

LAURE DE HERCÉ

**LE GÉNOME CHLOROPLASTIQUE DE L'ALGUE
VERTE *PABIA SIGNIENSIS***

Mémoire présenté
à la Faculté des études supérieures de l'Université Laval
dans le cadre du programme de maîtrise en Biochimie et Microbiologie
pour l'obtention du grade de Maître (Me)

DÉPARTEMENT BIOCHIMIE ET MICROBIOLOGIE
FACULTÉ DE SCIENCE ET GÉNIE
UNIVERSITÉ LAVAL
QUÉBEC

2006

Résumé

Les algues vertes et les plantes terrestres sont des organismes photosynthétiques qui se regroupent dans le sous-règne des Viridiplantae. Ce règne se divise en deux lignées évolutives que sont les Streptophyta (plantes terrestres et algues vertes de la classe des Charophytes) et les Chlorophyta (algues vertes appartenant aux classes : Prasinophyceae, Ulvophyceae, Chlorophyceae et Trebouxiophyceae).

L'embranchement des ordres composant la classe Trebouxiophyceae est encore mal connu. En étudiant les génomes chloroplastiques d'au moins un représentant de chaque ordre, il sera possible de reconstruire une phylogénie juste et de mieux comprendre l'évolution moléculaire de cette classe. La famille Trebouxiophyceae compte 4 ordres : Trebouxiales, Microthamniales, Chlorellales et Prasiolales. Pour chacun de ces ordres, excepté les Prasiolales, un génome chloroplastique d'une algue verte a été séquencé. Le séquençage du génome chloroplastique de *Pabia signiensis*, aussi appelé *Planophila sp* et appartenant aux Prasiolales, vient compléter cet échantillonnage. Selon les analyses phylogénétiques du gène codant pour l'ARNr 18S, *P. signiensis* aurait divergé relativement tôt durant l'évolution de la classe des Trebouxiophyceae et formerait avec *Chlorella ellipsoidea* (Prasiolales) le groupe sœur de *Chlorella vulgaris* (Chlorellales).

Le génome chloroplastique de *P. signiensis* est, avec ses 236,463 pb, un des plus grands génomes séquencés dans la classe des Trebouxiophyceae. La présence de régions inversées répétées, lui confère la structure quadripartite caractéristique de la plupart des génomes chloroplastiques. Sur les 111 gènes trouvés, seul le gène *ycf47* est propre au génome chloroplastique de *P. signiensis*. Les principaux regroupements de gènes ancestraux chez les génomes chloroplastiques d'algues vertes sont présents et 13 d'entre eux sont entièrement conservés. Aucun intron n'a été mis en évidence. L'ordre des gènes et leur répartition dans le génome sont peu réarrangés par rapport aux génomes ancestraux de la classe Prasinophyceae. Seules les longues régions inversées répétées (27.3 kb) dérobent à la règle, puisque celles-ci possèdent les gènes *psbA* et *trnS(gcu)* en plus de contenir un

opéron *rrn* brisé en deux segments qui sont inversés l'un par rapport à l'autre et un taux élevé de SDRs.

Bien que plusieurs des caractéristiques mentionnées ci-dessus se retrouvent dans les génomes chloroplastiques des Trebouxiophyceae, le grand nombre de caractères ancestraux présents dans le génome de *P. signiensis* semble confirmer la position basale de cette algue verte et appuie l'hypothèse selon laquelle les Trebouxiophytes seraient à la base du clade UTC (Ulvophyceae-Trebouxiophyceae-Chlorophyceae).

Remerciements

Je tiens à remercier les Docteurs Monique Turmel et Claude Lemieux de m'avoir permis d'effectuer cette recherche au sein de leur laboratoire. Je remercie Christian Otis pour sa gentillesse et son aide quotidienne. Je salue tous les étudiants du laboratoire et particulièrement Jean-François Pombert pour son éclairage tout au long de ce travail.

Je remercie également le Docteur Jean Bousquet ainsi que les membres du CREFSIP et du Pavillon C-E Marchand.

J'envoie un petit clin d'œil à toutes les personnes que j'ai rencontrées pendant ces deux années sur le sol québécois et les remercie des beaux moments d'amitié partagés en leur compagnie.

Table des matières

Résumé	III
Remerciements	V
Table des matières	VI
Liste des tableaux	IX
Liste des figures	X
Liste des abréviations	XIII
I. Introduction	14
I. 1. Les Plantae	14
I.1.1. Le règne des Viriplantae	14
I.1.1.1. Les Chlorophyta	17
I.1.1.1.1. Les Trebouxiophytes	20
I.2. Le chloroplaste	20
I.2.1. L'origine endosymbiotique	20
I.2.2. Le génome chloroplastique	25
I.2.2.1. Généralités	25
I.2.2.2. Les gènes chloroplastiques	26
I.2.2.2.1. Les gènes photosynthétiques	26
I.2.2.2.2. Les gènes de l'expression génique	26
I.2.2.2.3. Les gènes d'ARNs	26
I.2.2.2.4. Les autres gènes	29
I.2.2.3. Les régions inversées répétées	29
I.2.2.4. Les introns	30
I.2.2.5. Organisation génique	30
I.2.3. Evolution des ADNcp	31
I.2.3.1. Perte des gènes	31
I.2.3.2. La variabilité en taille de ces génomes	32
I.2.3.3. Absence et présence d'intron	32
I.2.3.4. Autres particularités évolutives des ADNcp	33
I.2.4. Les ADNcp de la classe Trebouxiophyceae	33
I.3. La problématique	34
I.3.1. Le choix de <i>Pabia signiensis</i>	35
I.3.2. L'utilisation de la génomique	35
I.3.3. L'utilisation de l'ADNcp	36
II. Méthodologie	37
II.1. Culture de <i>P. signiensis</i>	37
II.2. Extraction de l'ADN chloroplastique	37
II.2.1. Lyse des cellules	37
II.2.2. Extraction de l'ADN cellulaire par un mélange phénol/chloroforme	37

II.2.3. Purification des ADN sur gradient de chlorure de césium	38
II.3. Construction d'une banque d'ADN génomique	38
II.3.1. Nébulisation	39
II.3.2. Réparation des fragments et électrophorèse	39
II.3.3. Electro-élution	39
II.3.4. Electrophorèse sur gel d'agarose 0,8%	40
II.3.5. Clonage	40
II.3.5.1. Ligature	40
II.3.5.2. Transformations des cellules compétentes	40
II.3.5.3. Mise en culture	41
II.3.6. Sélection des clones par hybridation	41
II.3.6.1. Repiquage des clones	41
II.3.6.2. Transfert des clones recombinants sur membrane de nylon	42
II.3.6.3. Traitement des membranes et préhybridation	42
II.3.6.4. Préparation de la sonde	42
II.3.6.5. Hybridation et lavage des membranes	42
II.3.6.6. Radiographie des membranes	43
II.4. Classification de clones recombinants	43
II.4.1. Purification des plasmides	43
II.4.2. Dosage des plasmides	43
II.4.3. Digestion des plasmides et électrophorèse	44
II.5. Séquençage et stratégie	44
II.5.1. Amplification par PCR	46
II.5.1.1. Purification des produits de PCR	46
II.6. Caractérisation des génomes	47
II.6.1. Localisation des gènes	47
II.6.2. Assemblage des séquences nucléotidiques	47
II.6.3. Analyses de la composition et de l'utilisation des bases	47
II.6.4. Analyses comparatives	48
II.6.5. Construction de cartes, graphiques et autres schémas	48
II.6.6. Les séquences répétées	48
III. Résultats	50
III.1. Généralités de l'ADNcp de <i>P. signiensis</i>	50
III.2. Contenu génique	50
III.3. ARNt et utilisation de codons	54
III.4. Répartition des gènes le long de la structure quadripartite	57
III.5. Organisation génique et réarrangement de gène	66
III.6. Les séquences répétées	72
III.7. Les régions inversées répétées	79
III.8. L'opéron <i>rrn</i>	82
IV. Discussion	88
IV.1. L'ADNcp de <i>P. signiensis</i> : un génome très conservé	88
IV.2. Un génome de grande taille	90

IV.3. La prolifération de SDRs dans l'ADNcp de <i>P. signiensis</i>	91
IV.4. Les IRs de l'ADNcp de <i>P. signiensis</i> : les régions les plus réarrangées	92
V. Conclusion	95
VI. Perspectives	97
Bibliographie	98
Annexe I	101
Annexe II	201
Annexe III	202
Annexe IV	203

Liste des tableaux

Tableau 1 : Comparaison des principales caractéristiques des ADNcp des algues vertes de la division des Chlorophyta.

Tableau 2 : Contenu en gènes des ADNcp des algues vertes du sous-règne des Chlorophyta.

Tableau 3 : Utilisation des 29 ARNs de transfert de *P. signiensis*, codant pour 61 codons

Tableau 4 : Comparaison des différents contenus en regroupement de gènes entre les ADNcp des Trebouxiophytes et ceux des Prasinophytes

Tableau 5 : Nombre minimum d'inversions lors des réarrangements géniques entre les différents ADNcp de Chlorophyta

Tableau 6 : Abondance des SDRs au sein des ADNcp des algues vertes du sous-règne des Chlorophyta

Tableau 7 : Désignation des répétitions de plus de 15 paires de base et les plus fréquemment retrouvées dans l'ADNcp de *P. signiensis*

Liste des figures

Figure 1 : Schéma de l'évolution des Viridiplantae. L'ensemble des trois classes colorées est appelé clade UTC

Figure 2 : Schéma des différentes phylogénies du clade UTC selon les analyses effectuées. D'après Lewis et McCourt 2004 ; Pombert et al. 2005

Figure 3 : Phylogénie des Trebouxiophytes. Chaque organisme dont l'ADNcp a été séquencé ou en cours de séquençage est indiqué sous l'ordre auquel il appartient

Figure 4 : Schéma de l'endosymbiose primaire à l'origine des plastes à deux membranes. Les noyaux, les mitochondries et les chloroplastes sont respectivement coloriés en noir, jaune et vert

Figure 5 : Origine des gènes codant pour les protéines impliquées dans la membrane photosynthétique des chloroplastes. En bleu, les protéines dont les gènes ont une origine nucléaire, en rouge celles dont les gènes sont chloroplastiques et en vert les protéines dont les gènes ont au moins été une fois séquencés dans un génome chloroplastique. D'après Race, Hermann et Martin 1999

Figure 6 : Carte génétique de l'ADNcp de *P. signiensis*. Les gènes (boîtes vides) à l'extérieur de la carte sont transcrits dans le sens horaire tandis que ceux positionnés vers l'intérieur de la carte sont transcrits dans le sens anti-horaire. Les gènes codant pour les ARNt sont indiqués par la lettre du code des acides aminés à laquelle ils correspondent et sont suivi entre parenthèse de leur anticodon. La partie en gras de l'ADNcp de *P. signiensis* et les flèches délimitent les régions inversées répétées. *P. signiensis* est désignée par la même abbréviation que celle décrite dans la légende du tableau 1.

