC
 59101 AGGCCCTAAC CGAAAGGAGG AGCAAGGAGT GCTCCGTGT GGCAAAGCAG GTGCTCTGG
 59161 AAATCGATT CAGGAAGAGA GAGTGGGGAG GAGACAGAGG GCGAACAGAG CTCTGGGTG
 59221 CAGGCACAGA CTGGGGGCA GGTAGCGCCG AGGGAACAGC ATGAACAAGG GGCCTGGAAG
 59281 AGTGCACAGA ACCCCAAAGT CTGAAATACA ACTTGAGTAT AGAACACACA TTCGTGACA
 59341 TACATGCACA CACATACACG TATACACCAT GCACGTATAC GCGTCCGGCC GCACACACAC
 59401 ACACGGATCC ACCTGACCTC TTCTCCAGCC AGGGCCTCT CCATCCAGAT CCGGTAAGCA
 59461 GAGGAGAAGG AGAGAAGTGA GGTGAGTACT CACGATATAC AGCCTGCCGG ACATTTTCA
 59521 TCCGGATGAC CCTCACTGTC AACAGGTGGC ATGGAGACAG CCCCTCTGT GGAAGGAGAG
 59581 GACAGAGGAC TCAGTCTTCA AGCATTACCC CAAGGCCATT GCCCCAGGGG AGACTTTCCC
 59641 GCTATGCCTC CTCAGGCCGG CCCAGTTGCA TTACAACAAAC CCCTAGGCTG GTTTTTTTT
 59701 TTTTTTTTT TTTTTTTTAC AGAGTTTCAC TCTTGTGCTC CAGGCTAGAG TGCCATGGCG
 59761 CGATCTGGC TCACTGCAAC CTCTGCTCC CAGGTTCTAG TGATTCTCT GCCTCAGGCC
 59821 CCCATGTAGC TGGGGTTACA GGTGTCTGCC ACATACCCAG CTAAATTCTT GTAGTTTTAG
 59881 TTGAGACAGG GTTTCACCCAG GTTGTCCAGG CTGGTCTTG ACTCCTTATG TCAGGTGATC
 59941 CACCCACCTC GGCCCTCCAA AGTGTCTGGG TTACAGGCAAC GAGCCACCCAG GCCTGGCTCC
 60001 CCATGCTGT TTGCTATTGT TTGCTGGGGT TTCTGCTGCTC TCTTCAGCCT GGAAGGTCTT
 60061 GGAAGCTCCA GAGACCTGG GACCCTGACT GGTTTTCAT CTTTGACCT CCAGTACCTA
 60121 TAGCATGCTC ATATCTAGAA GGTGCTAATA CATGTTCAT AAGTGAGTGA ATGAATGTT
 60181 CAGCCCCATG CACAGAATGA AGCCCATGG CAAGGGAGTC CTGAGGTGTT CTCCCTCTCC
 60241 TTTCTTCACT GAATGCTGC TTATTGCCC TCTCCTCTA CTGTCACACA TGCAAGATATT
 60301 TGCTTGTGCC CAGGCAAGAT GGCAAGCTGC TTGGGAGGCA GGCACAGTGT GGCTGCGGG
 60361 AGGGGGTGA CAGACACTGA TCACGCACGG TGACATCTAC TCACTGGGAG ACCATGTGCC
 60421 AGGGGCTCAC CCCACTGGT ACTTACTTAG CTCCAGGCCA GGCCTTGCA TCTCCCTTG
 60481 CCTTGGGCC TCTTCCCCAG GGTATCTCCC CCAGGCTCTT CTGGTCACCT TATTTTCTT
 60541 TCCCAGGAAA CAGTCTGCC CTCTCCACTC CACTGTTCT CGCAAGAGGT CCTTGTGGAG
 60601 TCTTGGCTGG TCTGCGATGA CCTCTCTTC CTGATGGCAA AAGCTCAAGT AGTCTTGGT
 60661 TTGATGACTG GCTCTGTCA GATGCTCTG CGGGTCATGA TAACCCCAAGG GGACCAAGAGC
 60721 CCTGAGGTGT CCCCTCTTCCC TAGCACCTTG TCCTAACAAAC TGGGGTGTGG GAACACTCCC
 60781 AGCCCTGCTC ACGGGGTGTCT GATGGCTCTC CCTTCCGTCC TGCCCTCTT CTAGTTCTGA
 60841 TCACCTCTG GGGGCTCTGC CATCATCTCA AGGCCAGGTC CTCAAGGAAC ACCCAAGAGC
 60901 CTCCCTGGGC TGCACTGTCT GGGGCACTGG CTGTGGTTCC TTAAGGCCCA ACCTCCCTCC
 60961 CGAGCTGCC CTGCACACTG AGTCTCGAG CAGCCACTAC TCTACTTTAT CCCTTTGG
 61021 CCCATTACTT CATTACTTC TGCCCTCCAG AGGCCCTGAA ACCCCCTTCC CAAGCCGCAC
 61081 CAGCTCTCT GGACCTACA GCAACAGAGT GTCGCGGAT CATCTTCT GCCTACCGAG
 61141 GGACCGCAGC GCTCTCCAG TGCCCTCTG CCTGTTGCC TTGGCCCCAG GCGTTAAGCC
 61201 TGTTGGAGCC GCTCTCTGTTT CACCCCTCTC CTGGCCAGGC TCTCTTCCAT GGGCTGTTT
 61261 CAGAGCTCAG GGTCTGTGT CTCGCTCGTA GGGCTCCCTG CCTCTGTGCTC TCTGAGAGTG
 61321 CCAGGCTCT GATTGTTCTCT CTGAATGAAT GAACCATCTT CTAGAGTATT ATTTTGAGC
 61381 ACCTACTGTG TTCCAGGCC TATGCTAGAT GCTGGGGGGC CAGGGGTGAA TACGAGGAG
 61441 AGACAGAGGT TAAAGGCAAC AGCGGCCAGT GGCTTGATG GGGAGAGCAG GTCAAGCTGG
 61501 CTACAGGTTG TTGCTTCCA GGCTTGGGGC CTCTGGCTGG CTCCAGAAC TTCCCCGCCA
 61561 TCAGTCATGC TCCTGGACTC TGAAACCTC TTCTGGTCCC TAGGGCCCA AGGCTGCAGA
 61621 GGTGAGGCGC GTTCTACTCT TGAGCCCGCA CACCCGGGAA GGGGCAGCCT CGGCCGGTT
 61681 CTCTTCAGGT GTCTGGTTC TGTTCTGGT TCATGCTTGGAA CTCATGACTT
 61741 GAAGCCCTT CGCTGAGAGT TTACCCACG CTTAGAAGAG TTTTGTGTT CGTCATTCC
 61801 CAGCCTGAGC TCACGGTAA TGCCCTACCAA CGGAATGTCT GTATTATAG AAGCTCTGA
 61861 ACCGGCACCC TCGACATGCC ACCCTGAATG TTCTACTCCG TGTTGCCAGA ATCACTTT
 61921 ACCTTTTAA GAGGATTCAA TTCTCTCTC TTTAAATACT ATTGGGATTG GGGCAGGAG
 61981 AAACAAGATA CGAAGAAAAA AAAAGTTCT ATAATCCAC CACTCAGAGA CCATTTTT
 62041 ATCTTTACTT TTTTTTTTTT TTTTTTTTTT GAGACAGGGT CTCACCTGTT TGCTTGGCT
 62101 AGGGAACTT TTGATGCTGG TCTTCTCAGA CTTCCTGGG GCTTAATTCT TTTTCTTTG
 62161 GAGACGGAGC CTCACCTGT CATAGAGCT GGAATGCAGT GGCAGGGATAT CTGGCTCAC
 62221 TGCAACCTCT GCCTCCAGG TTCAAGCAAT TCTCCCTCT CAGTCTCTG AGTAGCTGGA

62281 CTACAGGTGC TCACGACCAC GCCCGGCTAA TTTTGTATT TTTAGTAGAG ACGCAGTTTC
 62341 ACCGTGTGG CCAGGCTGGT CTCAAACTCC TGGTCTCAGG TGATTCTGCC TACCTCAGCC
 62401 TCCCAAAGTG CTGGTTTAT AGGTTGAGC CACTGTGCC GGGCCCTGG GCTTAATTCT
 62461 GATCAGCAAC AGAAATGCTGA TTCTGGGAAG GTCACTTAT TTAGCCCCCTT TAAAGGTCTC
 62521 TTCTCCTCTC CTGAACAGCT GCCTTAAAAC CAGGACAAA GTCAAACTCA CAGCCAATT
 62581 TAGGTCTGGA GATTATTT GTGAGATCAA CCCTGAAATA TTGAACATCT CATGGTCAAT
 62641 AACTTGCTCT AACAGCATTC TGTTGGAGC CGGAGACTTG GGCCAAAAGC CATGCACACT
 62701 CTCTCTCTC TGATGCTGGA GACACGGCTG TGGCCTCTGG CCAGACCTCC AAGGTGCACT
 62761 CCTCCTCTC CACTGTGCC GAGTGCAGCA TGCAACCTGG GCTGGCTGG TGAGGAGGCA
 62821 GGGTCAACTT TCACCAAGCA AGCAGTTTG GCAAAACCTG TACCCACCGG CCCACTTCT
 62881 CAGAGGGCAA CTAGAGTCCA AGCTGTGGAG GCTTCAGCCC AGGCAGGAGT CCGTCTGAG
 62941 GGATATTCGA CAGTGTTCGA GGGCTCTGAA CCCAGGTCGG GTGGTCTCTC CGGCCTCAC
 63001 ATCCTCTCC TGCCTCCAGG AGGGGGTGCCT GGTGGCCAG CCTGCCTCTC TCAATATCTA
 63061 CTCTGGCAGG GAAGGCGAG CCACACTCTC CCTTGGGACCT TCACTCCAGG TGTCCTGCC
 63121 TCCAGGACAA GTTCACTCCC ATCATCTAAC CTCACTGCC TCCAATCCAGG GTGACCATTT
 63181 CTTCGGCTC TAACCTCAGCA ACAATAACAG CTAGCATCTG CTGCATGCCA GCCATGCCA
 63241 GGTGCCATGT ACAGGTTCTC TGATATTTAT CTTGCTTAAT CCTTACAACC GCCCCACGAA
 63301 GCAGTTATTA TTAGCTTCGT CTTTATAGAT GGAGAGTCAG AGGCTAAGTC CTTTGACAC
 63361 GGGCTCACAG CTAGAATCAG AGACATCCAA TGCTACTTC CAAGCTGTG TTTCTCCAC
 63421 TGAGAATCTC AGTTTACTTG GCAAAACAGGA AGACAAATGT TGCTTGCTT CATCTCTCT
 63481 TCCTTGCTT TCTTCTCTT AAAACAGTT TTCTGACAGA TTCTTCTTT CACTCAGCGC
 63541 TCCCCAAAG CCAAAAAATA ACCACCAAAAC TCTGACAATT CTTCTATGGC CTGCACTCTT
 63601 ACTGGCTGTT TCTACCAAA CATTCTAGGA TATCCTCTC TTTTTAAACA TCTAGAGCAC
 63661 ATCAGAGCCA TAAGGGTCTT TGAGAGAAGTT CCCATTTCAG AGTAGCCACA GATTGCTCCA
 63721 CTTGAGCTT AGATCACACC CTCCAGCTCT CGAATTTCAGA GGAGAGACCC AGTTTCAGG
 63781 AGACAACCC ATTGTCAGCACAGAATTG GGACTGGAA GGGGTTTGA GAGACATCCC
 63841 AGCCGGATCA TGCTGGGAGG AAGGTGGCAG GAATATGGGA CTTAGCTCTG GGGACAGAG
 63901 GCATTCCAGG AACCAAGGTCT ATGAGAAATG GCCTTCTGG GTTGAATGA ATGTACCTAT
 63961 TTTGCTTCTG ATATAATCTG TTTGATCAGT GTTGCTAGAG ATGAGCAAA ATTCAATGGCC
 64021 TGGAGGTAAT TAATCTACAG TGAAAGATG CAATCGAGCA AAGCACCGTG TCTCTGATGT
 64081 GCGTGTGGG ACCAGAAAG AGGGAGGCTG TGGAACCCAA GTTTCAGGG TTGCTTATGG
 64141 TGTGGTGAGG GCCAAGCTT TAAGGACTCA GATAAGCTCA TCCCTAGAAC AGGATGCCCT
 64201 GGCTGGGTTT GGGTATCAGG GTGAAGAGTC AAACGTTTTA AGAACAGACA TGATCAGGCA
 64261 CCATGCCAGC TGGGCTGAGC CCTCTGGAGTA GGACAGGGCA TTCATCTTC ATGGGGCAT
 64321 TGGGAGGGCC CTGGGTAAG GACCAGGGAG CACTTGAAGA CTTGGAGCTC AAGGCAGGAA
 64381 CTGTTGGCA AACAGTACAG AGGTATATTT AGCCACTTTG CCAATCAGTA CAGCTGGGCC
 64441 AGGGAGGACA GGCTGGACAG CAGTTTCTG GGAAGCCAGG GGAGAGAGAG GGGCCCACTG
 64501 CTCAGGCTGC CGACACACCA TCCCAGCCCC CTCCCCAGCC TGGATTCTGG GGTGATCTT
 64561 GCCAGGACTC CACTGCCCCC TGGGGACCTA CTGGGGACTG GGGGACCTTG CCAGCTCATG
 64621 TGGAAAAGAG GCAGCCCCCT GAATCTCAGG ATTGGCTATG ATGCAATCCA ATGCTGCC
 64681 TTCCCCACTCC TACTTTATA GATATGACTT TTGGCAGGGG TGAGGGTGGG GAACAAAATG
 64741 TATGACTCTG CTCTGGCAGG CCCAGCTGGG CTGTGCTCTA GCTGCTGTGC ATATGCTCTT
 64801 ATCTAACCTT CCAACCACCG CCAAGGAAGG CATCATTATC CCTGCTTAC AGAGGGAAAG
 64861 GCGCGTTTA GAGAGATTAC ATCTCAAACA GGCTTACACT GCCCAGGAGT GGCAGAGCCA
 64921 GGACTTGGGT GCAGGTCTTC AGGGCGCATG CACTCCATCC ACCTGGCTCT CACTATGAAT
 64981 GTACAGCACC AGCTAGCAGC TTGCTCCAGG CCATGTGCTG GACTAGTCAG TAGCAGAGAC
 65041 AGGAGGAAAA GCCTCAGCCT CGCATTGGG GCTCTCCCTA TACCCCAAGA ATCTAGCCGT
 65101 TACCAACAAAT CAGTCAAAAG AATATTAAT TACCAAAGGG CTCCAGGGT CAGGCTCTG
 65161 GGTGGGAGTG GGGTCTGGG AAGGGTCTGA CAACACATT TCAGAGGTGG TGATCCGAT
 65221 GCCCATCGCT GGGCCAGACA TACCCCTCAGC AAACCATCCC ATGAAGTCTG GGGAGCCTGG
 65281 AGAGAGGGGA AGATGGCACC TGGGGAGCCT CGAGGAACCTC CCTCCCCAGG GCTTCAGG
 65341 TCAGGCCAGG CAGGTGATGC CACCTCCTCC TCCCTAGAGA GCTGGGTGA CCCTTCAGTT
 65401 CCATGCAACA CCTGTGTCAG AGCTTGGACT AGCCAGGTGC TGTCTTCAG AACATGTGGG
 65461 CCCTCTTTA AGGAGACCAT TGGGAGGGGA CATGATGGGA CCCTAATACA TGTCCTAGCC
 65521 ACTCTGAGCA ACATCCAGA CTAAGACTCA CCTGGGGAGGA GCTACCATGG CTGTTTCCT
 65581 TTTTATCACC AGGCTAGTT AGTGGTCAAG GCAAGAACAT CCGGCACCAA ATGGACCCCT
 65641 CCAGGCCCTGA CAGGGCAGA TGCAAGCTCAT GGAGAGGCTC CTGAAGCCAC ATTGGGGTC
 65701 CCGGGCTCTT AACTGCTCA CTGCCCCCTT TGAGGTAAC TCTGTCACCT CTGTGGAAGA
 65761 CCTCTCCTGG CCCAAGCTCT TGGAGGCTCA AAAAGATTGG CAGGAACAAAT AGGTGGCCAG
 65821 ACGCCCTTCC AGCCACCCAA GTGCTGTGGG GAAACCAGGG CACACCCAG GAGGAGTGT
 65881 CTACCTGGCT GGGCAGCAGG GGCATTGGCC ACCACCCACC ACGGAGCTCC CTGACTTGGG
 65941 CTCTAGGAG ACTGAACAGG GTCTTCATCT CCCCCACCAA GGGGCAAAGT ACTTTCAAGG
 66001 CATCACCCAT GTCATTAAT GATTTAACCT GAGCCGGGGG GGTGGCCCTG GGAAAGGTAA
 66061 TGAGGAAGAA GTGCTGCTT CGAGGAACCA ACTAATGATT CTTCCAAGG TGAGCCTACA
 66121 GCAGATTGAA GCATTGAGGT CATCTCCCTA TTGCTTAGAT GGGGAAACTG AGGCAGAGTT
 66181 GGTCTAGAGC CCTGGCAACC TGATCCATGA ACCCAGAGGT CTCCCTGTGA GGCTGTGAGG
 66241 AGTTGGATGG ACCCATGAGC CTTGACTGTA AACCCAGGTG GGCAGCTTGTG TGGCTGATGG
 66301 CAACAGATTG ACAATGCCAC CAATTCTCAC CATAACTAC ACCATACTCT TGCACAGCAT
 66361 AGAAGCTTAG GATGTTTGCAT GGGTTAAATC CTGCTTGCA TTAGGGCAGG CAAATGTGG
 66421 ATCCCCCTGG CATACTCTC ATCTGGGTG CTAGCTGCC CCCAGGATGG CCCACGCCA
 66481 CTTCATGATG CCTTGAAT CTTGAGGGTA CTCTCAGCAG GCCCACCAA GCAGGAGCCC

66541 TGGCCCCAAC TGGGATAAAAC CACCTCCCCC AGGTTTTAAG ATGAGGCTGC CTGAGCCCAG
 66601 GTCCTTGGGA TGTCATCAATGC CTACCCAGCCC TCCCTGAGGC CGGTCTCCCTG CTCTCAGCTC
 66661 CCTGCCCTG CCCCTCACTT CCTGGCCATG CCTCCCTGGC CCTCAGCCTC TGGCCCCAGC
 66721 CACTGAGCAC CCCTCCCTCA CCTTCAGGC ACAGGGTGGCC GAGGGCCTGC TGTCTTCTGC
 66781 TTGGCATGGG TCTGGCAGGG AGTGGGAGAC CCTGGCGTGC ATCTGCCTCA GGAGGCATCAC
 66841 GGCAGCTCTC CAGCTGGGTG GCAAGATCAG GAACATTATAAC TCTGACTCAG AGCTCCACCT
 66901 TCCTGGCCAC AGCTTGGCAG TAGCTGGGAT TGCTCTGGGA GTGGTTTGAG TGGGTAGACA
 66961 GATAGGGTGC ATGGGCCTAG ATCTCTTCTAT ATCAAAGAAA TCACAGGTCT TCAGGTCTATA
 67021 GAAATACAGG GACCCCGAGG GCTCTGGCT TTGGGGTCAAG CAGTCTGTAG CCCCCAGCCA
 67081 CTCAGAATGAG CAGGCCCTCA AAGGGCACAG TGGGTGGCAA AAAGAGGACA CTGGACCCAC
 67141 ATGACTGTAG GACTAGCCCG CCATAATTG GGAAAGCTCT ATCTCTTCAG CTTATAGAGG
 67201 TCATGGCTCA TAGATGCCCG TCTAGGAGGA TGCGACCGTC GCTCCCAACT CCACCTTCAGT
 67261 TTCATGGCTT GGTACCCCGAG GAGCTGTGGC CAGGGACCTT CATGGGGTCC CTGGCCACAA
 67321 CTCCCAAAGA GACCAGGGTC AGTGCTCCCA AAGCTAGAGG GTGAAGAAAG AAAATAATTG
 67381 CCCAATACTT CGTGAGAACGG GTGCTGGGC AACCATCCCC TGTCCTCCCT CCTGGGAGGC
 67441 CAGGTGGAGA GAGAGCGTGT CCAGGCCCTG CCTGGCTCTCA GGACAGCTTC GAGGATGAC
 67501 CCGCCCCGCC CAGATGATTCTC ATATCTCTAG CCTGGTGGCT CACTGAGGGG CCTGACAAAG
 67561 GCAGCTTTCG CTGCTCTGCT GCCTGACTCT GTACACTGGG ATTTTCTGC CTGAGCTATT
 67621 GCCTGGGGCC CCACAACCCG CTTCCTGCG CCATAGATAG AAGTGTCTTT TTGAAATTGA
 67681 GGATGAGATC TGAACATGGG ACAGACTGGG CCTGAATGCA GGGATGGAAC CTGGTTGGAG
 67741 GTACCTGTGG GAACTCTTCAG CGTGCCTGGC CCTTCAGAAA GTGGAGGGC CTTAGCCTGC
 67801 TGCTCTCTG GCTTCCCTCC TTGCCCTGCC CTCCCTCTCT CGTCCAAGGA TGGGGACTCA
 67861 GCACCACATC CCTCAGTAA CCTCCACAAAT TCAATTTTTT TTTTAGTTC ATGTAACCTAG
 67921 TAAGCATCTG CTGTATACAT CCTACTGAGA GAGGAAATTC TAAGTGTGAA GTTGCTGACA
 67981 GTCTGGCTGG AGAGCCAGAT GCATCCAACCC CCTCAGAGGC TCAGGGATTC AAAGTTCCT
 68041 GTGGTGGAGAG GGGGCATCTT AGAAGGATCG ATAACAATCC TTCCCTTCACA TTTGTGAAGA
 68101 TCCATGGTTT ACAGCAGTAG TCCCCATCCT TTTTTAGTGC CAGGGACTGG TTTCACCGAG
 68161 GTTAATTTTT CCGCGGGTGG GCGGGAGATG GTTTGGGAT GATACAAGTG CCTTACATTT
 68221 ATTGTGCACT TTATTCTAT TATTATCATA TTGTATCATA TAATGAAATA ATTATACAAAT
 68281 TCACCATATAA GTAGAACATCG TGGGAACCC GAGCTTGTGTT TCCTGCAACT AGATGGTCCC
 68341 ATCTGGGGGT TATGGGAGAG AGTGACAGAT CATCAGGCAT TAGATTTC TAAGGAATGT
 68401 GCAACCTAGA TCCCTCGAGT ATACAGTTAA CAATAGGGT TGTGCTCCCTA TGAGAATCTA
 68461 ATGCCACTG TGATCTGACA GGAGGTGGGT TTCAAGGTGGT AATACAAGTG ATGGGGAGTG
 68521 GCTCTAAATA CAGATGAAAG TTCACTCACT CGCCCACTGT TCACCACTTG CTGTGTGGCC
 4 68581 TGTTCCAAA CAGGCCACGG ACTGGTACTG GTCTGTGGC AGGGGGTTGG GGGCCCCCTGG
 68641 TTTACAGAGC TCTGCCACAC AACTAAGAAC TGGCTAGGGG ACACAGAAGA TGGAAAATGG
 68701 ATTCCAGCTG GAGAGGGAAA GAATGGATCT GAAAAGGCCT TACAGAGGAG GTGGCATCTG
 68761 AATTGGCCTT TAGCAGGAGA ATGGGGATTC AACAGCTAGT GAGCCTAGGC ACCACTAGGA
 68821 GGAAGGAGAGA GGCATCTAG GAGGAAGAAA CAGCATGAAC ACAGGCCAGGG AGGCCAGAAA
 68881 GCAGGGCAGC CTTATGTATG CAGGTGGTCT GGTGTCTGT CAGTGAGTTG TGTGGGAAGA
 68941 GATACTGGGG TGAGAGAGTGA GAAACACAGGG GTGGTAGAAA TTGGGCAGAA AGAAGGAGGC
 69001 TTTGTGAAAG AACAAATCTG GAAAACACCC AGGTAAGGCC CCTCTGCTCT CTGGAGGGCC
 69061 AGGCACATGA AGAGGTCCCA CAGATACCC CAACTTGGTT CTGGCCCTGC ATTTGGACCC
 69121 AGTCTGCCCT GTGTTCAAG GTCACCCCTTC ATTCCATACT AGATTAGGGT CCAAATAATC
 69181 ATGACCTAA AAGGCCATAC TTATATTG GACGTAGTCT CTAAGCAGTA GGGGGCTATT
 69241 GATGGTTTT GAGCAGGAGA GCTGAACATT AAGAAGGCAGA GGATGATTGA TGTGAAACAA
 69301 AGGGATCAAT TAGGGAGTGA GATGTGAACC ATCTGGGAT GGGGAAGTAG GGTGGAAAGT
 69361 TAAGGGAGAG ATAAAGGCCAG AACAGGCCAT ACAAGCTAAG TCCAGGAGTT TCTGTGCAA
 69421 CTCACCCAC AGCACTATAC TGCTCTGAAA GTCATGGGTT CCCATCTGTC TCCCCATCCC
 69481 TCCCACCTCCC CCACCCCTTC CAATCAGATA GGCCAGGTT AGCCCCAAC ACCAGGGCAT
 69541 TCCTCAAGGC CTCCTATTG GCCTCTTCTC ACTGAAGAAC CCCAGCCACA AACCCCCACT
 69601 CCAACTCATA TGTTTCTCA CTGTCTATCA ACTCCTCAGA GCAGGGTCTC AACTCATTT
 69661 AGTGTGAAATA AATCAAATTG AATAGGGGTC TGATGGGAGC TTGGTGGAGG TTTCTCCCC
 69721 CACATCCCTCT GTGGTGCCTC TGCTCCAGGT GGGAGCCAGT GTGACCTTGA ACTAAGTGG
 69781 CTGATCTACT GGGAAAGTATT TACCGACTGC CTGCTGGGT CCCATCTCTG TGCTGACCAC
 69841 TCAGGTGGGG TCTGAAATAA AACAAACAGA AATGTAAATAAT CTAGAAAGGG GAGGAAAGGG
 69901 GGTAAAATAA AATGTAAAT TGCCAGTGT GGGCTCTAAAG TGGGATGGGA GGGGAGCAGA
 69961 GCAGAGGACT CGGGGAGTTG GAAAGTGGACT GGGACAGTGA GTAAGGAGTG TCAGGCAAC
 70021 AGTGCACATCC CCACGCCAGG GTGCGGGTGC CAAAGCAGCA ATCTTGGTAC ACAACTGCTG
 70081 CTCCAAGTCA AGTGTCTTT TGTTAGAGTT CCTCTCTGTG GAAAATTCTG AACTTGCTC
 70141 TGCAGCAAGG GGAGGAACA GTTCTTTCC ACCCAGCACT GATGACCAAGG GAGCCCCAGG
 70201 ATCCTCAGAT CCCCTGGGCT ACTTGGGCCCC AGCCTTGTG TGAGGTGCGC TGAGGCTCC
 70261 GACTCCAGGC CTGTCGCCAG CCTGTGTTACC TGACACTTCA GATGGGGTGA CCAGGCACTG
 70321 CCTCTGCCAG CAGACTCTTG GCCTCAGGTA ACCCAATATT GAAAAGCACA TTCCCTCCAG
 70381 AAGCCCAGTG GACGCCACTC CCCTGTGGGC CCAGCAAGGA TGACTCCCTA CTGCTCTCTT
 70441 TCCAGGTGGT TCAGGGCCCT GTCCCTTCTC TCCCACTCC TGCCAGCAGC ACCTGTCTCC
 70501 AGCCTCCCTCT CTGCTCCCTA CAGCTGTGA GCTTTAGAAA CACTCCCTC AACCTCTCGT
 70561 TCTTCTGCTG AAAAGCCACCT AACGCTTTTC CAGCTCTGC CAACAAACTC TGGATTCCAA
 70621 TACATCTCCAT AAACCAAGGTG GGACCCAGGA GCAGCTCTAC GTATATACGG AATTTTAGCA
 70681 TACAATAAAAG GAAGTATGTG AAATCAGTCG TGAAACTGTG AGAATTAAG AGGAGCCAGG
 70741 CGCGGTGGCT CAGGCCGTGA ATCCCAGCAC TTGGGAGGC TGAGGTGGGC GGATCAGGAG

70801 GTCAGGGAGAT CGAGACCAC TCCTGGCTAACAA CAGTGAACCC CCATCTCCAC TACACAAAAA
 70861 ATTAGCCGGG CATGGTCGGG GGCAACCTGTA GTCCCGAGCTA CTCAGGAGGC TGAGGCAGGA
 70921 GAATGGCATG AATCCGGGAG GCGGAGCTTG CAGTGAGCCA AGATTGCACC ACTGCACTCC
 70981 AGCCTGGGAG ACAGAGTGAAC ACTCTGTCTC AAAAAGAAA AAAAAAAA GGAAAAGAAA
 71041 AGAAAAGAAA GAAAGAGGAA AGAAACATGA AAGGTGACTC GCCAGTAAGA CAGGTTACT
 71101 TAAGGGAAAA CAAACCTGAG AGGGGCTCT GGCCAAGTTA GGTCAAGAGC ATACTCTCTT
 71161 ACAGACTAAG AGTTTTTAAG GATTCAAGT GGGAGAGTTT ATCAGAGGCT TGGACTGCTT
 71221 CTGTCGTCCT TTGTTGTCG TATCTGGGAG GGAGAGTTGT GTGTCGTCCT CCATACATCT
 71281 TCCCTGTAC GCAGGCATAC CTCCCCCAC CCCCCACTGG AGTCTGCTTT TAGCTTCCCT
 71341 ATCTTAGTGC ACCTGAAGGG AAAGGAATGT GCTTATTAAAG CCCAAATGTT TTACTGGGAC
 71401 CCATTGTATG CAGGGGAAGT TTGGCAGTT CAAGAGACCT TCCCCCTACC TCCCTCTGTA
 71461 CCCAAGCTAT CTTATCTGTG TTTTACTGCC TGCTCTTCT GGCTGCTTGT AGTTAGAAGA
 71521 GAAGTGATT CCTTGAAATG CATGAGGCTA GAAAGGGAGC TGGAACTTAG AGTGGTGGTG
 71581 TTTGTCTGAG ATGACGGTGC TCCTGCTCTG TCAGAAAGGA TGGACCATTC ATTATTGGAA
 71641 AAGCTAGCTT ACCATTGGAA ATAAAATGAA ATTAACCTCA CAAAATAAT TCCAGATAGC
 71701 TTAAATATCT AACCACTAAA ATAAAATCT ATATAATTAA AAATGAAATG TGAGAGTACA
 71761 TTATATTAAT GTTAGGTTGGA GGAGGATCTT CCTAAACAAG GTATAAAAC TAGAAGCAA
 71821 TAGGAAATA TTGCCAGTT TGAAATACACA GAAATTCTTAA ACTTCGTTAA AGCAAAAGAT
 71881 GCCATAAGCA GAGTTCCACC TCTACAAATAG ATCTCACCGT TCCCCTCTTA ATAAATATT
 71941 ATGCATAACA AAATCACTTT AAAAAAAAG ATATAGAGGA GAGTCTTGCC CTACCAAGGTA
 72001 CTAAGATAAA TTACAAAGCC ATACTAATAC AAACAGTATA TTATGGAAC AAAATGGAGG
 72061 GAACAGAGAC AGACCCCTGC TTTCAAGGG AACTTAATAT ATGATAAAAT TGGCACCACA
 72121 AGTCAATGGG AAAAGAATAA ACTGTTTAAT GGGTGACATG GCAAAAGTGA CTTACTATAT
 72181 AGAGTAAAGG AAAACAGCAT TCCCATTTCA TACCATATAA AAAGGCAGGC TGAAAAAAAT
 72241 TAACACAGGG AAAAGCAGG AAATGCAATTCATCAGAAA TGCAAAATAA ATGAAATG
 72301 AGATTTACT ACAGCTACTA AATTGGAAAA AATTGAGAAA GTTGGATAAT GGCAAGTGT
 72361 GGTGGAAATA TGGGAATAGA GTCAACCATCG TGCTCTGCTG TGGAAAGAGGA GGCTATCAGG
 72421 TCACTGAGGA CACTCTGTG TCACCTAGTC AAAGTAAGAA TATCAGACAG CAATTCTGAT
 72481 TTGGATATAG AAAGATATGT ATGAGGGTGC GGTGTGGTAG CTCATGCCTG TAATTCCAGT
 72541 ACTCTGAGAG GCCGAGGTGG GAGGATCATT TGAGGTCAGG AGTTCGAGAC CAGCCTGGCC
 72601 AACATGGTGA AACTCCATCT CTACTAAAAA TACAAAAAATT AGCTGGCGT GGTGGTGGGT
 72661 GCCTAGAACAT CCAGCTACTC AGGAGGCTGA GGTAAAGAA CGCTTGAAC CCAGGAGACA
 72721 GAGATTGCGAG TGAGCTGAGA TCACACCACT GCCCTCAAC CTGGCGACA GAGCAAGACT
 72781 CTATCTCAAA AAAAAAAAG AGAAAGAAAG AAAGATGTGT ATGAGGGCAT TCACTCCACA
 72841 TCACTTACAG CATGAGAAAT GTGGAACCA CCCATGTGAG AGCATGGCTG GAAGCATGAA
 72901 TTGGGAGTGA AGGATGTGCA ACCTGCTGCA GTCACTCAA GCAACACACG AGACCGCACG
 72961 AGAACAAACAT GATGGCTG TGAGCTGAGA ACACCAACTC TCAACACTTC AATATGGCTA
 73021 AAATCTATCA AGGACATCTA CAACAACTTC TCTCTGGGG AGGGGATGGG GAAAGTTTG
 73081 TATGCTGAGA ACACAGTAAA ATGATTGCTC TTCTGCTGGG AGGGGATGGG GAAAGTTTG
 73141 TGGGTATAGG GAAATAGCAA TCAATAGAAGA AAGCTGGAAAG GGGCCCCCA GGGACCACTG
 73201 AGGATTTAGG GCCATGCTT GGAGTCTCT GCACCTCAGG TGAGATATT TTTTCCCTCC
 73261 AGCTTTTCC CAGATGTTGC TGCCACTGCT CTTGTCTAAAG CCATTATCTC CACTTAACCTC
 73321 CTATATTGCC TAACGGATCT CACTGCTCT ACCTTGCCT TCCCTATAAGC CACTCTCTAC
 73381 AGAAGAGCAA AAGTGATCTC CTGAAAACAT AAGCGAGATT ACGTGACTGG CGGGATTCAA
 73441 GCCTGCCCT CACTCTGGC GGCTGACTTC AGACCTTCTC TGAAGGCTG TGTGGTCGG
 73501 CTGCTCACGT TTGTCCTACT TTTCCTCTCA CACCCAAAGCT CCAGCTACAC TGGTCCCTCT
 73561 TGCCAGGCT CTTCCCATG TTGGGAATTG GGCCTGAAGT CCTTGACCGC ACTCTGTTTA
 73621 TAGCTGCTACTCTACATCCCA TCACAGCCCT GCTCAATGTC CATTTCCTGT GGCAGCCCTC
 73681 CTGGGCCACC CTCCACCTCT CATCTACCTA CTTGCTTCTG TGTTGCTTA TTAACATAAT
 73741 TCATCATATG CTCACTAATC TTGTTTTCTT ATTCTTTTG CCTTTTCCA CCAGACTGTA
 73801 ACCCTTTCA GGACATGGAC ATTGCTGTC TTGTTCATCA GCTGGAAATA CCTAATGCC
 73861 ATCAGTAAAG GAATGGCTAA ATTGTTACG ACACATTAC ACAGCTACA CAGATGTCAA
 73921 CAGGAGCAA GTGCTCTCA TTATTGAAAT GTTACAAAG TAGCAGGGAC TGCTGCTAAGT
 73981 ATTTTATATT GCCTGCCAAAG CAGATACTAT TATTATCTCA TTTAAAGAA GCAGAATTGA
 74041 AACACAGAGA GGATAAGAAC CTGCTCAAG ATCACGATGG TAAGTGACAG AGCTGTGACA
 74101 CTAGCTAGTT TGGATTAAAC CACTATGATC CCAGGTGAA AGCTGCTGAC CTGGAAAAC
 74161 GAGATTCTGT ACTCGCTTTT TATTATTTT ATTATTTAC TTATTATTT ATTATTTCT
 74221 TTATTTATTT CTTTATTTT TTATTATTT TTGAGACGGA GTCTCACTCT GTTGGCCAGG
 74281 CTGGAGTGCAG TGGGTGAAT CTCGGCTCAC TGCAACCTCC ACCTCCTGGC TTCAAGAACAT
 74341 TTTCCTGCCT CAGCCTCTG AGTAGCTGG ATTACAGGAG TGAGCCACCA CACTGGCTA
 74401 ATTTTGTAT TTTTGTAAGA GAGGAGTTT CACCATGTGG GCCAGGCTGG TGTCAAACCTC
 74461 CTGACCTCAG GTGATCTGCC CACCTGGCC TCCCAAAGTG CTGGGATAT AGGCATGAGC
 74521 CACCGTGCCCG AGCCTGACT CTCTTTAA AAGGCTTCTT CTATGCTCT ATGTTCTTG
 74581 ATAAATAAAA TCCAATTCTC TAAACAAAT GTACAATGAA GAATATAATT TTCATGACAT
 74641 AAAGTAAAC AAGCCTTACA GCTTCTCTGG GTAAGATGAG ATTAAAAAAAT CAACCATGGA
 74701 AGTTAATATG TTGAGATGTC CCAGAACTTC TGGCCAGCAA ACTAGATGTG ACTCTATT
 74761 ATGATGACAA GATAACAAAG TATAAACATT ACAAGAAAGG ACCTTCACTC CAATGTAGGT
 74821 AATGTTGTA TTCTTAGTGT CATCAATGTA TCTGGCTCAT GAATATTTC TGCCCTGCCT
 74881 TATCTTTACA AAATTTAAAG GGCCAGCTTA GTGCTTTCTT TTCTTTCTT TTCTTT
 74941 AGACAGTCTC GCTCTGTTGC CCAGACTGGA GTGAGCTGGC GCGATCTCAG CTCAATTGCAA
 75001 CCTCTGCTTC CCAAGCTCAA GCGTGATTCT CCCATGTTAG CCTCCCGAGT AACTGGATT

75061 ACAGGCACAC ATCACCAACGC CTAGCTAATT TTTGTATTTT TTTTGTAGAG ACAGGGTTT
 75121 GCCATGTTGC CCAGGCCAGT CTCAAACACC TGGGCTCAAG CAATCTGCC ACCTCAGCGT
 75181 CCCAAAGTAC TGGGATTACA AGCATGAGCT ACTGTGCCGT GCCCTTTCA TTCTCAAAGC
 75241 ACTTGATATCC ACAGCTCAAT ATTCTCTTC CTGCCATTAA AACTTGTGGC CTCTGGTTTC
 75301 AGCTAGATGC CAGGCTTAT TCCCCTGCTC TGACTCTCT TGCTCATACA GTGAGGTGCC
 75361 AAAAGCCTTC TGACACTGA GTTACTCCT TATCACCCAT CACTTTGGTA AGGATGACAT
 75421 CGTCATTGTC TCAAGACTGA GTATGGTATA GCGGTCTAGA CTTTAATAGA CCTGGGTTTG
 75481 TACCTTCATT CTGTTACATA TAAACTACGT GGACTTGGGT AAGTTTATCC CTCTAACGCT
 75541 GAGCTTCCTC ATCTGAAAG TGGAAGCAT AACAGTTCT ACCTCATGGG ATGGTTGTCC
 75601 AGAATAATG AGGTAAGAGA CTTAGACAT GACGGGCACG TGCTGATT ACTATAAGTG
 75661 ATCCACAGTA CAGTAACATC AGGGATACAA AGACATCATC GCTGCCCAA GGACCTCCA
 75721 ATTTAAGTTG GATGAGGCAG GAAGGTTACA CAGATAGTT CAGGACAATC ACGTCATATA
 75781 TGAGAAGAGA GGTAAAAATA AGCTGCTCAG AGAGCTTGGC CACTCAGGGC TGGGGAGATC
 75841 AGAAGAGGGC CTGGAGGAGG AAGAATTGAG ACTTGCCCTT GAAGAAGGAA GGATATGGC
 75901 CTGACAAATG AAGGAGGGG GAGAGGACAT TTTTCTCAGG TGAGGTGGCA TGAGTAAAGG
 75961 AATGAAGGTT GGAGAAGGGG ATTAAGGAGG TTCATGGATA TGGAGTGAAGA GAATCACAAAT
 76021 CTTCTTCAG TCAGCAGTGG AACTCATGGC AGGTTTTGA GCAGAGAGAG TTGAAGAAAG
 76081 AGAGGAACAT CTGCTTCATC AGTAGAACAA GTGACTCTCA GCACTGTG TGTTAGGGT
 76141 GAGCTGAAAG GCCTTCATCA GCCATCATGG GGCTCCTGG GCAAGTGGTA GCCTGCAAAA
 76201 CGCCCCAGAT CAATGCCATC CCCCCTTGT CTGCGAGGAA CAGCCACAG GTTATGTTT
 76261 CTCCAATAC ACGATAAGGCT CCCAGAGCAG GGCACACTT GCATCAGGCT CTCCAACATAT
 76321 CACTCTGGC CCCAGGGGA TGAGGCTCAG GAAAGAGAAA TCTGCACGTT TGTGGAGGAG
 76381 GAGGAGAAAT TTGGAGCAGG GCCTGCTGAG GAGGGAGATG GTGGGAGGCC AGGTGGCAGC
 76441 CTTCATTCTAT TTCTCTACTT TGCCCTGCGC CATCCCCCAT TAAGCTAGAA CAGACTCTG
 76501 CTCCATCTGT GTTGAGGTGG TCTCACAGCC TTCTGCTTCC CCCCAGGCC AAGTAACCAC
 76561 TCATTGCTC CCTGTTGCTG TAGACTAGGT GTGTCTTTC TGAATTCA CATAAATGGA
 76621 ATCATGCAGT GTTTGCTCTT TTGTGTGTGG CCCCTTCAT TTAGCATAAG GCTTTGAGA
 76681 TTGGTCCAGG TTTTGTGTG TAACAAGAGT TTACTTTTC ATTGCTGTGT AGCATTCCAT
 76741 TGTAATAATT TTACAATTG TTATCCATT CATCAGTGA AGGGAAATAA TTTTCTTT
 76801 CTTCTCTTCTT TTTGAGACAG AGTGTGTTG TGTTACCCAG GCTGGACTAC AGTGGCACAA
 76861 TCAGGGCTCA CTGAGCCTT GACCTCTCG GCTCCAGTGA TCCTCTCACT TCAGCCCTCT
 76921 GAGTAGCTGG GACCACAGGC ATTCGCAAC ATGCCAGCT AATTAaaaa AAATTGGTA
 76981 GAGTCGGGGT CTCACTATGT TGCCACAGCT GGTCTTGAAC TTCTGGCTC AAGAAATCCA
 77041 CCCGCCTCTT CTTCCAAAG TGCTGGGATT ACAGGTATGA GCCACCCAC ACAGTCACAA
 77101 AGTTCCCTCC CTTCATGCTT AGCAATCATA AATAAAGCTG CTGTAAACAT TTGCACACAG
 77161 GTTCTGCAT AAATATACAT TTTCATTCT CTTGGCTAAA TCCCCAGACG TGAAATGGCT
 77221 GGGTCTGATT GATGTTTTTA CAGAGTGTGATG GCATGCTG TGCTGCTAA AAATCCACG
 77281 CTCCCTGCACA GAGCCCTCCTC TCTTGGCAC TTTCCATT TTCCCTCTCT CATTTCCTC
 77341 TCTCCCTCTC CTTCTCATCCT CTCCTTCTC CCATAGGCTT CCCCCTTCCC TCCTCCATCC
 77401 CACAGTCCTC CTCCTGCCAGG CTCCCTGGTT CTCCTCACAGA GGGGGTATTG TAACAGGCAG
 77461 CCCTCACGTT TTCTGGTCC AAGGGCAGCT GACCAAGCAC CCACAACTAG ACTCTGCCAA
 77521 GCTTCTTGAC ATCAGCTCCT GAGCCAGGGC ACCTCTCGGA GACTCCCAGA CTTCACAGGG
 77581 AGAGGGAAAGG GGAGGTCTTC TAGCAAGTTC AGCCCTGCAG TAATTGAGCA AGGCTGGGTT
 77641 TCTCTAGTT CCACTGTCTT AAGAACTGTG TCAGATGGTT CTAGTCTACC TGATATCATT
 77701 TGGAATGTTG TCTCTCCAA ATCTCATGTT GAAATGTGAT CCCAATGTT GGAAGTGGAG
 77761 CCTAGTGGGA AGTATTGGG TCATGGGGAT GGATTTCTCA TGAATGGCTT GGTCTTCTG
 77821 GTGGAATGAGA GTGAGTTCTC GCTCTGAGTT CATGCAAGAT CTAGTTGTTA AAAAGAGTT
 77881 GACCTCCCAA CACCTCTTGT CTCTCCCTCT TGCCACCCAA CATGTTGCT CCCTGTCAGC
 77941 CTTCTGCCAC GAGTAAAAGC TTCTTGAGGG CTCACCAAGA GCCAAAGCAGA TGCTGGTGT
 78001 ATGCTTGAT CGCCTGCAGA ACCTCTTTTC TTTGTAAATT ACCTAGTCTC AGATATTCT
 78061 TTATACAATG CAAAATGGAC TAATATACTA CTCCATTCA CTTGTGCTCTA TGTTTGACT
 78121 GAATTGAGTT CCAGCCAAGT TCTAACAGAG AAAATAATT CCCCATGGGC ACCATTCACT
 78181 TCAAAACCCCG GAGTCACAGGG TCCCTGCCCA CCACCCACCA GGGAAAGGGG GGCAGAAGGG
 78241 GCAGTGTCA GCTGGGGAGG GGCATGAGCA GGAGATGCC TCAGATCACG TCCTGGGTTT
 78301 GACAAGGACA GCACAGCTGA GGAAGGCCTT GGCTCCCTAC TGACTTTAGG TCAATGTCCA
 78361 TTGCTATTTT GGAAGGCAGT GTGGGGTGG ACAGAACACC GGGGGGGAG ACTTCTGTAT
 78421 ATTTTCACCA TCTCAGGTGA GTCTGAGAGG CAGCCGTGGT GTGACCATC CAACCAAGTG
 78481 GCCTTAGAGG CTTCTCTCAC TGCTGATTTC TGGGGCTGG CTTGTTCTT GCTGCCCTGGT
 78541 GAGGCTATGC CATGCTTGGG CACAGAGAAG AGAATAAGGT GAGAGATTG TGGGGTAAGA
 78601 CACAGACTAA AAATTGTGTT TGATGTCAGA AAGGAGGAAC ACAGCCAGGA AAGAGCTTA
 78661 AGTAACTAAA AATAGGCTGA TATTTAGTGA GGGGACCGAG ATTGGGGTTA CGTAAAACA
 78721 GAATTGTTG CTCCTCTTG GCAGGATCCA GGGCAGGGGA CTCCCCCACC CCTCCCCACT
 78781 GCATGTTGTTG TGAGTGTG TGAGTGTGAG AACAGGTATG TATGTGAGTA
 78841 TGCTGTGAGC ATGTATGAGT GCTATGTTG TGAGTGTG TGAGTGTG TGAGTGTG
 78901 TGCTGTGAGC ATATATTGTG TGAGTGTG TGTTGTGAGA GAGTACTTGG GAGCGGACGC
 78961 TTATGTGTTG TTGCTTGCAC CTTAGGAGG AAGAAGGCTC CTGAGCTCTT GCTGGAGAAT
 79021 AAAGTCCAGT TGCACCGAGA TGATTAAAG TCAGCAGCCC TGTTGCTGTG GGCCTGTGGT
 79081 CCTGGAGCCT GTGTTCTGGC TGCTCCGATG AGCCTTGGCC CTGCTTGGCC TGCCAGACAA
 79141 GAATGAGTGC AAGTCTTGTG GCCAGTAAAG GCAAATAACCT CTGGGGATT TTTCAGTGC
 79201 ACACATGTTGG TTTAAGGTT TGAGTTAAAG AATTTTCACT GAGGAAACAG AGGAGTGGCT
 79261 GAGCTGGGCC TTATGACACT GGGAGGGCTT CTGGAGGGGG CCCTGCCAG CCCTGTCAC

79321 CTGCTGGTT GGTGACTAGT GGGCTCCAGG GCCAGCATGG TGAGATGCCA GGACCTGG
 79381 GGAGGCATTG GGGCAAAGGA CATCCTGGGT GGTTGGGGTG GGAGGGCAG GTCCCTGGG
 79441 GTGACTCTGC AGTCCTGGTT TCTGTATCT GGGCAGCAC CTTCTGGTCT GTGAGGAGCA GGAGAGGGC
 79501 GATGATCAAG CACCAAAGGC AGAAAGAGGGT CCTCTGGTCT GTGAGGAGCA GGAGAGGGC
 79561 AAGAGCCACC ACCTGGCACC CTCTCCCTGC TGCGCCCCG CAGCCTGTGA AAGCAGCCCC
 79621 TGGCCCTGGG TGGAGTGGGC CCTGGTGGGG ATGGGCTCAA GCCTCCTGTC CTGTGTAGGG
 79681 GAGCTGGCAG CATGGCAGAG ACCCCTTGCC AAGGTGCAA TGACACTTCA CCTCTCTCAA
 79741 CTGCTATTTT CCTCTGTGCA CATGGATGTG TAGGGATTGG GTGGAGTGGG AGAGAGGAGT
 79801 TCAAGTCTT TTTCGGTGA TGTAGAAATT CAGAGGGAAAG GACTCTGGGC TCCAAGAGCA
 3 79861 CAGACCTGAC CCCCTGTAGT ATCTGCTCG TAGGCGCTCAG CAACGTCTG CTGCAACAG
 79921 TAACCAAAGA CTCATTGAA AGTAACCAAA CATCCTGGTT TGTCCAAGAC TAAAGGGTTT
 79981 CCTGGGACTC ACGACTTTA GTGCTAAAAC TGGAAAGTTC CAGGCAAAC GGGATGATCA
 80041 GATCACTTAT TTCTAATTTC CTTGAAAGT ATTTTGGAAT GTAGAGACAT ACCTGGTATG
 80101 CCAGACCCGG AGTACTCCTA GGCTTGTAAT TTATTGAGG CAAAATAATT TGATTACTTT
 80161 CTGATTTGT TTTGGAGAAT ATGAGAAAAA TCTGTATGTG CTGATTCTGA GTACATAAGC
 80221 CAGTCTGGAG GTTAACCGGTT GACTGTGTCT GTATGTGGGA TGCTGGTATA GTGGACACAT
 80281 GACCCCTAGAG TCCCCTTCCC CTTCCTGTGA TAAGGAAAGCA CGGTTTCCCT TTAGGATCTT
 80341 ATGGAGCTGT CATCAGGGAA TCATCCTCTT TTCTCAAAGG TGGGCCCCAA CCAGTCTGGC
 80401 CAGTCTCATT CTCATCCTCT TAGACTCAGT GATTGGTCA GGAACAATTA AATCTCTCC
 80461 TCTGCCTCTT CCTCTCTAT GTCCCTTTT TAAGGATTAT ATGAATATTG AGAAAGAGAG
 80521 AGAAGAAGAG GTCAGCTTG AGTTATTGGT GGCCATCTT GTCACCATGT GGAAGGACTC
 80581 AGCCTTAAAAA AGTTGAGAGA AGGAGAAAAA AAGACAAAAT CCTGAGGACA TGATCTGGAC
 80641 TGCTAGATCC AGCTATGCCT GAAGCCTATA TGCCCTTGTG ACTTCCCAGT TACATGAGAC
 80701 AATAAAGGCT TTTTATTTCG TTAAGCGTAT TTGAAATCAG CTTAGTTAC TCTGCCACTT
 80761 GCGCCTGAAAC AGAACCTCAC ACAGACTCTC GATGTGGCTG GCTGCCACAT CTGCAGGAGG
 80821 GCTGTACTCT GGGCTCTGTG CAGGGGATCA CAGGGGTGAT GCCTTCACT GTCCCTGAA
 80881 CTGAGTCCTT TATGCCTTA ATTAGGGCA TATGCTCCTC CAGGGCACAT CTACGGTGGG
 80941 AGGCACATGT CTGAACATTG GCCTTCAGGA GCCCTCTTCT CCAAATCGTG GGGGAAAGAGA
 81001 TCTTCCAAGT TCCTCTCTC ATCTCCTGT TCTCTCGCC TCATTTCCAG GATTGCTCT
 81061 GATCTCTCAT TTGTCCTTGAG CTGGAAGTCA ACACGTCTAT ATGTCCTACCC ACCACTCACA
 81121 TACCTTAAGT CCTTGGTGCAT ATCCCCTGCA TAGGGAAAGC CTCTCCCCAT CCTGGGAGGA
 81181 GCAACTAGCA CTCGAAACAT TACATTCTCA CACCTGTAAA CTGGGACAGA TGAAAAGCGT
 81241 TGGTTATTGT CCATGCCAG GGGCTAGAGA GTGGGGACTC ACCCTGTTTC TCTGCTGGG
 81301 AATGGGAGAG GCAACCATGC TTTCCCAGCG AGTAAGGCTC AGTTGAAGGA AGGGAAAGGC
 81361 AGGCAGGATG GGGCAGGCC AGGGGATGTT GCAAAACAGC CACAGGAAT GGAATCATG
 81421 TCATATGCAA ATCAGGATAA GAATTGAGCAG ATGGGCTCGG CACTTGGGCT TCAGGTTAAT
 81481 TTTTTAACCA AAGGTAAAC ACTATCCCTA GGAAGAGAGA GAGGGGAAGT GATTGGGTA
 81541 GGAGCACCTT GGGGCCTTTG ACTCTGTAC CAGTGTGATT TTTTAAGGCT GGGCAGTGG
 81601 GTCATAGGTA TCCATTTTAT TATTCCTAT ACATTTGTA TATCGTATAT ATTGCTTAAT
 81661 ACTTTTTAAAGC ACACCCCTGA ATTCTGTCT ATCACTCCCA GGCATTTCTT TACTCTTTCT
 81721 CTACCTATGT GTGTTCTCT GAGCTATACT TGGAATTATT TCACGTTTTT AAACCTTATA
 81781 TAAGCTAGGT AACATACAAA TATCCATTCT GCTATAAAACA TGTTTTTTT CTCTACTTTA
 81841 CATTAGGCTT GTAAGATTCA CCTATATTGG TTAACAATAG ATATGGTCA TTCATTTCA
 81901 CAGTTGTAC CTATCCACA ACTTATGATT TTGTTGGTTA GACATTTAAG CTGTTCTAA
 81961 TGATGGCAGC GGCGGACCAT CTGGAGCGGC CGCTGCCATC ACCGGCCGCA GCAGGGAGGG
 82021 TGCGGGGAGG AGGGCAGACAG CCTCTCTGTG AGCCCCACCG CCGGGGGGCC CGGGCGATGG
 82081 GGCTGGGCA GGTGGCCCGC TGGCAGAGGA GCAGCGCAGC TGGGGAGCAG CGCGACTGGG
 82141 GAGAGCAGGG CCCTGAAGTG GAACTGGGCC CAGAGCGGTG CTGCTCTTGC ACTGGGAGCA
 82201 GGAGCCCCAG GAGGGGAAAC TGGGGATGCG CTTCTGGGGC CCGGGCTGAA AAGTGGGAGC
 2 82261 GGTGCCGCTT TGGGGACAG CCAGCGGGGC AGCCATTGCT GCCATGCCG AGGATGCCGG
 82321 GTTCTGTGCA CTCGGGTGGA GGCTCTGCTC GGCGCTGCTC AGGGCCGCGT CCCCAGGTCC
 82381 ACCCTGCATC TGGGTGACCG CCTAGCCTGA CACTCCAAA GGCTGGGCCA GGGCCTGCCA
 82441 TCCACACACC CCTGGACCGC CCCTGGTTGC CAGGGCAACA GAAGGGAGCT CATAGCCGTG
 82501 TCGCCCCCTGC CCTGGACACT GGGCGGGGCC CAGCCAGGAC CTGGAGCCCC TGCCCCAGGC
 82561 TGCAGAAGGAG GGGCGGATGG AGCTGCATGC TCCGGGGAGC TGGCGGGAGC CGGGAGGCC
 82621 GGAGCAGGCA ACAGCCCCAGG TTGGGGCTGT GGGCCTGGGC CTCTGTGTG TCTTGGGGC
 82681 CCTGGGAGTC CCCCCCAGCC CCGGGCCTTG CAGGCTTGGA AGTGCCTGCT CCCGTGCT
 82741 GGCTTCTCCC CGCTGCCGGC GCCCACTCCG ATCTCAGAGC ACCGTGGGGC CGAGTCCAGG
 82801 TGCTGTTGCA GCCTGCCGA GGGTGCACAT ATTTGGGGCA GTGCTGGTAT GCCAGCCCC
 82861 TGCCGCTTCA GCCCCCTCCA GACTTTGGGT GCCAACGAGC ATGGGGTGGG GGGTGAGGCC
 82921 AAGGTGGGG ACTGAGGGCAG GCTCCGTGCT GGCCTGTAGG TGCCCTTGG CACCAGCAGC
 82981 CTGGGCAACA TAGGTGGCAG CAGAAGGGCAG ACAAGCTCTT GGGTGAAGG GGGTGGGTCC
 1 83041 CTGGTGAGGC CCCACCTCA GGCCAGGCTG CCAGTCCTGT AGATCAGAAAT GGGAACTTGT
 83101 GGTACCTTT CCAGGCTCAC CCATGGCTGC CCACGGACCA ATGTCGCCGC ACTTCCCTCCC
 83161 CTCGAAGCCC TGGGCTCAGC CAGAGCAGAG CGGTGCTCAG GATGACCAAGC TGCTGAGAGG
 83221 AGCTTCCAC TCCAGGGCT CCTCTCTGCT AGGAGTTGAA CACTAGTAGG GACGCCCTGG
 83281 CTGGGAAAG AGCTACCTAC TGCGGGTCTC CTCTGAGCTG TTCTATTGCT GGATAAAATCT
 83341 CCTCATCTTC TTGCTACCTT CCCACTTGTC TGTGTACCTT GTTCTTCTG GTCACAGGAC
 83401 AAGAACTTGG TACCTGTCTA ATGGCGGGTC TAAAGAGATCT ATAACACAAA CAGGGCTGAA
 83461 ACATGCCCT TGCTCACTT GTTGTGGGT AGAGAGAGGA GAGAAGAGCT GCGGCCCTC
 83521 AGGGAAACCA AACCTGGAG CTCATGGAGC CACATCTGTG ACTCCCTCTT CGGGGCCCTG

83581 TGGTTCCCGG CATCTCCAAG CTTCTGGGCA CCACCACATT CCTGGAGGC AGCTGTGGAA
 83641 GCTGCTTGC GTGCACCTGG AGCTGCCAC CTCACTGGG CAGCTGGCAT GTAGCCAGAC
 83701 CCCACGTTTG CTCACACACC CATTGCTGCT CTATGCAGTC TCCCTGGCA GGATCCAGG
 83761 CTGGTAGTGT GAGCTGAGTG CAGCCTGCCA GGCTGAGTGG GCAGAACGAA CCCAGCGAC
 83821 CCAAGCAAAA CTTGGGCAA GGTGCAATA GCCACAGGTT TCTGGCCAGA AAAGTGACAC
 83881 CCCAAAGATC CCATAACATT AAGTTTGCT GTTACTGTAT TTCACTCAGTT CCAAGGTACA
 83941 CATTTTTCC CATTTTAAAAA TCTGTGAAT AAGAGATACT TTTTAAAAAC TTATAATGTG
 84001 TAATATTCTG ATGTTGAAATA TCTAAGGTAT TGGTATATCT TAGATCAGTG GTCCCCAACG
 84061 TTTGGCACC AGGGACTGGT TTGTTAGAAG ACAATTTCATGGATGGT GGGGGCAGGG
 84121 ATGATTCAG TGCTTACAT TTATTGTCAT CTTTATTCTT ATTATTATTA CACTGTAATA
 84181 TATAATAAAA TAATTATACA ACTCACCAT AATGAGAAC ATGAGGGAGCC CTGAACCTGT
 84241 TTTCTGCAA CTAATGGTC CCATTGGGG GTGATGGGAG ACAGTGACAG ATCATCAGGC
 84301 ATTAGATTCT CACAAGGAGC GTGCAACCTA GATCCCTTG ATAGGCAGTT CAGAATAGTG
 84361 TTTGCACTCC TATGAGGATC TAACGGCATT GCTGATCTGA CAGGAGGCAG AGCTCAGGCG
 84421 CTAATGTC CCAGCCTGCT GCTCACCTCC TGCTGTGTT CCCAGTTCTT AGCAGGCCAC
 84481 AGACCAAGTAC GTACCGTGG CTTCTTAA ATTGACTATG GTAAGGGAGA ATAAGAGGCT
 84541 TGAAAGCAGT GTGGGGCGAG GGAGTACAAT GCTACTGTG ACTTTCTTGT ACATGTCATG
 84601 TTTCTTCTAGT GTCTATATTCT CTAGGAGTAT AAATTGCTGG TTATGAGGG ATATTATCAC
 84661 CAATTTACT AGATATTACC AAAGTGTCT CCAAAGTAAT TGTTACTGAT TCACACTTCC
 84721 AATACTATGT ATGAGAATTCC CATTGGGCT ATACGATGGG GTATCATCTG ACTTTTGAG
 84781 TTTGCTAAT CCAGTGGATA TAAAATGGTA TCTCATTGTG GTTTCATAG GCATTTCCCT
 84841 TATTATTAGG GATCGTGGGC ATCTTTCAA GATCTTTAC TTTCACTAAT CCATGTTCTT
 84901 GGGACTGACT TTTCTATTGAC TCCACCCACT TCCCCCACCC ACTGCCTTCA AGCCTCTTAC
 84961 CCTAGTCAAA TCTAACGACC ATCAACACTT ATGAAATAC AGCAGTAGTT ATCCAAAGCT
 85021 ACAAGACTAT TTATGCCCA GTAAATTACAC TCTTGGAAAC TCACACCGAG GAATACAAAT
 85081 GCTATCTGCT CAAAAATTTC ATTAAAATT TTTGATGGTG AAGAATTGAA ATCACACATAA
 85141 TATGCTCAAT ATTGGTAGAA ATATTAGTAA ATAATTGCAAT ATTCTATGA CTGAATTCTA
 85201 TGCAGTCTG GAAAGAGAAC TATAGAGAGA ATGTACCAAT ATGGAAAAAT ACTTTGTTA
 85261 ATGGTGTGAT ACTGGTATAC TGATTAGCAG ACTGATCAAC AGAATAGAAT GAATATCTCC
 85321 AAATAGAAC CAGTATACAT ATATACATAC ATATGTATGT GTACGTATAT ATATATATGA
 85381 ATACTGGAAT TTGCTATGT ATGTATACAT ACATATACCTT GGATTTAGGA TATAAGAAGA
 85441 CTGGATTATA AAGATAATAA AGTAAAAAAT ATTAATTTTT CTATAAAAT CATAGCCACA
 85501 GAAGAAAACA AGATCACATG CAAAATAGCA ATGAAAAAAT AAATATATAA ACCTAAGGAG
 85561 AAATAATAA GAAAATCACA GACACTATGT AGAAAGCTAT AAAACTCTCC TAAGAGACAT
 85621 AAAACAGCTT GAAAAATTAA CATAACATACC TTGTTCTTAA ATGGGAAAAT AAAATGATCA
 85681 ATTCTCCTA AATTGATCTA TTATTTAGT ACAATAACCA ATAAATTTC CAACCAGAAG
 85741 TTTGGGGGA ACTTAACCAA AAGTTGTAAA ATGTATGCAG AAAAATGAAA ATGAAAAAAT
 85801 GTTCAGTGAAT TTGTAACAC AGATAAAATT AAGGAGAGAA AACTAATTCT ATCAATAAAAT
 85861 GTACTTTAA GCCACAGTAT TTGGCAGTT GTGGTGGCTT CTGCCTGTA TCCCAGCACT
 85921 TTGGGAGGCC GAGGTGGAGG ATGCTTGAG GCCAGGAGT CAAGACAGT TTGGGTAACA
 85981 TAGTGAGACC CCCTCCACCA TTCTACAAAA CTTTAAACACA CACACACACA CGCACACACA
 86041 CACACCTACA CACACACTAC AGTATTTAAG GCCTAGAGCA TCCAGAGACA AATCCAAGTG
 86101 CATAACAGGAT ATTACATGT ACAAAATGACA ACATTGCAA TTAACAGGGG AAGGACAGGT
 86161 TATTGATAA ATTATGTTGA AAAAATAAGT TTTCTGAAAA AAATAATCAT TTTGTCATTT
 86221 GGGGAAAAT AAGCTGGTC CATATCTCCT TGTATTAAAAA TAAATTCCAG ATGTAATGTT
 86281 ATTGTTTTAT CATTATTAGT GAGAAAGCAA GATATGAAAT TGAGATGCAA TAAAGGAAA
 86341 TGAATGGATA ATTCTGACTG TACAGTTTC AGATATTAGG TTCTCTATG TCAAAACACG
 86401 TACACAAATA TACAAGACTA ATTGTTAAAAA ATGTTACTA AACTTATGAT GAAGGGTTAA
 86461 TTTTCCCAAT AAACAAATCTTCA TAAATCTGAA ATGCAAACA ACATAAAAC
 86521 TGACAAAAGA TATAAACAGG CAATTATGG TAGTAATGT TATGTCGCA GTTCCAGAAC
 86581 TTCGCTGAAC ATTGGAATCA TATGGAGATC TTAAATATAA GTATTGTTGC CTGGCTCTCA
 86641 GATACTCTGA TGTAATTGGT ATGGTGTATA ACTTGGCAGC TCCAATTTTT AGAAGCTTCT
 86701 CTGATGATTC TAAAGTTCAA GAAAGGTTGG GAATCAATGG ACTATGCAAT AAGCTATAT
 86761 AAAAGATTAT TGCTAATAAT TAAAGACTG TAAATTTAAA CAAAAAAGAG ATGCCACGTT
 86821 TCATTTATCA GATTGCAAG GATAAAAAT TGATAAGACG TAGTGTGGT GAGGGTGTGA
 86881 GGAAGAGCTA TAAAGATGTT GTGTATGAGA TGAGATTAG AGCAAAGTTT TTGGAGGGCA
 86941 ATTGTTGATAT CTTATGAATA TGCAAAACCTT GCACTGCTT TGACTCAGCA ACTGTATTC
 87001 TAGGAATTAA TTTTACACAA ACACCTGCTC AAATCCTCAA AGATATTGAA TTTGGTATTA
 87061 TTCCCATAAT ATTAAATTCTT TTGTTGTTGT TTTTGCCTG TTTGGTTAA TGAGGGATAG
 87121 GGCTTAGAAG TTGTTGATGG GTAAAAGGAA AAAAAGATT GTTGATGAAAT GGAAATAGCA
 87181 GTGCCCTCAT TCCACTCGCA CCCACCGATG TATCCATGGC ATCTAGGCGT TGGATACTGG
 87241 TGAGTGGAGC TGGGCAGCAT CCTCCCTTAC GAATCCAGT GCATCCATAA ATTACATAG
 87301 GAGCAGAGCA GCCTCCCTGGG GTTGGGGAAA GAGGTTGGGA CTCGGACTCC TCATTGAGGA
 87361 TTCAAGTTAT TGCACTGAT ATGATGAACT AAAGAGACCT CCCCAGGAGA CCCATTGAG
 87421 TCAGATGACA AGGAAGGACA CTGGGAGATC CAGAGAGGAA ATATGCCCTA AAGAAGAGGG
 87481 AAGTTTAGAC CAGTGTCTG TACTGAGTAT TGGGGTTAC CTTATCAGAA CCACTCTTC
 87541 ATGCTTATTC ACTTTGCTAC CGTGGGCTC TTTGAGGGAA ACTGGCTCTG TGTGTGTG
 87601 TGGAGGGACA GAGCTGCATC ATTATCAACA GACACTTGGT CACCAAGTCC CTTTGCTGAG
 87661 GGACCCCTCT CATTGCTCTC GACAGGGCAG ATGCAATTCTT TAGCCTGGGCA AAATCAGCAG
 87721 AGAGGCTTT GTAAAGTTTC AAGTCAGCAG ACAAAATGATG GGGGTCTGTC TCATGGACGG
 87781 GGACAATGGC CATGCAATT CTCTGGCTA TTTGGAGGGAA AATCTGACT TCTCTGCCAA