Figure 7 : Comparaison entre la répartition génique de l'ADNcp de *P. signiensis* et celui de *M. viride*. En bleu sont représentés les gènes appartenant à la LSC de l'ADNcp de *M. viride*, en rose ceux de la SSC et en jaune ceux des IRs. Les gènes absents de l'ADNcp de *M. viride* sont matérialisés en gris. *P. signiensis* est désignée par la même abbréviation que celle décrite dans la légende du tableau 1.

Figure 8 : Comparaison entre la répartition génique de l'ADNcp de *P. pseudoalveolaris* et celui de *M. viride*. En bleu sont représentés les gènes appartenant à la LSC de l'ADNcp de *M. viride*, en rose ceux de la SSC et en jaune ceux des IRs. Les gènes absents de l'ADNcp de *M. viride* sont matérialisés en gris. *P. pseudoalveolaris* est désignée par la même abbréviation que celle décrite dans la légende du tableau 1.

Figure 9 : Comparaison entre la répartition génique de l'ADNcp de *C. vulgaris* et celui de *M. viride*. En bleu sont représentés les gènes appartenant à la LSC de l'ADNcp de *M. viride*, en rose ceux de la SSC et en jaune ceux des IRs. Les gènes absents de l'ADNcp de *M. viride* sont matérialisés en gris. *C. vulgaris* est désignée par la même abbréviatiion que celle décrite dans la légende du tableau 1.

Figure 10 : Comparaison entre la répartition génique de l'ADNcp de *L. terrestris* et celui de *M. viride*. En bleu sont représentés les gènes appartenant à la LSC de l'ADNcp de *M. viride*, en rose ceux de la SSC et en jaune ceux des IRs. Les gènes absents de l'ADNcp de *M. viride* sont matérialisés en gris. *L. terrestris* est désignée par la même abbréviatiion que celle décrite dans la légende du tableau 1.

Figure 11 : Comparaison de la conservation en ordre de gènes entre les différents ADNcp des Trebouxiophytes et celui de *P. signiensis*. En vert et rouge les regroupements de gènes communs. Les algues vertes sont désignées par les mêmes abbréviatiions décritent dans la légende du tableau 1

Figure 12 : Contenu en paires de gènes ancestrales ou dérivées entre les différents ADNcp des algues vertes du sous-règne des Chlorophyta. Les cases noires correspondent aux paires de gènes présentes dans le génome tandis que les cases blanches indiquent une perte de gènes. En gris sont matérialisés les gènes présents dans le génome mais non regroupés en de tel paires. Les algues vertes sont désignées par les mêmes abbréviatiions décritent dans la légende du tableau 1

Figure 13 : Analyse PIPMAKER de l'ADNcp de *P. signiensis*. La séquence du génome a été alignée contre elle-même. Les gènes sont représentés par des boîtes noires lesquelles sont surmontées de flèches indiquant leur polarité. Chaque point correspond à une séquence répétée dans le génome pour laquelle l'homologie dépasse 50%. En lettres colorées sont représentées les séquences répétées directes de plus de 15 pb les plus fréquemment trouvées dans l'ADNcp de *P. signiensis*. Leur signification est donnée dans le tableau 7.

Figure 14 : Analyse PIPMAKER des IRs de *P. signiensis* et de *C. ellipsoidea*. Dans un premier temps les IRs de chaque organisme a été comparés contre eux-même, puis l'un contre l'autre afin de mettre en évidence les séquences communes. Les algues vertes sont désignées par les mêmes abbréviatiions décritent dans la légende du tableau 1

Figure 15 : Comparaison des différents opérons *rrn* au sein des ADNcp des Trebouxiophytes. Lorsque ces opérons sont présents dans les IRs, l'ensemble du contenu génique des IRs ont été rajoutés. Il en est de même pour l'opéron de *C. ellipsoidea* bien que la présence d'IRs dans cet ADNcp n'ait pas été confirmée. L'opéron *rrn* de *C. vulgaris* est identique en composition et en répartition génique à ceux des ADNcp ancestraux de *M.*

viride et de *N. olivacea*. Les algues vertes sont désignées par les mêmes abréviations décritent dans la légende du tableau 1

Figure 16 : Les différentes structures secondaires des ITS $\gamma\delta$ de l'ADNcp de *C. ellispoidea*, de *P. pseudoalveolaris* et de *P. signiensis*. Les algues vertes sont désignées par les mêmes abréviations décritent dans la légende du tableau 1

Liste des abréviations

ADN : Acide DésoxyriboNucléique
ADNcp : ADN chloroplastique
ADNmt : ADN mitochondriale
ADNr : ADN ribosomique
ATP: Adénosine Tri-Phosphate
ARN : Acide RiboNucléique
ARNt : ARN de transfert
BLAST: Basic Local Alignment Search Tools, logiciel de recherche d'homologie
CW : ClockWise, sens horaire
DO : Direct Opposed, opposé face à face
dNTP : DésoxyNucléotides TriPhosphates
dATP : DésoxyAdénosine TriPhosphate
dCTP : DésoxyCytidine TriPhosphate
dGTP : DésoxyGuanosine TriPhosphate
dTTP : DésoxyThymidine TriPhosphate
IR: Inverted Repeat, région répétée inversée
ITS : Internal Transcribed Spacer, espaceur transcrit interne
kb : kilobases
LSC: Large Single Copy, long segment unique
ORF: Open Reading Frame, cadre de lecture ouvert
pb : paire de bases
PCR : Polymerase Chain Reaction, réaction en chaîne de la polymérase
SDR : Short Dispersed Repeat, petite répétition dispersée
SSC: Small Single Copy, court segment unique
UTC: Ulvophyceae, Trebouxiophyceae et Chlorophyceae

I. Introduction

I. 1. Les Plantae

Le règne des Plantae comprend les Rhodophyta, les Glaucophyta et les plantes vertes. Il est caractérisé par la présence de chloroplaste à deux membranes (de Reviere 2002), issus d'un ou plusieurs phénomènes endosymbiotiques primaire entre une cyanobactérie et un procaryote (Nozaki et al. 2003). Les autres eucaryotes photosynthétiques se différencient de cette lignée par leurs plastides issus d'événements endosymbiotiques dits secondaires ou tertiaires

I.1.1. Le règne des Viridiplantae

Les algues vertes et les plantes terrestres sont définies comme des organismes photosynthétiques, se regroupant sous le sous-règne des Viridiplantae de part leurs caractéristiques communes (Lewis et al. 2004) :

- la présence de chlorophylle *a* et *b* et d'autres pigments accessoires tels que la β -carotène et la xanthophylle
- la présence de thylakoïdes dans leur chloroplaste
- la présence d'amidon intraplastidiale
- la composition cellulosique de leur membrane cellulaire lorsqu'elle existe
- leur organisation flagellaire

Les Viridiplantae (figure 1), divisés en deux grandes lignées, comprennent les Streptophyta et les Chlorophyta. Ces algues vertes forment un groupe d'eucaryotes photosynthétiques uni- ou pluricellulaires, voisins des plantes terrestres ayant pour ancêtre commun *Mesostigma viride* (Lemieux, Otis, and Turmel 2000). Malgré les controverses, cette minuscule algue verte situe la séparation entre les Streptophyta et les Chlorophyta.

Les Streptophyta englobent les plantes terrestres et les six groupes monophylétiques d'algues vertes rassemblés sous le nom de Charophytes : Mesostigmatales, Chlorokybales,

Figure 1 : Schéma de l'évolution des Viridiplantae. L'ensemble des trois classes colorées est appelé clade UTC.

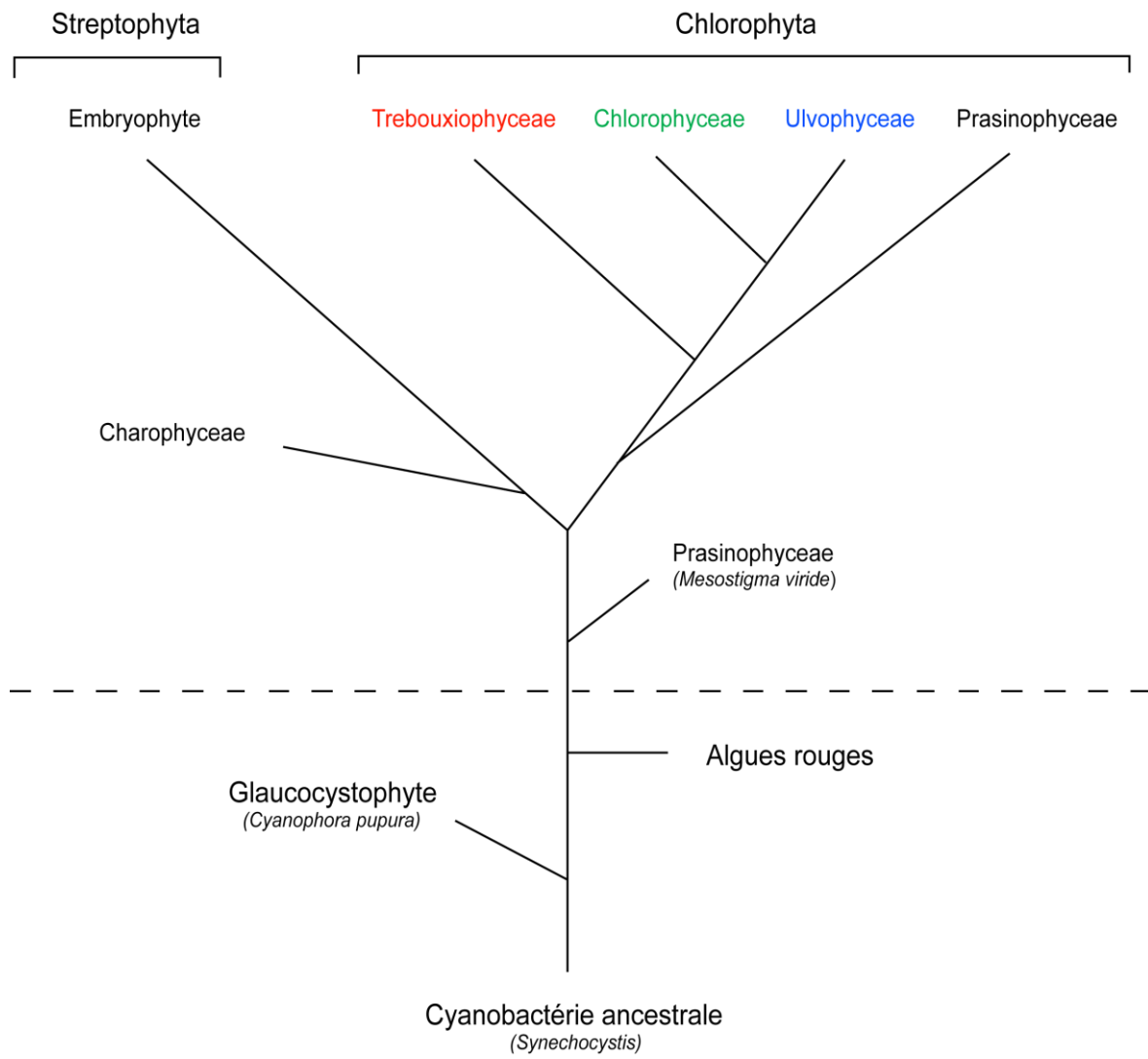


Figure 1

Klebsormidiales, Zygnematales, Coleochatales et Charales (Turmel, Otis, and Lemieux 2003).