87841 CTTGACAAGG AGTGGAGAGG AATCATTGTA ATTTGCATAT CAATGACATT TTCTAGGCTT
 87901 TTGGGTGTC ATAGCACACT CCAAATGGTC ATAGCATCAG AAATTGAAA CAACCTAGAT
 87961 GATATGATGT TATGTTATGT AGCCATTTC AAAAATCTGG TTTTAATAG CAGAGAAAAA
 88021 TTAATCACTA TTCACTGATA TGAGAAATG ACAAGTTTA CAATACGAGC TTAATTTC
 88081 ATATATCAAT ATACGTATAT ATAAAGAAA TTTATTTCTA AATATATGTG CCTATAAAA
 88141 AGACTGGAAG GAAGTATATT AAGGGTTCAAG GAAGTTATT CTGGATGATG TTATTGGCAA
 88201 TGTTTATTT CTCTATTACA TTTTCAAAAT TTTGTAATGT CTATACTGAA TACATATATT
 88261 ATTTATAATC AGAACAAACCA TTTGTAAAGA AGAACATTAT TATAAATAAA GTTGCCTTT
 88321 AAATAAACGA AAAATGCAT TTAGAGACAG GGAAAGAAA CTAACATTTA TTTCATGCCT
 88381 ACCATGTCAGGCC ACCAGATATT TATATATGTG AGCTCTGTTA AAAGTATCTT
 88441 CTCCCCTGG CATCTCTGG GAGTAGCAGG AGGTGAGCT TGAGGTGGG TCAGGCAAAC
 88501 CCACCCAGGG AAGGAATTT TTCTCAAGAG CTGGGAAGCC CTTTCTTCCT TCTACTAAGC
 88561 AGCTGGCAT CTGCCACCA ACCAGATCCT TCTCTTGCCT TTGGGTGAAA CATCCACTCA
 88621 TTGGGCAAGG AAGGGGGCAG AGCCGTGGCT CACTGAAAAT GTGACTGAAG CAGATACTGG
 88681 AGCCCAGGCC AGCAGGGCTG TGAAGACTATA GAGCAAGGA GGCAACTTCA GGTATGAAA
 88741 GGGCACTCAA GCACAGAAA GGTTGAGCT GATATTGTCG CTGACCTGAT GGTGAAATA
 88801 GGGCCTTCT GAGTGGTAG ATTGACCTT GGTGTAAGAG AGAACTTAGT CACTTGGAAAG
 88861 GTAATAAGCT TCCATTCTTA GAATTATGCA AGGGGTTCTT GCCTGGAAGG TGCAAGGAG
 88921 ACTCTACAC TGGGCAAGCG GGGGTGGCTC TGGAAACCTA ATGATCTGTT CCAGTTTG
 88981 GAGTCTGTTA CTCTCCATCC CACCATGCC CAGATTAGA AGACAGCTG CTCACCCTG
 89041 CAGCTGCTGG AAGCATGTGG GAGAGGCCCTT CCGGAGGCCA GGCCCACCTT TGTTGCCTAA
 89101 GAGGTAGGTG GCTCTGACGG TGGAACCTT ACTAGGCAGG GAGGAGGGAA CAGGGAAAGG
 89161 CCAGCCTGGC CTAGTGGAGG AAATATGTGA GGAAGGTGGG CCTGTGGCA CCCGTAGGTC
 89221 AGGACCTGTG GGTGCCATC ACTCAGGCCCT TTCTCAGAA CAGCCCTGAA ACTCATCTG
 89281 TGGGCAAGT CTGATGCCA CATGGACTCT TTGCTCTGCT CTTGCTCTCA CCTGCTAATA
 89341 GACACATCTT CAAACATGGT AGGTATTAG GGAGCATGGT GATGGAGGAT ACAAAAAAAA
 89401 ACAAAACCAAA CAACACACAA CCAAGAGGC CAGTGACAAA AAGCTTACAA CAACACAGA
 89461 AGTCACAGG CAAGACCCAT AAGACAAATG TAGAGGAAAG AGCAGGAGAT CTAGAGTTAG
 89521 AGAGACCCGA GTGCGTATCT TGGCTCTGGC ACTTACTGGG GCAGCAGTC GCAATTACTT
 89581 TATCTCTCTG AAACTCACTG TTCACTTTCC TCATCTGTA AATGGAGTTA ATAGTACTGA
 89641 TTTCACAGGA TTGTTGTAAGA GATTATATAT TATATATACAA TAAACACCGT ACATATACAT
 89701 CTGGCACACA GAATCTGGCA TGTCATAAAC CTTAATAAAAT TGTCTGTT ACTGTTATTC
 89761 TTTTGCTAC AGGACAAATAA AATTGGAAAG AACTAACCGT AAGTGTACTT GGACCGAAA
 89821 GAACCTCAAG GTTGCATGG GCTGGAGAG TCAGTGAAGG CTTTGTTCAC TGATGACAGA
 89881 GACATTGAG CACCTCTAA ATGCTAGGCA CTGTGCTACA GGTTCCAGT ATCAAAATCA
 89941 ACATGGTGCCTG TGCTCTGGG AAGCTCTAG TGAAGTAGCT TTGAAAGAAA ATAATTCTA
 90001 ACCATCAACT GCGATAAGAA TCAGGAAGGA AAGTAACAGG GTGCTACGAA AGACGGGAAC
 90061 GGATTCTGAT TTGGGAGAC GTCAGGGATG ACCTCCAGA GGAAGCAGCC TTGAAGCTGA
 90121 AACCTGTAGG AAAATCGGGG TTAGCTACAT CGGGCTCTGC GGGAGAGTCC CCCAGCTGA
 90181 AGGTACAGCA GTGCAAGGC TTGAGGAGTG GAAAAGGTG CGCTGTTCAAG AAACGTGACA
 90241 CCTGTGCTG GGGCACGGT AGGAAGGGAT GATGAGGTCA GAGAGGCTT GGAAGCCATG
 90301 GTCAAGATTGTTT CTTAAAGAAC TGGATGCCAC CGAAGGGTAT TAAGCAAGGA
 90361 CAAAAAACCA AACACCGCAT GTTCTCACTC ACAGGTGAGA ACTGAACAAT GAGAACACTT
 90421 GGACACAGGG TGGGAACAT CACACACTGG GGCCTGTTG GGGATGGGG GAGGGGGAG
 90481 GGATAGCATT AGGAGATATA CCTAATGCTA AATGATGAGT TAATGAGTGC AGCACACAA
 90541 CGTGGCACGT GTATACATAT GTCATGTTGT GCACATGTAC CCTAAACACTT
 90601 AAAGTATAAT ACAAAAGGAAAGT GATGTAATGT AGATCATCTG GCTGCTGGGT
 90661 GGGGACAGAA TTGGGAGGA CAGGGACAGG AAGATGTTGT GTAGTGGCTC AGATGGGAAG
 90721 CAAATGGTCA CTTGGGCCCCG GGTAGGGGGC AGAGATGGAA GCAGTGGTGG ACTCGAGAGG
 90781 TATTCTGAG CAAACCTAC ACAGGCTGGT GATAGATTAG ATATGAGGGC ATCCAGTGGG
 90841 CTTCAGTGCAC CCGGAGGAT AGATTAGATA TGAGGGCATC CAGTGGCTC CAGTGCACCA
 90901 GCAGGTGAAG GTACACAGGG CTGGCTTGTG AGGGAGGCT GTCAAGATCC CACGTGGAGG
 90961 CGGGTGGGGA GTCTCAGGGG GCTGAGCCTT GAGTTAGGGG CAAGTTTGG AAAGGTTGCA
 91021 GCTGGGAGG GGAGACCCCT TACGATGGGA GTTACAGGAA CTGGGAAAG AAGTAGGAAT
 91081 GAGCTTGCCT CACTCACCGT GGGTGGCACA GAGTGGGCA AGCCCTGCTG AGAGGGAGG
 91141 GCATGAGCTG GGGGCCAAGA GTGTGAAACC GGGAGGATGG GGTGGGGTGG GGCCCTGGA
 91201 GTGGGCGGGAG GCCTCTGCCT CTACATGAGAT GCAAGAGATA GTAGGGAGCT GTTAGGTTCC
 91261 CAGGCCAGCA GCCAGCCACC CAGGCAGCAC CGGGACTCCA CGGCAGTGCC GGGGGATGAA
 91321 AAGGCATGAA CAAGCAGCTG CAGAGCTTGT GCTTCTGAG AGTATTACT GACCACCTGC
 91381 TCTGTGCTCA GAAAGATAAC GGGATGTTTA TGAAAAGT CTCTGGCTC CCAGGACACC
 91441 ACTTTCTCAT GGGTGTCTC CAGCCTCACT GCCTGCTGCT TTGCTAC TCTGATGGCT
 91501 CTTCATCCTT GCCGGGCATC TGAAAGGTGG AAAACGGGAG ACCACAGACT TCTTCCCTT
 91561 CGCTTCTGTC TCTTATTCCT TTAGAGATGT CATCCAGCTC CATGGCTTAA AAACCTCTA
 91621 CATTCTGCTT TATGATTCTA AACAGATTTC CTCTCTCAACT TGTCTCTGA GTTCCAGACT
 91681 TGTATTCTCA ACGGCCCACC TGACATCTCT CTTAGATGT GAATCTCAAG TTAAATG
 91741 CAAAGACAGA GGCACCCAGA CCGCTCCACC CCAACCAATT CTATCCTAGC TGAGTAAGGG
 91801 GCATCTCTCA GTTACTAAGG CAAAAACTCA CTAGGCATCT GTGATTCTC CTTCCATATC
 91861 CAGTGTGTC CAAAGTCTGT AAATTCGACC TCCAGGAAAC CCAGAATCCA TCCACTCTA
 91921 CTGTCCAATG CTACCAACCC TAACCAAGGC ACCTCCATT CTCTCTGGAC CACAGCAAAT
 91981 GCCTCTCAC TTGGCCCGCA CTTCCTCTC GGCCCCCAGG AAATCCATGC TCCACAAAAC
 92041 AGACAAAGTG ATGCAGGTAA AACATAAAAAC AGATCAAGTC AGTCTCAGG TTAAAGCCCT

92101 CCATGGGCTC CCCCTGTCTT TTGGAACAAA ATCAAAGCTG CTTCCGCAG CATCCGGTCC
 92161 ATAAGGCCCG GACTCACCT CCCAATGCTT TCTTCCTTGC TTTCCCTGCT CTCTGCTCTC
 92221 CAGTCACCCCT GGTTTCTCC CTGTTTCTAC AAAAGGTCAA CCTTGTCAG ACCTCAGGGC
 92281 CTTTGCACTT AGAATGTGCT CTTCTGGAA TACTTGCCAT CAATCTCTC ATCATTGGA
 92341 TCTTAGTCA AATGTCACCC TCTTCTGATG TCCCCTCCCT GCTGCCGTGTC ACTCGCTGTC
 92401 ACATTTGTC CTAATAGCAC CAATATCAGA AATAGTCAGA TATTTATTTA CTGCTCCCCA
 92461 CATTGAAATG TAAGTTCTG GAGAGCTGGA TCTTGTCTGA CCTGCTCAGT GCTCTCTCTC
 92521 CCACTCTTAG AACAACTACTA GTTGCATAGT GGGAGCTAG TAAATATTTG TTCAATGAAT
 92581 GAAAATGTC TTTGGATGTG GTCCCAGGTG TTCTTGAGGG AGTGTGACAA GCCACATGAG
 92641 AAGTTATCTA AGAACACAGG GCAGGAGATG GTAACCTCTC CACAGGGCAG GAGATCGGGT
 92701 GATTCAAAGG TAATCTGATA AGCTGATG TGCTTTAGGG TTAATGGATG ACAATATGCT
 92761 ACTCAGTAAAT GGGAGCCTCC AACTGGCCCC AGAAAAGAGG AAGGCCTGAT CCATAGGGG
 92821 GGTCTGGAG GGAATCCTTC AAGGACAAA ATGGAGGGCA GATGAGGAGC TATTGAGA
 92881 CTGGTGAGTG TCTCAGTCCA TTTTATGCTA CTATAACAGA ATATCTGAGA CTGGGTAATT
 92941 TATAAAAGAC AGAAATCTAT TTCTTTAAAA TATGTTTTG GCAGGTAGAG GTGGGGTGTG
 93001 ACTATGTGA CCAGGCTGAA CTCCGTACCT CAAGTGTAGA TCCTGTCTTG GCCTCTCAAA
 93061 ATGCTGGGAT AATGGGTGTG ACCAACAGCA CCTGACCAAG AACAGAGATT ATTTCTTATA
 93121 GCTCCAGGAG CTGGGAAGTC TAAGGTCAAG GGGTCCACAT CTGGCAAGGG CCTTCTTGCT
 93181 GTGTATTCC ATGGTGGAGG GCAGAAGGGC AAGAGAGCAC AACAGAGAGC AACAGGGTGC
 93241 AGAACTCACT TTTATAACAA TCCCACCTCTC AGTGACTCCC ATGAAAACGA CATGAATTCA
 93301 TCCATGAGGG CAGCAGCCTC ATGACCGAAT CCCCTCCCAA ATGCCCACC TCTCAACAGT
 93361 GTTGCCTTAG GGATTAAGTT TCCAACACAT GACCTTTGGG GGGCATATT AGACCATAGC
 93421 GGTGGGTCTC TGCAGTCTCC CCAAATCTGC CACGTAAGT ATCGAAGATT CCACGTGTTAC
 93481 CCAGTGTGCA CACATTTCCC AAACCTACAT GTCCCCCAT CATACTATA TACACACCC
 93541 CTTGCAGAGC TGGTGGTGCAC TCTTTTATC ATGTTCAAGT ATTTCACAAA GGTCTCAGGG
 93601 AATGACCAAT TGGCCAGTCT CTGAGAAATT CCATGGAAAT TTGCAAAAT AAAACACCTA
 93661 TAAAAACTGA CATTTCATCA CAAAAATCCT AATAAAATT TTGCACTACA TATGCACAAAT
 93721 AGAATATTAT TTAGCCACTA AAATGATACG AAGTTTAGGT AGCAAGATAG AAAATATTAA
 93781 TGAECTATAAT TTTAGGCTAA CCCCCCTCCC GGGCATAGAT ATGGTATAAA TATTATGCCA
 93841 TATTGTTATG CTCTGCATAA TGACGGTTG GTTAACAATG GACCACATAT ATGATGGTGG
 93901 TCCTGTAAGA TCAGAAATTA CTTTTTTTAC TGTAGCTCT CTATATTAG ATACATTAG
 93961 GTATGCAAT ACTTACTGTG TTACAATGTC CTACAGTGT CAGTACAGTA ACAGGCTGTA
 94021 CAGTTTGCA GCCTAGGAGC AAAAGGTAC ACCATATAAC CTAGGGTGTG AGTGANCTAT
 94081 GCTACCTAGG TTTGTGTGCA CACACTCTG GATGTTGCA AAATGATGAA ATTGCCTAAT
 94141 GATGCATTC TCAGAACATG TCCCTTGTGTAAGCAACAC ATGACTGTAG TTATCGTATT
 94201 GGGTTAATGG GATTATAGTG GATTGGCCCT TTATGTTTC CAAACATTAA TAGAAATTAC
 94261 TTTACACATA TCAAAATTAA AATATTAAAA TTAAATACAA TGTCTTTTG TGGTCGTCTG
 94321 TGTCCCAAGTC ATGTGGGCTT TTCTGCTCCC CCTCAAGCAG CTGGAGAAAG ACCAGGGGGC
 94381 CATGGGTACA GCCTCATATG TGCACTGTGAGGAGCATGGG GATGACGGGG CGGAGGGGG
 94441 AAATCAGCGT AAAGTCATAG TGATAACAAAT GACAGCGAGC CGCATTGCTC CAACCTCATA
 94501 TGCCGAGGC ACTGTGCTAA ACATGTTGTT TACTCCGCAC ATCATTTCC CCATAGAGGG
 94561 AGAAACTGAA GTTCAAAAGAG GCTAAGGAGC TTGTCAGGG TTACAGGGT GGCAAATGAC
 94621 AGAGCAAGGG ATTGAAGCTG GTATAACTCC AGAGCCCCGTG CTCTTCACCA CTTGAATGAG
 94681 GCCAATAATT TCACAAATAT TAAGCCATTA TGTTCAGGGT GCTCCAAAAA CCTTCAGCT
 94741 GCCACTCTA GGTGTGAGTC AGGAGGCTGA AGATGCCAGG GCAGGCAAGG GCAGGCCATT
 94801 AACCTCGTTA GCAAGACAGC AGGGGAGACA GCTCGCCCCA GTTCCCATCC CCCTCCAGGG
 94861 TATCAACAGC AGGCCCTAGGC GATGAGACAC AGAGACTCT AGAACGGAGG CAGAGTCAGG
 94921 AAGAGGAAGG GGAGAACGAC CTGCTCACCG GAGACAGGGG GACTGCAGAA GATGAGCTCT
 94981 TGTGGGATATG GTGGGATAGT GTGGGGCAGA GACATCAAGG GAAACCTGG GGCAGGGCTC
 95041 TTGACCCCTC TTCACATGTG ACTCTGGGAA GTGCCACAG CAGAGGGGG AGCCTCCAGT
 95101 CCACCTGGGT TAAGGGATGG AGGCTGAAC CACTTTATA ACAATCTCAC TCTCTCAGTT
 95161 ACTCACTCCC ATGATCATGA CATGAATTCA TCCATGCGGG CAGCGCCTC ATGACCAAAT
 95221 CCCCTCTAA GGGTCCACAG TCTAACACT GTTGCAATTG GGATTAAGTT TCCAACACAT
 95281 GAATGCTGGG GGACATACTC CCCAGTATG GAGGATGGTC CCCAATATG GAGGACGGGG
 95341 TCATCCCCAA GCAGCCCTCA GTTACCTGTG TGATGGATGC TGTGGTAGGT ATGAAGTGGG
 95401 ACCTTCTGGA TCCAGACAAG GTGACTCCGT GGTTGGGGCT CAGGATGGAC CCTCCAAGGC
 95461 AGGGGCCAGA TAGAAGGGAGG CAGAGACAG AGGGGTCCAT AGGAGGTGAG GGGATCTGG
 95521 CCAGTTCTCC CTGTAAGCCA GGCCTCGGGT GGCAGGAGGA AGCGAGGCTC ACTCCCTGCA
 95581 GAGCTCAGTC TGGTGCCACC GCTAAGACCA TAGCAGCCAG AGGGGAGAG GAATGCTGTC
 95641 ACCTCCCCAA ATCACTGAGA GCCTCCAGA AGCTCACACG CCCGCCAGA GGAGAAAATG
 95701 CACGAGGAGC ACCTGGATCA CCTGATAATT ACCTTCAGAT GCTGTTGAT TTCTGTAGCC
 95761 CACTCAATT ACCAGATGGG AATGCGTAT TTGCTCATAG TGAGGGTTTG CGAGGTAAGC
 95821 CAGATCTATG TGTCCCTCC AACTCTCTC GACTCGATG GAATTACACT TACGACCCCTG
 95881 CCAGGCCACT ATGCTAGACT GCATGGTTCAG AGAGCTATG TGACAGGCAC TCTGCCAGGG
 95941 GCTTCACATG CATGGTCTTG TTCACCTTC TCAGCATCGT GTTGCAAAAT ATCTTCATC
 96001 TCAGGTGCCCTC ATGAAACTAG CTATGTGGCC TTGGGAGGT CCCTTGACCT GTCTGTGCTT
 96061 CAGTTCTTC ATCAAATGCA GATACTGATT CCTGCCTAGC AATATCAGTG CATAGAATGC
 96121 ACTTGATAAA AGGTCAACAC TGTGGTTAAC AGTCTTTTG AGCACAGAGG GAGGTAGGA
 96181 CAGATAGGTG ATGGCAGCAG GGTAAAGGGG CTCCCTCGGA GAGAAAGAAA TGAACAAAAA
 96241 CCACAACCTC TTAACCTGTC AGCAATGCAAG GTGATCTTAT TCCGAGTCAG GAAAGTGAGG
 96301 TAAATTTA AGGGACCAAGA CCCCCCTCCC AGGAACACAG ACATACCTCA GAGCCCCAAG

96361 GAGCCATAGG ACAGAAAGGT GGGGAGACAG GAGTGTCCAG GTCTGTGTA TCCTCGAAGT
 96421 CACTGAAACT TCTTCCTGAC CTGCTGCCTT CTTCATCCGT TTGTGGGCTC TGTGGGACAA
 96481 ACACATTAGT TCCCAGGCCA GGACAGCCTT CCGAGGCTG GAGACTCATC CCAGCCTCCT
 96541 CTCTGGGACT CAGGACGGGC TGAGTGGAGC TGAGGCTGGT CCTCAAAGTA AGGTGTTGTT
 96601 GGAGCCTCCT GGGGTAAATT CAAGGGAGTT TTGTGTGAAA GGGAAAGCAGT GCCACCAGAC
 96661 TCCTCCTGAA GCTAAGCCAG GATGAATTAT TAAGGAGCTG AAGGAAGCCA GGGCTTAAC
 96721 AAAGGAGGCC CACGTCTATC TCTGTGAATA TTTCAAAACCA TCTGGACCTA TTTGTGGAG
 96781 AGAAATTCTT TCGCTGCTC GAAAGACAAG CAGGGAGCT GCATTATGCT CACATAGGAC
 96841 AGTTCTCGT CAAGGGAGGC CTGCGTTCA GGGCTGGAT CCACGCTAG AGAAAGCTTC
 96901 TTAGTGTGTC AGAGCTATT TAAAGTGCAG AACACTAATA TGTCAAGGG AGAGGGAGGG
 96961 AGAGAGGCTC GGGGGTGAAGA TGAGCTGAGGC CATCCAGGCC CCTGAAGGAG CCTCACTTC
 97021 CTGGGTGGGT GAAGGCCACTC ATTGGGTCT TGCCGTGACAC CCCCAGATCT CTATTTAAC
 97081 CCCTGAAAT AAATACCTAG CCTTGTTGTC AGAGCTGATG GTGGGGCCAC AGGCCAGAA
 97141 AGAATTAAAT CATTACTCT GGCTCTTT GGTTTAATAG ATGCCATATC TGGCAGCTTA
 97201 ACAGGGATGTC GTCCTAGTC CATGTAAAAA ACAAAATTA AAAAATTA AAGAACGCTT
 97261 CTTTATGTC TGTTTACAAAC TGTTTAAAC ATACATTGA ATGTTCATC TTCTGTCTT
 97321 TCTGTGTC GGCTCTGAGG CACCCAGCAT CCTCTCTAGT TGGGAGTATC CCGCTGTCCC
 97381 CCTCTCTTC TGAGGCTGCG TTGTTCCAGT GGGTGTCTG GGCCCAAATA GCCCTATCAG
 97441 AAATGAAGCA ACAGGCTGGG TGCACTGGCT CATGCCTATA ATCCCAGCAC TTTGGGAGGC
 97501 CACGGCGAGT GGATCGCTTG AACCCAGGAC TTGAGACCA GTCTGGCAAA CCTAGCGAAA
 97561 TCCTGACTCT ACTAAAAACA CAAAAATTAG CTGGGCATGG CGATGGTGC CTGAGTCCC
 97621 AGCTACTGG GAGGTGAGG TGAGAGGATC GCTTGAGGCC AAGAGGTTGA AGTTGCAGTG
 97681 AGCTGAGATT GTGCTCTGC ACTCCAGCCT GGGCTACAGA ACAAGACCT GTCTCAATCA
 97741 ATCAATAAAC CAAGAACAG CCAGACCCAC ACTTAGACTG ACCAGTCATG TGTTGGGTA
 97801 GGAGACGCTT ATAAGATGTC ATTGTTCCGT TTGTCATA TAGAAAAAA TAGCTTGAAT
 97861 GGCAGTGTAA ATAAGGATAA AAGGAGTGA CATTGCTAA GCACACACTT GGTTCCCTA
 97921 CAAAGCAAG GCAGTTGGA ATAAAACAAT AAAAAGTAA CTTGAAATA AAATAATGG
 97981 AAAATGCATT GTAAAAATAT CAGTTGGA TTGACTTGA GAGTTGCCAG ATCTTCCC
 98041 TGCACCACAT TTCACAGTTT ACAAAAGCTTG TTGGTTTACG TTTCCATT CTTAGGGGCC
 98101 TTGTTCTGGA GGTCTTACT AGAAGCCAGT GCCGGCAAGT GACTGGGAA AAGGCAACTC
 98161 TTCATGGAGG GAGGTGTCAG ATGCCAGACA TGGCTATAAT ATTTTACTTT TTATTTGCC
 98221 ACAAGGGAGA AAGGAAGCTT GACATTGGA GAAAAAAAT TATGTGGATG GAGATATGTA
 98281 GATAACAAAGG TGTTGAAAT CAAGCTTCCGT TGAAGTCAAA GAGAAAGTTA TGTTTAAGA
 98341 CCCAGATAA TTTCTGATTCA AAGGTGAAGT TTAGAATGT GAGAGGAGA CCTGGATTAG
 98401 ACTAAATATT GATAGGTGAT GGTTTATTGT GCAATTAAAA AAAACACAAA TTATCAGTTC
 98461 TTTATCCAGT TTCCAGTGG AATTACGTTG GTGAATAATC GGTGCAAAT AAAACCAACA
 98521 CCAGCTAATA TTTCACTGG TAAACTTGA AAGAATCGTA TCTTGGGAT AAAAACCAA
 98581 CCTGGCAAA AAAAAAAA AACCCAAAAA CAAACAAACA AACAAAAAA CAAAATTAGA
 98641 CTGCTAACG AATACTAGCC ACAAAAGAAA ATATACGAAAG GTCAAGCTAT AATGAAAAC
 98701 TAGTGAGGAT TAGTGTGTTG AATAACCAAG TGTACCTCGG TATTCTCATT TTTCTACATC
 98761 TTACCTTGTG TGTAATTAA AAAAATAATA CAGAGATCAG AATGGGGTCG TTCAAGATT
 98821 CTTTGGTAG ATGTGTTTG TGCTTAGTCA AGGTGAAAGG ATGCCAGGC ACTGCCACTC
 98881 AGAGGTCTCT GCCCTGTCCG GGGCATCTGG GGGGCACCTG GAAGAGGTCT TTCTCTCATA
 98941 GCTCTGCCCA CCCCCACTGTG CAGGTGAGAC CCAGAATGGA CGCTGCTTG AGGCAGCTGA
 99001 CCAGTGGGAT CTGACCTGGA GCAGCTGGT GTGGCCAGGG CTGAGAGCCA GGATGAGTGT
 99061 GGAGGTGTTA GCACCCCTGGA GAGTGAGCAG AGAGGGCAGC TTGCTGTTA GACTGTGTGT
 99121 GGGGCCAGG CAGGGAGGAG CCAGGCTCC TTTCCCTGTG TTGAGGACAT CCTCTCTCCT
 99181 GTTCTTCTT CCTGCCCTTG AAGCCTCTC TTGTCCTGG TTATTTTC TCTCCCTGG
 99241 TCTGACTTCA GGACTTCTC AAGCCAGGAA CTATGTCAG AGATAGACAG GCAGACAGGG
 99301 TGAGAAAGGG GACCCCTCTC GCAGACCCCTC AGGAACCTAG TGCCAGGGGA GGGGGCTCTG
 99361 AAGTCAAGCT TCCCTGAAGT CAAAGAGAAA GTTATGTGTA AGACCCAGAA TAACTCTGA
 99421 TTGAGGTAA ACTTTAGAA TTTGAGAGGG AGACCTGGAT TAGACTAAAT ATTGATAGGT
 99481 GATGGTTAT TGTCATTTG AAAAATACAC AAATTATCAG TTCTTTATCC AGTTTCCAGT
 99541 GAGAATTACT TTGTAATA ATCAATGCAA AAAAAAACCA ACACCAGCTC ACATTTACAA
 99601 TGGTAAACTT TGAAAGAATC ATATCTGGG GATAAAAAA GCAATCTGGC AAAAAAAA
 99661 TTAGACTTGG GCACAGGGGC AGGAGCAGAG ACAGGCTCTC ATGTCATTT CAGATTGGT
 99721 TCAGGGAGTA TCTGAATGTT TATCTTTTA AAATGTTCT AGGGGATGTC TCCAGGCTT
 99781 GTTGAGCTTA GTCTCCTGTA AAATTCACCT CCTACAAATT AAAAATAAAAA GCCAAACAAA
 99841 GAATTTAAC CTCAGCTGAA ATAACAGCTC CAACAAACAT TCTGCACTG AGCACCAGT
 99901 GCCCGGAATC CCTCAAGGGC ACCTGGCCAC TGGCTCTGCC AAATCCCCA CTTCCCTGCC
 99961 CAGCCAGGCC CTGGGTGGCC CCACAGTTGG CCACCTCTAG AATAGGCTCC ACAGGACTAT
 100021 GTGTATTGTT TGTTTGTGTT AGTTTTTTT TCTTATTCAAA TTCTAAAT GGAAGATCTA
 100081 CAGGTGGATA GAACGCTCAC TAAGTAAGTC CTCAGCTTC AAAACCTAGA GGCACACTAT
 100141 TACCCCTACAC ATTAGTTAA GGGAAATGG AAAGGAGAAA AGGAGGGAGG GAAGGAGGAG
 100201 ACAGAGGAAA GAAAGAAAAG CAAGGCAAGG CCTGGAAAGA TTTGAGAGA TAATGACAGC
 100261 TTTTTCTTT TTTTTTGGG GACATAGTCT CACTGTGTC CCCAGGCTGG AGTACAGTGG
 100321 CATGATCTCA ACTCACTGCA ACCTCCAGGT GCCAGAGATT CTCTACCTC AGGATCCGA
 100381 GTAGCTGGGA TTACAGGCAC ATTGTTATAT TTGAGTGTAGA GATGGGGTT TGCCATGTTG
 100441 GCCAAGCTGG TCTCAGACTC CTGACCTCAA GTGATCCGCC TGCTGCAAC TCCAAAGTG
 100501 TTGGGATCAC AGGCGTGAGC CACCGCACCT GACCAAGTAAC GACAGCTTT AATGGTGA
 100561 AACAGAACAA GCTGAGAAGG CTACAGTGA GACTGGAATG TTGGCTCTG ATGTTGGCT