I.1.1.1 Les Chlorophyta

Selon la classification basée sur des analyses ultrastructurales, tel l'orientation des appareils flagellaires des algues vertes (Mattox and Stewart 1985), la lignée des Chlorophytes se compose de quatre principales classes : les Prasinophyceae, les Chlorophyceae, les Ulvophyceae et les Trebouxiophyceae.

Les Prasinophytes sont une classe paraphylétique qui contient majoritairement les Chlorophytes les plus basales comme *M. viride* (à la base de la divergence Chlorophyta et Streptophyta) ou *Nephroselmis olivacea*. Des analyses de la petite sous-unité de l'ARN ribosomique nucléaire démontrent que cette classe émerge avant le clade UTC (Bhattacharya and Medlin 1998)

L'ordre d'embranchement des trois autres classes (figure 2) est encore mal connu et diffère selon les types d'analyses effectuées (Lewis and McCourt 2004).

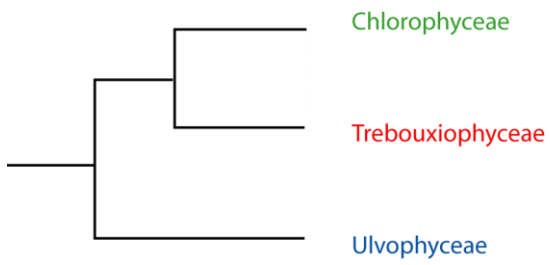
Le groupe non-monophylétique des Ulvophytes comprend les algues vertes présentant une grande variabilité en taille de leur ADNcp (Lewis and McCourt 2004). *Codium fragile* qui possède le plus petit ADNcp avec 89 kb en fait partie.

Les Chlorophyceae, groupe monophylétique, sont regroupés selon l'orientation CW ou DO de leur appareil flagellaire (O'Kelly and Floyd 1984; Mattox and Stewart 1985). Cette classe comprend les algues vertes microscopiques les plus familières. L'ordre des chlorellales appartenait à cette classe avant qu'il ne rejoigne la classe des Trebouxiophyceae (Lewis and McCourt 2004).

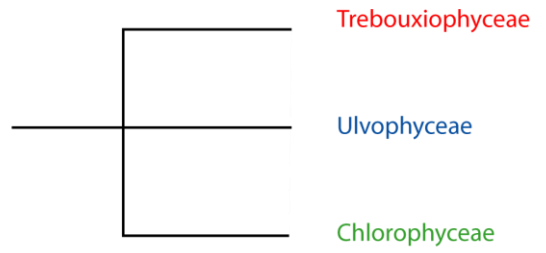
Tout comme les Ulvophyceae, les Trebouxiophyceae sont un groupe non monophylétique ayant le sens anti-horaire comme sens d'orientation de leurs appareils flagellaires, soit le sens ancestral (O'Kelly and Floyd 1984).

Figure 2 : Schéma des différentes phylogénies du clade UTC selon les analyses effectuées.
D'après (Lewis and McCourt 2004; Pombert et al. 2005).

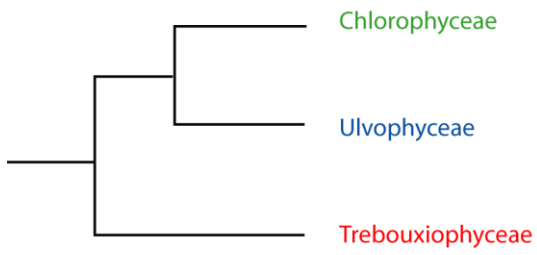
Analyses Ultrastructurales



Analyse Nucléaire



Analyse Mitochondriale



Analyse Chloroplastique

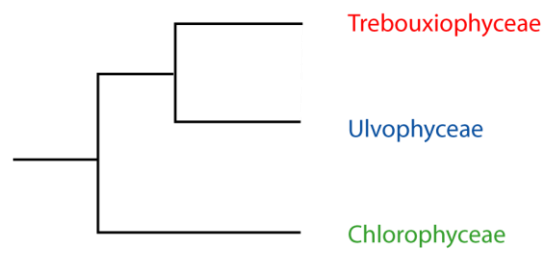


Figure 2

1.1.1.1 Les Trebouxiphyceae

Cette classe a pour caractéristique un mode de reproduction asexuée utilisant soit des autospores soit des zoospores (Lewis and McCourt 2004).

Ces algues vivent majoritairement dans des habitats terrestres et sont repartis en quatre ordres : Trebouxiales, Microthamniales, Prasiolales et Chlorellales (figure 3).

Les Trebouxiales désignent une partie des algues formant les lichens, tel *Trebouxia* qui peut vivre avec des champignons en symbiose.

Les Microthamniales comprennent des algues filamenteuses (*Microthamnion*) et unicellulaires (*Fusochloris*) proches des Trebouxiales mais qui se différencient par la présence de cellules biflagellées (Lewis et McCourt, 2004).

Les Prasiolales regroupent des unicellulaires, à courts filaments ou à petits thalles, ayant une reproduction asexuée par zoospores tel *Pabia signiensis* (Friedl and O'Kelly 2002).

Les Chlorellales se composent de deux familles d'espèces à autospores : Chlorellaceae et Oocystaceae. Certains organismes de ce groupe (Chorelles) ont la possibilité de vivre en endosymbiose avec des animaux et présentent des formes cocoïdes.

Beaucoup de représentants de la classe des Trebouxiphyceae étaient, avant l'analyse de leur ADNr 18S, classés parmi les Ulvophyceae ou les Chlorophyceae (Friedl and O'Kelly 2002)

I.2. Le chloroplaste

1.2.1. L'origine endosymbiotique

Les chloroplastes sont des organites cellulaires présents chez les organismes photosynthétiques. Ils sont le siège du processus de photosynthèse. La théorie endosymbiotique primaire de leur origine au sein des eucaryotes est maintenant bien acceptée, surtout au sein des Plantae. Ces chloroplastes sont issus de l'absorption d'une

Figure 3 : Phylogénie des Trebouxiphytes. Chaque organisme dont l'ADNcp a été séquencé est indiqué en italique sous l'ordre auquel il appartient.

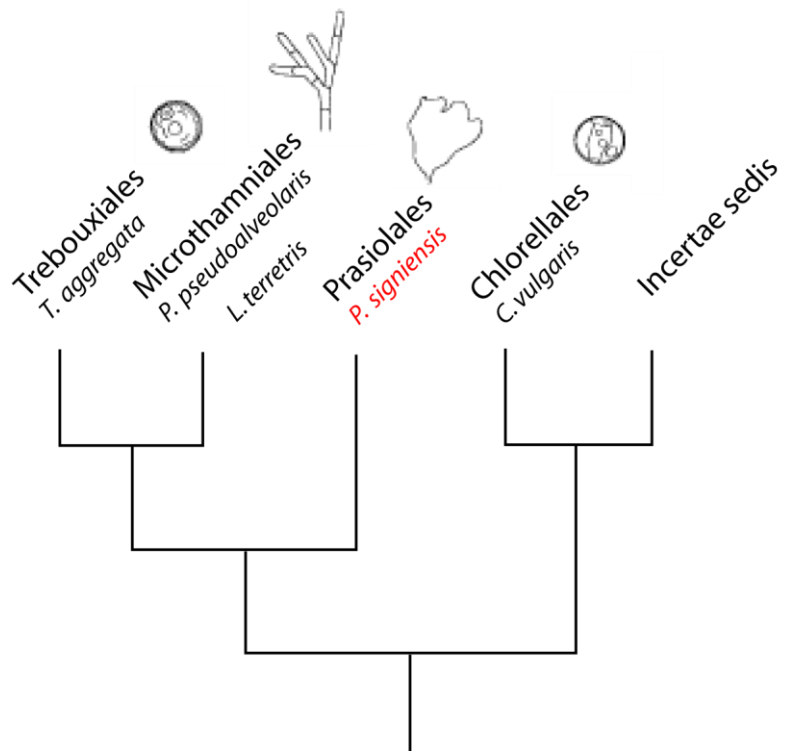


Figure 3

Figure 4 : Schéma de l'endosymbiose primaire à l'origine des plastes à deux membranes.
Les noyaux, les mitochondries et les chloroplastes sont respectivement coloriés en noir,
jaune et vert.

Endosymbiose primaire:



Endosymbiose secondaire:

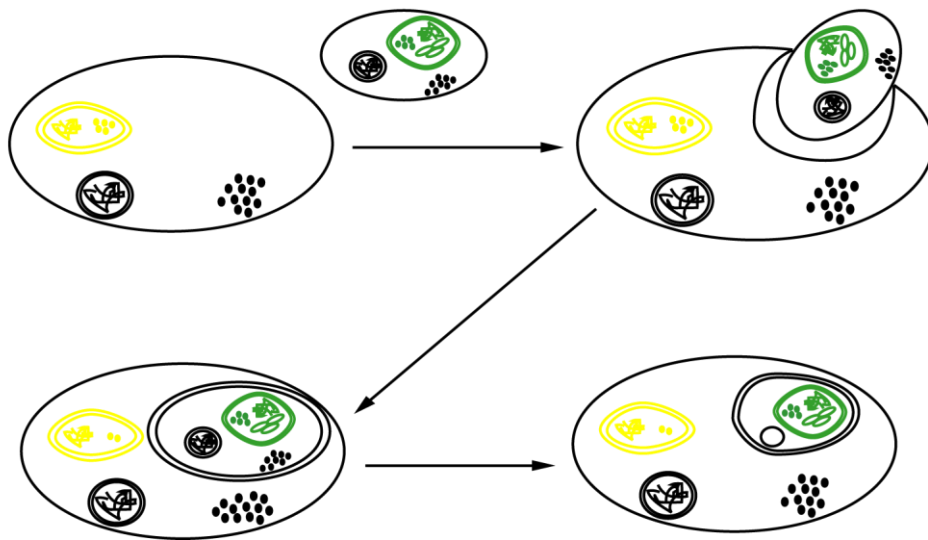


Figure 4

bactérie photosynthétique (cyanobactérie) par une cellule eucaryote hétérotrophe (figure 4). Mereschowsky, le premier a supporté cette hypothèse par des études morphologiques et biochimiques prouvant l'existence d'ARNs, d'ADNs et de ribosomes au sein des plastes (Raven and Allen 2003). Des analyses phylogénétiques des gènes *tufA*, *atpB* et de l'ARN 16S sont venus confirmer cette supposition (Delwiche, Kuhsel, and Palmer 1995), tout comme les études comparatives des génomes de la cyanobactérie et de la plante évoluée *Arabidopsis* (Raven and Allen 2003). Seul les cyanobactéries et les chloroplastes possèdent deux photosystèmes et utilisent l'oxygène de l'eau comme source de pouvoir réducteur (Raven and Allen 2003). Si la théorie de l'endosymbiose primaire est maintenant acceptée, il reste à déterminer le nombre de ces événements.

Suite au phénomène d'endosymbiose, un important remaniement génétique a diminué de 90% le contenu génique de l'endosymbiont (Palmer et Delwiche, 1998). Ce remaniement permet de retracer l'histoire évolutive des Plantae dont les Trebouxiophytes.

1.2.2. Le génome chloroplastique

1.2.2.1. Généralités

Les chloroplastes possèdent leur propre ADN et peuvent se diviser indépendamment de la machinerie nucléaire selon le même processus. Leur ADN double brin est circulaire. Les ADNcp des plantes terrestres ont une taille moyenne de l'ordre de 100 à 200 kb. Aux extrêmes, *Codium fragile* possède le plus petit ADNcp des plantes vertes avec 84 kb. A l'opposé, *Chlamydomonas moewusii* présente le plus gros ADNcp des algues vertes avec 292 kb (Sugiura, Hirose, and Sugita 1998). Les ADNcp de plantes vertes sont généralement constant mis à part ceux des algues vertes (Simpson and Stern 2002). L'ancêtre des plantes vertes présente un génome dont la structure est dite quadripartite ; soit un génome circulaire découpé en quatre sections : une grande région simple copie (LSC), une courte région simple copie (SSC) et deux régions inversées répétées (IRA et IRB) (Palmer 1985). L'organisation génique est dense et compacte avec

un pourcentage important en bases A+T. Les ADNcp contiennent entre 87 et 183 gènes (Sugiura 2003) répertoriés selon leur fonction au sein de la cellule.

I.2.2.2. Les gènes chloroplastiques

I.2.2.2.1. Les gènes photosynthétiques

De l'héritage cyanobactérien, on retrouve les protéines impliquées dans la construction des deux systèmes photosynthétiques. Ces protéines sont soit codées par le génome nucléaire soit par le génome chloroplastique (figure 5). Ce dernier comprend les gènes *psa* et *psb* mais aussi les gènes *pet* codant pour le cytochrome *b6/f* et les gènes *atp* pour les sous unités de l'ATP synthétase. A ceci s'ajoute l'enzyme ribulose-1.5-biphosphate carboxylase/oxygénase dont les grandes et petites sous unités sont codées respectivement par les gènes *rbcL* et *rbcS*.

I.2.2.2.2. Les gènes de l'expression génique

La majorité des gènes sont impliqués dans l'expression génique. Ce sont les gènes utilisés lors de la transcription et traduction, soit 20 à 40 gènes *trn* codant pour les 30 à 35 ARNs de transfert, ainsi que les gènes *rpo* codant pour les sous-unités α , β et β' de l'ARN polymérase et les gènes *tufA*, *infA* correspondant aux facteurs d'élongation EF-Tu et d'initiation IF-I. Les ribosomes 70S impliqués dans la traduction utilisent les gènes *rps* pour les protéines composant la petite sous-unité (30S) et les gènes *rpl* pour les protéines de la grande sous-unité (50S) (Sugiura 1992).

I.2.2.2.3. Les gènes d'ARNs

Les ARNs ribosomiques 23S (ou LSU) et 5S, associés à la grande sous-unité du ribosome découlent respectivement des gènes *rrl* et *rrf*, alors que l'ARN ribosomique 16S (ou SSU) associé à la petite sous-unité ribosomale provient du gène *rrs*. Ces gènes sont

Figure 5 : Origine des gènes codant pour les protéines impliqués dans la membrane photosynthétique des chloroplastes. En bleu, les protéines dont les gènes ont une origine nucléaire, en rouge celles dont les gènes sont chloroplastiques et en vert les protéines dont les gènes ont au moins été une fois séquencés dans un génome chloroplastique. (D'après (Race, Hermann, and Martin 1999))

présents dans les IRs des ADNcp, lorsque ceux-ci existent, et sont regroupés en opéron (Palmer 1985).

I.2.2.2.4. Les autres gènes

L'ADNcp contient les gènes de division de l'organite appelés *minD*, *minE*, *ftsH*, *ftsI* et *ftsW*. Il compte aussi plusieurs cadres de lecture ouverts (ORFs). Certains ORFs codent pour des protéines de fonctions inconnues dont les gènes sont appelés *ycf*.

I.2.2.3. Les régions inversées répétées

L'origine des IRs dans les ADNcp est encore inconnue. Les IRs semblent soit avoir une origine parasitique (Glockner, Rosenthal, and Valentin 2000), soit avoir été acquis durant la réduction du génome (Sugiura 2003) ou soit être issu de l'ADNcp de l'ancêtre des plantes terrestres (Palmer 1985). Leur présence reste très variable au sein des algues vertes (Yamada 1991) bien qu'elles apportent une stabilité aux ADNcp.

Les IRs séparent la LSC de la SSC avec lesquelles elles forment la structure quadripartite propre aux ADNcp des plantes vertes (Palmer 1985). La taille des IRs varie entre 5 et 76 kb et semble à l'origine des variations de taille des ADNcp (Sugiura, 1998). Malgré leur taille variable, le contenu en gènes de ces structures est bien conservé au sein des plantes vertes et présente quelques modifications chez les algues vertes (Yamada 1991). L'ensemble des gènes d'ARNr se retrouvent dans cette région (Palmer 1985) et la présence d'autres gènes dans les IRs est limité aux plantes vasculaires et aux algues vertes (Simpson and Stern 2002).

Chez les Trebouxiophytes, les IRs sont très peu conservés. Ces régions sont absentes de l'ADNcp de *C. vulgaris* et *L. terrestris*. Elles sont réarrangées chez *C. ellispoidea* (Yamada 1991) et conservées chez *P. pseudoalveolaris*.

1.2.2.4. Les introns

L'ADNcp des algues vertes contient entre 0 et 155 introns majoritairement des introns du groupe I. Les introns du groupe I et II se différencient par le processus autocatalytique servant à les éliminer des ARNm lors de l'épissage. L'ancêtre commun des Chlorophytes et des Streptophytes ne comporte pas plus d'un intron, lesquels prolifèrent indépendamment des deux lignées (Turmel, Otis, and Lemieux 1999).

Chez les ADNcp des Trebouxiophytes, la présence d'intron est restreinte. *P. pseudoalveolaris* n'en contient aucun contrairement à *L. terrestris* qui en compte quatre et *C. vulgaris* trois. Les introns des Trebouxiophytes sont tous des introns du groupe I et s'insèrent dans différents gènes. Seul le gène *chlI* contient un intron à la même position dans les ADNcp de *L. terrestris* et de *C. vulgaris*.

1.2.2.5. Organisation génique

Dans les ADNcp, la plupart des gènes sont organisés en opérons ou en regroupements de gènes.

Les opérons correspondent à des unités de transcription constituées d'un promoteur, d'un opérateur et d'un ou plusieurs gènes de structure co-transcrits et positionnés sur le même brin d'ADN. Les gènes codant pour les protéines impliquées dans la transcription et la traduction sont souvent organisés ainsi et présentent une organisation semblable à celle des opérons bactériens (Palmer 1991; Stoebe and Kowallik 1999). L'opéron *rrn* est un bel exemple. Chez les embryophytes, le prototype de l'opéron ARNr est de 20 kb et comprend les gènes *rrf* et *rrl*, séparés par une région codant pour deux ARNt et suivi du gène *rrs* (Simpson and Stern 2002).

Contrairement aux opérons, les regroupements de gènes ne sont pas nécessairement positionnés sur le même brin. Les regroupements conservés parmi les Prasinophyceae font office de caractères ancestraux et sont des indicateurs du degré de réarrangement des génomes (Kelchner and Wendel 1996). Parmi eux, certains privilégient des régions du génome, tel le regroupement *psbB-T-N-H* majoritairement localisé dans la région LSC (Dixit et al. 1999). D'autres se retrouvent régulièrement au voisinage d'autre regroupement

comme *atpA-H-F-I* et *rpoB-C1-C2*, lesquels auraient fusionné dans les lignées chloroplastiques (Stoebe and Kowallik 1999).

1.2.3. Evolution des ADNcp

Le contenu et l'organisation génique ainsi que la faible vitesse d'évolution des ADNcp permettent à l'ensemble de servir de modèle lors d'analyses phylogénétiques. Les changements de taille et du contenu génique sont souvent expliqués par l'expansion et la contraction des IRs (Boudreau and Turmel 1995). Le remaniement génique fait appel aux mécanismes d'insertion/délétion, mais aussi aux phénomènes d'inversion et de transposition.

1.2.3.1. Perte de gènes

Le génome des plastides actuels code pour 5 à 10% des gènes du génome de l'ancêtre commun. Seul 18% des gènes de cyanobactéries ancestrales sont présents chez *Arabidopsis* (Martin and Herrmann 1998). L'importante réduction génique au cours de l'évolution des plastes s'explique selon trois destins géniques possible (Race, Hermann, and Martin 1999) :

- Perte irrémédiable de gènes (ex : ceux utilisés par la cyanobactérie libre pour assurer son autonomie)
- Transfert de gènes vers le noyau en deux étapes (Martin and Herrmann 1998)
 - Gène présent en deux copies : une dans le génome chloroplastique et l'autre dans le génome nucléaire
 - Le gène chloroplastique devient défectueux et est perdu au profit du génome nucléaire
- Rétention de gènes dans l'ADNcp

Certains groupes de gènes peuvent entièrement disparaître selon les lignées. Les gènes *ndh* codant pour des protéines semblables à celles de la chaîne respiratoire NADH hydrogénase des mitochondries sont un bon exemple (Sugiura 1992). Ces gènes sont retrouvés chez les plantes vertes mais disparaissent dans la lignée des Chlorophytes. Ils sont toutefois

présents chez les Prasinophytes basales (Wakasugi et al. 1997). Le transfert des gènes vers le noyau est fréquent. Les photosystèmes sont un bon exemple de partage de gènes codant pour des protéines photosynthétiques entre le noyau et le chloroplaste (figure 5), (Race, Hermann, and Martin 1999).

I.2.3.2. La variabilité en taille de ces génomes

Par rapport aux génomes des cyanobactéries, l'ADNcp des algues vertes présente une taille 20 à 30 fois plus petite (Palmer et Delwiche 1998). Au sein même des plantes vertes, ces génomes varient de 84 à 292 kb et ce à cause des phénomènes d'expansions/contractions des IR, de l'augmentation de régions intergéniques et de séquences répétées dispersées. Ces phénomènes sont à l'origine des réorganisations génomiques (Pombert et al. 2005)

I.2.3.3. Absence et présence d'intron

La présence d'intron témoigne également de l'évolution moléculaire des espèces. Leur présence dans l'ADNcp vient de la capture d'éléments génétiques mobiles capables d'auto-épissage et de leur intégration dans ces génomes. Les ADNcp d'algues vertes présentent moins d'introns que ceux des plantes vertes et contrairement à celles-ci, ils possèdent plus d'introns du groupe I que du groupe II (Palmer et Delwiche 1998).

Au sein même des différentes lignées d'algues vertes, le contenu intronique est variable. Les Prasinophyceae ne possèdent pas d'intron. Les ADNcp séquencés des Trebouxiophyceae présentent moins de 5 introns du groupe I, *Pseudendoclonium akinetum* (Ulvophyceae) présente un maximum de 27 introns du groupe I tandis que les algues chlorophycéennes présentent moins de 10 introns des deux groupes. La présence d'intron est considérée comme un argument de divergence (Kelchner and Wendel 1996).