100621 TTGATGTGTA TATGTTGAG CTTCTCCCTT TTTTACATCC CTGGGAAAT TTTAACTTCA
 100681 TTCTGTTAAG CTTAAATGCT TTATGAGAA GTATCCTAAT TTAATAAAGT CTTATCTAGA
 100741 AAACGTTTA TGAAAATAC CTAAGATGTA GAATAGGGAG TAGAATAGAA TCTGGTGGT
 100801 AAAAGTGGGC TGTATACATA AGTTAAAGA AAAGAAAAAG AGAAAATGT ATTTCTGGA
 100861 GCAAATTCAA TTAAGCCAAT TTAATTCTT GATCTTCTT ATTATTTTA ATAGAATAAT
 100921 ACTTAGCAAT TTCCAGTTT GCTGCTTAC TTCGTTGGG CCAAGTCTT TCCACACACC
 100981 ATGAAATCAT ACCCTTATG ATGGGTTTA CCTTTCTT GTGGCCGCTA TCTCTTACA
 101041 AGTCTAATT TTGCCAGGAC AATTTTTTA TTCTTAAGAT GTGGTTCTC AAAAGTTTAT
 101101 CATGTGAAAT TTGGTTGGT TTCTATTGTC ATTGGTTATT CACAAAAGTA ATGTTCTATTG
 101161 CATTGGATAA AGGACTGATT ATTGTTTT ATTTTTATT GCAAAAAGAA AACATAAAGA
 101221 AAAGGCATT TCCATGGTAT GGTTAAAATG TCAGGAAGCT TTGTTACATT AAAATAGTTG
 101281 TTTAAAATGT GGGAAAGCCT ACAGATAGAA AAAAACAAA CAAAAACAT CCTTAGTTT
 101341 CAGGGCAGG ATCTTAATA AGGAAAGATT TTCTCCAATA AGAGAGGCC TTCACGGGAA
 101401 GAAGAGTGGG GAGATCACAT AGGGAAAGGG GCTGTTGGAC ACACCAGAGA TCTGAAGCCA
 101461 CATGGGAGA ACCATCCCAC TGGGGCTAA ATTAAAATGT CCAGATTGGC TCTGGGCACA
 101521 TTTAACTAAA TGTTAGGGT AATAGTGTGCT CTTAGTTT GTGGAGCTGA GGACTTACTT
 101581 AGTGAGCGTT CTACCCACT GTAGATCTTC CATTAGGAG ATTGAATAAG AAAAAGAAAT
 101641 TAACAAATAC AAAAACATAC ACATAGCAAT TAATGTAGGT TGTTTACACA TGATAATGTA
 101701 TGAGACCATC ATAATGGTTA TGGTCTCATC TCATTAGGAG GACTCTATT AAGAAAGGAG
 101761 GATTTAACAG AGTATACTG GGGACAGTTA GTAGAAAAAA TAAAATTGTT TCATGTTGC
 101821 CCCCGAGTG AGCTGAGGCT CTTCAATGAA CTGGTTCAGG TTGTAAGGA CATGGAGTAG
 101881 AGACCGTGC TGGCAAGACC CCAGGGACAG CAGCTGCTGC TACCAACCTT AGGTACAAGC
 101941 AGGATCTGC CAAGGCACTT GCACCTATGT TTCAAGGGAGC AAAAGTCTAT TTAATTGTC
 102001 TCTACAAAAG GGAGATGGGG CTTTATCAAA AGGAGCTGAT TACCTTCAGC CCTGGATCCA
 102061 GGTGTATCTC CTTCCAGGTC CATGGAAAGCAT ATACTGGGTG ACATCAGGAA GACATGAGAG
 102121 GCTAGGCTCA GTGCCCTCTCC AGTGGTTCTC GATTCAAAA CGGATGGTGC CCACCCACCC
 102181 AGGAGGGTGT GGGGTGCTGT GGGAGGGTGG GGCTCCCACA GCTCTTAAAGA GTATTATTGG
 102241 CAAGTTATCC AGTCTTCTG AGCCTTAGTT TCTTCCTGG TAAAACAGGG ATAATAATGTA
 102301 TACCGACTTT ATAGAAAGTGC GAGAACTAAC GCCTGCCAA TAGTAGGGCA CAATTAAATGT
 102361 TACTGTCTTT TCTGCCAAAT TGCCAGTAC AGAAAAGTGT GCTTCACTAA CCTAGGCCAA
 102421 TGCCAGGACT GATTCTAGG AATATCATAG ATATATTCGC AACGTCTAAA ATGATCTCGT
 102481 TATAACAATAG ATCGTTTTA CAAAATAAG ATATATTGAC TGTTTTGTAT CCCATGAAGG
 102541 TGAGGGGAGG GATTTCCTA ACTTCTCTTC CAGAATCCGC TATTAGGGT ATGGACTCAG
 102601 CACCTTCTG TCTCCCTCTA CCCGCTTAA GAGTTCCTCG TCACAGACTC GGAAATGGGG
 102661 ATAGAGCTTG GCCCAAGTTA GAATTTCCAC CACTCCAGGG GCAGAATGCA GAAAGAAGAG
 102721 AAAAGTAGCG GGGGGCCTGG CGCAGCTGCT AATGCGGTGC AGAAACTCTT TGAGCTGCCA
 102781 CCCTGGCAGG GCCGCCCTTG GGGGATGTGC CTGAGGATGG TGGGTGAGCA AAGCCACCTG
 102841 CCCCGGCACA GAGCCTGCGT TTGTTGCTCT GCAAAGCCCT CCCGGTTCA TTCCAGAGA
 102901 CGGGCCGAGA AAGGGCAAGC CAGGCCAGC ACCTGGGGG CGCGCTGGCC AGGCTACCCG
 102961 GATCGTCCAG CGCGACATGG AGTGTGTGTC TTCTCACCG TGCTGCGTGA CAGGAGCCGG
 103021 GCACCCACCGA CGCCATAGGC TTGTCGGTGT CTCAGCACTG ACACCGTGTG GCCCTGAGAA
 103081 AAAGCAGTGA GGGGTCGCAGA GAAAGTCTTG GAGACGGGAG GTTGTCCCAT TTCTTTCTC
 103141 CGGGATGTTA ATCAGCATCC CTGGGACTTT CCTGGTTATC GCTGATGGCA GAAGTCCAGG
 103201 AGGGAAAGGA ATGTGTCTGA GCTCAAAGGC TGAAGGCTTC AAGACTTAGA AGTCAAATTG
 103261 ATCACTAATG AGAAAAGAATG TAATCTCAAG GTGAACCTCCA GAGTAGGGAG TGAGTGTGTA
 103321 TCGTGACCTT GGGGGCTAGT CATCACCCCCA GCCTAGACAA AGATGATGTA AAACACCCGC
 103381 GTCAGGGTTG TGACATCTC CTGACGTGCA GGAGCGACCC CCCCTCCCG GATGATGTGG
 103441 CTGCGCACAGC TTACACTTTC TCCACCTCCG CTGGCCGTAC CTGTTCACT AAGTGGCCT
 103501 GCCATCCACC CAGTTGCGTA ATGGGATTAGT CAAGACTCTC TTAGTCATAG GTAGGAGCCA
 103561 AACTGAACAA GCTGAGGCAT AAAAGCAAT TTACTGCCTC ACCTAGCTGG GCAAGAGGCC
 103621 GAACTAACTC ACAGGTTGAG CACAAGGAAG ATTCTGGGAA GTTGGAGAG AACAGAACCT
 103681 GGGATTGGCA GCGGCCAGGG CTGTTGTCC TGCCCTGTCC TCCTCTGTCC CCCTTCACCC
 103741 ACTGGCCCTCA CTCTCTCTT CTGAGACAG CCCCTCTCCA GTGGTGGGA TACAGCCCCT
 103801 GGCAGCCCGAG GCTCGGCCAC CCCAACATCA AGGGAGAAAT TTTTCAACA GGAATCTTTC
 103861 TAATAGGGAA GGGCTCAGAT GGGCTGCTC AAGTTAGGTG TCTGCCCCC GAGGCAATCC
 103921 CCAGTGGCAA GGAGATGACA CACTGAGATA GGCCAGTCCT GGTACACACCC TCTCTCTAG
 103981 GGCTGCTGGA CAGGAGTTGT CAACAGAGAA GAGTGACAGG AACACTTACT TGGCAGGCAC
 104041 AAATCACAGG TTGCTGTGCT CCACATGCA GATGATGACA CCCTCAGATT TTTCTCTT
 104101 AGTAAACTCT TAAAACATAGA CATATACACA GAAAATGCA TGAATTATGA GTGTGCGATG
 104161 TTTATTCAAGA AGAGCAAGGA CATAAGATCA ACCCAGGCAC CCATCAAAG TAGATTAAG
 104221 AAAATGTGGT GCATATACAC CATGGAATAC TACACAGCCA TAAAAAAGAA CAAATCGTG
 104281 TCCTTGCAAG TACCATGGAT AAGGCTGGAG GCTGTTACCT TAATTGATT AACACAGAAA
 104341 CAGAAAACAA AATACTACGT TTCTTACTT GTAAGTGGCA GACACAAAGA TGGGAAAG
 104401 AAACACCGGG AATTCCAAA CGGGGGAGGG AGAGAGTGGG AGAAGGGTGG ACAAAACTGCC
 104461 TATTGGGTAC TATGTCGCT ACTTGGGTAG GGCCCAAACC TCAGTACCAT GCGATATACC
 104521 CATGTAAACCC CAAACCTGTA CATGTACCCCC CAGAACTAA AATTGTTT AAAAAAGAAA
 104581 ATAAATATGT ATAGAAATAG CAGAATGTA GATACTTTA AGTCACCTT TCAGAGATAA
 104641 CACTGATGAT GTTATCAGCA TTATTATTT TCTTTAAGGA TCTCATTCTC ATTATTTTAT
 104701 TTATATATAT ATATTTTTA GAGAAGGT CTCACTGTG TGCCAGGTT GGAGTGCAGT
 104761 GATGCAGTCA TAGTTCACTG CAGCCTCAAA CTCCCTCAGTT CAAGTGTACCC ACCTGTTG
 104821 GCCTCCACCA GCGCTGGAAT TACAGGCATG AGCCAGCACA CTCCGCCAAG AATCTCATTT

104881 TTTAGCCTCA CTTTTAAAAA ATAAAATAGA AAGTATACTG CATAAAAGAAA AAAAATGTGC
 104941 AGCTTCATTA AGTTCCCTT TTGAATACAC CCATGTAAACA AGCATCCTGA TGGAGAAACA
 105001 GAACGTCACC AGCATCCTAG AAACCTTCCT TGCGCTTCCTT CCAATCATGG CACACCTGTC
 105061 TTCCAAAGGT AAATGCTATC CTGATTTCTG CTACTGTAAAT TAGTTTGATC AGTTTTGGCA
 105121 CTTTATACCA ATGCAACTGT TCAATATTT ATTCTTCCAT GCCTTGCTTC TTACATTCAA
 105181 CATTATGTTT ATGAAACCCA TCCATACTGT GTGTAATACA GAGTGTGGT TTTCACGGCT
 105241 GCATAATATT CCATTCATG AATACTCCAA ACATTTTAC CCCATTTCAC TGTTGATAGA
 105301 CACCCGGCC GTGAAGAGGA AAGTCCCAAG TGCCATTGGG TTGTCAGACT ACTGTGAGAG
 105361 ACTTTACCTA GGCACTTATG TCTTTGAATT TTGCTTGAATC TCATGTTCTT CTTGAAATAT
 105421 ATGTGTAAT ACGTGGCTGCT AATTAAATA TAAGGTTGTC ACATGAGTT AAAAGCATTAA
 105481 TTTCAGCCAA GCGCAGTGGT TCATGCTTGT AATCCCATCA CTTGGGAAG CTAGGCAGAA
 105541 GAATTGCTTG AGCCCCAAGAG TTCAAGACCA GCCTAACCAA CATAGCAGGA CTTCTATAAA
 105601 TAAGAGCGTA TTTTGTTGAA AGTTTTACGG AGTGACGAGT ATGCTCTGT CTCTTGGGAA
 105661 CCACTACTCC AGGCAATTAA TTGCTCTGGC AACACCGTCA CTTGGCAACA CAGCCCCCTG
 105721 GTGGCAGGCC CCTGAACTGC AATCTTACCT CAGAAAAAGT TAATTGCTGT TTTCAAACCC
 105781 GGGCATTATA CGAACGTCA CAAACAGGA CATAATTAT ACTTAGACCA GAAATCATT
 105841 ATGTGATAAA AAGTGTCCA TCTCCATGGC ATTGTATCTA ATCAGGATAC AAGTCCATCT
 105901 CTAAAATAAT GATAAAAATT TTTTAAAGAA CTACCATTTA TCAAGTACTT CCTGTGTCTG
 105961 GTACCGTGC ACACCCCTCA CATAAGTAT TCAGCGTCAT CCCCTCAGCA GTGCTGCAGGA
 106021 GTGGCCATTG TTGTCACCTT TTAGAGACAA GGAAGCAGAG GCTTGCAGAT GTTCAGGAAT
 106081 CATCTCAAGG TAAATGGAGC CAGGATGGCT TCTCCACAA TCATGATCTC CTCTGCC
 106141 TGCGATTGGA GGCACCTGGC TTGTTCCACGC TCACTCAAGG TTCACTCCCA CAGACCA
 106201 CCCACCCCGA GGAGGGCTGG AGTCAGAGAA GAGCTCACT TTCTCTCTGT GGCTCTCTGC
 106261 TGGGAGGCAC GTACTCTAG TTCCCTTCTC CTCGCTTATA CCGCCCAAGGT GGACAAGTCG
 106321 TTTTCTGCT AGTAAAGGCC CAAGGATTG TTCGGGGCAC ATTGTACTAG CGCCCTGGGAT
 106381 TCATGTGGGT GTCTTCCCA TCTTCTGTG CCCAGGCAGT GTCAGGTGG CCTGCTCTCC
 106441 GCCCCTCCTT GACTGCTCAT GACCAGGTGG CAGGTAGAAT AACCAACTTG CCCCAC
 106501 CCCAGGACCT TCTCCATTAA TGCACTGAAA ATCTCCATC TCAGGAACCC CTTGGCTCC
 106561 AAACAAGCTG CCTTTCCCCC TGATCTTGG GAAAGTCAGG CCCAATTTG ACCACAGATT
 106621 CAGCTGCTC AGATCTGTG AGAGGATGAG AAACCTGTTA GGCTCTCTGG CTTATGATGC
 106681 ACTGGCCAAT CCAGCAAAAC AGAGCTCTGC GTTGGGATTG AGGATTAACA GGTGCTTGTC
 106741 CTGCTGTGA AAACAGCTCT TGCGAGCATG TAGGGTAGTC CAGCATGTGA CTTCTGCAGG
 106801 CAGGTGGCCT GGATTCACT CTCGCTTCA CCTCTGCTG GCTCTGTGG CTTGAC
 106861 TTATTCACC AACCTGACCT CAGAACCCAA CGCTTAGCAC AGCACATACC TCACTGTGGC
 106921 TGTTGTGAAA ATTCAATGTG AAAATGCAGA TGCTTGGAAAG CATTCAAGCCC ACAGGTC
 106981 CGCCACAAAT GACAGCCCTT GCCATTCCCT TTGCTTCTTC AGCTGCCAA CTGCC
 107041 AGTTCTGATC CTAAGACTTA CTTATTTTA GTATGAAATG CTTCTATAAT ACAAAAGGAAT
 107101 ATATATCACA CATTCACTA TAAATGTGA AGATCAACAA ATATATATACC AATGTA
 107161 AACCTCTCT CTTCACTTAAAGAAT ATTACTGACG AGTCATAGT GTTCACACCT
 107221 GTAATCTAGC AGGTTGGGA GGCAAGGCA GGAGAATCAGC TGGAGGCCAA AAGTTGAAA
 107281 CCGGCCTGGG CAACATGGCA ACACCCCAAC TCTACAAAAA AATACAAAAA TTAGCAGATG
 107341 GTGGGGGGGTG GTGTGCACCT GTGATCCCAG CTACTTGGAA GGTTGAGGTG GGAGGATCGC
 107401 TTGAGCCTGG GAGGAGGAGG TTTCAGTGA CGAGATCAT GTTACTGCAC TCTAGCTGG
 107461 GTGACAGTGA GACCTGCTC AAAAAAAACACAAAGAAG AAGAAGATTA CCAATGCTT
 107521 TGTATATCCC TCTAATCATA ACCTTCCCTG TAATGATAAT GAATCTGAAT TTTGCC
 107581 CCTCACTTGA ACGTATCATT TCTTCATT TTGTTATATG TTACTACAT ACGTATGTAT
 107641 TGCTGATAAC GTATTGTTAT GTTGGAGATA TCTGGCTCTT GATGCAAATG CAATC
 107701 TATTTCTT TTTCACCTAA CATTGTGTTT TTGAGATTCA TCCATGTAA TATAGAAGC
 107761 TGTAGTTCAT TATTTCTT TGCTGTACAG CATTCCACTG TTTGAATGTG TGTCATTCA
 107821 TTGATCTACT CCCCTATGCA AACAGTGCAC CCACAAACAC TTTCATAGGT TTCCAGGTGC
 107881 ACGTGTGCAC AAATTACTCT ACAGTATCCA TCTAACCCGA AAAGTCTCA CTGATGGCTC
 107941 AGTGGCAATG GCCTGCCCTC TGTTTGGTT TTCACTGTGA CCAGAATTCT GAACCC
 108001 CTCTCTGCTC CATTAAATCTT TTGTTCTGTA TTTAAAGTA GGATTCAAGGA TTATGAGTGA
 108061 CAAAAAGCTC AACTCAAGTG ATATCAGGCA AAAAAAAAGG TGGGGTAGAG
 108121 AGAACTGTAT TGGGATGAAT GATGGAGCAG TTAGGAGGGG ACTGTTGAA TCTTAGACTA
 108181 AGCACTGTCC ACACAGCCCT GTGGATCTCT CGGTACAGGG GCTGGCCAC CCCC
 108241 CTTCCATCTT TGGACAGACT CTTCCCTCAT CGGAAACAGA TAGTGGCAGT CACTTCTGC
 108301 TGCAGTGCCTT AGGTCTGAGT CCAGTGGGAG GGAGCATCCA TCTCTGCAGC CTTAA
 108361 CTCCCTGGGTG GTCACTGGTC ACCTGCTCAT CCCTGTGCTA ATCTCTGTGG CCGGAGGA
 108421 GCAGTCCATC CTCGAGTCA GGGATTTCTT CGGTACAGTC ACCTGCTTC GGCTGAGAG
 108481 TGAAGTCAAG GGGATTTCTT CCCAGGGAAA TAGGGTACT GTTCTAAAAA GAAGGGTA
 108541 GACCACAGCG GGGCCAAACAGTAAAGTGC ATACAGGGTGT GGTGGCTCAC ACCTGTA
 108601 CCAGCACAAG GAAGCAGG TGAGCAGATT CCTTGAGGCC AGGAATTGAA GACCAAC
 108661 AGCAAAATAG TGAGATCTG TCTCTACAA AAATAAAATT AGCTAGGTGG GGTAA
 108721 ACCTGTAGTC CCAGGTACTC TGGAGGCTGA GGCAGGAGCA TCACTGCAC CTAGGAGATC
 108781 AAGGCTGCAG TGAGCTATGA TCATGCCACT GCACCTCAGC CTGGACAATA GAGTAAGAGT
 108841 CTGTCCTCAA CAAAAACAA ACAAAAGTG AGTGTCAAC GCGCCCTGGG CTGCTGTGT
 108901 GACCTTGAT AAGAGACTTC ACATCCAAA TGCCCAAGTC TTCACTATTG ATGAGTATGG
 108961 CAGAGCTAAA GGTCCCTGAG ATTCCATCTG ACCTTGACCT CTGATTGTGT GGGCTCTGGC
 109021 TGCCCTGTCACCCCTTGCACCTCT GTGCCCTTTA ACATCATAGC TTGTCAGTC
 109081 CTCCAGATCT TTCCACTATA ATGTCGAGTG CATACTCTG GGATCACCTC TGTGAGCTCA

109141 TTCCCTCTTGC TGACAACATG GCTTCTCAGA GGTAGCACAC AGGACAAACAG CTCCCTCAAAC
 109201 CGCACCCCT GCACAAACTG GGAGCCTCTC CAGAGCGCT CATCATGCTC GCCACCACCG
 109261 GCCTTCAGG CCACAGCCTT CAGAGCTTCTC TGTGGATCTC CATCTCCCTT AAGCCCCCTGA
 109321 ACCTGCTCTC TTACGGAATT TTGGCCTTTG TTTCCTCCTC TCAGAGTTCT TCTTGCTTGC
 109381 TCACATCAAT GAAGCCAGCT GCCATATTGT GACCACCCCG TGGATAGGCC CATGTGGCAA
 109441 GGAACGGAGG GTGGCCTAGG GACACATGGC CCACGAGGAG CTGACCCCTG CCAGCAGCCA
 109501 CATGCATGAG CCTGGAGACA GTGCCCTCG CAGCTGAGGC TCAGGGCCAG GTCAGTCCCTG
 109561 GCCAACACCT TGATTGAGA TGCCAGGTTA AGTCATGCC CAATTCCCTGA CTCACAGAAA
 109621 CTGATTGAGC TCCAGAAAAC CATTGTCAG AACAGAGAT AGTGGGCATC CTGATCTTGT
 109681 TCCTTAGCTG TAAGGAAATG AATCTAGGAT TTCACTTCAC TATTCAAGGAT GACAGGTTT
 109741 GTTTTTCAT CAAATATATA CTGTATTGT CAATGCTCTG CAGAGAAACA GAACCGAGAA
 109801 AGAGAGAGAG AGAGAGAGAG AGAGTGTGTG TGTGTGTGTG TGTGTGTGTG TGTGTGTGTG
 109861 AGTGAAGCGA GAGAGAGACA TTTTGTGTTT GTTTTGTTT TTTGAAATGG AGTTTGCTG
 109921 TTGTCACCA GGCTAGACTG CAATGGCGCA ATCTCGGCC ACTGCAACCT CCACCTCCCG
 109981 GGTTCAAGCA ATTCTCTCTC CTAGCCCTCC CAAGCTGGGA TTACAGGCC ACACCAACAT
 110041 GCCC GGCTAA TTTTGTTATT TTAGTAAG ACAGGGTTT ACCATGTTGG CCAGGCTGGT
 110101 CTTGACTCC TGACCCCTAGG TGACCCACCC ACCTTGGCCT CTCAAAGCAC TGGGATTATA
 110161 GGC GTGAGCC ACTGCACCCCA GCCAAGACAT GTATTGTAAG GAATTAGCTC AACGTGGTGG
 110221 AGGCTTGGTG AGTCCAGAGT CTGATGGGGC AGGTTGACAT GCTGGAGACT GAGGGAGGAG
 110281 TTGTAGTTGG GGTCCAATGG CACTCTGCCA ACAGAATTCC TTCTTGCTGC GGGGAGGTCC
 110341 GTTTTGTTCT TATTAAGGTC TTCAACTGAT TGGTTGAGGC CAACCCACAT GGTGGAAGAT
 110401 AATCTACTTT ACTCAAATC CACTAATTAA TTATTAATTAA TTATTATTGT TATAGAGACA
 110461 GGGTCTCACT TTATTCCTA CTC TGGCTT GAACCTCTGG CTC CAGCGA TCCTCTCCCC
 110521 TCAGCCCTCC AAAGTGCAGA GATTACTGGT TTGAGCCACT GCATTCAAGC CCAAATCCAG
 110581 CAATTAAAC GCTCACCTCC CTTCACAGAA ACACCTAGAA TACTTTAAA CCAAATCTCT
 110641 GGGCACTGTA GCGCAGCAA TTGACACAT AAAATGAACC AGTACAGTTA CCCTACATCA
 110701 GGTTATGCGAG ATGCCCTAAT TTATTAAGAG CTGTTCATAA CTTCCGAGTT TGTGTTGAAT
 110761 TTTATCAAGC AATTACTTT ATATTGTTG TCTCATTGT TTATACATTG TTTTAAAAAA
 110821 TTGTAATTT TCTAATATAA AAGCATTCT GCATAGTGGA GATAAATCCA ACTTGATGAG
 110881 GAAAAGGGAG CATCTCTCAG TTCACTCTGC TCGCCTTTCC CCCGGCCACA GCCTGTTGGT
 110941 AGCAGAGGGC AGAGGCCAG CCTTGTCTG GGACTGGTGA CAGACAGCCT GCTGCTTCC
 111001 TGTCTCTGAG CTGGCCCGAG GAAGGGCCCA CAGGCTCGAG GAAGAGCACT GTGTTGATTTA
 111061 CATAGATGCC TTTGTGTGGG GGCCCTGACAG CTGCTGGAGT CACGGCTCTC ACCCGGGTCA
 111121 GGGTGGGCT GAGAGGGCCT CCGGGTGTG GGGTAGTCAG GGGCTGGGT GACAGCAGCT
 111181 AGTGGGGGCG GGGTATATGA GACAGATAAA CATCCACACT CACATGCCAG GTGTTGCTT
 111241 GAGCTCAGC CCAGCCCCT CTCCTTGCTC GTCTGGCTG CTTTATTGAA AGTTCGCCAA
 111301 ATGGAAGGG TCTTTTTTTT TTTCTACCAT GCGCCCTCAC TCTCCAGAGC ATCCGTGCCG
 111361 GCAGTTGCC AAACCTGAG GTTGGGAGCA CAGTCTCCCA GACCACACA TCTTCTGGGA
 111421 CGCCTGACCT CTGCGACAGA GGTCAACCAA CTGTCAGCT CGAGGAAACA GTCCCCACAA
 111481 GACTGCCCA ACCTCAGCA CCAGCCCAA GTTCAGACT CCCAGGCCA CCTTCACCTC
 111541 AGACCAGCTG GCTGCAAATT TGGGGTGTCT TACAACCACCT CTCAGCTTCA GTAATTTCT
 111601 AGGGTGAECT ACAGAACTCA GGAAAGCACT AACCTTATGA TTACAGTTT ATGATCATGA
 111661 AAAGATACAA ATTAAAACCA GCCAAGGAA GAGAGGCTTA GGGCAGAAC TGGGAGGGT
 111721 CCAGTTGCAA GCTTCCAGGG AGGTGTCTC TCTCTTTGGG GTTCTGGACA GCATTCGCTC
 111781 CTCTGGCCG CAATGTTGTA GGATACGCAC AGGGTGTGTG CCACCAAGGA AGTCTACCTG
 111841 AGACTTTGGT GCCCAGAGC TTTACTGAGG CTTGATCAGA TGCTGCTGT GTGACCCGACC
 111901 TTTAGTCTCC AAAGCCCCC AGCAAAACAA GACACTCTG TCAGACAGGA CATTCCAGGA
 111961 CCCAGAGGCT ACCTCCACT AGCCACGGGC AGAGACTGGA CCTCTCTTG GTGAGGTTAA
 112021 TTCTAAACTA CGAATAGGGC ACGTTACAAA GTCACTGCC AAAGACACAG AGAGGTGTGG
 112081 CTGTCAGAAG CTCCAGTCTC GGATTGCCA CCTTTGAGGT TAGCAAGTTA CTTAATCCCT
 112141 GTGGACGCC ACCAGCCAGA GGGCAGCAGG AACAGCCAG TAGCCAAACA ATTGAGTT
 112201 AACTGGTACCT CATTGCTACT AGAAGACTGT ATACTAGGAG GCAGTGCAGT CTCAGTAAGT
 112261 GTGTTGTGT GTGTTGTGT GTGTTGTGT GTGTTGGGG TGCTTATTGA AGGATTGGG
 112321 CGTTGTAGG GTGACATTAC GAAATGTTCA AGGAAGCAA GTGTTGTGT GGATTGGGTG
 112381 CTGTCAGGG A GTGGGGAG GAGCTGGTAG AGAAACCGCA GTCTCTCACG TGAACACAGGA
 112441 GAAGAACCGG AGGGACGAAA AGCCTGGTAG AGAAACCGCA GTCTCTCACG TGAACACAGGA
 112501 GAGGGCACTG TTTGGCCATT TTATGGTTT CACAGTGCC TCATTTTG TCCATGCTCA
 112561 GGCATATTAA TGGAGTGGCC TTGATTCTGT CCTGCTCGT CACGGGGCG CGGTGACCC
 112621 GCCTGACACA GGTGCTCTGT GAAGTTGTG TGCACTAGGC GGACACCAAG GACCTGAGA
 112681 CGAGTCACAG GATGCTTCTGT TCTGTTCTCT TCCCAACCC GAGTTCTAC ACTGATAACA
 112741 AGGAGACGGT AATTCTACCT AACACATTAG GCAAATGTGA GAATTAAGGG GAATAATGTA
 112801 AAATGCCAG CATAAAATAA AACACACAG CTAGTCATT CCCATATCAC TGTTAAAAAA
 112861 ATAAGAGTCC ATGATAAAGG GAAAAAGTGA AGAATCAAAA TAAAGCCAAA TGCAGAACAG
 112921 GCACAGGGAA GGAAGCTGAG AGAGGACATC AGATGAGGGC AGTCCCTGAG CCAAGGTCTG
 112981 TCCCGAAGCT CTCACTACAA CCCCCACCTG GGGCCAGGTG TGTGTGTGT TGTATGTGTA
 113041 CACCGCACCT TGCACTGGGG AGGTAAAGAAA TGTGCTAGCA GCTGGCCGTG GTGGCTCAC
 113101 CCAGTAATCC CAGCACTTGGAGGCTGAG GCAGGAGGAT TGCTTGAGTT CAGGAGTTCA
 113161 AGACCAGCCTT GGGCAACATG GTGAGACCTC ATCTCTACAA AAAACCAATT TTTTAATT
 113221 GCACAGTATG GTGTTACATG CCTGAGTCC CAGCTACTTG GGAGGCTGAA GTGGAAGGAT
 113281 CACTTGAGCC CAGGAAGTTG AGGATGCACT GAGCCAGATT GCACCATGAA ACTTGGCTG
 113341 GGTGACAGAG TGTGATGAGA TTCTGCTCA AAAAAAAA AAAAGAAAGA AAAAGAAAAA

113401 GAAATGTGCC AGCAATAAAC AAAAAGGTTT ATATACCATA ATCAAATGGA AACTGCCCA
 113461 GGAATGCAAA GTGGTTAA TATCCAAAAA CCAATTATG CAATAGAGTA CATTAATAAA
 113521 ATACACGACA AAACCACATG ATCATGCCA TAGACACAGA AAAATCACT GACAAAACCT
 113581 AACACCGTT CATCATAAAA CACTCAATAA ACTAGAAATA AAAGAGAACT TTCTCAACCT
 113641 GATAAGGATG TCCAAGAAAA ACCCACAGCT AAAATCACAC TAAATGGTGA GAGACTTATG
 113701 CTTCTTGTC AGAGTCAGGA ACAAGACAAG ATGTCGCTC TTGAAACCGT TGTTCATCAT
 113761 TGACTGCTC AGGGCAATTG GGGAAAATAA ATAAAAGGCA TTCAGATCAG ACAGGAAGAA
 113821 GTAAAACCAT CTCTATTGG AGATGGCATG ATCTGCTAC AGAAAATTCT AAGAAATCCG
 113881 CAAAGTATT AGAGCCGTA AACAACTTCA GCAAGATTG GGAATACAAG ATCAACATT
 113941 AAAATCCAT TGTATTCTA TAAAATAGCA ATGATAAATC TAAATGAAT TAAGAAAGTA
 114001 ACTCCACTTA TCAAAAAGAA TAAAATATT AGGAATAAT TTAAACAAGT ACAAGACTTG
 114061 GATACGAAA GCAGGAGGAC AAAGAGGAGA AGGACAAGGA GAGGAGGAGG TGGAGGAGCA
 114121 GGAAGAAGAT GAGGGAGCAA TGTTGTCAGG ATGTCCTGTT ATTCCCGAGG CCTGAAGCCA
 114181 CCTGTCAGGT CAGGCATCCC CTGCTCTTC TGTCCTACT CCTGATCTGC TGCCTCCGGC
 114241 TTGCACTTCA AATCTATCCT GCTCAGGCAT GGAAGTAATT GTGGTGTGAA AGTCTGTCTG
 114301 GTGGACACAG CTGTTTACAA CACTGGCCCTG CGCAAGATTAT TTCAACTCA TCTCAAGCCA
 114361 GCCATGAACG TAAGAGAGAA GTCTGTGTT GCAGTCCCT CTGGCAGGT AGACCAAGCTA
 114421 CAGAGGCCAC CCAGGGCTGG GAGAGCCCCAG AACTCCAACC AGGAAGGCC GTGGCTGAGC
 114481 CCAGCTACAG ATCAGGTTAT GCCCGCAGGC CCTGGAGGGT CCTGCAGCCT CTGAGCAACC
 114541 TCAGGTCTGT GCCCTGGAG GCCTCGCCTC TAGAGTCGG TGCTTCAGCG CGGTCCCTCAG
 114601 GGCCTGCAGG ATGGCACCCCT GGCTGGCTCG CACGTGTTAG TCACTGAGCC GCAGCAGGCG
 25 114661 CTCGAAGTCT TCCTCCCTGT AGGTCTATGTT GGACAGGGTG TAGGGCCAGG TGGCCCCGGT
 114721 GAGATCCACT TGGCCACCCCT GGAGCTCTGC GGGGCTGCGC TGGACACCTG CCCAGGGGT
 114781 GGAGGGGTGT CAGGAGCAGG ACCGGGGGGG TTCTACAGCC TGGCAGGGC GCACAGGGAG
 114841 GGCTGCCCCC GCTGGAGGGG CCCCTGCTTT CCTCCAGCT CTCCTGTGTC TTTGCCAGG
 114901 AATGTAATG TGGGTGGCAG GGTCTCTTT GGATGGGTAC TATTGGTGG TAGAAAATAC
 114961 AATTTCATAT TTCATTGCT TGTGGCGTGA CACCCCTAG GCTATTCAAGG CAGCCCCACA
 115021 TGTCTCTACA TCGCTTCAG GCAAGGTGGA AACAGCCAGA AGGGGGAAAGT CAACTGGTT
 115081 TCGGAATGGA TTCTTAGAAC TCACGGTGGAA AGCAAGGGGG CAGCGAGAAC TTGCTGAGAG
 115141 ATGAAAGGTTC TTCTAGAATTG TCAGATGTTG GCTGAGCATG TTCAAGAGGAA AGCTGGGAGA
 115201 TCTAGAATGGA TGTTAGGAAC TGAAAGCTG GGAATATCGC ACAAAGGCAG CAATGAGTTT
 115261 TTCCCACCAAG AAGGAAATAA AATTCTGACC AGTTGGTGAAG AAGGAAAGAT AGGAAGGAAT
 115321 GTGATGGGGG TTAAATTTC GGGAGCCCCCT GAGGAGAGGC CCAGCTCCCT GAGAAGGGGC
 115381 CCAGCCTGGG CTGGGGAGG ATGGTGGAGA GGCAAGTGA GTCCCCCACA GAGCAGGAGA
 115441 GCAGAGGAGC CCTGGGCAGC TGGGGTACT AGGGCAAGTG GAAGGCTGCA GGAGCCGGC
 115501 ATCATGACAA AGGGGACCCC TGCACTAGCGA CCCTGCAACAC AGAGGGCTG CAGGGATGTG
 115561 GGCATCAGCA AGGAGGCCAG GGTGGCTCC AAGGCTGAG TGGGCTTGGGA TGGGCTTAGC
 115621 CTGAAGCCTT CTCTGGCAG ACTCATTTCC CTTCCCTCTG CTGGCAGGCC GGGGGCCCCA
 115681 GGGCAGGCCCT CTGAGGGGGC AGGGCAGCTG GGTGCTTGGG AGGGGCTGCC TCACCGGGGG
 115741 CTGAGTGGCTC CTTGAAGGAG GCATTGACCA GCGGGAAAGTG CAGCAGGATC GGGGCTCGG
 24 115801 GGCAGGGGG GTCTGAGAAG AGGTGGCATT CCCTTGGCTG GTGCTGGTCC TGAGGGCTGG
 115861 GTTCCACCCG GGGGAAGGGC AGCCCCCGGG CCCGGCAGTA CAGCTCCGTC TGCTCAGTG
 115921 CCTGGTGGGG AGAAGGTGGC CCGGAGAGAA CCAGCCGGC CCAGGTCAAT GCTGGGCCAG
 115981 TTCCAGCCCA CACCAGCTGC TTCACTGTCA AGACAACAGC CTGGCTCTGT CCTGTTGTG
 116041 GTCGGGCAAA GCTGCAAGAGT GACTCGGAGG GAGACCCAGG GGGAGTCAA GGACTCTGGT
 116101 GGTCTGCAAT GTTGTGTTGGT TAAAGCTCT AGGTGGGGTCA AGACCCATCC CCCAGCCCCC
 116161 CAAAAACCCCA ATTCTGGATT GCTATTGCTG GACTGTTCAA ATCCCAGAGC CGAGAGGAGA
 116221 CCCACATGG CACAGCGTCT CAGGGAGGT CTGGGGATG GACATGGAGA TGCTAGGTC
 116281 TCCTCTCAGC TCCGTCAGC TTGTTTTGG AGTAGGTGTG GGGAGGTTCC TCCCTGTTCA
 116341 CCTGTCCTGG GAGCCCAAGG CCCCTGCACC ACCCCAGCT TCCCACCTAA GGACCTGGC
 116401 AGGAGCCAGC AGATCAAGGC CTGACACCTC TCCTTCCCTT CGGGACCCCG GCGGTGTTGG
 116461 GCAGGGGCTC AGTGGCCCTG CTGCTCGCCT CTGGAGCTTG GCCTGGCTGG GCAGAGGGGT
 116521 TCCCCTGGG TGACCAAGGGT ACCTCGAAGG CGCCAGATAG GGAGTAGTCG AAGGAGAGGA
 116581 TGAGGTCCAG CCTGCCCTG GGCGGAACA TGGGGGAGA GCTGGTGTG ATGAAGTAGG
 23 116641 CGGGCTCCCA CAGGCAGAGC CGGGGCTCT TGGGGTCAG CTGGCTGGGC ATGGAGTC
 116701 GCTGGTAGTC TGGTGGAAAT CGCAGGATGG TCAGAGGCCA CCACCTGCGC ACAGGTCACT
 116761 CATGTCCCCCT TCTCTCTCT GGGTCACTCC TTAAATGCCA CCCAGCCCTT CCCCAGGACC
 116821 TTCCAAGGGG ATCGCTCTCA GGCCACACCC CAGAAGTGGA TGCTGACCTC CTAGGCTCA
 116881 GCCAGCTCCCT GCCACACCAC CCAGAGGCC AACCTTGTGA CCTAGGGACC CCTGGCCACG
 116941 GCCCTGACCT GCCCCAGGTGG AGAAGTCTTT GTGGCTACAG TAGTCCTGGT GCAGCTGGAG
 117001 GCCCCTGGAGG AAGTTGGGGC TGCGCTGGTG GAGGGGCCCTG CCTGTCAGGA AGCCTTAAA
 22 117061 TGCGCTGGGCC AGGCCCGTGC CTGGCTGGCAG CCACGGGCC TCCAGCCCG AGGAGGTC
 117121 CGAGGTGGTC AGGGGCTCT TCTCTGAAAGC CAGAGAAACA GAAATTGGTC CCTTCTTCCC
 117181 CTTACTCTT CCCCCTCAGAC ACCGCACACC CCAAGCATTC TAAAGGAACA GCGTCATCCC
 117241 AGGCCCTCA GTGCCCTGC CCTTCCCACC TAAGCTCCTG GTCTGCTCT TGATGTGCTG
 117301 TTTCCAGGAC TCCCCAGAAC TGGTGAGGTC ATACCAGGCA TCCAGCAGGT TCAGGGAGAA
 21 117361 AATGTTGCTC CAGATGGCTG AGGAACACCA AAAGAGATCA GCCTCAGGCC CCAGCAGCG
 117421 GCCCCACCCCT CAGGTACTGAA TGAAAATGG GAGTGGGGCG AGGGCTACCC CTACAAATGCA
 117481 TCCCCCCCCC CACCCCTCCC CCTGTCTCTG AGGTGCTGCC CCACCTGAGC TACAGGACCC
 117541 ATGTCCTTCA AGACACCTGG CCCAGGGCTG CCCAGGGCCT TCAGAGCCAC CACCCTCACC
 117601 TTCCAGAAAG CAGATCCGGG GCTCCGGGAT CCTCCTCATC AGCCGTCCA TGAAGAACTC

20 117661 GGAGCCGAAG AGCTCAGGAG GGACGAAGGC CCCGTACTTC AGGAAACCGA CCTCATAGGG
 117721 GGAGAACTCA ACCCACTCTA ATGGGGTGGG AAGGAGAGGC AGGAGTGTGA GGGGAGTGG
 117781 TTGCGGGAGT CCCCCAGCTCT GGCGAAAGAC ACTGCCCTG CCCCCAGCAG GCCTGAGCCC
 117841 TGCTGTGCCT GGCACCAATG CAGCAGATGT GAGGAATGGG CAGTCCTGGA CTGACCACAC
 117901 CCCTCCCAGC CCTTGGTCCA CACCCCTCAAG CACTACCCAC TCTCAAGCAC AGAGAGATGC
 117961 CCTGAGCCCC TCAGCACCAC TGCCCTTCCG GAACATTGGG CAATCTGTG GGAGCTTCCA
 118021 GCATGAATTCT TCTGGGGCT GCAGGGTGC AGAGGTTGGG GGGGGTGGAG TATGTGGGAG
 118081 GGGAGTAGGT GAAACTGTAC CCTTGAAGTC CAGTGTCTC AGATTGTCT CTTTGACATT
 19 118141 GAGGCTCAAG TAGAGGGCA GAGGGTCTG ACCCCGTTCC AGGGCGGCTC TCTGTCTGA
 118201 CAGCCTCTGA TCCATCACCT GGGGGCAGAG GGCATCAGGG CCTAAGTGAG GCTGGGAGTA
 118261 CCCTGGAGGC AGTGGCTCCG CCAGGACAGG GGCAGGATGG GGGACCTGGC TGGGGGTGAG
 118321 GAGGCAGCTC CTCCCCTGGG GTCTGGGAC CATATGTGGG TCAGTTCCA CTCCCTGCC
 118381 TGTGTTCTGA ACAGCCTAG ATGCAGGGTC GCTGGACAGG GCATGCAAGC TGTGTACTAT
 118441 GAATCTCCAG CGAGGGCCAT TCATAGGGAC CACAATGCCA GCGGTGCTTC CAGGCCATC
 118501 TTCGTGGATT TAAGGGTGAAG TTGGGTGTGG ATACTGACAC ACAGACACAA GGACACATGC
 118561 ACACATCCAC GTATTGCTT GTAGAAAAGGA TGCAAGATG CCCTTTAGC ATATACTGCA
 118621 ACCTATGCCA GGGGATGGAC TGCCCCCTCT TCTAGGTTCA CCTTAATGCA GTTCAAGTTT
 118681 CAGCTCCCTC CGAATGCTT CCTGACATCC TTGCTCCCCC CTCATCGTCC CTCTCTTTC
 118741 TTGCTTTGCA TAACATGAGG TTGACCCCCA TTGATTATTC TTGCTTCTG TTCTTAATT
 118801 CCTCCTCAAT AGAGCTCCAA GCTCTGCCT TGTTTATT TATTTTGAGA CAGGGTCTCA
 118861 CTCTGTCAAC CAGGCTGGAG TGCACTGGTG CAATCTGGA TCACTGCAAC CTCCACCTCC
 118921 CAGGGCCAAAG TGATCCTCCC ACCTCAGCCT CCCGAGTAGA GTAGCTGGGA CTACAAGTTA
 118981 CACTTGGCTA ATTTTGCTA GTTTTGAGA GATGGAGTTA CATCATGTTG CCCAGGCTGG
 119041 TCTCAAATTC CTGAGCTCAA GCAATCCACC TGCCCTGGC TCCCAAAGTG CTGGGATTAC
 119101 AGGGTGTGAGG CACCATGCT GGGCAAGCT CTTTGATGGC AACGACTATA GCTTGCCTT
 119161 TGGGGGAGCC TCCACAGAGC CTCCCTGAAT TGTAAGGGAA CAGAAAGGCC TCATTCTATG
 119221 CTGAGCTAGA ATTGCAAATA GAGCTCTCC CAAGCACCAG CAAGGCGAAT GGTGGTGCCT
 119281 GCCTGGAGCA CCACCTTGGG AAGGATTCTG CCGCCAGAGC TGAGGTTGAT GCTGTAGCC
 119341 CTGGGCTGTT TCTACTGGGC AGTATGTGTA CCATGAATAAC GTCATCCTC CTTTGGCCTC
 119401 TCCAAGATGG AAGAGTGTCT AGTCTAGAG GCTCCCAACC TGCCCAAGCA GCCCCTATCA
 119461 TCAGAGTGCAC TGAGTGAATG TGACACGTG AGAGCATGTG AAGCATGCGG ACTTCTGGT
 119521 CTCCCTGAG AATCAGCCTG TCAGAAGTCA TTGAGGTTA TGCCCTTTAA AAACCCCTAG
 119581 ATGATCTGA TGCCCCAGCCA GGTGTGAGGA TTCCCTACCCA GCCTACCCAA GGAGGGTCCC
 119641 TCTCCCACAT CCATGGCAGA ACTCAACACA CCCAGTGTGT GTGCTCCTC TGCTGGCAGC
 119701 TCTTCTGTCC CTGACGAGAA CACTCATGTG TCAACCACAG TCATCCTCC ACCCCATT
 119761 CCAATGGTCC CCAGTATGAA TTACAGGGAA TTTGCAGGAA ATCTCTAAAT AGGGCTCGAT
 119821 CCCTCTGCAC TGATGGTGGT GGTGTACGTA CGAGTCAGTC CTTAGGTACA AATGTGTGTA
 119881 CGTACCCCT TCTCAAATTC TTCTTATTTA GAAACTCTGT TGCTCTTTG ATACATATGA
 119941 CTAGGGAAAATG TGTCCTTT TACTATAGTT GCACAAATTGA TGCCCTAGGG AAAGTCTGG
 120001 AAAAAGGTCA GCACACCCCT CCATCTACTG AAACCCACCC GGATATCTA ATGGCTCACC
 120061 AGTAAGCTGC GGCAGGACTC TCAGTCTCTA ACCCCCCAGT TTCCAAGCAA GTAACCCCAA
 120121 CCCTGCAGAA CACCATTCTG ATGATTGAG GGTCCCATGG GAAACACTGG AGCTTGGTGC
 120181 AGTACCGGCA ATGCAGACTA AAGGGAAGGA AAGTGTGTCA GCTCTTAAC CAGGAAAGGA
 120241 TAAGAGAACC ACTCTGAGTT TCTCCTTGAG GAATGATAAG CAGCTCTAA GCAGTTCACT
 120301 ACTTGCCAGC AAAAGACTGA AGAAAAAAAGT TTCGAAGGAGA TGAGAGTTG AGAACACTG
 120361 ATTATAGCTA GTGTATTTA CGAGCGGGTT TTATATAAAA CCTAATCAAT ATTAAGATG
 120421 TATGTAGACT AATAAAGTCA TTCTGAGCAA AGTTTCAGGG TATGATATT ATTAATATTA
 120481 ATGAAATTGT GATTGGTCAT AAATGGAGTT ATATATTCTG ACATTGTAAG TGATTTGG
 120541 TATTTAGAT TATTACATTA TAGAGGTAAT AAATACCTGT TTTTATCTAA AGTCTTAA
 120601 AAATTAGAAA TTGACACTAA AATTGCCGA AGAAAACATT ATTTGTAAA GTCTACAATG
 120661 AATGTTGGCT TTAAGAGGTA AGGTATGATA ATCAGAAGAA AAAATAATA AATTAAAAC
 120721 CTTTCATATT GGGACAAAAT TTAAAAAAAT GCATAGGTGC TTTTGTGAG ACAGAGTCTC
 120781 ACTCTAACGC CCAGGTTGGA ATGCACTGGG GTGATTACAA CTCACTGCAG CCTCAACCTC
 120841 CCATGTCAA GCTATCCTCC CACATTAGCC TCCTCAGTAG CTGGGACTAC AGGTGTGCAC
 120901 CACCATGCC AGCTAATTTT TTAATTTTT GTAGAGATGG GGTCTCACTA TGTTGTCCAG
 120961 GTTCATTTT AAGAATATAT AGGTGTCTAA TAATACCACA AGCATTCTA TTTTGTGAA
 121021 CAATTTAAAAT ATGCTCTCTA AAAAATGTCT ATAAAGCTTG CACATAAAAC CGTGCATTAA
 121081 GTCATTACTG GAGTATCCAA TGGTAATTAA ATAAAAGTCT GCTTACAGC CTAGGATCAA
 121141 ACTGCGTTTT CTAGCACCTT GTAAAGTGAAG CTGTTAAAAG AGCTTCTCCC ATTCTGTAGT
 121201 AACTAATTC TGCCACTCG CAGGAAAGTC AGTTATTAAC ACACAAACACC ATCAAAACTG
 121261 AATATGAAT TAAACAGTAG ATCCAAAGAC TACATAGCAC ATAGTAATAA AATGAAAATA
 121321 TTCTAGTAA TTAAGTTGGC TTCAATAATT CAATGCCATT TCAATTTAACAC TCCCAACAAAC
 121381 ATTTATTGCA AAACATGACT AGGTGATTCT ATCATATAAT TACAACAGCA GGCCAGTCAT
 121441 AGTGGCTCAC ACCTGTAATC CCAGCACTTT GGGAGGCTGA GGCAGGCAGA TCACCTCAAGG
 121501 CCAGGAGTTC AAGGCAAGCC TGGCCAAAGC AGTGAACACAC CATCTCTACT AAAATAACAA
 121561 AAATTGCCG TGCAATGGTGG TGCAATGCCCTG TAGTCGCAGG TATGCGGCAG GCTGAGGCAC
 121621 AAGAATCACT TGAACCCAGG AGGCAGAGGT TGCAGTGAGC CGATATTGTG CCACTGCAGT
 121681 CCAGCCTGGG TGACAGAGCA AGACTCTGAG GAAAGGAAGG GAGGGAGGGAGG GGGAGAGGGG
 121741 AGGGGAGGGG AGGGGAGGGG AGAGGAAGAA GGAAGGAAGG AAGGAAGGAA GGAAACAGCA
 121801 AAGGGTCAAG TACAGCAAG AGAATTCTGA AGAACAAATT AAGTGGACTG GTAACCTGGC
 121861 CTATCAGATA CCAAGATTAA TCACAAAAT TTAGTCAGTG TGCTGTTGGC ATAGGGATAG

121921 AGAATCTGTC CTGTTGGAAC AGAATAGGGC ACTCAGGAAG AGTCACAGGA ATATAAATAC
 121981 ATGAAAAATA TATGAAACCT GATACATGAC AGAATTAGCA ATGCAGATCA GTGGAGAAAA
 122041 AATTCAATAG ATGAGACTGG GAAAATTGT TATTAATGTG AGAAAAAACAA AAATTGGATC
 122101 CCTATCCACA TGATACCTGG AGATAAATTG TGGGTGGATT AAAGTCTTAT TAGTAAGAAA
 122161 TAAATCTTA AAAATTTAT AAGAATGTAT AGGAAGATAT TTTATAACTT TGGGGGAGTG
 122221 AAAAGGAAGA CTCCCCAAA CACAAACCAT AAAGGAGAAT GATAAATTAG ATTATATTGA
 122281 AATTATGCA TAAAACCATT AGAACTGGTA AGGATTCAACA TCGAGAATAC ATAAGAATT
 122341 CCTCACAAAC AATATGAAA AGACAAATTAA AACAGGAAAA TGAAAAGGGG TTTGAATAGG
 122401 TAATTCTGG AAGAGGAAGC CCAAATGATA AATATACTTA AAAACATCAA CTGCTAACT
 122461 AAAAATCAA AAAATGCAA ACTATGATAA ATGTTCTTT CACACCCAT AGTTAAACAA
 122521 AACTTTAAA GTCTGCCAA GATATAGAGA AAAAAGCAAC TATTATAACAC TGGTAGTGGG
 122581 AGTGTGAATT AACACAACCT CTTGGAGAA TGTTTGGCA ATACCTAGGA AAGTGAAGG
 122641 TGCACAAAC CAGTGGCCA ACAATTTAAC TACTGGGTAT GTATCTTGGG GGCTTAGAGA
 122701 ATCTTAACTG CAGGTGTGCA AGGAGACTTG TACAAAATTG TTCATGCAA TAATATCATA
 122761 TTTTCTAA GGATATATAT GTGAAAATAT TGTGACTGG AAAACCTAGA AGAAACAGAT
 122821 AAATTCTGG ACACATTCAA CCTACCAAGA TTGAACCATG AAGAAACAGA AAAGCTGAAC
 122881 AGAACAAATA CAAGCAATG GATTGAAGCA GTAATAAAAT ATCACCCATC AAAGAAAAGT
 122941 CAAGGACCTG ATGGCTTCAC TGCTGAATT CACCAAACAT TTAAAGAACT AATACCAATT
 123001 CTACTCAAC TATCCAAA AATTGAAGAG GAGGGAAATAC TTCAAAGTC ATTATCCAAG
 123061 GCCAGAACTA CATTGATACA AAAGCCAGAT AAGGACACAA ACAAAAAAAG AAAGCTACAG
 123121 GCCAAGATCC CTGATGAACA TAAATGCAA AGTTCTCGAC AAAATACTAT CATACGAAAT
 123181 TCAGCAACAC ATCAAAAGA TTATTCACCA TTGATCAACT GGAATTCTAC CCAGGAACGC
 123241 AAGGATGGT CAACATACAC AAATCAATAA ATGTGATACA TCACATCAAC AGATATAGTT
 123301 GGGATCAGTG CCCTGGTTAC ATAATCAAAT ATGCCCTCCC AAGTCTCACC TCATTCTTAC
 123361 CTCTCTTAGG ACCTTTGTTA GCTATTGTG TATTAACTA CTTTGGCCCT TGTGATTA
 123421 ACCTCTTAGT AAAGTTTGTG TCTTTTAGAT TACAATAGT CCAGGAAAG ACAATGCTGG
 123481 CACAAGGCTT CCAATCCATC CTGCCTACTG ACCAGGAGAA TTAAATGTT CTGCCTATGG
 123541 GCCCCTTAGA TCAGGTATCC AGAGATTTT ACTCCTCTGT TGCTGGCCA GGGCCTACGC
 123601 CCATAACATA AGCAGGAAGC AGTTACAGAA GACAGATTTC TACCCCTCTG CAGCCCTGTT
 123661 AAGATTAAGG AGGAGTGTCT AATCTCTGAG ACAGGAATGA GGTAGGAGGC AGGGCTTGAC
 123721 TCGAAGTACA TTGGAGACTG TGCTGAAACAG GGAAGGGCTG AAAGCACCTC TCTGTAAGGC
 123781 ATGCTCACCA GTATCATGAC AGTTACTCA TCCCTGGCA ACAACCTGGA AGTTACGACC
 123841 TTTTCTAGA AATTCTGAA TAATCTACTC CTTAATTGCA ATGTAATTAA AAATGGGTAT
 123901 AAATGTGACT GTGGAACCTG CCTGAGCTG CTGCTCTGGA CACACTGCTG ATGGGTGGCC
 123961 CTGCTCTGCA GGAGCAGTCA CAGAGCTGTA ACACTGTGCC CTTGATAAAAG CTATTTTCC
 124021 TACCACCAGC TAACACTGAA ATTCTCTCTT GGGGGACACC AAGAATCTCC ATGGGTGGAA
 124081 CCCAGTTTT GAGGGCTCCCT TGCCCTGCCTA CATGATGATA AACCTAAATT CAGGATGTTG
 124141 TTTGCCCTG GGGCAGGAGT GAAATGAGAC TGGGGAGGGT ATACAAAGGG CTTAACTGT
 124201 ATCTGTAACA TTTCATTCA GGTGCTGGTT GGTAACTTC TGCTATTC TTATGTCATT
 124261 ATGCTTATT CAATGCTTA GGTGTAAGA TTTCTAAATAA AAATAATAAA GATTGAGATT
 124321 TGCTAAATGA ATATGTTAG CTAATAACAA ATAAGGATGG TAGTTCTAA GCAGTGTGTT
 124381 ACTTAACCGG TAAGAATAAT ATTAAAGAA AAAAGTTGAA GGCAAATTCA TTTTGGTTA
 124441 GTTACACGA TAATCATTCT TTCAAAATTC CTTCTATATA AATTATTTT TATACCCATT
 124501 AACGAATTTC TCTTCTCCC CTCCCATGAT AATCATACTT GATTAACCCCT CTGACCCAA
 124561 GATAAAGATG AGTCCTCTG CTACTAATAT ATTCCATACTT TTGCTTTTC ATTGTTTCAA
 124621 ACGCTTTGA AGCTTGTGTC AGATAGTTAT GTGTTGAAACA ACAACACCGT GAAAAAAATAG
 124681 CTGATTTAA TAATTGTATC TGTGTAAGT TCAGCAGTT ACTTACTAAG CTTTCTCTCA
 124741 AGCAGCATGTA TAATTTTT AAAAATGTA TGTTTACTTT TGCTTGTG TGTTTAGAT
 124801 ACTAAAGAAA TATAATAGCT GACATTTAAT GCTTATTCTA TGTTGAGGTA TCACTGAGTA
 124861 TTCAGAACAG CTCTATGACT TAAGAACTAT TCTTATCCCC ATTGTACAGA TCAGAAAATG
 124921 GAGACACAGA GAGGTTAAGT AACTTGCCTA AGGACACACA GCTAGTAAGA GGTACTGGTG
 124981 TTGAAACCCA GGAATCTGG TTACAGAGTT TGTGCTTTA ACAGCACCT GAAGATGAAG
 125041 AAGAACAGTA AATTTCATAT TGCAAGGATAA ATAAAAACGG TTCTAGCAAT CTTACCCATT
 125101 CAAAACAAA GCTTGTGTT TGCAACTGAC TTTGACATCG ACTCATTCCC AGCTTTGGG
 125161 AGATCCAGAA ATTCTGTT TACAGCAATT AATTAAAGAA TCTGCTAACT CAGAAAGGCA
 125221 GCATTCCAAA GTGAGGGCAGA CACTAAAACA CTGGCTCTTC CACGCTGCAA GAAATACGCA
 125281 CGCGCATGTC AGCCGCTTGG GAGCCCTGCC TTGCCCCCG CGGTGCACAC ACGCGCTCTC
 125341 CCCACGTGCG CAGGGCCCT ACCTGGCCGT GCAGCATGGA CTCCAGCACT AGCGCCCA
 125401 GGTCCACAAA GGTGTTGGGG TGCCCTGCT CAGCCCGAG CTCCAGCTCC CGGGGGTAGC
 18 125461 TCGCCAGGCG CTCTGGGAA AAGACCTCCA GCTTGCTCTT GGCCAGGTGC TCCCCGGCGT
 125521 ATCTGATAGG TCCCTCCAGG TCCCTCTGCG ACCACTCAGG GTCCCCGTAC AGGTGGGCCA
 125581 TTGCTCTGGA AGAACGGAGGG GACAGAACAG TAGGAGCCCG AGGCCAGCA TGGGGCGCTG
 125641 CAACTCCCTGC TTCCCCTACTG GCAGCTTCTC CTGACACGTG GCTCTGCCA CCAAGGCTCT
 125701 GCCGATCTGG TAAATGGGTCA TCAAGGGCTTT CTCTCCTTCT AGATCAGCCG TGTCCAATAT
 125761 TAATACGATG TGAGCCACAC AGGAAATTAA AATTTCTTA GTAACCACAT TAAAAGATA
 125821 TACAAGCTGTG TAAAATTAA TTTAATATT TTATTTAATT CGCTATATTCA AAAAATATCA
 125881 TTCAACAAAG TAGTAATAC AGTTATTGAT GTAAATTCTA TTTTTTCAA ACTAAGTCTT
 125941 CCAAATTCACTG TGTTGATTTA TAACACATCT CAATGTTGGAC TGGCTGCATT TCAAATGGTC
 126001 AATAAAACAT GTGGCTAGTG GCTACCATAG TGGCTTACTG TACTGGACAG CCCAGTTCTG
 126061 GACCCACTC CCATCTTCAC GGTCTGTCC TGAAGCTTT TCACACTCTG CTTGTAACA
 126121 GCTGTCACCT AACTCCTCTC GAGGCGTTTC CTCTTCTTT ACTCCTCAAA GTCGCATTTC

126181 ATAAGACTC AAAGGAATGA ATCTCTAAC A GCGGGATGCT GCACCTCGTC ATGCTGCAGG
 126241 GTGAGGGCGT AGGGAGGTAA TGCAGCTTT AGGTAGAAAT AGTATCTGA CCTCAATGTC
 126301 TACACAGTCT TAAATGGACA ATAAGGGAAT CTACCTCGAG TGTGGTGAAG ATTAAATGCA
 126361 ATAATAACAA TGAAAGATAA AAAGAGCTAC CAATTAGGAA TTCCTGGCTG TGCTGGCAT
 126421 TAAGTTAAGT GCTTACTTG GGTGATCCCC ACACCCAGT CCCACGCGGT GTGTGCTAGG
 126481 ATTGTCGCGT TGGAGCTGCA TCAACTGAGA ATCAGAGGAG ATAAGTTGCT CAGTGTACG
 126541 CAGCCAGTAA GTGGTGCCTGC CAGGACTCAA ATCCAGCTCA ATCTGAGTCA GCTACCTCAC
 126601 AGCCTTCAG TAACATCTGG AATGACCTC TCAGGCCGCG TCAGCTCACGG CAGGGCTTCT
 126661 GTAAACGCT TGTCTCTTC TTGTTGGAAA GCTGCTTGGC GCCCATCAGA TCACATGAT
 126721 CACACCTACT TGCCCTCCCC ACACAAAGTG TCCCTCTGGT GCCCACTCTG CCCCCTATG
 126781 GCCCCTGATT GTCTCTGCCA CCCAGTCCCC CAGGATCCTC ACCACGTAGA GCCAGAGATG
 126841 CCACTGAAGT AGGTACACACA GTCTAGGAGG CCCAGCTCTC GCAAGGCCAA TAGGTGGCCG
 17 126901 TAGAGTGAGG TCATGGCCCCG GGCACCTCCT CCTGTGGCCA TGATGCCAC AACGGTAC
 126961 TGGACCACAC ACAGACAGGC TGAGATTAACA AGGAAAGGGG CCACCTGCCT CCTGTCTCAC
 127021 CCAGGGCCCT GCCCCGCTCC TTGCTCTCCAG GGAAGGGAC GCATAGTTGG GTCCGTATAG
 127081 AGTCAGGGTT AGAGCTGGGT CCAGCCCTC TCTGGAAGCA ATGCTCTCA GCTCAGGATG
 127141 GAGCTCCCAA GGGAGTTAGA ACCACAAAGC CCACTCACGC TCCCATGTCT CCACACAT
 127201 GCCCACTCTC CACTGACTGG AGTTCTGGA GAGGCATAGG AATCAGGTAC TCCCTTATGG
 127261 CCCGTCCAGG ATATCTGCAC ACACATCCCT CTGCACCCCC AGACCTCATC CTCCTGCAGG
 16 127321 TCTCTGTCCA GTCGCAAGGG CTGCTTCAGG GCCTTGGCCA CCACCTGCTT CCTCTGCTC
 127381 AGGAAGGCC GCTCCTCTGC ACAGAGATTG AAGCCCAGGT GCACGGCCAG CTCCCTCAGGG
 127441 CTGTGGCAAT GGAGGATCCA GGGGTCAGGG GACACCTGCA TAGCTCAGCC ACTGCTGGC
 127501 CAAGCCCCACA ATGAGCTGCC AAAGAACCCC TCTCCTCCAA TCTCAAGGAC CACCTTGTC
 127561 CATAGCCCCC CCCCCTCTGC TGCCGCCCTC TCACTGTCCTC CACAGATGT AATGCCCC
 15 127621 ACCAGCCCCC TGCCTTGAGC TGCAGCCTCA CTCTGGGGC CTGAAATCAA AGCCAGAGAC
 127681 TCTGCTCAGA CCCTCTAGA CCCTAAGCGG CACATGGGTA TCCCTGGGGA CTTGGTCCCC
 127741 TTCACTCTCA TTCTTTTTTT TTTTTTTTT TTTTGAGACA GAGTTTCACT CTTGTTGCTC
 127801 AGGCTGGAGT GCAATGGTGC GATCTGGCT CAATGCAACC TCCGCCTCCT GGGTTCAAGG
 127861 GATTCTCTG CCTCAGCCTC CTGAGTAGCT GGGATTACAG GTGTCGCCA CCACACCCGG
 127921 CTAATTTTTG CATTTTTAGT AGAGATGGGG TTTTGCCATG TTGGCCAGGC TGTCTCAAA
 127981 CTCCCTGACTT CAGGTGATCC GCCTGCCCTG GCCTCCAAA GTGCTGGGAT TACAGGCTTG
 128041 AGCCACAGAG CCCAGCTTA TCTCCTATTCT TACTGGCATT TCCCATCCAG TTGAGGCCAC
 128101 TTACATTGG AGCAGGAACA TCCATAGTCA CCTCCTTCCC AATGGTCAAG GGCTCAGGG
 14 128161 GCACAGTGAG GTACCCAGCT GAGTTGTCCTC CATTCCAGGC ATTGCTTCTG GAGCTCTGA
 128221 GGGAAAGCAT CATTCAATTCA TTCAATTCTT CATTCAACAA ATCCCTACTA AGTATCTAGA
 128281 CTACAATCCA GACAGACACA TTCTCCCTGA TAATACAGAA TATTATAATA TTGATAAGTA
 128341 TCATAATTAT AATTAAAAAA TTATAAGAGC TAATATTATAT GTAGTACGAA GGCAATTGTC
 128401 TCACGGAATA GCATGTATAA ATTATATAAG ACATGGGTA TGAACTGTAT TTAGTAAGTG
 128461 TTGTGAAGGA AATAAGTGGT GTGATATGAC AGTTTCAAT GCAACCAAAG GTGGAGTGG
 128521 TGGGGTTTAT TTGAGGTGGG GTGCAAGGAG GACCCCTCTGA AGAGCCACCA TTTGAGCTAA
 128581 GATGTTGAGG AGGAGGGAGGA GGAGGGCAGCC TTGGAGGAAG CCAGCTGGGG ACAGGAGAGC
 128641 ACTTGGCATA TTGGAATGA GAGAGGAGAC AAGGATGGCT GAGGCATCGT GCATGGGGGG
 128701 AGTGTGGAAC AAAATGAGGC GGCAGAGGTG GACGGGTGCA AGGCTGTGCA GGGCATGTG
 128761 GGCACAGAT TCATGCTTA TTCTAAGGGT ATCACTGGAG GGGCGCTGG AGGATTCA
 128821 GACTGATATG GTGCAAGCAG CCCATGCCCG CTCTGGCCAC TGGTGGTCTC CTGTGCTGC
 128881 CAAGAGTGAG GTGCTTAAC AGGGGAGGCT GACAAGGCCA CAGGGGCCCT GGGCCATGGT
 128941 GGCAGGGAG AGCAGGGAGG GCCCCGCTGA GCTGGAGCTG CCAAGCCTA GGATTGCAAA
 129001 CGAGGTCTAG AGCCAAGACC CACCTCAGG CGCCCGCTCA GCTCTGTCTC TAGGGCTGCC
 13 129061 ATGTAGTGG ACGGGAAGGC AGAGGCTGTG CCCAGGAAGG ATGTCTGTG TGTCTCATAG
 129121 GACCCCTTCA GCACCAAGCTC CAGCTCCAGC TTGTCCTGAT CTAGGGAGAG AGTAGGGCGGG
 129181 TTAGCATGAG ACAAGAGCC CGGAACCCCA GCTCGGACCT GGGACAGCAG CACCACCA
 129241 TGGATTCTAGA GGATGCCCTGG GAGGCTGGGC CATCGTCAAC AGCATTCTGA AAAGGGCCCT
 129301 TCCCCAGCCA AGGGGCCCCA CAACCCCCCA TCCCCGACTC CTGCTCTCTG CAACATGCTC
 129361 TCCCCCAGGT CAGGGGACCC TCCTGCCAG CGGAGATGCG TGGCCCTCTG TGCCATGTGG
 129421 GTGAGAAGAG AGGAGCAAGC TCTACCTCTG TGCTGGCCTC TGCCCTGCG CAGGCTTCCC
 129481 AGCAATCTC CTCACCCCCCA CATCCAGGA GAGTCCAGGC CTGGTTCTAA GCTGGTCTCA
 12 129541 CCTGCAACCA CAGCGGTGCT CCTGTGCTG TCCAGATGCA CATCCAGGC TGACAGCTCT
 129601 CGGGCCTGGG GAAGACCACA TCATGGGCAG GGGCCTCTGC AGGGTAAGCT GAGGTCCCG
 129661 TAACCTCAAC CCTCTGAGAC CCACGTGCCC AAATCCACCA CACACACAAG GAGTTGTGTG
 129721 GGCTCAGGAT GGGGATGGGT GCAGAGGGCC TTCATCCCCAT TGCCGATCCCT CTGCCCGCGC
 129781 TGAACCTGTAT ATTCCCGAGC AGCCAGGAGG GCCCTGGCTC CACACCTCCC CTCCAGCTCC
 129841 CCTGGCTTGC CCAAACCTCCC GAAACCTCTC AGAGCGCCCA GCCCTCAGAC ACAGCTGTT
 129901 CTTCTCTGT TTCCCTGTGG CTGGAGCTG GAGGAGGTAT TCTACCAAA CTCCAAATAA
 129961 TCTAAGCAAG CCGCCCATCC ATAGGAGGGT AGACGGGGCG TCAGGACAAAC CCTCAGAACG
 130021 TCTCCTGGCT GTCTCTGTG GCCTAACTGG GGGACAATGA CAAGGCCAGG CCCCCGGTGGC
 130081 CCTGCCAAGG GAGCCTCTGC ATGCTGTCTG TGAAGCATAAC CAGCAGATGC CCGCCACCCA
 130141 CGCTCCACAC AATCACTGAA TTGCCCCATC TGCTGTGTG CTCCCATCACC TCCTAAGAGG
 130201 CAGTGCCTAC GGCTTGACTT TATTTTCATG TGTTTTTTAA AAAGTAATAC CGGTCCATCA
 130261 AACAGCCCCCT AGGAGCTTAA ATGCTGGGCT CTAGGTGCTG TGCTGGCTGA CCCTGGACCC
 130321 GCCTGCTCCC AGGTCTGTG 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 GGCGTGGCCT ACCTGATTTC TTCACCGCTT CCCAGGGAAG CACTCAGTCA
 130621 ATCCCTAGAA CTATGAGTTC ACCCGAGGTG ACTAAACACC TTCCGGAAGC ACCCCGAACA
 130681 GCCGTGGGCC AGAACCTCCT CTCACAGAG CCCCTGCCTCC CCTCTCTCTA TGCGCCCTT
 130741 AGCTCCAGGC CCCTCCAAGA CCCACAGCCA CAGGTGACCA CCTGGCTGGA AGGCTAGGG
 130801 TGGTTTACG ACGAACCCCT GCCCCTTCCT TGTCCTGCTT GGTCCACACC CCAGGGCAGCT
 130861 GCCCCTCCTC CCCAACGCG TGCTCTGGG CCCCTCCCTC AAGGCTGGG AAGGTGCTTA
 11 130921 CCACAATGAC TTTGGTGGTG ATGAGGTTT CTGGGCGATC TGACCTGGAA AAATCAAACC
 130981 ATGCAAAGGC AAACCGCTTC TGCAATTGACC TCCCGCTCCC GCCCATGTGG GGTCTCCCTG
 131041 GGTGATGCTG AGGGCTCTGT AAGTCCTGGA TTTCGGTTCAAG GGACAAGGGG AGATGCTTTG
 131101 CAGGAGGAGA GGGGAGGGGA GAGAGCAGAG AGAGAAAACA TCAGGCTTCC TCTCTGTGTT
 131161 TCCCTAAAGAG AAGCCCCGCC CCACCTCCCGA CCTCTCTGGC AGGGGTAGGA GATGAGGGAC
 131221 AGTCAGGCTG GACTGACCCAG AGGGAGATCT GGGTCAGGGC CATTTCCTCA GAGCCTGAGT
 131281 CCTGAGACAC TCCCCTGGCT GTGTGACATT ACTCACGTTT CTTCCATCAG GAACCTCACA
 10 131341 TCCAGCTCCT CCTCTCCCTG AAATCAGAGA GCATGAGCAA GTCACTCAGCA CATACTGTG
 131401 TCTTCAGAGT TGACACAGTT CTTTACAGG TGTCATCTCA TCCTCTCAAG CGGTCTCCAA
 131461 GGGAGGCACG TTAGTCAAT TATTTCCTAAAT TTATGGATGA GGAAACTCAG GTATGGGGCC
 131521 AGATGGGCTA CCTGCCACAA TTGGCCACAG CTAGTAAGTC TGCAAGAAC AGGTCTCCCT
 131581 CCCTTGGAGG TCCCACAACCT CTGGGCAAGA CTGAAGGGGA AGACAAGGGAG CACAGCCAAA
 131641 GGAGGCATGA TGTGCAAGGAC CCATATCTTT TTCTTGAGGA GGGCCAGAC CCCGAACCTC
 131701 TGCGTCCCC ACCCCCCAGGC TGCAAGCAGGA AGGGACGTGA GGCCTTCACA AAGCAGGCT
 131761 TGCCACCAAC AGCAGGAGTG ATGGATGCTG AGCTGTGTT TGAGGGTGG CCTTCCTGCC
 131821 ACAGGTGGCT AAAATCCTC CACTGGGTGA TTCACTGCTG AACTGTAAAG TGAGCCTTT
 131881 TCCTTGCTTA TAAAAATGCA AATTCATTTAAGGGTTAGA TTTTTCTGT AATGAGTTCA
 131941 CCCCAGCAGC CTCCCAACAG GTGGGAGAAA TAGCTACTTC CTCAGCTTGG AGTTGAAAGA
 132001 AGTGGGGCCC ACGGGGACTT CCCACCCAC CCACCCACCT GGGGACTCTG GGAGAAGGTT
 9 132061 TTCCGGAGCA GCTTGCAGG GAGGACTTCT GAGATGTCA AGAGAACCTT GAAGCAGATG
 132121 TCATCCTCCG TGACTGAGTC CTCACTCATAG ATGCTAAGCT CCAGAACATT CTAGGAACCA
 132181 GGAACAGCGA CTTAGCATTT TACCTGCTCA TAGTGGAGGT AGCTCACCTC TGTTGAGCCC
 132241 TTACTTGATG TCTAGAGTC GAGAAGATCA TTATGCCACC ACACGCTG GCTTCAAAAT
 132301 AAAAAAAACTTAACTGAAAC TTAATATAA GGAAATCTTA TTAAATTCTCA AATTTCCTCC
 132361 TTCTAATTTC TGTTTTATC AAACATTTAA TTCTGAGCTG GGCATGGTGG CTCATGCC
 132421 TTATCCAGCA CTTTGGGAGG CTGAGGTGGG TGGGTACCT GATGTCGAGA GTTCCAGGACC
 132481 AGCCTGGCCA ACATGGTGA ACCCCATCTC TACTAAAAAT ACAACAAATT ATCTGGACAT
 132541 GGTGGTGGGT GCCTGTAATC CCAGCTACTC GGGAGGCTGA GGCAGGGAA TCGCTGAAC
 132601 CTGGGAGGCC GAGGTGCGAG TGAGGCCAGT TCGGCCATT CAACTCCAC CTGGCAATA
 132661 AGAGCAAGAC TCTGCTCTAA AACACAAAACA AAACACATAA AAGTGTCTCA AATAATTCCC
 132721 TCTTCTCTC TCCCTGCTG TGTCCTCTC TCTCCCTGGC TTTCTGGCT CCTCTGAGGTG
 132781 CACCTTAAGG TATGTGCTGG TGATCTTCA CGGTCTTCCA CGCATTATCT CATCTCACCC
 132841 CACAGTGAAC TGATCAGGAG GGGGCTTTA CTGTCCTCTG TTCACTAGCTG GGCTTCAGA
 132901 GGAACCCAGC TTGTGGGAGT CAGTCCCACA GCCAGTTGGT GAAAGGACCA TGTCATTCT
 132961 TTCATACATT CCCTTGACCA CAGCACAGAC AGGCCAGGG ACAGAGACCC CCTACTGCTG
 133021 AGAGTTTCTC TGACAGCCCC AGAGGAAGCA TGCTCTACCC CAGGAACCCCT TGCAATGGGG
 133081 TCCAGCCCGT TCACAAGGGA GTGCCAGGA GTCCCGGGCC TTACCTTGA CCTGACTTTG
 133141 GATAAGGAAA CGGAAGGCCT CATTCCACAC AGGATGACTG GTGTCGGTGA GCGTCTTGGT
 8 133201 CTTAAACTTC ATTCCAGGTG CGTCAGACAG CTGAGGATC ACGTAAGGGT CGGCCCACT
 133261 CACTGCCAAG GTAAAGGAAG TCTCTGAG GGAGTACTCC CCACACCCCT CTCCATGAGC
 133321 CATGCCAGTT TCCTACCAAG ACCCTGGCG TCACCGCAGG TGACCCAGCC CAGCCCAGG
 133381 GCACTCGTCT TGGCCCTTCC TATGGGATCT GGTCTCCCT GCTCCCGACA GAGGCCACAG
 133441 GGCGGTTACT CACACAGTC AGCCCAGCGC AGGTTCCCGC CCTCCAGGAC CCTCACTGTG
 7 133501 AGCTGCCAGC AGGTAGAGGC CTCCCCCTGA AGAGAAAATC AAACCTCAGA GTCCTCTG
 133561 CATTCCCTCC CTTATGCCCTG GACCTGCAGC CGGGCTGCG GACACCCCTC ACCACTCTCC
 133621 TGGTACTTGG TGGTGTGAC ACCCCCTGCCA GGCCTTCCGG GGACACCCCTC TGGGGACCAA
 133681 AGAGCCTTGG TCCACTCTG TTTTGGGGGT CATGGTTCTG GGCTGAGCA GTCAAGGAAT
 133741 AGCAGCACAC CCAGATGTGG AATTGTGGA ACCCTTCCC CTTCGGGAAC AGACAACAT
 133801 GGGTAGTGGG CAGAGCAAGG AAGTGAGGGC CAAGGAGTGG CAACCACAGA GGTCAATTG
 133861 ATGTCAGGCCA TTCCCAGAGC TCCCAGGGAG GGCAGCCCTT TGGCAAAACA AGGGTACCT
 133921 TTGGTAGGCC TGCTGCCAGC ACTGCCCACC CCTCAGCCCT TCCGGACGGG AAGCAGGCAT
 133981 ATGACTTAAG GGAGGGTGG CTCAGCAGC CCAGTCCACA GGCCTACCC TGATCTCAG
 134041 CTATCTGCAC TGTAAGGGGT ATGATGAGG AACTCTCAAG AGCATCCAGC CATCACAGGG
 134101 AGCAGTGGGT GCTGTGGGT AAGGGCCACT GTGTATTAAG ATGCTGATGA GGCACATCA
 134161 GCCACAGACA GCCCAGGTG GAGGGAGAAG CAGCTGAGTG GGCAGAAGCT GGAGGAGAAG
 134221 ATGAGCCGAG AGTAGGGTGGC AGGAGGGAAA GCAGATGGCC AGAGGGCGT GGCTGAGGCC
 134281 TGCTGGGTC AGAGCACTGG GGGTGTCCCC ACACCCACT GCTGAGGTCC CGGGCACTGC
 134341 CACGATCAGG AGGGCCCTGG AGTTCTGTT CTTCTGAGC CCTGCAGCAC GGCTTTCT
 134401 GGGCATCTGG CACAGTTTC TTATAAGCCC CTTTGATGTT AGGTTACTCA AATCGTTTG
 134461 GCAGGGCAAA TGGCATTNTT TCTGTGGGC TGACCCCAAG GAAGGCTGTT GGCCTTGAT
 134521 CCTCTGGAA GGGAGCCCTC CCAGGAGGAC CACCTATGAG GCACGTGGC CCAGCAGAAA
 134581 GAGCCCCATGT CCAGAGGGTC TGTTAGAAC GGGGCTGTCA GGAAGGAACA CTGCTGTG
 134641 CGGGGAGTGG GGCAGTGGGA GCCCCAGCCA GGAGGGATC CTGAGGGCC ACAGCATGGC