1.2.3.4. Autres particularités évolutives des ADNcp

Outre les introns, les espaces intergéniques considérés à haut taux évolutif (Kelchner and Wendel 1996) contiennent aussi des séquences répétées dispersées à fort pourcentage en GC (Kusumi and Tachida 2005). Ces séquences augmentent la complexité des génomes (Palmer 1985) et certaines sont capables de se replier sur elles-mêmes formant des structures en épingles à cheveux (Boudreau, Otis, and Turmel 1994). A l'exception des IRs et de l'ADNcp de *Chlamydomonas reinhardtii* (Maul et al. 2002), les génomes chloroplastiques contiennent peu de répétitions (Palmer 1991). Ces dernières sont à l'origine des réarrangements dans l'ordre des gènes tout comme les phénomènes d'inversions (Jansen et al. 2005).

1.2.4. Les ADNcp de la classe Trebouxiophyceae

Les ADNcp des trois Trebouxiophyceae séquencés (*C. vulgaris*, *P. pseudoalveolaris* et *L. terrestris*) sont tous circulaires avec des tailles variant de 145 947 pb à 195 081 pb (Wakasugi et al. 1997). Malgré sa petite taille, seul l'ADNcp de *P. pseudoalveolaris* possède la structure quadripartite de la plupart des ADNcp des algues vertes. Ces trois ADNcp sont riches en base A+T. Leur contenu génique est bien conservé et contient entre 106 et 112 gènes selon les génomes. Ce sont à quelques exceptions près les mêmes gènes qui se retrouvent d'un génome à l'autre. Très peu de ces gènes contiennent d'introns, l'ADNcp de *C. vulgaris* n'en compte que trois et celui de *L. terrestris* quatre.

Les grandes différences de ces ADNcp résident dans leur répartition génique et l'organisation de leur opéron *rrn*. L'ADNcp *P. pseudoalveolaris* qui est le seul à présenter une structure quadripartite, a une répartition génique semblable à celle du Prasinophyte *M. viride*. Certains gènes (*trnMe(cau)*, *trnS(uga)*, *trnG(ggc)* et *psbM*) ont toutefois migrés de la LSC vers la SSC. Malgré l'absence d'IRs, la répartition génique au sein de l'ADNcp de *C. vulgaris* est bien conservée puisque les différentes régions géniques de *M. viride* sont visibles contrairement à l'ADNcp de *L. terrestris* qui a été entièrement remanié.

L'opéron *rrn* n'est présent dans son organisation ancestrale que chez *C. vulgaris*. Celui de *P. pseudoalveolaris* présente une inversion dans l'enchaînement de ces 2 gènes d'ARNt alors que celui de *L. terrestris* est scindé en deux fragments éloignés de moins de 60 kb.

I.3. La problématique

Chez les Chlorophyta, la position basale des Prasinophyceae est bien documentée et maintenant approuvée (Bhattacharya et al. 1998; Turmel, Otis, and Lemieux 2002; Martin et al. 2005). A l'inverse l'ordre d'embranchement des trois autres classes, regroupées sous le terme UTC (Ulvophyceae-Trebouxiophyceae-Chlorophyceae), est moins étudié et présente des résultats divergents selon le type d'analyses effectuées (Lewis and McCourt 2004).

Dans la figure 2, les analyses ultrastructurales donnent la classe Ulvophyceae à la base du clade UTC alors que les analyses phylogénomiques basées sur les ADNcp et ADNmt réfutent cette position. Les phylogénies basées sur les séquences en acides aminés des ADNcp montrent que la classe Chlorophyceae diverge en premier (Pombert et al. 2005) tandis que les phylogénies mitochondriales et les analyses structurales des ADNcp et des ADNmt donnent la classe Trebouxiophyceae à la base (Pombert et al. 2005; Pombert et al. 2006; Pombert, Lemieux, and Turmel 2006). Cette dernière hypothèse semble se confirmer par les dernières phylogénies basées sur les séquences en acides nucléiques des ADNcp (résultats non publiés).

Cette hypothèse peut être confirmée ou réfutée en augmentant le nombre de taxons dans les analyses phylogénomiques et dans les analyses comparatives effectuées sur les ADNcp. Pour chaque classe du clade UTC, un ADNcp d'un représentant de chaque ordre est ainsi séquencé et étudié. L'objet de ce travail a été de séquencer l'ADNcp de *P. signiensis* représentant de l'ordre des Prasiolales au sein de Trebouxiophytes. De plus le séquençage de ce génome doit permettre l'étude de son architecture et de son organisation afin d'apporter de nouvelles connaissances quant à l'évolution moléculaire des ADNcp au sein des Chlorophyta par comparaison aux autres génomes d'algues vertes disponibles.

1.3.1. Le choix de *Pabia signiensis*

La classe des Trebouxiophyceae compte trois ADNcp séquencés jusqu'à présent : *Parietochloris pseudoalveolaris* et *Leptosira terrestris* pour les Microthamniales et *Chlorella vulgaris* pour les Chlorellales. Pour les Trebouxiales, l'algue verte représentant cet ordre : *Trebouxia aggregata*, est en cours de séquençage quant à *P. signiensis*, elle a été choisie pour représenter les Prasiolales (figure 3).

Les analyses de l'ADNr 18S de *P. signiensis*, place cette algue dans le clade monophylétique de *Chlorella ellispoidea*, lesquelles avec *Chlorella mirabilis* forment le groupe sœur de *C. vulgaris* (Friedl and O'Kelly 2002; Pombert et al. 2005). Le choix de cette algue verte fut donc déterminé par sa position phylogénétique mais aussi par sa disponibilité et la possibilité d'isoler l'ADN de ses organelles.

1.3.2. L'utilisation de la génomique

Afin de répondre à l'ordre d'embranchement au sein des Trebouxiophytes, l'approche génomique a été utilisée. Cette technique intègre de nouveaux outils, comme le contenu génique et l'ordre des gènes au sein d'un génome, à ceux de la génétique classique. Tout comme la génétique, elle se base plus sur des liens héréditaires, facilite l'estimation d'homologie et permet l'intégration de l'outil informatique par l'utilisation de données moléculaires.

D'une part comme les données sont issues de génome entier il n'y a pas de confrontation arbre espèce et arbre génétique, permettant de refléter l'évolution d'un organisme et non d'un simple gène (Moret and Warnow 2005). Au niveau des nœuds cette approche ajoute un support statistique jusqu'alors impossible avec l'étude d'un seul gène (Delsuc, Brinkmann, and Philippe 2005). Le passage des caractères moléculaires aux caractères génomiques réduit encore la source d'erreur des classifications et apporte un maximum d'informations.

D'autre part la génomique permet d'avoir en plus de l'aspect phylogénie, une approche structurale des génomes ; soit de nouvelles informations comme l'ordre de gènes,

l'organisation en opérons ou en regroupement de gènes, la présence d'IRs, d'introns, de SDRs.

1.3.3. L'utilisation de l'ADNcp

L'ADNcp a été préféré à l'ADN nucléaire en raison de sa taille plus restreinte. Sa faible vitesse d'évolution en fait un outil adéquat pour la compréhension de l'évolution des plantes vertes. Son contenu et son organisation génique permettent de servir de modèle d'évolution lors d'analyses phylogénétiques.

L'ADNmt de *P. signiensis* a aussi été envisagé pour cette étude mais en raison d'un manque de temps, l'ADNcp a été privilégié.

II. Méthodologie

II.1. Culture de *P.signiensis*

Mise en culture dans le milieu C de la souche SAG 7.90 (Sammlung Von Algenkulturen Universität Göttingen), sous alternance de 12 heures de luminosité suivi de 12 heures d'obscurité le tout à 18°C.

II.2. Extraction de l'ADN chloroplastique

II.2.1. Lyse des cellules

Chaque culture est centrifugée à 5000 rpm pendant 10 minutes pour précipiter les cellules. Celles-ci sont broyées avec de l'azote liquide, à l'aide d'un mortier refroidi à -85°C, jusqu'à l'obtention d'une fine poudre. Cette poudre est incorporée petit à petit à une solution de tampon A (25ml) (Tris-HCl 10mM pH 8, EDTA 10mM pH 8, NaCl 10mM) à 50°C, contenant 1.25% de SDS, 100µg/ml de protéinase K, puis incubée à 50°C pendant 2 heures. Le lysat obtenu est stocké à 4°C.

II.2.2. Extraction de l'ADN cellulaire par un mélange de phénol/chloroforme

Le lysat, préalablement placé deux heures à température ambiante, est ajouté à un tube de 50ml contenant un volume équivalent de phénol, et placé sous rotation pendant 15 minutes. Le mélange est ensuite centrifugé pendant 15 minutes à 6000 rpm. A l'aide d'une pipette de verre stérile, la phase supérieure est transférée dans un autre tube de 50ml contenant un volume de phénol/chloroforme : alcool isoamyl 24 : 1, et mise sous rotation pendant 30 minutes, puis centrifugée comme précédemment. La même étape est répétée à

l'aide d'un volume de chloroforme/isoamyl alcool 24 : 1. L'ADN de la phase supérieure est ensuite précipité en ajoutant 200mM NaCl (volume final) et 2 volumes d'EtOH à 95% froid. La solution, une fois mélangée, est stockée à -20°C pendant 24h. A la suite de cette incubation, la solution est centrifugée à 4000 rpm pendant 10 minutes. Le culot, une fois séché, est suspendu dans 5.5 ml de TE1X (Tris 10mM pH 8, EDTA 1mM).

II.2.3 Purification des ADNs sur gradient de chlorure de césium

Dans le but de séparer l'ADN nucléaire de l'ADN des organites, l'ADN total est soumis à une ultracentrifugation sur gradient de chlorure de césium en présence de bisbenzimidazole. Le bisbenzimidazole se lie à l'ADN et change sa densité apparente selon la conformation et/ou le contenu en A + T. Les bandes d'ADN sont visualisées sous ultraviolets.

Pour cela, 2.5 ml de la solution d'ADN précédente subit une ultracentrifugation dans un gradient de CsCl d'une densité de 1.67 g/ml en présence de 100 µl de bisbenzimidazole à 10 mg/ml. L'ultracentrifugation est de 50000 rpm pendant 20 heures à 20°C à l'aide de l'ultracentrifugeuse Beckman L8-70M avec le rotor NVT 65.2.

Afin de débarrasser l'ADN purifié des traces de CsCl, l'ADN récupéré est précipité par ajout d'un volume de TE1X, et par centrifugation à 35000 rpm pendant 20 heures. A l'aide d'une lampe UV et d'une pipette de 10 ml, la solution saline est éliminée. L'ADN est suspendu dans 100 µl de TE1X.

II.3. Construction d'une banque d'ADN génomique

L'ADNcp a été fragmenté pour que la taille des fragments soit compatible avec le vecteur plasmidique choisi : pSMART HCKan de la compagnie LUCIGEN. Une fois inséré dans les vecteurs de clonage, les vecteurs recombinants sont intégrés dans des cellules compétentes d'*E. coli* afin de les multiplier.