134701 AGGACCCCCA GGAGGCACCA AGCACATGAA GGCCTATGAA ATGCTAACCC CAGGAGCCAT
 134761 GCAGTGTGGC GGCCCTGCTG AACAAACCCAG AGACCACCTT CATTCCAGAA AAAAACCGT
 134821 GCTGGGGCTT CTGCCATTCA GCGGCTCTGC CACCAGGTTA GACTCCCCAT TAAGTAAGTA
 134881 CGATCTTCTC AGACCCATG TCCCTTCACC CGCTCCTCCC CGACCCCTGAG GGAGCCAAAA
 134941 GCCAGTGCAG TGTTTGACAA GTGAACATT GATATCGCAT AGCTGGGAGA CTGTTGTGAT
 135001 AAAACCAAGAG GTGTTTCTG CTTTGCCTC TCCCGTGAAG AGACCCACTT CCATGCATGT
 135061 GGTTGGAAGT TCTCGTGC CACAAGGACT CGTACCCCGT CTGACAGAT GCTGCTTGCT
 135121 AGGGGCTGGC TCAGCACCA CGGGGACCTG AATCTCCACT CCCCACCAAA AACTGTAAG
 135181 GACCATGGC CACAGCAAAG CCATATGCC GGTGGCCCTG CCCCTGCATC TTTCTCTCT
 135241 GTGGTCCAT GCCAACGTCC TATCATTCCA GGCAGCTTCC CAGCAGAGCA CATGTGACCT
 135301 CCCCGAGCAA GGTAGATTG TTTCTCTCTT CTTGGGGATT AACTCCAGTG GCACCAGCAA
 135361 CATGTCAACT CTATAAACCT TAAGCCTTAT AGAGGTTGTC AGCAGCAGAG ATTTCATGAA
 135421 CTTTGGGACA TCTAGCTCA TTTCCTCCCC TCACACTAGAT TCCACCACTG CGGGCCACAG
 135481 GGAGGTGATG GCTGGGATGC TGGGGCCTG CAGGGTATTG AGTTGAATG TCCCGCTCCA
 135541 AGACAAGACA GGTGAAGGGC AGAGGCAGCC CAGCCTCGCT CCCCTGCTCC CGGGCCAGG
 135601 CTGATACCTC ACAGGTCCAC AGCTGTGGG CTGCTCTAA GTCACCTGGG CTACTGAGAA
 135661 GTGGCCCCAG TTACTCACTG TCTAGCCAGT CACCATCCCG GGTAAAGCA CTGTTGTTCA
 135721 TTTCTCCCTC CCATCACTGA GAGCTCTCA ACCTCAGAAA GCCCACCAATT CCCGGCAGG
 135781 TTATTTTCCC TAGAGAAAAA TCCCGATTCT CTTCACCTCT AAGATCATCG CTAACATGGA
 135841 GTGCCAGAA CATGGTATTAA TTGGGACAGT GATTACAATG ACTGGTCTTT ACGGGGGCTT
 135901 AAAGGCAGAC TCTGGACCA GGTGCTGTT ATTAAACCCC TGGCCCTGCC ACTTACTTGC
 135961 TGTGTGACCT TGGACAAGTC ACTAAACTTC TCTGTGCCTC TGTTTCTCTA TGTGTAGTGG
 136021 AAATAATAGT TTGGACCTCA CATAGTATT TGAGGCTTAA AATGTTAATG CCTGCAGACT
 136081 GCTTAGGAAT AAAGTAAGCA GTTTTTAGGT AACTGCTATT GCTACTACTG TTATTAAGT
 136141 ATTTGCCATG TTTTCAGGCA CTGTGAATGT TGAGATACCC AGGATGAAAT CACTTTGTC
 136201 AGACCCCGAC AAATTAGGGC TGGGAACGCA TGAAAGAGCA GAGCTCGTAT TTACATGTCT
 136261 GAGATGAGAG TTGTTCCAA GGACTTCCTA CAAACCCAC AGAAAACAGC TTCACATCC
 6 136321 TCATGCCAA GTCCCTCTTT GACAGGAAGA GGGAAATGAG CAGACCCGCC CTGTTCCCTG
 136381 CTAGCACCGG TGCCCTGTT GGTTGTAATG CAGGAGGTCAGAGAGAG GAGTGGAAACA
 136441 CCTGGGCACA GCCCCCTGGTC CTGGGGACTC AGTCTGTGA ATGCCGCTG GAGGTGAGGC
 136501 AGTTTCCAGT TGTGGGCGAG ACTCAGGCTC CCAAATCGCT GACACTTAGG CTTTTGTTTT
 136561 ACCTGGGAGC TAGACCTAGA GGAGGCTGAG GAAAAGGGG GAAGGGGTAC AGCCTTCGCC
 136621 AGGGAGATGA GTTGACAAGG TTCACTACTA GACATCCATT AGGACTGCAA TAATTCAAGAT
 136681 AAGATGCTCT CAGAAAACAC GTGCCAGTA ACAGCATCTC CACCAATGAA CTGACAACCTC
 136741 TGGCTTGAG CTCTGGAAC AATGAACCTCT GTTTCCAAGC AGTTTCTGTG AGCCTCCCCC
 136801 TTTTTTCCAG TAAAAGCTTC CCTTTCCCT TCCCTCACAA TTGCCCTGTC GGCCTGTAT
 136861 TCGACACATT CCGAGCCACA GTCACTTCTC ATTCCAGAGT AAACACAAACA TAGAGATCAT
 136921 TTCTCTCGT GTCTTCTGT AGGGTGAACCA TACTGACTA ACCATATTAC CTACCGTGTG
 136981 GGTGGAAGC CACCCAGGGC CCGAGGCAAG AGACCGAGGA TACAAGCTGT TCCAGTATAA
 137041 TAAAATATAA AACAAAGATA GTTATACCAAG ATATAGATCT TAGATACCAT TATATATGAA
 137101 TATCATTAAAT CATTAGTTTG TAGCAATTAC TCTTCAGTCC AATATTATAA TAATCCTCGC
 137161 TCTATAATCA TAACCTAGGA AAAACCAAGGC CATAACAGAGA TAGGAGCTGA GGGGACACAG
 137221 TGAGAAGTGA CCAGAAGACA AGAGTGCAGG CCTCTGTGTT TGCCAGACA GGGCCACCAAG
 137281 AAGGGCTCT TGGCTAGCG GTGAGGCCAG CGTCTGGAA GATGCCGTTT GCCAGGCCAG
 137341 CCATGGCTCA CGGGTAGCGA AAAGTGTCAA GAAAACACAT CTGCTACTTA GCAGACCGGG
 137401 AAAGGGAGTC TCCCTTCCC CGGGGGAGTT AGAGAAAGACT CTACTCCCTC ACCTCTCGTG
 137461 GAGGGCTGTA GCTTAGGCTC GCCCGCAGT ACCTGGAGGC CTAACCATCT CCCTGTGATG
 137521 CTGTCTCTCA GTACTCACGC TCCTAGTCTG CCTTCATGTT CCATCTGTG CACCTGGCTC
 137581 TGCCCTCTAG ATAGCAGTAG TCAATTAGTG AAAGTACTAA TAGTCTCTGA TATGCAGAAA
 137641 TAATGGCGTA AGCTGTCTT CTCTTGTCT CCTCTCTCTC TCTGCCTCGG CTACCAAGGCA
 137701 GGGAAAGGCC CCCTGTCCAG TGACACGTG ACCCACGTGA CCTTACCTAT CATTGAAGAT
 137761 GACTCACACT CTCTACCCCT GGCCCTTTGC TTGTATCCA ATAAAACAACA GCACAGCCAG
 137821 ACATTCCAGG CCACTACCCG TCTCCGCGA TTGGTGGCAG TGGCCCCCG GGCCAGCTG
 137881 CCTTTCTCTT TATCTTTTG TCTGTGTCT TTATTTCTAA ACTCTCTCGT CACCCACAC
 137941 GGGGAGAGAC CCACCGACCC TGTTGGGCTG GTCCCTACAC TACTGTAACT TTATTTATCC
 138001 TCCCACAGCT TCAAAAGACA GTACTCCCAT TTCCCTCCCAT TTTCTACATG AGGGAATTGA
 138061 GGCTCAGAGA GGCTAACTTG CCCAAGGTCA CAGAGTTAGC AGGGACAGAG TTAGAATTCA
 138121 AACTTCCAGG CCAGAGCTCA CAACCCCCGT CTATTCTACC TTCCCAAGGAC TTTGCTTCT
 138181 TTGGAAAAGC ATGCAACTCT TCCCTCATTG CCCACAGGGAT ATCTAGGATA TCTAGACACT
 138241 AGATTTTCAT AGATTGTCTT AATTGACTA ACAGATATTG ATAGATTGTG CACAGATTG
 138301 ACTAATATGA TTACTGAGAT AACCGACACT TTAACTGATA TGAATCTTT CCCCTGAGGG
 138361 CATAGGGCG TCACATTCCC TTCTGAAAGC TCCCTCCCATC TGAAATACCT CCTTAATTGA
 138421 TGACTAACCC GGAGGCTCT GGAGGAAATT CGTCAGCCCA GTCACCTGTA GGATCCCCCG
 138481 AATGCCCTCC TGCTGAGGGT CTCAGCCACT CTGAACCTTG CGCAGGGACC CCAGAGGCTC
 138541 TGGGCATCC TTCTTCCGGT GGAGAGGCC AGGGCACACG TTTGCCAGGA AGTTTGTGA
 138601 GAGGGCTCTGT GCTGCCCGT GCCCATCTTG GGAGATGTGG CCCAAGAGAG CTTAAGGGGG
 138661 TTGGGCTCCTCC CGAGCTGGGG TCCAGAGTAT CCTTGCCCCGT GATCCTGGGT GTTCGGCAGG
 5 138721 AAGAGAGAAA TGAGCAGACCC CACCCCTGTCG TCTGCTGGCA CTGGCGCCCT GTGGTGGGGC
 138781 AATGCGAGGAG GTCAGAGAGA AGAGGAGTGG AACCCCTGGG CACAGCCCCA GGCCCTGCAC
 138841 CTCCCTCAGG CTGGGCTGAG TCTCTGGAC CACACGGCAG GCACGTTCCA GAACAGGTCC
 138901 CGGATGGCCA GCCTAGCCTG CTGGCCCAGG TCCCGCTCTG GGGCAGGGTTC CAGGCCCTG

138961 TTCCAGCAGA AATGAGGCTA GGCTAAACCG CTGCTGAAGC TTGCCAACCA CAGGAACCGA
 139021 CTTGAAGCT GGAGCTTGA GACCTGGTCC TGGGGACTCA GTCTGTGAA TGCCAGCTGG
 139081 AGGCAGGTG CTTTCAGGT GTGGGGCCGA CTGAGGCTCC CAAATTGCTG ACACCTAGGC
 139141 TTTTGTTTA CCCAGGAGCT AGACCTAGAG CAGGCTGAGG GAAAAGGGGA AAGGGGCACA
 139201 GCCTTCGCCA GGGAGATGAG TTTAAAGGGG GAAATATTCC AGAAGCCTCC ACAATGGCC
 139261 GTTCTGGCT CAGGGCTCT GATGGGACGG GTGGTGGAGC CAGGGCCGCC AGCCCTCCA
 139321 CTGGCTGCA TCCGTTGACT CCCCGCATC CAAATGTCG CTGATGTCG ACTCGTGGGT
 139381 GGATCCACCT GACCTCAGGC TGGGGCAACA GTGTTGCTC TCACTGAGGG GTCCCTCTCA
 139441 TCTTCCCCCA GGGCTCTTA GTGTCGCCC TGCCCCAGGC TGAGCCTCTT CCAGGAGCAC
 139501 CAGTCTGGAG CTGCTGTCT GGATCCAGCT TGGGCCCTGG TCTCAGGGCC ACAGCTCTAG
 139561 AAGCCCAGCG AGGTCTGCAT ATGAGGCAA GGACCTCAGG ACAGTGACCT TGACACACTT
 139621 AGCCTCATTC CTGGCCACCA CCTGGGGCAC CGGGGGGAGA CCTCAGGAGC CAGCTCTGCT
 139681 GGCCTGGTC CCTGGCAGAG GATGAAAAGT CTGGGCTTTG GAGAAACACA TGACAGAGGG
 139741 TGGGGGGGTG GCGAGAAAGTT TGCAAGCCTC CTTGGCCTGG ATGCAGCCCC ACAGCTCTG
 139801 GGCCTACTAA ATACTTGGCA GTGGATGGGG ATTGGGCCCT CCAGAGCTCT GGAGACACAC
 139861 AGGGTGATCC CAGGGCTCCA GCCAGTTCTG GGAAGTCTGC CGGGCTGTGC TGGTGGGGCT
 139921 GTGGGCTTG TGGGCTCTCT TTTGTCAC ACCCAGGCCAC TATGTTCTCC GATGATGGTG
 139981 GGGGAGAGGT CCCTAAGGAC ACTGCTGTGG TCGCTGGTGG TGGGGAGAGG AAGCTCTGCT
 140041 GACCAGACTC ACGTTTCCA GCCCCCGCAT CTACGGATGC TGCTCACCA CCCCACCCAC
 140101 TCTGCATCCC AAATCTCAGA CTCCCTCGCC CTCTGGCCCA GTCTGAAATC ACCCTGATT
 140161 CAGCTATCAC ACTTCCAGA ATGCACCTG GCTCCAGGCC TGCCCTGACT GGTGCCAC
 140221 TGCTCTGCAG ACACGTTAC CTGTAAGGG TGGCAGGTG GTCCCCCAGG TGACAGGCTC
 140281 TCCATGCTAG CCCCCCAGC TTCACTCCAG GCCCAGCCCT TGCCAAGATG CTGGCTCT
 140341 GGCTGAGTGG CAGCGACGGT CCCAGTGGGA TAGGACAGC ATGAATCTGC CAGCCTCAA
 140401 TGAGCCACGA TCCTATAAAC AGGTTGGGG TCCACCCCCA GGTCGCTCT TCAGCCCTGC
 140461 ACCTGCCCT CTTCCCAAG AGGCTGGTC AACAGCAGTC TCCAGAAACC TCTGGCAGGG
 4 140521 CTCCAGGACC TGCAGGCGGA GCACCTCTTC CAGACCTGGA GTTCAGGAC AGAAGATGCA
 140581 GCCAGGGGA GGCAGACGAG TCCCAGGAGT GGTGGCAGGA AGCTGGTCC CCTGAGGCTG
 140641 AGGGCAGAGG ATGGGGTGGG GTAGAGCTG AGGAAACCCC CAGCTCCACT GGAGCCAGGG
 140701 CCAGCCTCTA GAGGAGGGAGG GCTGAGCACC TGAGACTAGG CAAGGAGCCA GCCCTCTGG
 140761 GCAGCTCTGT GCAGCAGGGG GTTCAGGGGA TGAAAACAAG AGCAGTCTG GTGACAGGCC
 140821 CGGGCTAGTG TATCCTCTC ACCTGACCTG AGCAAGCTCC AGGGAGCAGC ACAGGGGTCT
 140881 GGACATTCCC CACCAACAGC CCCCTACCTC CATGCGGGGG GCCCTAACAA TTTAGGGAT
 140941 TTCAATTAGA CTTTGTATAA AAAGTAAAG AGACCAATCA GGAGATTCTT AGGTTATGAA
 141001 AATTGTAGAA GTAAACTTTG TGAAGCCACC TCACCGCTTA AAGCATTCA GGTACACATT
 141061 AGCACTGAAC GCTCTCTAA AATGACATGA ATTCACTCAT TTTACCATCA ATTCTGTAAA
 141121 CAGTTGAACA GCGATGATGT TGTGTTCAA AGCAAATGAA ACTGAGTTAA TCTTAAATA
 141181 CTACAAGTGG ATCTAAATG TATTATTC GAGTTGGCCA AAACAGGAAT CTAATGGTCT
 141241 ATGAAACAAGT AAACATATT TGGGGAAAT TATCTGTCA TTGCACTGATA TAAAATTAA
 141301 TAAAAAAATG AATGTAGTGG TTGAGCATGG CAGCTAACGC CTGTAATCTT GGCACTTTGG
 141361 GAGGTCAGG AGGGAGAGTT GCTTGAGGCC AGGAAATGGA GACCAGCTG GGCAACATGT
 141421 TGAGACCCA TCTCTACCAA AAAAATTTC TAAATTATG GGGCATGGTG GCACGTGTCT
 141481 ATGGTCCCAG CTACTCAGGA GGCTGAGGTG GGAGAATTGT TTGAGCCAG GAGTTCAAGG
 141541 CTGAGTGGAG CTATAATTGC ACCTTGCAC TTCAGCTGG GTGACAGAGT GATAACTGT
 141601 CTCTTTAAAA AAAGGGGTGG GGGGAAAGA AGTAAATGTA ATATGTAGAA AACTACAAAG
 141661 AACAAAGTCAT TTCAAGTGGG TCACAGACAA AGCAATACAG AATCACCCAG CTGGTACTG
 141721 GAAAGGCTGG GATTGACTG TCTGCTCTGA AGCCCCACCA CAGAACTTC CCTGGTCT
 141781 GGTGGCTCTC ATTGTTTAGG CCCGTTCCAG AGGAATCTG GGACCTTAGG AGTCCCTGATT
 141841 AACCTCCCTG CCCCTGTCA CCCCCATACGC CCGCCGGGAT CCGCTGGGG AACCTAGGG
 141901 GGCAGAGCCG CAGCAACACG TTCTTGAGG TGACCCCAGC CCCAGCGCC CACTCCAGTA
 141961 CGGGAGGGCC AGACACCCCC TCACCCCCAG AAGACAGCTG GCTGTGCCTG GAGGAGGAGC
 142021 TGGTGTCTAG TTTCCCCACC TGTCATGAA TGGTGTGTAGA GATGCTGGC CTGGTGAAAA
 142081 CTACGAAATT CCTAATTCA AGATCTTAA TTTGGGTCTA GGGTCCATGG ATGGGAATT
 142141 GGATACTTT AACCGTCCCT GAAATTGTAT ACAAAATTAT GTGACTATGT ATGCATGT
 142201 GAATGTTTC AGAGAAGGGG TCTAGAACAG CACAGTCTG TAGGTGACTA TGTATGCATG
 142261 TGTGAATGTT TCCAGAGAAG GGGCTAGAA CAGCACAGTC CAGTATGTA CTATGTATGC
 142321 ATGTGTGAAT GTTCTAGAG AAGGGTCTA AACACAGCACT GTCCAGTAGG AATGTTCTGG
 142381 GAGCTCTGCA GATGTTGTAT TAAACTTCT AGTGGGCCACA AAAAAAGAGG TAAAATGATA
 142441 CAGGTGAAT TAATATATAT TTAGCCAAT AATCTAAGGT ATTGTCAATT CAACATGAA
 142501 TCAACATATA AAATTATTAAG GGAGATATTT TACTTTTTT CATACTAAGA CTTCAAAATC
 142561 CACTGTGTGT TTTATGCTTAGA CAGCACATCT TAGTTCAGCC CAGCCGCATT TCCAGAGTCC
 142621 ACCAGTTGCA TGTGGCTAGA GCCTACCTTA CCAAGAGCT CAGGCTATA GCTTTAATCA
 142681 GATTCTAAA AAGAGTTGAA AAAAGATGGA AAACATTGC TCGATGGTC ACACGTGGAT
 3 142741 ACCCCCCACT TGAAGAGAGC CGCAGGGTTC TGCACTGGACG AGAGGAAATA TTTAGGATGC
 142801 TCCCATCACA GTTTCATGA ATTATGTGCC AAGACCTAGG GTTTGACAC ATGGTGGAGA
 142861 CCCAAGGTGG GAAGGGGGAA GGACACGGAG CAGGGAAGCA CAGTGGCTGG CCTCGGACTG
 142921 AGGTCTAAAG ATCTCCCCA CTCCAACCCCT CCAAGGGCTA TGAATAGGCA GCTGGGAATA
 142981 GGGATCAGAC CAAGACAAAG GTCAGGACAA GAAGACAGAA GACAGGTAGA GGTGACCTGG
 143041 AGGCAGTCAG CCTGGAAACAG GGCTGGCAGG CACAGGGGG CTGGGCTGAG GCCACATGCT
 143101 GTCACCCCGA GCCACGGGTG GAGGGCTCTT AAAGGGGGC CATGCTAGA GAGCCCATCA
 143161 GAGCAGACCT GACCTAAGCA GCCATGCTGG AGGAAGTCTT GGGCCCATG TTTTCACTC

143221 TCTCATCCCT CTCCCTTGCT GGGCATTCTT GCCACCACTG TCTCTTCTCA ACGTTTACTC
 143281 TATACTCACA GAAAATGAAC ATCTATTTT TCATTTGCA CTCAGGCTC CTTTAATTTC
 143341 CCCAGATTT CTCTGGAAAC CAGCCCTCCC CGACTCTCTT CCCCCTGGT TTGGGTGCTG
 143401 TTGACTCCAC CTCCAGCTGC AAGAACACAG CATGTGGTCC AAAGCTGGCC AGTCAGTCC
 143461 TGGAAGCTT CTAACCAAGG CATTGGTTC ATTGGACTCA GGACTCTGGC TGGTGCACATC
 143521 ACAATAAATG CTGGTGCATG AAGTTCTAAG AAAGACACTT TGTCTTCCCT GCTGGCAGA
 143581 ACTTGGGAGC ATATGAAGCT GGGGCTGCTA AATGTATCCA TCTTGCACCC ATGAGAGGAA
 143641 AATTGTCAG GAAAATGGC CCAGTAATAA TAAACACCTA TATAACTAAT ACCTCCCTATG
 143701 TGCCAGGAG AGGGGGTTT ATTGTGCAAGC TAATGATAC TAAAGCTGAGT CTTCTCCTTG
 143761 CACACACACA TCCTTCAAG GCTCCAGGGG AGCCCTAGCA GTAGGTTCGT GTGATCATAT
 143821 ATTTTAGTAA AATTACACAGT AGTAAACGAT TTTATCCACA ATTGGTAAAG ACTGCTGTCT
 143881 TCCCATTCT ACCCCCTCTGT TAGCATTGGT GTGGCTATAG ATATTTTAG GATCTGGTTA
 143941 AGGAGAGGTT GAGTTAGAGA TATATTTTAT TTCAGTTAG TGAGATGTAT TTATGGGGTT
 144001 TGCAGTCAGC TTCACATCA GTGAAGTCAG CTTCATGTCAG AGTCAGCTTC ACATATAGT
 144061 AAGTTCACTG TGATGTGGGT GTGCTCTTG GTGCCAGAG AATTGGATAG AGGAAGAGAA
 144121 ATGAGATCTG AAATATCCAG AGCCAGAAC CAATCTGAA AATTCCCTTA ATCCAACAC
 144181 CAACATTGTA AGTGGAGAC TTGGTTTCC TTGACACCTG ATCAAAACAG AAGTTCTCTC
 144241 CTGAAAAGGG TGTTTCCCAT CATGTCTGTA TTGCACTTT TCCTTTATT GAGGGTAATT
 144301 TTAAGAAAATA TTAATAGAAT ATGTATCGT ATTCCCTGTC TACAGCAGCT AGCACAACAC
 144361 CTACAACAGC TTATACACCA AAGAAGCTG ATGGAGGCTG CGGCAGATT TGTTTCCAA
 144421 TAATGTTGC ACCGATATCT CCTGTCCATG CTCTTCCAGC ACCTTGCAC TCCCCATCAA
 144481 AAAATAGAGG CCAACTCCCC TTCTCTTAAAT TCTGATGGG TTGTGATTGG TAACCTACAG
 144541 GATGTGAGAA AAATGACTGC TTGACTCTCA AGGTTAGATC ATAAGGTGAT GCAGCTCTCT
 144601 TTTTTAATGG CTAGAACACT CCCCTTCTAGA GCCCTGAGCC GCCAGGGCAA ACACCCAACT
 144661 CCCCTGAGGC GGCCATACTG TGAGGAGGC CAGCCATAAG GACAGGCCAC CAGTGGGTGC
 144721 TGGAGTCAGC AGTCTTAATC TCCAAGTGGC CCAGGCCAGG CTCCAGACCC AGGGACGAAT
 144781 GAGCTTCCA ATAATTTAG CCCAGTGGT TGGGTCTACC CTAGCCAGGG CCCTTTCTGA
 144841 ATTTCTGACA CACAAAATAT CTGTGAGCCT AATAGAGTGG TTGCTGTGTT ATGCCATTAC
 144901 ATTTGGGTG ATTTGTTACA TAGCCATAGT AAGTAAAGCA GAGGTTTCC TAAATTTGTA
 144961 ATGATCTGA AGTTTATGT AACATTACCA GTAAAGAATT ATGAAGCTGA AAGAAACAAT
 145021 CAATAATGAA CTACTATTAA CTATCAATAA TTTCTAAATC ATTCCAATCA ACCAGACTGG
 145081 AGGAAAGATT GCTACCTATT TGACCAAAAG AAAGTGTAT TACAAAATG GTATTGTTAT
 145141 ACACAAAGA AGAAATGAAA AAAAACAAACA ACTCACAAAAA TGGGTGTGTA ACAAAAGGC
 145201 AATCAAGGTG TATTCAGGCA AGCAGTGGAG GGAAAAGCAT GGAGGTGTGT CTGGCAGTTG
 145261 ATTCACACAA ATATCATGGT GTGTTTTTT TGTTTGTGTT TTGTTTTGTA TATGGAGTCT
 145321 TGCTCTGTGG TCCAGGCTGG AGTGAATGG CATGGTCTCA GCTCACTGCA ACTTCCACCT
 145381 CCTCGGTGCA AGCAATTCTC ATGTCCTCAGC CTCCCCAGTA GCTAGTATTA CAGGTGCCA
 145441 CCACCATGCC TGACTAATT TTGTATTTTT AGTAGAGATG GTGTTTCAACC ATGTTGCCA
 145501 GGCTGGTCAA GAACCTCTGG TCTCAAGTGA TCCACCTCCC AAATGGTGG
 145561 GATTACATTTG TGAGGCACT GTGCTGAGCC CATATCAATG TATTTTTTA AATTTGTAA
 145621 TATTTGGAA TTTATCAACC TTTAAAATT TAAATTTTG CCTGATTCT TTTTATCATT
 145681 TACAAAGAAT ATTAACCTTT GTACCTTAAC TTGTGTCAC TGTGATAAC TTCAGGCTCC
 145741 ATAATACACG CCCTTGTGTC ACACATTCA AAGCACTTA TATAGAGAAC ACCGTTGTTT
 145801 TATCTACAAAT TGCCAAAGCT GCCTGTCTTC CTGTTTCTGC CTGTCAGCC ACATTGGTGT
 145861 GGCTATGGCT GTTTTAGGA TCTGGCTAAAG GAGATCCCCC AAGCCCTGTC AGGTGAGGAA
 145921 CTTTTATATT CCTCATTCTA CAAATGAGAA CACGGAGGC GGAGTTGGT GATGAACCTGC
 145981 CCCCAGGTCA CCTGGAGGT GGAGCCAGA GGCCAGCTA CAGTCTCTT CCAGAGGAAG
 146041 TGCAAGGAAA AGGGAGCAGA GAGAGCCCTT GGATCGAGCT TTGCTGAAG GAGTAGATCC
 146101 AGTCCTGGGC TTTTCACTTG GGTCTCAGCC CCTCCCCAC TACTTTAACCC CCAGGAAAGA
 146161 AGGGTTGGT CACTCCCTTC CTTTATCCT TGCTTAGAAC ACCAATCGAT GACAATGTCA
 146221 CAGAGAAGAG GAGGCAGCTG ACTGACAGGG GTGAGGAAG TATCAGGGCT GATGACCTGC
 146281 TGAGTTTCCA TTCCAGGCA AAGCCCTGGA GCAGGTTGTG GGTAGATAG CAAGATATGT
 146341 AGCATGGGC AGGCATGACA CATGCTTCGG TGTCAGCC TGACTACCAT GGGAAAGGCAC
 146401 AAGCGGGCCT GTACTCTTG CTCATGAGGA AGTGGTTATT TACCTGTG TG ACATGAGGG
 146461 AAGGACCAT TCCCTGTCAG GAGTGTCTAG GACAAGCAG GAAAGGGACA CCTCACCCCT
 146521 GCCTCATTC ATTCTAGGGC TACTGGAAAG GACAAGCCTT GCTGTAAGCA AGGGACAGGG
 146581 CACCACCCAG CCAAGCCATT CATTCACTCC ACAAAATTGTT ATTGAGTGT TACCGTGTCC
 146641 CAGACAGAGA GCTGTGAGGG TAAAAACAGG CCTGGTCCCC ACCTCCAGTG TACAACCCAC
 146701 CTTCCACTCC ATGTCCTCA GCCAACGGGG TAAGTCAGTG TTCCAGAACCC TGGGACATGA
 146761 CACTTCAGGC CCTGGAGTGC TGGGGCAGAT ATACAATGCA GAACAGAGTC TCTGGAATT
 146821 CAAGGTAAAG AGCCATGGAA GAGCTGAAA CAAATATTAG TGAGCTAATT AACTTAGTT
 146881 GGAGGCAGTC GTTGCAGGTC CATGAGCTT GGATAAGAA ATAGAAACAT GTTTGTTT
 146941 GTTTTCTC TGAAATTGTTT GTTGTGTTA ATACAATTGA ACATCTCATT CCCCCCAAC
 147001 AGTAATCAAA GCAAAACATT TATCATTATT TTTGTCTAT TATTCTGGT ATATCAAAC
 147061 AAATTCCAGG AATGTAAGAA AATCTGACTT GGAAATATTA CATAGAAAAA GGCTGAATTA
 147121 ACACATTACC TTTTTAGTT TTCCCTGTA AGTATACATT GACATCTACT TATCCGATAA
 147181 CAGCACTGTG GTTTGAGTG AGTTGGTTT GTTTGGTTT GCTTCCCTC CAGAGGAAAT
 147241 GCAGTATGGC AGTATTATCT TTGGAACAAA TTAAAAAGGT ATTAGCAAAT AACAGAAGAG
 147301 TTACTCCAC ACATATATT TCCATGGGCA ATAAGTAGAGA AAGTATATTA TGAAAAAATA
 147361 TGCCTCACT ATGCTTTTA AAATTCTCG AGGTTACTG GTTTGTTT CTTATTA
 147421 TATAATGCACT TATCACTACT GTCTGTCCGG GGAAAACAA ACCCTGAGGT TTGAGTTCC

147481 TCCCTTCAGCC CATAAAAGTGA TATCAAGAGT AAACCTGGAT TGAGGTCACC CTGTTTGAGC
 147541 CACCAAGAGG TTCTGTAGTA TATCTGGTTA TTTTACTAAT GTATACATAT TCTAGGTTAT
 147601 AATCAAGATA CTCATCAAAT ATTTTCTATG AAAGTTTAT ATCCCAAACA GGTTCTTGA
 147661 TGATTTATCA TGCTACCATA GGAAATTGAGA AGAAAATTACA AATTGTGTAAG CAGAAATAGA
 147721 CATGCAATAT AGAGGACAGA AAAGATACAG CTTGACCTCG GGTTTTCCC TTAGAATAAT
 147781 TCCTTCCATA AATCAGACTT TATTTCACAA TATAGACTCA TCCTATAACT GGTCTGGAT
 147841 CATGAAGTTA ATGACATAGT TGAAAATAC TAGGCCACAC ATTTTGTTGG TCTTTGTC
 147901 CCCTAGACCA CACACAAAAA AACATTCTCC AAGTGGATGC TCAGGGTTAGA TTATGATGCT
 147961 CTTACATTTA AAAATTGACT GAAGCTGAGA ATGTAAGCA CACAACATT TGCAGCAGCA
 148021 ATAGTCTTC AAAATTGACT AAAACTTTA AATACAAGGC ATACAAACAA GACTATTAAG
 148081 GAGAGTTTC CAAACTATAT CTAGAAAATAC AGGAAAATAT GTATTGCCTC ATAATAAAAGT
 148141 TAAATAAAATA TCATTCTGA GTTCTATAAA CACTATAAGA GGTTATCGTT TTATCGGAGA
 148201 ACCATGTTT AAGACCCCCCT GAGGTACCTG CTCCGACCTC TCCCAGAAC CCTCTGCCTC
 148261 TCCTCACCAG GCCTCTCCG TCTTGAGGGG ATCTCCCGC TCCCCACTCC GCCCTGTTGG
 148321 CTTTATCCTC TGATTCTGT TTTCATCAAAG CTTCCAAGCA ATGAAATGAG TATGTCGCGT
 148381 TTTAAATGTA CTGTATTGT CACTTTATA TTCTAGATCG TAATATTTT TAAAAAATAAA
 148441 ATTGAGACTT TCATACATTT TCAAAACATA AACCAAAACAA AACATGGGTT AAGGGTGAAT
 148501 AGGGCCCTCT AAAAGATTAT TTCTGTAGGC TGTTGTATCT GAATAGAAGA CACATTTTC
 148561 TCTAAACTCC AGTAGTGACC TGCTGACCAC AGATTTGGCA AGAGAATCCT CAGTCAGAC
 148621 GTATGTCAG AGTTTCATCC TAATGCTCAA TTGCCAGGAG TTCTATGAAA TTTGCCCTC
 148681 CTCTCCAC ATGGCACTGA CTCATCTGAA AGGCAATAGG TTTCTGTAA ACACCAACT
 148741 CACCAACAGC TGCCAAGGGG GGGCACTGTG AGGTAATGAT CTGCTGCC TGGGTTCAAGG
 148801 AGGGTGTGGG GTACACCTAT GCCACCTGCA AGGGCCAGGG CCCAGCCAG CTCGCTCCTT
 148861 CCTGCCCTCG AGGAGGCAGC ACCTGAGTGC CCCGCACTG CGGAGGAGAC TTCCCTCTG
 148921 TGTGCCACCC GGGGGGGGCC AGAGGGATTG GAGGTGAAGT CCAGAGCTGA GGGCAAGCC
 148981 CAGCCCTCCC CATCACACTC CCTCAGCACC AGCTATAGTT GATGACTTCC CTCACATGAC
 149041 AGGATTGAA ATGCATATGT GAATGTGCTT TTAATGTTT GCCCCCTCT CCTATTCTT
 149101 TCTCTCGTCT TCAGCCAATG GCACCAACTG TTAAAGCTGG AGATGAGAGA AAGAAAGGGG
 149161 CCCTCCTGGC CGGGTGCAGT GGCTCTCACC TGTAATCCCA GCACCTGGG AGGCTGAGGC
 149221 TGGCGGATCA ACTGAGGTCA GGAGTTTGAG ACCAACCTGG CCAACATGGC AAAACCCCGT
 149281 CTCTACTAAA AATACAAAAA AAATTAGCCA GGTGTTGGTGC CGCACACCTG TAATCCAGC
 149341 TACTCGGGAG GCTGAGGCAG GAGAATGCGT TGAGCCATAG AGGTGGAGGT TGCCGTGAGC
 149401 CAAGATTGTG CCATTGCACT CCAACCTGGG CAACCGAGTG AGACTCTGTC TCAAAAAAAA
 149461 AAAAAAAA AAAAAAGGG GAGCAGGGGG CCTCCTCTG CCTCAAATCT CAGAGCCAAC
 149521 GTTCCAGAAA CAAGGCTAA AGGCTGACCA CTGTGAGGAA AATGAAGACT TTTCTAAATT
 149581 CCCAGAGTTT CTTTCACAG ACCATAAGCA GAGCTTGAA ATGTGGTCC TGTGAAGAGA
 149641 GAGTCTTAGA GAAAGTGCAC ACCCAGCATA GGATGAAAGG AGAAAGCCTC TGACTTCCA
 149701 GGAGGGAAAGC AGGGAGAAGG ACTCACATGT CTCCCCAGCC CAGCCCAAGC CCTTGGGGGT
 149761 ATTCTGCACG AACAATCAG GTGGGTGAAA TCCCTTAGGT GTGGGCTTAG CTCCTGCTCC
 149821 ACAGGCACAT CCTGGGAAAA GTCACTATCCC TGACAGGAG CACTGTGTCC CGCATAGTAA
 149881 TGCCTGCCA TCGTGGGTGT GGGAAAGTGT TGCGGAAGGG CCTAGGTAGG ATGGGGCCT
 149941 GTGGTGTGTC CCTCTCATCT TCCTAGGCC CTTGTGTGA CAGAAGGATC CAGAAGTGT
 150001 CCCAAGCTCC CACTGGGGGT GTGGAGGTAG CTCAGCTGAC AGCAAGAGGC CTGCTGAAG
 150061 CTGCCATGGG TGGCGTGAGG TGAGGCAGAA GCTGGGATG GTCAGCTTGA GCCAGGGCCT
 150121 GCCCCGACCT GGCCAGCGTG GAAACGAAACA CGGGTCCCGT CAGCTGTGGC TTGGCTGTG
 150181 CATAACAAGCA GGAGTCCAAA GGCTCTGTCA GAAACCCCCA CATAGATCCT GGACTACAAC
 150241 CCTCCCAAGCA CCACTGAGGG TCCCCATGC TTAGCTGCCA TCTCGGCCTC AAGGGTGGGG
 150301 TAGAGAGAGC TCCAAATGAA GCATTTCTA AGAGAAGCTG AGTTTAATCC AAAAGTGAAC
 150361 CTGAATTCTT TTATGTTAAG TGGGAAATTG TAATTCCTCC CAGCCTATTT GAAACCTCCT
 150421 CTTTACACAA CTGTCCAGAG GTCTGGGAC TAGGAAATAA GATGTCGCAA GTTATGAAGC
 150481 AGAAAAGGAA TCTAAATGGG ATATCTGGAA AGAGATTCAAG AGAGGAGGCA TTTCTTAAGA
 150541 TGAGCCTTA GAAAATGGGT TGGTGCTTAC AGATTATTGC AACACCAATT TACAGTATT
 150601 TGTGCTATT AAAAGTGGAG TACCACTATT AGGTATAAGC TGTTACTTCA GAGGTCTAC
 150661 AGCAGCGCA TTCAACCTT TTGGCTATCT GTAAACCCCCC TCCCCCCTATA TTCCCGCCTA
 150721 CCCTCTTCAGCAGCTCTC ACCACACTCC ATCCTAAGCC CTCTCCCTCA CTCAGCTACT
 150781 GCAAAACCCCCC ATCTCCTCCA TATCATATTG CCAGCCCTTC TCAAAAAGC CTGAGTTGGG
 150841 TCAGGTGCGG TGGCTCACGC CTGTAATCCC AGCATTGGG GAGACTGAGG CAGGCAGATC
 150901 ACCTGAGGTC AGGAGTTGCA GATCAGCTTG GCCAACATGG TGAAACCTG TCTCTACTAA
 150961 AAATACAAAA ATTAGGCCGG CGTGGTAGTG CACACCTGTA ATCCCAGCTG CTTGGGAGGC
 151021 TGAGGCAGGA GAATCACTTG AACCCGGGAG GTGGAGTTTG CAGTGAGCCG AGATCATGCC
 151081 ACTGCACTCC AGCCTGGGT ACAGAGCAAG ACTCTGTCTC AAAAACGAAA ACAAAATCAA
 151141 AAGCTGTAGT AGAGGAGGT TCCAAAGCAG ATGTTCCACAG AGGCATGGGA GGGATTAGGC
 151201 ACCAACACCA GCCCCCTTCAC CAGGGCAGAG CAGCAGCCCA GAGGCAGAAAG TCCTGGGAGT
 151261 TCCATTCTTA GAAGCCTCAA AGAGCAGGAA GGAAATCAGC CCCTTCTATC AAGGGCTGCC
 151321 CCTCTGAGGA CACCCAGGG AAGATGGCACC AAGATCAGAG GTGTCAAGAG CAGGAGGACA
 151381 GAGGGACAAG GCCCTAGAAT CCTTTCTTTT GGCATTTAAA ATATGGAATC AGAACATCT
 151441 TGAAGCCAGT ATCAGTGTACT CTATAATTCT ATGTAAGCTG AAAATTCTTA CTGAATACAA
 151501 CTCTCCCCCTG TGTAGACAGG CAAGCAGGGC CTGCAATAAA ATGTTAAACC AACTAAGGC
 151561 ACATTTGTT TTTAAACCT CTCAGGAAAGG TTCTTCTATCT TTACTTCT TTAAGCTTTA
 151621 AGAAAAGAAA GAAAATTATA ATCAACAAGG CAAAGCTCTC ACTGCAAAAAA TAATGCATAC
 151681 TTTAAAGAAT TTCTGGGGGG GGGCGAGAGG CTCACGCCCTG TAATCCAGC ACTTTGGGAG

151741 GCCAAGGCAG GTGGATCACC TGAGGTCAAG AGTTGGTAC AAGCCTGGCC AACATGGTTA
151801 AACCCCCATCT CTACTAAAAA TACAAAAAATT ACCTTGGCCT GGTGGTGTG GCCTATAATC
151861 CCAGATACTT GGGAGGCAGA GGCAGGAGAA TTGCTTGAGC CCAAGAGGCC GAGATTGAG
151921 AGAGCGGAGA TCACCCCACT GCACTCCAGC CTGGGACAAG AGAGTGAAC TCCGTCCTAA
151981 AAAAGCGAGT ATCTGGCA TAAATGAGC AGTATTGAAA TTAGCAAAGA TCATTCTGAA
152041 CAGAACATGC CATCTGTATA CAATGAAATC AAAGGACAAG ACAATGCTT TTGATGATT
152101 AAAAGATTC AGGCTGGAT CGGTGGCTCA TGACTGTAAAT CCCAGCACTT TGGGGGCTG
152161 AGGCGGGAGG ATTGCTTGAG CCCAGGAGTT TGAGACCGC CTGAGACAAG TAGGAGACCT
152221 AGTCTCCACA AAAAATTTC TAAAATTAG CCAGGCATGG GAGTGTGCG CTGTTGTCAC
152281 AGCTACTCAG GAAGCTGAAG TGGCAGGATC ACTCGAGCCC AGAAGGTCAA GTCTACAGTG
152341 AGCCGTGATT GCACCACTGC ACTCTAGCCT GAGTGCAGA GTGAGAAAAA AAAAAAAATC
152401 ACAACTTCTA CAAAAGATA TTAAGTAAAT TTATTAAGA AGATGCCATC AGTTTCAAA
152461 TAAGGGGTG CTTTATTGAA ATATGAAGAA CTGAGAGAAA AGTTAGAAGT TTGAAAGTT
152521 GAAGTTAGT TCAAGTTAGC TCTTTACAT TCCCAACAGCT CATATGGATT CAGAACAGCT
152581 TTCTCTAGCG GTAGAATTAT TGGCTCTCAA CTCCAAAGATA ACAGCTTGA CACACTCGT
152641 TTCTTCACAA GCATACTTT CAAAATATC ACATATTACT GTCATTCTGA CCCAGCGGTG
152701 TCATGCCTGC CACCCGTGACC ACGCTGAGCT GCCCCAGTA AGGCTGCTGT GTGGCTGTG
152761 CTGCTGTTCC CACTGGTCTT TAAAAGGGG TGTTGGTCC GGGCACGGTG GCTCACGCCT
152821 GTAATCCAG CACTTTGGGA GGCTGAGGCA GGCGATCAC GTGAGGTCAG GAGTTTGAGA
152881 CGACGGTGGC CAATATGGTG AAACCTGTC TCTACTAAA ATACAAAAAA TTAAACCGGAT
152941 TTGGTGGTGG GCACCTATAA TCCAGTACT TGGGAGACTG AGGCAGAGA ATTGCTTGA
153001 CCTGGGAGGC GGAGGTTGCA GTGAGGAGC ATTGTCGAT TGCACCTCGAG TCTGCGCAT
153061 AAGAGTGAAG CTCTGCTCA AAAATAATA AATAAATAAA ATTTTTAAAG GCGACGGGTG
153121 GGCGGGGGGG GGGTGTGGGG AGGAGACAC TGGGTCAGAA TGAGTCTCCT TTTCACAGCC
153181 TAAAATCTAC TGACCCCTAAA GAAACTCACA ATATTGGTG GATCCAAGGG GCTTGAAGGA
153241 ATGAGAATCA GAGGGACAGA GAAACAGACC TGCTCAGATC AAAAGATGCA TCGTAAAGA
153301 GCCAAGAATT CAGATCAGAT TTACATACCC CGCACACAC ACTTTCTCAC AGACACTGTT
153361 TAGCTTACAG GAACAATGAA CGTATTAAAT GTATTGTCAT GGGGTTGGCAG CTGTTGACTC
153421 GGCCCAAGGT CAAAGAGCTG ACTGTGGATA ATTATCACCG AAACCTGCA CACCCACAG
153481 CTGGCGCCAT GCGGGGATGG GTCCGTCCT TACAGGCGAT CAGGAAGCTG ATCAGAGGAG
153541 ACGCTGAGTT GTGAAGAGGG ACAGGGGAGG CCACAGAGTC AAGGGGTTGA GAAAGAAGGA
153601 TGGTGAATCA GAAAGCAAGA CTAGCAACTC AAAACAAAAAT GCCTGTACAG ACATAATGTA
153661 CCCTGAAAGA AGCCAGGCAT AGAGGAGTGT GATTGTCAG TTCCGTGGAT GCCAAGCTCT
153721 AGAACGAAACA CAGTTAATCA GTGGTACAG AACTAAAAAT TGTGGTTACC TTTGGTTGAG
153781 GGGTATTTTC TGGTAAAGAAT CACAAGGGAA TCTTCATACA TGATGAAAAT GTTCTGTATC
153841 TTGATCAAGG TGGTGGTTAC AGGGTAGAG ACATGCAAGC ATCCATCCAG CTGCACACT
153901 ACGGTCTTGT GCTTTCTGT CTATGTTATA CTCAATAACA TTAAAAAAA AAAACACAA
153961 ATTAAGAGTC TGCTGTTGGG GGTGGAGCA AAGAGCTATA AGCAAAACTGA TTGATGGTTC
154021 AACTCTTAGG AGAGCTCAGT TACCATGAAA GGTAAGGTG CAGATTAAACC TGAAGGATAG
154081 ATTAGTTGGA AGTTGGCAGA AGTGTAAATAG TCCCCAAAT CACCTGTCCA GTCTCAGCAG
154141 AAGGTTAAAG GTTACTCTC TGGGGAGGAG CAAGCAGAAG ACATGGGAGT CCACATACTG
154201 AAATGGGAGT CATCACCTTC AGACCTTTT CCCCACTCAG CCCTGAGAAT GCTGGCAGCC
154261 AGTCTTACAC CTCCCTGCT GCCACCCACCA CCAGTATCAT CAGGAAGAG AACAGAGGAT
154321 TCCCTTCCAG GGAAACACAC CAACATAAAA GGAAAACAA AAACAAACAA AAAACTACCA
154381 ATACTAATAG GTGATTGCTG TATTCCAGAG TTCCAGATCC TGCCTAATC AACCTACAT
154441 GAAGTTCTCA TTACGGCAGG TCAGACCATG TGACAGAGC TTCCAGTGT CACTTTAGT
154501 CCTCACTTAA AAATGTGACC AGACACCTG GGATCACCAA ACATTTGAAAG AAATCCTCTA
154561 ACATGATTAA TGATTAACAC TACAAACCAA ATAAAGAGAC ACAGAAACAG AGACAGTGT
154621 GGAAGTATGA GAAAACCTT TACAAAGCTG TAATTATGT TGTTAAAATA AAAACTTCAG
154681 CCAAATTAAA TGTAAGGAG TTTAATTGAG CAATGAACAA TTGCAAAATT GGGCAGCCCT
154741 CAGAATCACA GCAGATTCAAG AGAGAGTCCA ACAGAACCCAC GTGGTAGAAG AAGATTTATA
154801 GACAAAAATG GGAAGTGTAG TCCAGAAATC AGAAGTGGAG TACAGAAATA GCTGAATTGG
154861 TTACAGCTCA GCATTTGCCA TATTGAAACA CAGTTAAAC GCTGAGACTG ATGTTGAGTGG
154921 TTGAGGTATG GCTCTGGGA TTGGCAGAAG CCCAGCCACT GTTAAGGCAC ACACTCCCCA
154981 GTAGGTTCT CAGTCTGTC TACATTAAAT TTAGGTGCA GTTCTCCCAC AAGGACTCAA
155041 ATATAGAAGT ACAGAATCCT TCTCAAGGCC TATTGTTGTC GCTTCAACAA TGCTTTTACA
155101 TGAAAAGACA TAAAAAAGAC AATGAAGATA AAAGATGATG GATCTTTGCA AAAAAAAGGG
155161 TGGGGATGCC ATGAAATGGA ATGTTCAAGAG TATAAGAAC CCTCGTACAC TAAAATATG
155221 ATGGCAGAAA TAAGTAAAT TAATGGAAA ATTAGATAAT AGAGTTGAGA AAATCTCCCT
155281 AGAAAAGCAA ATCAAAATT CAAAGAAATG GTCTAGGCAGC ACTGGCTCAT ATCCATAATC
155341 CCAGCTCTT GGGAGGCTGA GGCAGGGCAGA TTGCTTGAGC CCAGGAGTTC TAGACAGGCC
155401 CGAACACAT GGTGAAACCC CATCTCTACA AAAAATAC AAAAATGGC TGGGCAAGCAG
155461 GGCATGTGCC TGTAGTCCCA GFGACTTGGG AGGCTGAGGT GGGTGGACTG CTTGAGCCCT
155521 GGAAGCAGAG GTTGCTGTA GCCATGATGG TGCCACTGCA CTTCAGCCTG GGTGACAGAG
155581 TGAGACCTTG CCTCAGAAAA AAAAAAAAGA AAAAGAAAAA AAAGAAAAAA AGAAATGAAA
155641 ATAGGAGATA AGACTATGTA TTTAAAGAG CATGCCAGGA AGCTTAAATT ATTCTTTCTT
155701 CCCTGCTCTC CTCCCTTCTC CCCCTTCTCTT CCCCACCTCT CTCTCTGTC TCCCCCTTCT
155761 TCTCTTCTC CCCCTCTCTC TCTCTCTTCT TTCTCTTCT TCTCTTCTAG GATGTTGTTAT
155821 CACTTTGTT CTCAGCTGGA GTGAGTAAAT ATCTCTTCT TTGGGAGTTC AGGAATGGAT
155881 CAAAGAAATC ATTTAAATTA TCAAGAAAT ATGCAAGAA AGGATTGAGC AACAGAAAGA
155941 CTCTCTAGAT AAAAGGGGTT GCTGAGTACA CAGAACACCC AGCAAAACAG TAAAACAAAC

156001 AAACAAACAA AAAACCTCA TCAAAGGATA TTTATCATT AATTTTAGAA ATCTGGAAAT
 156061 AAAAAGGTCA CACACCAAGG ACTAAAATC ATAATGGGAT TGAACCTATT AACAGCAACA
 156121 TTGGAAGTTA GGAGACAATG GAAAAATGTC TTGAAATTTC TTACAGAAAA TTATTTACAA
 156181 CTTAGAAAGT TTTTCCCAA TCAATCCGCT AGTCAAAAGT GAGAGTAGTG TATAGGCATA
 156241 TACAGAAATG CAAGGACTCA AATTTTTTT AGCCTTCTC ATGAGCTACC AGAACAGTGC
 156301 TTTCTAGATA TTAAAGTGCCT ATGGATCACC TGGAACCTG TTACAATGCA AATTCTGGTT
 156361 CAGCAGGTCT AGAGTGAGGC CTGAGCACTC TGACAAGCTC TCAAGTGACA TCAGTGCCTC
 156421 TGATCTGTAG ACCCACCGGT GAGAGTGAGG TACTAGAGGA TACGTCTCAC CAGAATGAAA
 156481 GAATAAACCA ACAAAAGGA AGAAGTGGG ACCAGGAAAT TGAGGATCCA ACATAAAAGC
 156541 AAGTTAATAT CCAAAATGAT GCTGAAGGGG GCTCCAGGG CATTACCTAA AGGGAAAGCA
 156601 GTCCAAATTG GGACATATAA GGAGATATTA AAAGGAGGAG GAAGTTGAAG AGGAGAAAAT
 156661 AATGAAACTG ATGATTACCT GATATATCAA GAGGAGATT ATACTCTGG TGGAGAGTTT
 156721 AAAGACAAAT TACTGATAGT AACATAGAAA ACTAAGCTA TCAGGAAGTA AGGCAATTAT
 156781 TAACTCCAAG GAAAACCTAGA ATGACAGTAA CCATATCATA CTACATAATC ACAATGAAA
 156841 AAATGAATAT TGACCAAAATTATAACAT AACAAATTG AGGCACTGAG AGAGGATAGA
 156901 GCAAAAAAAA AAAACAAAAA AAAACAAAAA ATTGAGCCCT CATCCTCCAT AATAGGAAGT
 156961 CAATAAAAAA CAAGAAATAA TAGTATAAGC ATGTTATTAA GAAATCTGGA GATAGGGCTG
 157021 GCCACCGTGG CTCACACTG TAATCCCAGC ACTTTGGAG GCTAAGGTGT GTGGATCTG
 157081 AGATCAGGAG CTCAAGACCA GCCTAGCCAA CAAGTAAAA CCCCATCTC ACTAAAATA
 157141 CAAAAATTAG CTGGCGTGG AGGTGCACGC CTGTAATTCC AGCTACTTGG GAGGCTGAGA
 157201 CAGGAGAATC ACTTGAAACC AGGAAGCGGA GCTTGCGATG AGCCAAGATT GCGTCACTGC
 157261 ACTCCAGCCT GGGAGACAGA GCGAGAAAAA AAAAAGGAAAG AAATCTGGAG ATAAATACCG
 157321 AAAGACACAA CTAAAAGAGT CCTAGCAGT TGCTGCTGGG TGAATAGGAG CGGGGGAGGA
 157381 AGCTAATGTT TCTTCATTAT AAGATTGCA AAGCTGTTG ACTTTCTAAA AAATAATTAC
 157441 TTGGGCCGGG CATGGCTGG CATGGCTGTA ATCCAGCAC TTTGGAGGC CAAGGTGAGT
 157501 GGATCACCTG AGGTCAAGGA TTCAGGACCA GCCTGATCAA CATGGAGAAA CCTCGTCTCT
 157561 ACTAAAACA CAAAAATTAG CCAGGTGTGG TGGCACATGC CTGTAACCCCC AGCTACTTGG
 157621 GAGGTTGAGG CAGGAGAATC ACTTGAAACC AGGAGGCGGA GGTGCGATG AGCTGAGATC
 157681 ATGCCATTGC ACTCCAGCCT GGGCAACAAAG AATGAAACTT CGCCTAAAAA AAAAATTA
 157741 CTTGATCAAA ATATACATTT CAAATTAAAT GAAAATTATT TGAAAATGAG GATGTTTTGT
 157801 CATTAAATAAT AACTGACTT ACAACACGTA TCACAACTAG TTGCCACTCC CCAGTGTGG
 157861 GGGCTGCAA ACCACAGCTC AGAGGCGATT CATTCCCTTT AGATGCTCAA TTCTAGGAGA
 157921 AGCCATATGG GCAGGAAACA CGTGGAGTAC AGCAACGGCA AGTCAAGAG ATGCTGTGA
 157981 AACTTCTGA CATTCACTAG GTTCTAGAT GAGACAAGAT CATGCTGTAT CTACTTTTT
 158041 CAAAAGAAGA ACTTTCCCT CCCAGCTATA TCCACCCAGT TAGCCAATGG ACTAGTCAGG
 158101 ATTTTTTCAG TTGCAAGTAC CAGAAACCAA ATTGAAATG GCTTAAGCAG AAAAGCAATT
 158161 CATAGAGTCC CATAAAGCCC ATGGGAAGCC CTGGGAAAAA CCTAGGTCTCA AATCTACCT
 158221 TCAAATCCTA CCTTCAGTAT TTACCACTGT GTGACTTAAT TGCCCTAAAA ATGAAACAGA
 158281 TAATAGTATC ATTGTGGATA TGAATGAGT TAACATTTC TTAGTGCTG GTGCATAGTG
 158341 TATTAGTCTC CCCACTGCTA TAAGGAAATA CCCGAGCTGG GTGAATTATT AAAGAAAAGA
 158401 AATTAAATTG GCTCAAGATT CCACAGACTG TACAAGAAGC ATGGCTGGGA AGGCATCAGG
 158461 AAACCTACAA TAATGGCAGA AGGCAAAAGAG GAAACAGGCA CACCTTCACA TGGCCAGACC
 158521 AGGAGGAAGA GAGAGAAGGG AGACGTGCTA CACATTTTA AATAGCCAGA TCTCTGAGA
 158581 ACTCACTTAT TATCATGAGA ACAACAAGGA GGAATCTGC CCCATGATC CAATCAACTT
 158641 TCACAAGGCC CCTCTCCCAA CACTGGGATT ATAACCTGGAC ATGAGGTTG AGTGAGGACCA
 158701 CAAATCTAGA CCATATCACA GAGCTAACAC TTAACACATA TTATAACATT ATTATTGCCG
 158761 GGCGCGGTGG CTCATGCCCT TAGTCCCAGC ACTTTGGAG GCGGAAGCGG GCGGATCATG
 158821 AGGTCAAGGAG ATCAAGACCA TCCCTGGCTAA CGCGGTGAAA CCCCATCTCT ACTAAAATA
 158881 CAAAAAAAT CAGCTGGCG TGGTGGCAGG CGCCTGTGGT CCCAGCTACT TGGGAGGCTG
 158941 AGGCAGGAGA ATGACATGAA CCCGGGAAGC AGAGCTTGCA GTGAGCAGAG ATCATGTCAC
 159001 TGCACTCCAG CCTGGCGTC AGAGCAGAC TCCATCTCAA AAAACACAAA ACAAAAAAG
 159061 AACATTATTA TTATAAGCTA CTATTCAGGA ATCACAGCAA ATCTGGAGAA TAAATTAAAA
 159121 AGTTCATCTC GGACACATTG TAATGTAACT CGAGAACATC AAAGATAAAG AAAAATACT
 159181 AAAAGTCCTC AGAAAAAAAG ACATAAGTTG CCCTATGAAG GATAGACAGG CTGGAGCAGG
 159241 CTTCTCAACA CGACTAACTA AAGCCAGAG AAAGTAAAT ATCTTCAAAAA TGCTGGAGA
 159301 AAAAGACCTG CCAACTGAGA ATTATTAATCT CAGCCAAGCT ATTATTTAAG AATGAGCATT
 159361 ACATGAATAA TTATGGAT CAACAAAAAT GGAAGGTTTA CCACAAAGAA ACCTCACCAA
 159421 GCTCTCTAT AGATTACTC CAGGAAGAAG GAAATAGAAAT AGGTTATTAA ACATATGGTA
 159481 AACCAAAATA TTGACTATAC AGATAATAAT CATAATGATG ACTAATATGG GGAATAAAAAA
 159541 ATGAAGGAGG GAAGTAGAAC ACCAAACAGT GACAATGAA ATGGGAAGAT ATGATAGAGC
 159601 GAAGATTCTA AAGCACCTGG CCTTGTCTG GAATAGGGTA AAGAAGTGGG TTAACCTTAA
 159661 ACTTGTCTAA TTCAAGTAA CAGGTTAAAAA TTAGAGATA CTTACAAAAA GAATAAAATT
 159721 AGAACATGTA AATTCCAAGA TAATAAACAG GTTAAATGTT AGAGCTACTT ACAAAAAGAA
 159781 TAAAATTAGA GTATGAAAT TCCAAAGGTA TAAAGAAGGA GGGAAAAGAA AAGGAATCCA
 159841 AAAGTTTCA GGGGGAAAAA AAAGAAAACC TCAGAAAAAA AGAGTGAATA GAAAGCACAA
 159901 AATAAGATAG CAGGCCAGGC GCAGTGGCTC ACACCTGTAA TCCCAGCACT TTGGGAGCCC
 159961 AAGGCAGGCG GATCACCTGA GGTGGTAGT TCAAGACCAAG CCTGGCCAAC GTGGTAAAC
 160021 CCCATCTATA CTGAAACTAC AAAAATTAGC CAGCCATCGT GGCAGGCACC TGTAATCCA
 160081 GCTATTCTAGG AGGCTGAGGC AGGCAATTG CTTGAACCTCA GGAGGAGAG GCTGCTGTGA
 160141 GCTGAGATCG CACCACTGCA CTCCAGCCCTG GGTGACAGAG TGAGACTCCA TCTCCAAAAA
 160201 AAAAAGGAAAGA GATAACAGTA ATAAATCCAA ATATATCAAT AATCGTAATA AATGCTAATG