II.3.1 Nébulisation

L'ADN_{ncp} est nébulisé pour obtenir des fragments de 1 ou 2 kb qui seront par la suite insérés dans les vecteurs. Ainsi, 2µg d'ADN est ajouté à 300µl de glycérol 50% stérile et complété jusqu'à 600µl avec du TE1X. Le tout, passé à l'azote à 5PSI pendant 110 secondes, est centrifugé 15 secondes à 300g dans la centrifugeuse eppendorf 5810R. L'ADN nébulisé est séparé en deux tubes dont chacun contient 200mM NaCl (concentration finale) + 2 volumes EtOH 95% + 1µl glycogène (20µg/µl). L'ensemble précipite pendant 12 heures à -20°C, puis est centrifugé à 12000 rpm, pendant 45 minutes à 4°C. Une fois le surnageant éliminé et les 500µl d'EtOH à 80% ajouté, le mélange est centrifugé 10 minutes dans les mêmes conditions que précédemment. Le surnageant est éliminé et l'ADN mis à sécher.

II.3.2. Réparation des fragments et électrophorèse

Afin de ligaturer les fragments obtenus pas nébulisation avec le vecteur, les extrémités de ces ADN doivent être réparées. Pour cela, le kit DNATerminator®End Repair Reaction de la compagnie LUCIGEN est utilisé. Par la suite, les fragments d'ADN présentant une longueur comprise entre 1kb et 2 kb sont visualisés par électrophorèse sur gel d'agarose contenant 1,25 µl de Bromure d'Ethidium (10mg/mL) à l'aide d'une lampe UV. Les blocs d'agarose des bandes recherchées sont ensuite électro-élués.

II.3.3. Electro-élution

Une fois la cellule à électro-élution rendue étanche à l'aide de membranes, le bloc d'agarose est déposé dans la chambre de la cellule 1. Les deux chambres de la cellule sont ensuite remplies de tampon TAE1X, puis l'ADN est soumis à un champ électrique de 120 volts pendant 7 heures. Les liquides des deux chambres sont récupérés ainsi que la membrane de la chambre 2. Cette dernière est lavée à l'aide de 150µl de tampon TE dilué 100 fois et d'une centrifugation. L'ADN récupéré est précipité par ajout de 200mM de NaCl et de 2 volumes d'EtOH 95% froid, puis incubé 12 heures à -20°C. Le mélange est

centrifugé à 12000 rpm pendant 30 minutes à 4°C, le surnageant est éliminé et le culot rincé avec 500µl d'EtOH 80% froid. Le tout est centrifugé à 12000 rpm pendant 10 minutes, le culot est séché puis resuspendu à l'aide de 18µl d'H₂O bidistillée stérile et de 2µl de TE1X, après avoir éliminé le maximum d'éthanol.

II.3.4. Electrophorèse sur gel d'agarose 0.8%

Afin de visualiser les différents échantillons d'ADN récupérés après nébulisation, réparation et électro-élution, on réalise une électrophorèse sur gel d'agarose à 0.8%. Ce gel permet aussi d'estimer la concentration d'ADN électro-élué.

II.3.5. Clonage

II.3.5.1. Ligature

Le vecteur pSMART HCKan, préalablement linéarisé et déphosphorylé, peut au niveau de son site de polyclonage, ligaturer un fragment d'ADN, et ce grâce au mélange suivant : 3.25µl (~25ng/µl) *P. signiensis* DNA, 1.25µl 4X pSMART mix et 0.5µl T4 pSMART Ligase. Ce mélange réactionnel est incubé 2 heures à température ambiante et 15 minutes à 70°C afin de bloquer l'action de la ligase. Après avoir été déposé sur la glace, l'ensemble est centrifugé une minute à 12000rpm.

II.3.5.2. Transformations des cellules compétentes

Afin d'amplifier les inserts, les vecteurs recombinants sont intégrés au sein de cellules compétentes d'*E. coli* XL10-Gold de STRATAGEN (La Jolla, CA) par choc thermique. Lorsqu'un vecteur recombinant est intégré dans une nouvelle cellule, ce dernier est capable de se répliquer de façon autonome et ainsi de répliquer le fragment d'ADN inséré. Pour cela, des cellules compétentes d'*E. coli* sont transformées selon le protocole XL10-GOLD®Ultracompetent cells de la compagnie STRATAGENE. On ajoute 100µl de

cellules compétentes à 4µl de β-mercaptoéthanol, que l'on incube pendant 10 minutes dans de la glace. 2µl de produit de ligature (soit entre 0.1 et 50 ng d'ADN) sont ajoutés et mis à incuber pendant 30 minutes dans la glace. Du milieu NZY + broth pré-incubé à 42°C, 0.9ml est prélevé et ajouté aux cellules. Le mélange est incubé à 37°C pendant 45 minutes sous agitation et complété à 3ml avec le milieu NZY.

II.3.5.3. Mise en culture

Afin d'amplifier les vecteurs, 100µl de cellules compétentes sont étalées sur des milieux LB contenant de la Kanamycine (50µg/ml). Les boîtes de pétriensemencées sont incubées à 37°C pendant 12 à 14 heures, puis stockées à 4°C. L'ajout de Kanamycine au milieu LB, permet de sélectionner les bactéries ayant intégrées les vecteurs puisque ces derniers portent le gène de résistance à cet antibiotique.

II.3.6. Sélection des clones par hybridation

Le principe de l'hybridation moléculaire est de repérer ou de rechercher une molécule d'ADN spécifique dans une population hétérogène. Ainsi en marquant à l'aide d'un traceur radioactif la séquence d'ADN voulue et purifiée, on peut la repérer dans une population. Grâce à cette technique, il est possible de sélectionner les clones contenant les vecteurs recombinants intégrant les différents fragments d'ADN de *P. signiensis*.

II.3.6.1. Repiquage des clones

Chaque colonie est repiquée dans 100µl de milieu LB + Kanamycine (50µg/ml) puis incubée à 37°C toute une nuit.

II.3.6.2. Transfert des clones recombinants sur membranes de nylon

10µl de la culture de chaque clone est transféré sur une membrane Hybond-N. Le reste des cultures est additionné de 30µl de glycérol 50% et stocké à -80°C.

II.3.6.3. Traitement des membranes et préhybridation

Les membranes sèches sont traitées selon les recommandations d'Amersham Pharmacia-Biotech, soit dans des bains de SDS 10% ; NaCl 1.5M, NaOH 0.5M ; NaCl 1.5M, Tris HCl 0.5N pH 7.5 ; SSC2X, dans le but de lyser les cellules. Chaque membrane est exposée à la lumière UV pendant 2 minutes 30 afin de fixer l'ADN à celles-ci. Ces membranes sont préhybridées par une solution de préhybridation contenant du SSC6X, de la solution de Denhart 2X (BSA 0.2%, PVP0.2%, Ficoll 0.2%), d'EDTA 1mM, de SDS 0.1%, d'ADN de sperme de saumon dénaturé à 100µg/ml. La préhybridation a pour but d'améliorer les conditions d'hybridation afin d'empêcher les hybridations non-spécifiques.

II.3.6.4. Préparation de la sonde

La sonde est réalisée à l'aide du kit « NEBlot » de NEB (n°1550-25) et de l'ADNcp entier de *P. signiensis*. Cette sonde est préparée en effectuant six recettes, chacune avec 50µCi d' $\alpha^{32}\text{P}$ -dATP. Les désoxynucléotides non incorporés sont éliminés par passage sur une colonne Qiagen (QIAquick Nucleotide Removal kit, n°28304).

II.3.6.5. Hybridation et lavage des membranes

Les membranes sont hybridées par la solution d'hybridation et le mélange de sonde, et ce durant 13 heures à 65°C. La solution d'hybridation est de même composition que la solution de préhybridation. Chaque filtre est ensuite lavé par 2 bains de 20 min à température ambiante dans SSC 2x et SDS 0.1%, et 3 bains de 30 min à 50°C dans SSC 0.1x et SDS 0.1%.

II.3.6.6. Radiographie des membranes

Les membranes sont exposées à -80°C contre un film BIOMAX-MR de la compagnie Kodak et contre un écran intensifiant durant quelques jours.

II.4. Classification des clones recombinants

A partir des films radiographiques, les clones contenant les fragments d'ADNcp de *P. signiensis* sont identifiés en fonction de leur intensité radioactive. Afin de certifier la présence et la longueur de ces inserts, les plasmides sont purifiés, dosés, digérés avec l'enzyme de restriction EcoR1 et soumis à une électrophorèse sur gel d'agarose.

II.4.1. Purification des plasmides

Chaque colonie présentant une intensité radioactive est repiquée dans 1.5ml de milieu CircleGrow + Kanamycine (50µg/ml), puis incubée durant 24 heures à 37°C sous agitation. Les culots de cellules sont traités selon la méthode QIAprep 96 Turbo de QIAGEN, qui consiste à suspendre les culots à l'aide du tampon P1 et à lyser les cellules par le tampon P2. Après avoir neutralisé la solution de lyse P2 par le tampon N3, l'ADNcp retenu sur la colonne est lavé plusieurs fois par les tampon PB et PE. Les solutions plasmidiques sont chacune éluées par 125µl de tampon EB.

II.4.2. Dosage des plasmides

Chaque solution plasmidique est dosée par spectrophotométrie entre 260 et 280 nm contre un blanc de solution EB dans l'appareil Multiskan Spectrum. A partir des concentrations obtenues, ces solutions sont diluées ou non afin d'obtenir une concentration finale de 50µg/ml.

II.4.3. Digestion des plasmides et électrophorèse

Chaque plasmide est digéré par l'enzyme de restriction EcoR1 par le biais du mélange réactionnel suivant : 2µl d'ADN plasmidique (50µg/ml), 1µl de tampon 10XR, 1µl d'EcoR1 et 6µl d'eau bidistillée stérile. Une fois mélangé, l'ensemble est incubé 70 minutes à 37°C avant d'être additionné de 2.5 µl de Stop Mix et chauffé 5 minutes à 65°C. Chaque échantillon est déposé sur gel d'agarose à 0.8%. Il en est de même pour les marqueurs de taille λφχ et de quantité d'ADN. Les produits de digestion et les marqueurs sont soumis à un champ électrique de 100 volts jusqu'à ce qu'ils migrent au 4/5 du gel, avant d'être photographiés grâce à une lampe UV.

II.5. Séquençage et stratégie

Afin de caractériser la totalité du génome de *P. signiensis*, le séquençage aléatoire appelé « shotgun » est utilisé. Le shotgun consiste à séquencer des clones choisis selon leur longueur et leur intensité radioactive, dans le but d'assembler leurs séquences par recouvrement. Les séquences sont assemblées selon le chevauchement de leurs bases nucléotidiques par le biais du logiciel informatique, SEQUENCHER 4.2.5 (Gene Codes Corporation, Ann Arbor, MI), qui permet la formation d'îlots. Dans un second temps, les îlots sont reliés entre eux par amplification et séquençage de leurs extrémités.