160261 GACTAAACTA ATCCATTAAA AACTATAAT TTAATTTCT GTATGCTGA ATTTCTCTA
 160321 ATAAAAAAA TTTTAAAGAG AGATTGTCAA TCTAGTGT TTTTTAAACT CCACCTGTGG
 160381 CTGTTGTAA GAGCACACTG AAAACATAAG AACACACAAA AAAGTGTAAA ATAAGAGCCT
 160441 GATAAAGATA CACCAGGGAA ACACTGATCA TGAGAAAATT GGTGTACATA TAGATAAAAG
 160501 AGACTCTAAT GTAAACATCA TTAGGAATAA GGAGGATTGC TGTGTAAAGA TCAATGGAAT
 160561 AATTGCCAG GAAGATATAA CAGTTATCTA CTTGCATGT A TCTAAATAAC ATAGCCTAA
 160621 AATGCTATGC TCTAATTAAA ACTTGACAAA ATTAAAATGT CCTAACGAAA ACTTGACAAA
 160681 TCTACAAATC TACCTCTC ATCAATGGAT AAATAAACATA TATAAAATAT TATTGAGAAT
 160741 ATAGAAGATT TGAACAACAC AGCGAACAG CTTAATCTAC AGGACATATA CAGACCCCTA
 160801 CAAATACATA TACAGGACAT ATACAGCCAC AACATTAGG GAATACACCT TCTTCTCAA
 160861 CACAATGGAA CATTTTTAC AAACATTGGT CATGAACAGG GTTTATGAT AAGTCTCAA
 160921 AATCTGAAAA AGGTACATA CAGATAACAT TTCTTTAC AAGGTGCAA TAAGGTAGAA
 160981 ATCAATCACA AAAATATGCC TTACAAAATC CATGATTGAT GAATTATGAT ACACATTACC
 161041 AAATGGAAAC AGAATAAACT TACTTTTTA AAAAATTAAG CCTTTTTT CAGAACAGAA
 161101 AAATCTGAA GATAGTACAG TGTCCCACAC GACCTGTGG CCAGTTTCC TATTATTAAC
 161161 ATCAACATG AGTATATTG CTACAAATTAA CCAATATTGAT TATGTTATTA TTAACAAAT
 161221 CCATCCCTTA TTCAAGATTC CTTAGTTTT ACCTAAATGTC CTTTTCTTT CCAAGATTCC
 161281 ATCTAGGATA TGACATTTAG TAGTCACGTC TTAGGCTCT CTTGGCTGTG ATGGTTCTC
 161341 AGACTCCCT TGTGTTGAT GACTTTGACA GTCTTGAGAA GTACTGGTC GGTATTTGT
 161401 AAAATGCTCT TTGATTGAGA TGTGTCTGAT GTTTCTCA TTATTAGACT AGAATCATGG
 161461 TTTTGAGGAA GGAAGACCCAG AGAGGTTAAAG GGCCATTTC CTCACATCAC ATCAAGGGTA
 161521 CATGCTATCA ATGTGCCTTA TCAATGTTGC TGTTGACCTG ATCATCTGGA TGAGGTAATA
 161581 TTTGCCAATT TTCTCTACTG TAAAGTTACT CCTCTCTTC CATTCTGTAC TCTTTGGAAG
 161641 CAAGTCACATA AGTGGAGCTC ATACTTAAGG GGTGGGAGAT CATATTCTAC CTCCCTTATT
 161701 GAGGAGTATC TACATAAATT ATTGGAATT ATTCTGCATG AAAGATTGAT CTCTCTTTC
 161761 CCATTGATTT ATTTTATTAG TATTGTTTA TATTAGTTA TTTATATTAG TATGGACTCA
 161821 TGGATATTTT ATGCTTAGG TTGTAATCCA ATGCTACTTT ATTTCTTTA ATTTTTAAA
 161881 ATTTTTTATT TTATTTAATT ATTAATTAA AAAAATTTT AAACATTAT TAAATTTTT
 161941 TTGAGACAG GGGTCTGCT CTGTCACCCA GGCTGGAGTG CAGTGGCATG ATCTCGGCTC
 162001 ACTGCAACCT CGGCCTCCCA GGTCACAGTG ATTCTCCCGC CTCAGCCCTC CAAATAGCTG
 162061 GGATTACAGG TGGGTGCCAC CACACCCAGC TAATTTTTT TTGTTATTT TAGTAGAGAT
 162121 GGGGTTTACAT GTGTCGCC ACAGCTGGTCT TGAACTCTG ACCTCAAGTG ATCCACCCGC
 162181 CTCGACCTCC CAAAGTGCTG GGATTACAGG CATGAGCCAT AGTGCCAGT CAACTTTGTT
 162241 TTCTTTGTT GCTCAATTG TTCCACCAT GGCCATTGGG GAGAATTGAT TGTTGCCTCC
 162301 TGTGTCTCTT TTGACACGTA TGTTGTT TGTTTTGA GCACCTCCTT GCTTTTGAT
 162361 CACTCCCTTA CTTCTGGCA CTACAAGATG CTTTAGGATC ATCTTGATA CTGACTGTTT
 162421 CCACTCTAGA ATGAACCATC TCTCCAAGAA ACCCTGGTCT TTCTTGTTG AAGAATGCTA
 162481 TTAGAAGCCA AGATCTGGGC ACTGTGTGTC CTCAATGCTA CTGGGGTGTG ATTACTTCTA
 162541 GTTCTCTCA GCAACACAGAG CTAGCTT ATTGTGTGTGT GTGTGTAAACA CATATTAAT
 162601 ATATTATATA TGTTTATATA TATGTTGTA TATGACATAT ACTGATCTGT GTATACACAC
 162661 ATATATATAA TTATTTTTAT ATCTATCCAT CAATATCTAA ACTAAACATG GTATCTCTGA
 162721 CTCTAACCTA GCACCACATT GTTCATTCTA GTAATCCAAA AGCCTTTGA AAAACTGAAT
 162781 TACTGGCTGG GTGTGATGGC TCATGCCTGT AATCCCAGCA CTTGGGAGG CCGAGGTGGG
 162841 TGGATTGCCT GAGGTACAGGA GTTCGAGACC AGCTGACTA ACATCTGAA ACCCTGTATC
 162901 TACTAAAAAT ACACACACAAA AAAAACACAGC CACGCATGGT GGCACACGCT TGTAGTCCC
 162961 TCTACATAGG AGGCTGAGAC AGGAGAATTG CTTGAACCCA GGAGGTGGAG GTTGCAGTGA
 163021 GCTGAGATTG TGCCACTGCA CTCCAGCCTG GGTGATAGAG CAAGACTCCA TCTCAAAAAA
 163081 AAAAATTAAG AAAAGAAAAA AAAGAAAAAC TGAATTAGTA ATTCAAAATC TAACAACAAAT
 163141 AAAAATATAA CTCAAAGACCC AGATTACTTT ATAGGCAAGT GTTACCAAGT AAAGACTAGA
 163201 CACCATTAA AATGTGCATT TGAAAAACAT CAAGACCTCA ACAATGGAA AAAATATGCC
 163261 ATGTTCTTA GATAGAATCT TTATCTAAAG ATATCAATTCTC TCTCAAAAAA CAATCTCTAT
 163321 ATTTAATGCA ACTTTAAATAA GAATCCCCAA AGGGTTTTT ATGAAACTTG ATGACCTGAA
 163381 TCAAAATTTT ATATGGAAGA TCAAAAAAAT AGTAATAGCC AAGACAATAC TAAAGAACAA
 163441 GGAAGGAGGC TTGCTTACT GTAAATTAAAT TACAGCAGCT GAGTCTGGGA GCAAGAACAG
 163501 ACACAAACAT ATAACAGAAT ATAGAGCCCA GAAACAAATAC CTACACACAG CTGGAAACTT
 163561 GTTATATGAT AAAAGTGGCA GTGTCGATCA GGGCAGAAAA CATACTAGCT ATTGAATGGT
 163621 GCTAGGGAAA ACATAAAATC AATTCAAGGCC CTATATTAA TTTTTGAGA CTGGGTCTGG
 163681 CTCTGTCATA AGGCTGGAGT ACAGTGGTGC AATCTTAGCT CACTGCAACC TCCACCTCCC
 163741 AGGCTCAAGT GATCCTTCCA CTTCAAGCCTC CTGAGTAGCT GGGACTACAG GTGCAAGCCA
 163801 CCACACCTGG CTAATTTTA TATGTTTTGT AGATACGAAG TTGTCACCAAA TTGCTCCAGGC
 163861 TAGGCTTGAA CTCCCTGAGCT CAAGCAATCC TCCCAGCTCA GTCTCCAAA GTGCTGGGAT
 163921 TGCAGGCGTG AGCCACCGTG CCCGCCAGG CCTCTATATT AAAAGTAAA GATTAATTAAC
 163981 CTTAGAATAA AATATATGAT ATTGAGAACA TCAAGATAAG GAAGGATTAA TAAAAAAGA
 164041 AACAAAACCC ATAAATCACA AAGGATAACT ATTGTTTATT CAATTATATT AAAATATAAA
 164101 ACTTTAAAT AAAAGACACC TTCCAGTGTG GAAAAAATAA TTCACAGACT GAGGGAAGAT
 164161 ATTTGCAATA TATTTAACCA AAAATGTATT AGAATAAAAA AACATATTAA TATGTTTACA
 164221 AATCAGAAAA AGACAAAGAA CTGAATAGAA AAATGGGCAG AATATACAAA CCACGATATC
 164281 ACAGATGAGC AAACCAAGCAC AGACAAAAAC ACACCAAAAAT TATGCTGCAC CTCACATAA
 164341 TATGAGGAAA TGCAAAATTAA ATTCACAGTG GGATACATC TCGAACACAT CAGATTGCA
 164401 AAAACTCAAA AGCCTACAA GCTCCCTGAA CTCAGCATT TCCCTCTCT TAAGTCATT
 164461 AAAGGCAACA GGAGAAAATA ACAACAGAAA CACAAACTCT ACGTTGATG AAATTTAAA

164521 AAACATTG AATCCAATCC ACAAAATATAT GTGGACTACT AAAAGCAATG AAGATCAAAC
 164581 AGGGGTGTGG AGGGAAAGGG ATCAAGAATG CCTACAACCT CTGAGCTCAC AGACACCTAC
 164641 AGGGTGCCTA TGCCAAGGG AGGAACACAC CTTAAACTGC AAAGCCAAA TAATGGATGG
 164701 GGTGCATGGA CTTGTGGCAG GAACTTCTCT GGAATGGATT TAGTTATGGA AAATATTGGG
 164761 GGCGGGGCTG ACTGACAAAT CCCAGAAGCA AGACTCTTTA AGGAGAACG GTGTATTCCA
 164821 ACAGCTGCC ATCTCATTG GCTGAGCAAC GATAAAGCTA AAGGAACCTCA TGCACCCCTGA
 164881 AAGGACAGT CCTGCCAGC CACAAGCAGC CTATGCCCT CCGTGAGACT CTTATCTGA
 164941 AATAGCTGT GTCTGTAGCA GGGTCTTCAG CAAAGACAAA TAGGATGTGT ACATAGAGAG
 165001 AAAAGAGATT ATCATAAGGA AGCAGCTCAC ATTATGGAG GCTGGCAAT CCAGAATCTG
 165061 CAGCGTGGC CAGCAGGCTG AAGACGCAGG AGAGCTAATG GTGCAGGTGA GGCTAAAGG
 165121 CAGAGTCTGC TGCAGAACGG GCTCTGCTC GGGGAGGTCA GTCTTTTGC TCAGTTTCAGG
 165181 CCTTTAACTG ATTGGATGAG GCGCACCTAC ATTAGGCAGA GCAGTCTGCT TTACTCAGAG
 165241 TTCACCAATT TAAATGTTAA TCTCATCCCA AACACCCCTC CAAATTGACA CATAACATTA
 165301 CCCATTACAG TGTCTTAGGT CTGTTCCCTT AGAAGCAGAG CCTGAGGCAG GGACTGGGG
 165361 CACTGTTGT ACAGGGAGAA TACTCCCAGG AGACAGGGAG AGAGGGAAAGC AGGGAGGAAA
 165421 AATAGCTAA CAAACACGTG GTCTCAGCTG CAGCCTCGCT TCAGTCTGGT CCCACGGGG
 165481 CCTCTGCAGC ACAAAATTGCA AGAGGTGGTC CCACCTTGGA GCAAGGGGCC TGGCCTCTG
 165541 AGCCCCTCTG TCTGTCGGGG ATTGACTGCT GACGGCCTCA GGGTTACGGG TATGAGTGAG
 165601 AACTTAACCT CTCAGGTGCG ATGGTTCAATT GCATGTGCTT CTTGGCTAG GCTATGATGC
 165661 CCAGCAGTTC AGTCAAACAC TAGCATAGAT ATTGCTATGA AGGTATTTG TAGATGTGAC
 165721 TAATTTACAA TAAGTGTACT TTAAGTAAAG GGCATTACCC TCAGTAATGT GGGTCGGCCT
 165781 CGTTTAATCA GTTGAAGGAGC ATTAGGAGCA AAAACTGAGG TTTCTCAGAA AAGGAGGAAT
 165841 TCTGCCTAA AGACTGCAAC TCTGCCTGG GTTCCAGGC TGCCAGCCTG CCCTACAAAT
 165901 TGCAAGCTG TCAGAGCCTA CAGGTGCTAG AGCCAATCTC CTCCCTCTC CATTGATATT
 165961 GTCTCTGTA AGAACCTGAG GTGATAAACG AGGTGTGGGA CTCTCATCA GCCAAGCACA
 166021 GATCTTGGGAAAGGGTTA GCTGAGGGCC ATTAGCAACC AACCCCTCTCA GCAGCCAGGC
 166081 TGGTCAAAGG AAGCTGTGCA GGCCACCAAC AGCATAATACT ACAGTCCACC TCCCAAAACCA
 166141 CTGAGATCCA ATTGCTTCCC ATATTAACCT CATTCAATCT AGGCACAGCT TTTCCAGGAT
 166201 TCTGCTGTT CACAATTGTTG GGGAACATTT ACAGGGAGAAAG AGTTAGTGGG ATGAACATACA
 166261 GCCCCACACCA CCACAGTTG TCTCAATCTG TCACTGATAAC TCATCATCTT CCTCCCTCCAC
 166321 CACCCACTG AGGTTCTCTT GCCTCAGC GAGCACTCTC GCTGCTCTC GTGGCCTCTA
 166381 GGAGGGGTGT TCCAGACCTT TATCTCAGAG GAGTCTGAAG CTCAAGTTAT GCTTTACTCA
 166441 GACCATGGTT GCTATATTAA CCCATTTTA CATGCAAAT TTCACATAGA AAAAAGCAAA
 166501 CCCAGGATCA CCTGAGTGCC AAACATATTC CTCCCTGCC CCATCATCCC ACTGTATAGC
 166561 AGCAGCTCTG TGTCTCTG ATGATCAGAT CAATTACCC TGCTGGAGCT GTAACCCCTT
 166621 CCTTTGCTG CTGATCTAC AGCATGAGGA GCCAAAGTG GCCAGGCTGC AGCCATAGCT
 166681 GTAAGATTAT TGGGACTCCT ATGGCCTCT CCAATAGAAG CATTTCGGC TCCTGCCCC
 166741 AGGAACCCAGG AGCTCTAGAC CCACAGAGT TAAAGTTCT GTGAAGAGAA GCATAAATTC
 166801 CCCAGTGGG TCACCTAGAA TAGTGTGAG CAGAGCACT CCTACTCTA CCGCTTGGAC
 166861 CCTGGACCAT CCATTCTACC TAGTGGGGAC AAATGCACAT GCAATGGCCA TTGGCCTTCAG
 166921 ATGTGTATTG CATCCTGAAG ATAGCACACT ATCCCCAAAC TAGCTCTCC ATGCATCTC
 166981 AGGAGGCCAT CTCGCCACTC TATTGGGAA TAGTTGGTGG TTGATGTGAT ATGCACCACA
 167041 CTGCTCCCT TACTCTAAGT TGCTGTTATG CACAGTCATA TGTAGATTAG ACATTGTGTA
 167101 AGTCGTTGGA TTGTCGTGCT TGCTGAGGTG CAGGAAGCCG GAAAGGTAAT CCACATACAG
 2 167161 AATGTGTGTC CATTCTAGTA AGTATGAATT GTCTTTTCA GGACAGAAAGG GGTCTGTAT
 167221 AATCAGCTG CCATCAGCGG TCAGTTGAGC TCCTGGGTGC CTTATTGCAAG ACTCGGTGTC
 167281 AGTCTTGCA GCTGGCAGGT TCGACAATCA GCAGTGTAGC TAGCTAGGTC AGTCTTGGAA
 167341 TGGGGATCCT ATGCTATTGG ACCCCCTGTAC TGCCTCCATC TCTGCTCCAT TTGGGAGCAC
 167401 AAACATTAAG ATGGCTGAGG ACAGGGCTGA TTGACAACCA CTGGACGATC CATCCCACTC
 167461 ATCTGGTTGT TCAGTCTTC TGCAGTGGT CAGCACTGTC ATTAGCATAA AGATGCTCAC
 167521 CCTTTGTGCC TACTCCGAA AGGGCTCAGC TGCACTGAG ACTGGGAGA ATAGGCAAG
 167581 TGACAGGCCA GCCCTGGTAC TGGTACTAA CAGCCTGCTG GACCTCCCTG CCCTAGTCT
 167641 GAGCTTGAGC TCTGCCGGC GGCTCTGGC CTGGCCCTGG AGGGGTCAC AGGTTGCAAG
 167701 AAGTACCTTG TGCAATCCAC ACTGGAAACC TGTCCTCACC CAGATCAGGG TGGAGCTGAG
 167761 GGGGCCCTCA GGGTTTCAGG GCCTGGGTGT AACAGCAGCC AGTCAGATCA GGGTAGATAC
 167821 ACACAGATAC ACACACACAC ACACACACAC ACACACACAT ATTATCCCTG GGGCTCAGCC
 167881 CAGCTCACTC TGGAGGCTGA AGGCACTCT GTCAGCCCTG TATCTGGCTT CATTGGAAGT
 167941 TTCCCAAATG CAAAGACTT TTGTTGTTGT TGTTGAACCTC AGGCCCCCTTA CTTCTCCAC
 168001 ATTCTGCTT AAAACACAAG ATGCAAAAAA AAAAAGCCAA GTGCTCAAAG GCATACGGTA
 168061 GTATGGTGGC TAATTCTTTG GGTGTTGGAA TTAAAAAAA AAAAAGGCC CTGAGTTTGA
 168121 GTTCAAGGTTG TGGCTTGTG ACTTTGACT TTAATGTTCT TGGATAATA ACAATTCTCT
 168181 GAGCTCTAGT CTCTACATCT ATAAAACATG ATAATATTAA CCTCCCTCAC AGAATTGTGA
 168241 GGATAAAGGA AAACACATAT GTAAAATGTT TGACACAAAAA CAAGTATTCA ACAGAGGCTA
 168301 TATAACAATA TTATTAAGGA CAAGAGTTG TGATAAAAATA GGGAAACACAG AGAGAAAACC
 168361 CAAAATAAT GCAGAAGACA GAAAGAAATG GGAAACACCG AAGACCCGTG GGCAAGCATG
 168421 AAAATCAGAC CTCCCTCAAA AGAACCTGCT GCAGGCTTAT AGTGACCAAC AGTCCCTGGC
 168481 CACCCCTCTGG AGCTATCACT GCGTTCACAC CAAGGCCACG CTTTCTTGAG TTACTCCTAG
 168541 CCAAAAGACAG AACGTGGTGG GGCACACTGTG TCTGCCATT CCTGTTGGGAT ATGGGACCCC
 168601 TCTAATGGGC AACTATGGCT CCAGGACTCC CATCAGCCGG GCTGAAATT TCTTAGAACT
 168661 GAGTCTGAGG CCCCTCTTC CATCAGCCCT TCCCTCTCTG TCTCTTCTCA TAGGTGTCCA
 168721 AATGGTATTA CAATCTGAAG GCTCTCCTGC CCCTTTCTT CCTCATTCA AGGTGTGTTAC

168781 TTTAACAAAGT CTCTGGCACA AATAATCCA TCTTGGCAC AC TTAGAGCAC CTGAAC TAAC
 168841 ACAGGCCCT TCCCTAATGC TGGGCTTTCT TAGAGGGAGG GTCCCTGAGAG GATCTGTCCC
 168901 AAATCTGTGT TGTGCACCAC CTGGGTGAGG GCAGCCGAG TTCCAGGTT ACATGAAATT
 168961 CTTAGATCAT GGAGCTTGAA TGGGAAACAG GAAGTGAGT TCCTTGGGAC ATGCTATGAT
 169021 GCCTGATGGT TCATGTTCT CTCTGTACCT GAAGTCGCC CTCAGATTAA GCAATCCCTG
 169081 CCTCTTCCTG CCCCCACTCC CTAGGCTGGC TCCTCTGGC TGCACTTCAA ATTCACTCTG
 169141 CTTGTGCATA AACATAATTA GTGTGAATGC TTTGTTTGTG TGAACCAAGTA TTCTATATAG
 169201 GTCTGGGGAG GAGTTTATTG TGGCTTCTC TGAAGCCATA AATGTAAGG AGAGCTTTAG
 169261 GTGGATGTG GTCCTCGGG CAGGGAGACA AGCTGCAGAC GTGGCCCAAG CCTAGGAGAG
 169321 CTGGAAACT CCAGGCCAGG GAGGGCTGTG GCTATGCCA CTTGCAGGTT GGGTTAAATC
 169381 CCCAGCCCTT GCAGGGACAG AGTGTAGCCC TCAGGGCTCCT GGGGTTCTG GGTCA GATAA
 169441 GGGCTCTGG AGACTCCCT CCACCCCTCA TCCCAGGACC CAGTGCATCA GGGAGTCCT
 169501 CAAGGCTGC AGCATGGCGC TGTTGGCCGA CAGGTGCAAC AAGTGGTCGT AAAAGTCTT
 169561 CACCTTGAC ATCATGTAGT ACAGAGAATA GGGTGAGGA AAGTCACAAA GGTGTAATTG
 169621 GCTTCCCTGC CTCCAAGGAG AGAGGAGATT AAGAGCAGGT CACAGAAGGC TTCCACCCAT
 169681 CTAGTAAGCT GAGGTCAA CGTTGGGAAAC AAGCCAGGT TCATGCTTG AAACATAGAA
 169741 ATGTTGCATT TTCTGGATTG AGACACATGT AACCACATG AACCACATG AACCAGCACT ACAGTGTCA
 169801 ATAGAGTAGA TTCACTACCC TCAAAAGCTT CTGTGCTCTG CCTCTTCATC CCTCCCTCCC
 169861 ACCTAACGCT TGGCAATCAC TGATTTTTT TTTTACTGTC GTTTAGTTT TCCTTTTCA
 169921 GAATGTCAA CAATTGAAAT CATACTCTGT ACTCTTTCA GACTGGCTTC TTTCACGTAG
 169981 TAATACACAT TAAGTTTCTT CCATGCTTT TCAAGGCTTG AATGCTTATT TCTTTTTAGT
 170041 GCTGAATCAT ATTCCATTGT CTGGATGTAC CCCAGCTTC TCATCAATTTC ACTCACTGAA
 170101 GGACAGCTTG GTGTCCTA AGATTTAGCA ATTATAGCA AAGCTGCTAT AGACATTCA
 170161 GTGCACTTT GTAAAGAGA TAAGTTTCA ACTCTTTGG GTAAATATT CAAAATGAAA
 170221 TGAGAATAT GTTTCGTTT TTAAGAACCC TTCAAGATTG TATCAAAGT GACTGAATCA
 170281 TTTGCAATT CCACCAATTAA TGAGTGGAAAG TTCTGTGTC TCCATGTCTT CGCCAGCATT
 1 170341 TGGTGTGCTT AGTGTCTGG ATCTAGTCAT TCTCATAGAT GTGTAGTACT ATTTCAATTGT
 170401 TCCTTAATT TGCATTCCC TGATGATGTA TGATACGGAG CATCTTTCA TATGCGTACC
 170461 TGTCACTGT ATTCTCTCTT TGGTGGAGGT TCTGTTAAGG TCTTTGGCCC ATTTTTAAAT
 170521 AGGGTTGTTT TCCTTTTATC AAGTGTAAAG AGTACTTTGT ATTTTGATA ACAGTCCTT
 170581 ATTAGATATG TCCTTGCAA ATATTTCTT CCAGTCTGTG GCTCTCTTC TCATTCCTT
 170641 GGCAGTGTCA TTGAGGAGC AGAAAATGTT TTTTTTTTG AGACAAAGTC
 170701 TCACCTATT GCCCAGCTG GAGGTGAGTG GCACAACTTC GGCCCAGTC AACCTGTCC
 170761 TCCGGGTTCA AAGCAATTCT CCTGCCCTCAG CCTCCCAAGT AGCTGGATC AGAGGCATGC
 170821 GCCACCATGC CCAGCTAATT TTGTATTTT AGTAGAGATG CGGTTTCA CCGGTTGCC
 170881 GGCTGGCTC AAACCTCTGA CCTCAGGTGA TCCACCCGCC TCAGCCTCCC AAAGTGCTAG
 170941 AATTACAGGT GTGCAACCACC ACCGCTGGCC AAGAAATGTT TTATTTAAAT AAAATCCAGC
 171001 TTATCAATTA TTCTTTCTAT GGATGTTGCC TTTGGTATTG TATCAAAAAA GTCATTGCAA
 171061 AACCTTAAGT CATCTAGATT TTCTCCCGTT ATCTCTAAAG AGTCTATAG TTTTACATT
 171121 TACATTTAAAG TCTCTTATCC ATTGAGAT AATTGGTGG AAGAGTATAA GGTCTGTGTC
 171181 TAGCATTAT TTTTTTTT TTTTTCTATGG TGGATGTC TGCCTCCAG CATCATTG
 171241 TGAAAAGACT CGGCTTATT GTATTGCCT TGCTCCATTG TCAAAAGATCG TATATTATG
 171301 TAAGTCTATT TCTGGCTGT CATTCTGTTC CATTGATCTA TTTGTCTCTG TGCTGTGCT
 171361 TCATCCCTTC CTCCCCCTTA ACTCTGTCA CCAATATCAC ATTGTCTGA TTATCATAGC
 171421 TTTATAGTTA GTCTTTAAGT TAGATGAGT CAGTCTACCA ACTTTGTCT TCTTTAATGT
 171481 AGTTTGCGT ACTCTGGGTT TCTGCCCTCC CCATACAAAC TTTCGAATCA GTTTGTCAGT
 171541 ATCCACAAAAA TAGCTGCTG GGATTTGAT TGGAAAGTGC TAAAAACTAT AGATCAAGTT
 171601 GTGGAGAAGT GACATCTGA CAATATTGCA TCTTCTAAC TATGAAATAG GAATATCT
 171661 CCATTTACTT AGTTCTTAA ATTCTTTCA TTGTGGTTT ATAGTTTCC TCATATAGAT
 171721 GTTGTGCATA TTTTTAGCT TTATACCTAA GTATTCAATT GTGGGGATGC TAAATAGTAT
 171781 TGTGTTTTT AATTCAAAT TCCACTCATT CATTGCTGGT ATATAGGAAA GTGATAGACT
 171841 TTTAGATATT AAACTGTGT CCTGAAACCT TGCCATAATT GCTGTTCAC TTGGAGTTT
 171901 TCTAGTACAT GTATGGAGC ATGTAATCC AATTATTTT TATTTACAA ATAACATATCT
 171961 GTTATCACAA ATGCTCTTAT TAAAAAGTC TTTCTCTT CAGTGACTTA TCATGTATAG
 172021 TTTACACTA ATGTTCATAT GTATTGCCT ATTTTTTAT ACTGTGTATT CTGTTCTGTT
 172081 GGTTATGTGT CAGCAGCATA TAATTAAAT TATAAGGCCT TTATCATGTT TAAATTTG
 172141 ATAAGGCTCA TCCCCGCC CCCATTCA CCAATAATTAT TGCTTTCTA AGAGATGGGG
 172201 TCTCACTATA TTGCCAGGC TGATCTAAA CTCTGGCCT CAAGCGATCC TCCTGCCCA
 172261 GCCTCCAAA GTACTAGGAT TATAGGTGTG AGCCACCGTG CCTGGCTCT GCAAGTTAT
 172321 CTGTTCATAT AAACTTACA ATTAACCTTC CTAGCTTAA AAAATAGTT TTTTTAATTG
 172381 AATTACATTG GGGAAATTG ATGACTTTAT GACATCAATT TCCCCCAATG TAATAAATAT
 172441 ATACTTATT TTATTATTAT TATTTTTGAA GACAGAGTT TGCTTTGTCA CCCAGGCTAG
 172501 AGTGCAGTGG CACAATCTCG GCTCACTGAA ACCTCTGCC CCCCAGGTT AAGCAATTCT
 172561 TGTGCCTCAG CCTCTGTAGT AGCTGGCCAT CAGGCACCTG CGACCACACC CAGCTGATT
 172621 TTTGTATTG AGTAGAGACG GGGTTCGCC ATGGTGGCCA GCCTGGTCTT GAACCTCTAA
 172681 GCTCAGGCCA TCAGCCCAGC TCAGCCTCCC AAAGTTCTGG GATTACAGGC ACGAGCCACC
 172741 ACACCCGGCC TCTACTTACT TATTTTATTAA TATTTATTAA TTGAGATGGA
 172801 GTCTCGCTT GTCACCCAGG CTGGAGTGCA GTGGTGCAAT CTTGACTCAC TGCAACTTCT
 172861 GCCTCTGGG TTCAAGTTAT TCTCTGCCT CAGCCTCTCG AGTAACCTGGG ATTACAAGCA
 172921 CGTGCACCA TGCCAGCTA ATTTTTGTA TTTTAGTAA GACCGGGTTT CACCATGTTG
 172981 GCCAGGCTGG TCTTGAACTC CTAACCTGCG CCTCCGGGT TCAAGCAATT CTTGTGTGTC

173041 AGCCTCCCTA AGTGCTGGGA TTACAGGCAT GAACCAC TGC ATCCAGCCC A TTTATTATT
173101 TTAGATTCA GAGGTATAAG TATAGGTTT G TTACATGGG ATATTGCATG ATGCTGAGGT
173161 TTGGGCTTCT AACCCAAGTA CTTCGCCCAA ATGATATGGG ATGGGACAGG GAAGTTCTGG
173221 TTAGAGAAGG GCGGGGTCCC TGGCGAGGGC TCCATCCCCA TTTTGTGCTC AAATGTTGTG
173281 TTTCCCAAGA CCACCCCTGGC CTGCCACACT CCCATCCTGT GCCTATAAAT ACCCTGAGAC
173341 ACTAGCGGGC CAAGCAGCTG GATGTTGAGA GGAACCTCAGC CATGGAAGAA GACACAAGCC
173401 ACCTAAAGAT GGCAAAACTG AAAGAGCACA CCATAACACA CGCCCAC TGG GGCTTCAGCT
173461 GTAAACATTC ACCCCCTAGA CACTGCCATG TGGTCGGAAC CCCACAACCT GTCTGTATGC
173521 TCCCCTAGAG ATTGAGCAG CGAGCCACAC CCCCATCGCA CGCCCTGCAG GGGGATAAGG
173581 GAACTTTCC CATTTCACAA GTATTGAATA GGTAGTTGTT CAACCCCTTGC CCCCTTCCTT
173641 TCTCCCTGCT TTTGGAATTC CTAGTGTGTT ATTGTTCTG TCTTTGTGTT

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 INDUIENT QU'IL Y A UNE INCERTITUDE
SUR LEUR IMPLICATION AU NIVEAU DE LA SÉQUENCE CODANTE