Le séquençage, basé sur la méthode Sanger, repose sur l'utilisation de nucléotides particuliers appelés didésoxynucléotide qui bloquent la synthèse d'ADN par les ADN polymérases après leur incorporation. Ce blocage est dû à l'impossibilité des nucléotides à former une liaison phosphodiester avec un autre nucléotide en raison de l'absence du groupement hydroxyle sur leur carbone 3'. Pour connaître la composition d'un fragment d'ADN, on synthétise donc le brin d'ADN par une ADN polymérase. Cette enzyme commence son travail à partir de l'extrémité 3' d'une amorce, en incorporant des nucléotides complémentaires à ceux du brin qu'elle copie. Cette enzyme a pour substrats des désoxynucléotides triphosphates normaux mélangés à des didésoxynucléotides. Ceux-ci sont, par leur fonction, incorporés en dernier et marqués spécifiquement par des

molécules fluorescentes. Les fragments synthétisés sont ensuite séparés par champ électrique en fonction de leur longueur. Sachant que les électrophorèses permettent de séparer deux fragments ayant simplement une différence de taille d'un seul nucléotide, il est possible de reconstituer, à l'aide de ces tailles et de la fluorescence des fragments, leur composition en nucléotides.

Les fragments sont séquencés à l'aide de l'amorce SR2. L'amorce SL1 située de l'autre extrémité du site de clonage, a été utilisée pour prolonger les séquences utiles à l'assemblage du génome. Chaque séquence obtenue est comparée avec les banques de données protéiques et nucléotidiques de GenBank, par le biais du logiciel BLAST 2.1 de NCBI (<http://www.ncbi.nlm.nih.gov/>). Ce logiciel permet de connaître les gènes composant les séquences étudiées selon leur degré de similitude avec des séquences connues. Les séquences de références peuvent être soit chloroplastiques, soit mitochondriales, soit nucléaires ou encore d'origine inconnue (ex : contamination) permettant ainsi l'étiquetage des séquences recherchées. Une fois toutes les informations reportées, les séquences sont assemblées et éditées par le logiciel.

Ces îlots assemblés sont complétés soit à l'aide de clones si cela est possible, soit à l'aide de produit de PCR. Des oligonucléotides de 22 à 25 paires de base sont créés à partir de la séquence consensus de chaque îlots. Le choix de ces oligonucléotides est basé sur leur composition en GC, sur leur incapacité à se replier sur eux-mêmes et leur position dans les îlots (les oligonucléotides présents dans un gène sont privilégiés à ceux situés une région intergénique). Afin d'éviter les mésappariements, le logiciel FINDPATTERNS de GCG version 10.3 pour UNIX (Wisconsin Package Group) vérifie que chacun des oligonucléotides soient uniques.

Dans un premier temps, les îlots présentant les mêmes gènes sont reliés entre eux. Dans un deuxième temps, les regroupements de gènes les plus conservés parmi les Trebouxiophytes sont testés afin d'ordonner les îlots et de refermer le génome. Si ces combinaisons ne suffisent pas, chaque amorce, complémentaire aux extrémités des îlots, est alors testée avec une autre, augmentant ainsi le nombre de combinaisons possibles.

Les oligonucléotides sont aussi utilisés pour rééditer des séquences ou portions de séquences ambiguës. Les produits de PCR ayant des tailles variables entre 100 pb à 10 kb,

leur séquençage est réalisé à l'aide des oligonucléotides ayant servi à leur amplification, mais aussi avec des oligonucléotides complémentaires à des séquences interne au produit de PCR.

II.5.1. Amplification par PCR

La méthode « Hot Start » a été utilisée à l'aide de la trousse GeneAmp XL PCR (ABI) pour les amplifications. Elle consiste à séparer le milieu réactionnel en deux phases par une bille de cire. La phase inférieure est composée de 7 µl d'eau bidistillée, 6 µl de tampon 3,3 XL, 1 µl de chaque dNTP 2,5ML (dATP, dGTP, dCTP, dTTP), 0,5 µl de chaque amorce 50mM et 2,2 µl Mg²⁺. Cette phase surmontée de la bille de cire est portée à 80°C pendant 5 min. La phase supérieure est par la suite ajoutée et se compose de 19 µl d'eau bidistillée, 9 µl de tampon 3,3 XL, 1 µl de l'ADN étudié à 1/100 de l'ADN utilisé pour faire la banque, et de 1µl d'ADN polymérase rTth. L'ensemble est déposé dans le thermocycleur Gene Amp PCR system (Perkin Elmer) où il suit les cycles suivants :

- 1 cycle de dénaturation : 1 minute à 94°C
- 16 cycles d'appariement et d'élongation, de respectivement 15 secondes à 94°C et de 8 minutes à 65°C
- 19 cycles d'appariement et d'élongation, de respectivement 15 secondes à 94°C et de 8 minutes à 65°C
- 1 cycle de terminaison de 10 minutes à 70°C
- incubation à 4°C

II.5.1.1. Purification des produits de PCR

Afin d'optimiser le séquençage, les produits de PCR sont débarrassés des bandes parasites. Ils sont purifiés à l'aide de la trousse de purification Wizard PCR Prep (PROMEGA), avant d'être visualisés sur gel d'agarose à l'aide d'un marqueur de taille et d'un marqueur de masse d'ADN.

II.6. Caractérisation des génomes

II.6.1. Localisation des gènes

Bien que la comparaison des séquences avec l'outil BLAST 2.1 de NCBI donne grossièrement l'emplacement des gènes, celui-ci est confirmé par le logiciel FINDORF présent sur le site du Laboratoire des docteurs Turmel et Lemieux. Ce logiciel compare la séquence nucléotidique de *P. signiensis* avec la séquence nucléotidique et protéique des organismes présent dans la base de données. Pour chaque gène trouvé, il en ressort un fichier fasta. Ce dernier par le biais du logiciel CLUSTALW, permet d'aligner la séquence peptidique de chaque gène en la comparant aux séquences peptidiques de la base de données du laboratoire. Ces alignements sont visualisés à l'aide du logiciel CLUSTALX et permettent de corriger les séquences peptidiques de certains gènes et de vérifier la présence ou non d'introns.

L'emplacement des ARNs de transfert est identifié par tRNASCAN de NCBI. Les gènes codant pour les ARNs ribosomiques ainsi que les ARNs de transfert ont été construits à l'aide des gabarits issus d'*E. coli* afin de vérifier leur emplacement ainsi que leur structure secondaire (annexes II, III et IV).

II.6.2. Assemblage des séquences nucléotidiques

Afin de circulariser l'ADNcp de *P. signiensis*, les IR ont été localisées par le logiciel BestFit de GCG puis l'IR manquant a été ajouté à la séquence par le logiciel Assemble de GCG version 10.3 pour UNIX (Wisconsin Package Group).

Ce même logiciel a permis de reconstruire l'intégralité de l'IR de *C. ellipsoidea* à partir des fichiers Fasta récupérés sous NCBI.

II.6.3. Analyses de la composition et de l'utilisation des bases

Le tableau d'utilisation des codons pour les gènes codant pour des protéines a été construit par Codonfrequency 1.0 tandis que les pourcentages en bases A+T découlent du

logiciel Composition de GCG version 10.3 pour UNIX (Wisconsin Package Group). Les portions codantes de l'ADNcp ont été mises en évidence par le logiciel Sepextract présent sur le site du Laboratoire des docteurs Turmel et Lemieux.

II.6.4. Analyses comparatives

Les différentes analyses d'ordre de gènes, de regroupement de gènes ou de répartition géniques ont été réalisées par combinaison des logiciels Geneorder, Matrice, Genecontent et goListExtract développés par Patrick Charlebois sur le site du Laboratoire des docteurs Turmel et Lemieux. Le nombre d'inversions de gènes entre les ADNcp des Chlorophytes a été dénombré par le logiciel GRIMM 1.04 (Tesler 2002).

II.6.5. Construction de cartes, graphiques et autres schémas

Les différentes cartes ont été créées à l'aide du logiciel Physmap développés par Patrick Charlebois sur le site du Laboratoire des docteurs Turmel et Lemieux ainsi que par le logiciel Adobe Illustrator CS. Ces cartes circulaires se basent sur la position angulaire de chaque gène, intron et ORFs donnés par les fichiers Genbank des génomes.

Adobe Illustrator CS a également servi à la réalisation de toutes les autres figures et schémas.

II.6.6. Les séquences répétées

Les pourcentages de répétitions sont donnés par le logiciel Repeatanalysis présent sur le site du Laboratoire des docteurs Turmel et Lemieux.

Les séquences répétées sont cartographiées à l'aide du logiciel PipMaker (Schwartz et al. 2000), qui a également servi à la comparaison des IR de *P. signiensis* et de *C. ellipsoidea*. Les répétitions sont révélées par REPuter 2.74 (Kurtz et al. 2001) avec les options -f (forward), -p (palindromic), -allmax, -l pour la longueur minimale des répétitions souhaitées. Ces répétitions sont ensuite regroupées selon leur ressemblance pour identifier les patrons les plus fréquemment retrouvés, ceci à l'aide du logiciel BioEdit.

Le nombre de copie et la position de chaque répétition dans la séquence sont obtenus par FINDPATTERNS de GCG version 10.3 pour UNIX (Wisconsin Package Group).

PALINDROME et ETANDEM du pack Emboss ont été utilisés pour mettre en évidence d'éventuelles répétitions palindromiques ou en tandem dans l'ADNcp.

III. Résultats

III.1. Généralités de l'ADNcp de *P. signiensis*

L'ADNcp de *P. signiensis* est long de 236 643 pb (fichier GenBank en annexe I). Cet ADNcp est le plus grand jamais séquencés parmi les Chlorophyta (tableau 1), devançant l'ADNcp de *Stigeoclonium helveticum*. Cette molécule circulaire comprend deux régions inversées répétées longues de 27 336 pb qui séparent une grande région simple copie (LSC) d'une courte région simple copie (SSC), figure 6. La présence de ces IRs lui confère la structure dite quadripartie propre à la majorité des ADNcp chez les algues vertes. Cet ADNcp présente un contenu en base A+T de 66,6% et un faible pourcentage de régions codantes avec 45,9%. Il contient 111 gènes ne contenant aucun intron.

III.2. Contenu génique

L'ensemble des 111 gènes de l'ADNcp de *P. signiensis* se répartissent comme suit, selon leur fonctionnalité au sein de la cellule :

- 37 gènes photosynthétiques
- 27 gènes d'expression géniques et 29 gènes codant pour des ARNt
- 3 gènes d'ARNs (présents dans les IRs, soit en double exemplaire dans le génome)
- 15 gènes de fonctions diverses dont 7 codant pour des protéines de fonctions inconnues notées *ycf*

A ceux-ci s'ajoutent 24 ORFs uniques qui codent pour plus de 100 acides aminées et n'ayant pas de fonction connue. Tous ces gènes sont localisés le long de la séquence et, à l'exception du gène *ycf47*, ont tous été trouvés dans l'ADNcp de *C. vulgaris* (tableau 2)

Le gène *ycf47* n'a été trouvé que chez l'ADNcp de l'algue verte *N. olivacea* ainsi que chez les ADNcp des algues rouges *Porphyra purpura* et *Guillardia theta* et chez certaines cyanobactéries telle *Synechocystis sp 6803*.

Tableau 1

Comparaison des principales caractéristiques des ADNcp des algues vertes de la division des Chlorophyta

Caractéristiques	MESOvirid	NEPHoliva	CHLOvulga	PABIsigni	PARIpseud	LEPTterre	OLTMvirid	PSEUakine	STIGHelve	SCENobliq	CHLAreinh
Taille (bp)											
Total	118,360	200,799	150,613	236,463	145,947	195,081	151,933	195,867	223,902	161,452	203,827
IR	6,057	46,137	- ^a	27,336	6,786	- ^a	18,510	6,039	- ^a	12,022	22,211
LSC	83,627	92,126	- ^a	141,652	115,976	- ^a	33,610	140,914	- ^a	72,440	81,307
SSC	22,619	16,399	- ^a	40,139	16,399	- ^a	81,303	42,875	- ^a	64,968	78,088
A+T (%)	69.9	57.9	68.4	66.6	68.4	72.7	59.5	68.5	71.1	73.1	65.5
Séquences codantes (%) ^b	74.5	68.7	60.9	45.9	56.4	44.5	59.2	62.3	55.8	67.2	50.1
Gènes (nb) ^c	136	127	112	111	107	106	105	105	97	96	94
Introns (nb)											
Groupe I	0	0	3	0	0	4	5	27	16	7	5
Groupe II	0	0	0	0	0	0	0	0	5	2	2

Les algues vertes sont désignées par les abréviations suivantes : *MESOvirid* pour *Mesostigma viride*, *NEPHoliva* pour *Nephroselmis olivacea*, *CHLOvulga* pour *Chlorella vulgaris*, *PABIsigni* pour *Pabia signiensis*, *PARIpseud* pour *Parietochloris pseudoalveolaris*, *LEPTterre* pour *Leptosira terrestris*, *OLTMvirid* pour *Oltmannsiellopsis viridis*, *PSEUakine* pour *Pseudoclonium akinetum*, *STIGHelve* pour *Stigeoclonium helveticum*, *SCENobliq* pour *Scenedesmus obliquus* et *CHLAreinh* pour *Chlamydomonas reinhardtii*.

^a : ne possède pas de régions inversées répétées (IR).

^b : les introns, les ORFs et les ORFs introniques ne sont pas pris en compte.

^c : les gènes présents dans les IRs ne sont comptés qu'une seule fois.

Figure 6 : Carte génétique de l'ADNcp de *P. signiensis*. Les gènes (boîtes vides) à l'extérieur de la carte sont transcrit dans le sens horaire tandis que ceux positionnés vers l'intérieur de la carte sont transcrit dans le sens anti-horaire. Les gènes codant pour les ARNt sont indiqués par la lettre du code des acides aminés à laquelle ils correspondent et sont suivis entre parenthèse de leur anticodon. La partie en gras de l'ADNcp de *P. signiensis* et les flèches délimitent les régions inversées répétées. *P. signiensis* est désignée par la même abbréviation que celle décrite dans la légende du tableau 1.

Le contenu génique de l'ADNcp de *P. signiensis* est semblable à ceux des trois autres ADNcp de Trebouxiophyceae (tableau 2). Comme chez ces derniers, les gènes suivants présents chez *M. viride* et *N. olivacea* (les Prasinophytes de référence) n'y sont pas retrouvés : *bioY*, *ftsI*, *W*, *ndhA*, *B*, *C*, *D*, *E*, *F*, *G*, *H*, *I*, *J*, *K*, *odp*, *petN*, *rpl22*, *33*, *rps15*, *16*, *ssrA*, *ycf61*, *65*, *66*, *81*, *trnA(ggc)*, *S(cga)*, *S(gga)*, *V(gac)*. Il contient cependant les gènes *accD*, *rpl12*, *trnR(ccg)* et *psaM*, *ycf20*, respectivement absents de *M. viride* et de *N. olivacea*.

Le contenu génique de l'ADNcp de *C. vulgaris* est très proche de celui de *P. signiensis* qui ne diffère que par l'absence des gènes codant pour les ARNt *S(gga)* et *L(gag)* et la présence du gène *ycf47* cité plus haut. L'ADNcp de *P. signiensis* possède 4 gènes supplémentaires par rapport à l'ADNcp de *P. pseudoalveolaris* (*cysA*, *cysT*, *ycf20* et *ycf47*) et 5 par rapport à celui de *L. terrestris* (*ccsA*, *chlI*, *ycf12*, *ycf47* et *trnT(ggu)*).

On note une diminution de nombre de gènes dans les ADNcp mais ce nombre correspond à une diminution des gènes communs. Il n'y a pas d'émergence de nouveaux gènes hormis pour le gène *ycf47*.

III.3. ARNt et utilisation de codons

L'ensemble des gènes codant pour les ARNt se répartit le long de la séquence de manière isolée ou groupée. Ces gènes présentent tous la structure secondaire standard du trèfle à quatre feuilles. Les 29 ARNt transcrits suffisent, à l'aide de Wobble et de SuperWobble, à utiliser les 61 codons du code génétique universel, (tableau 3). Seul les ARNt *I(cau)* et *R(acg)*, possèdent une base modifiée au niveau de la première base de l'anticodon, puisque l'on observe la présence d'une lysidine pour *I(cau)* et d'une inosine pour *R(acg)*.

La majorité des codons initiateurs des gènes consiste en le conventionnel codon AUG excepté pour le gène *rps9* dont la traduction s'initie par le codon AUU. Parmi les trois

Tableau 2
Contenu en gènes des ADNcp des algues vertes du sous-règne des Chlorophyta

Gènes	MESOvirid	NEPHoliva	CHLOvulga	PABIsigni	PARlpseud	LEPTterre	OLTMvirid	PSEUakine	STIGhelve	SCENobliq	CHLAreinh
<i>accD</i>	○	●	●	●	●	●	●	●	○	○	○
<i>ccsA</i>	●	●	●	●	●	○	●	●	●	●	●
<i>chlB</i>	●	●	●	●	●	●	●	○	●	●	●
<i>chlI</i>	●	●	●	●	●	○	●	●	○	○	○
<i>chlL</i>	●	●	●	●	●	●	●	○	●	●	●
<i>chlN</i>	●	●	●	●	●	●	●	○	●	●	●
<i>cysA</i>	●	●	●	●	○	●	○	○	○	○	○
<i>cysT</i>	●	●	●	●	○	●	○	○	○	○	○
<i>infA</i>	●	●	●	●	●	●	●	●	○	●	○
<i>minD</i>	●	●	●	●	●	●	●	●	○	○	○
<i>petA</i>	●	●	●	●	●	●	●	●	○	●	●
<i>psaI</i>	●	●	●	●	●	●	●	●	○	○	○
<i>psaM</i>	●	○	●	●	●	●	●	●	●	○	○
<i>rpoB</i>	●	●	●	●	● ^a	● ^a	●	●	●	●	●
<i>rpl12</i>	○	●	●	●	●	●	●	●	○	●	○
<i>rpl19</i>	●	●	●	●	●	●	●	●	○	○	○
<i>rpl32</i>	●	●	●	●	●	●	●	●	●	○	○
<i>ycf12</i>	●	●	●	●	●	○	●	●	●	●	●
<i>ycf20</i>	●	○	●	●	○	●	●	●	○	○	○
<i>ycf47</i>	○	●	○	●	○	○	○	○	○	○	○
<i>ycf62</i>	●	●	●	●	●	●	○	●	○	○	○
<i>trnI(cau)</i>	●	●	●	●	●	●	○	●	●	●	●
<i>trnL(caa)</i>	●	●	●	●	●	●	○	●	●	○	○
<i>trnL(gag)</i>	●	●	●	○	○	○	○	○	○	○	○
<i>trnR(ccg)</i>	○	●	●	●	●	●	○	●	○	○	○
<i>trnR(ccu)</i>	○	○	○	○	○	○	●	●	○	○	○
<i>trnS(gga)</i>	●	●	●	○	●	○	○	○	●	○	○
<i>trnT(ggu)</i>	●	●	●	●	●	○	○	○	○	○	○

Les algues vertes sont désignées par les mêmes abréviations décrits dans la légende du tableau 1. La présence ou l'absence des gènes est indiquée respectivement par ● ou ○. Ces génomes ont tous en commun les gènes suivants : *atpA*, *B*, *E*, *F*, *H*, *I*, *cemA*, *clpP*, *ftsH*, *petB*, *D*, *G*, *L*, *psaA*, *B*, *C*, *J*, *M*, *psbA*, *B*, *C*, *D*, *E*, *F*, *H*, *I*, *J*, *K*, *L*, *N*, *T*, *Z*, *rbcL*, *rpl2*, *5*, *14*, *16*, *20*, *23*, *36*, *rpoA*, *C1*, *C2*, *rps2*, *3*, *4*, *7*, *8*, *9*, *11*, *12*, *14*, *18*, *19*, *rrf*, *rrl*, *rrs*, *tufA*, *ycf1*, *3*, *4*, *trnA(ugc)*, *C(gca)*, *D(guc)*, *E(uuc)*, *F(gaa)*, *G(gcc)*, *G(ucc)*, *H(gug)*, *I(gau)*, *K(uuu)*, *L(uaa)*, *L(uag)*, *Me(cau)*, *Mf(cau)*, *N(guu)*, *P(ugg)*, *Q(uug)*, *R(acg)*, *R(ucu)*, *S(gcu)*, *S(uga)*, *T(ugu)*, *V(uac)*, *W(cca)*, *Y(gua)*. Seul les deux prasinophytes possèdent en plus les gènes suivants : *bioY*, *ftsI*, *W*, *ndhA*, *B*, *C*, *D*, *E*, *F*, *G*, *H*, *I*, *J*, *K*, *odp*, *petN*, *rpl22*, *33*, *rps15*, *16*, *ssrA*, *ycf61*, *65*, *66*, *81*, *trnA(ggc)*, *S(cga)*, *S(gga)*, *V(gac)*.

^a. Chez ces deux organismes le gène *rpoB* est coupé en deux fragments nommés *rpoBa* et *rpoBb*

Tableau 3**Utilisation des 29 ARNs de transfert de *P. signiensis*, codant pour 61 codons**

ARNt	Acide Aminé spécifié		Codons reconnus
<i>trnA</i> (UGC)	Alanine	A	GCA GCC GCT GCG
<i>trnC</i> (GCA)	Cystéine	C	TGC TGT
<i>trnD</i> (GUC)	Aspartate	D	GAC GAT
<i>trnE</i> (UUC)	Glutamate	E	GAA GAG
<i>trnF</i> (GAA)	Phénylalanine	F	TTC TTT
<i>trnG</i> (GCC)	Glycine	G	GGC GGT
<i>trnG</i> (UCC)	Glycine	G	GGA GGG
<i>trnH</i> (GUG)	Histidine	H	CAC CAT
<i>trnI</i> (CAU)	Isoleucine	I	ATA
<i>trnI</i> (GAU)	Isoleucine	I	ATC ATT
<i>trnK</i> (UUU)	Lysine	K	AAA AAG
<i>trnL</i> (CAA)	Leucine	L	TTG
<i>trnL</i> (UAA)	Leucine	L	TTA
<i>trnL</i> (UAG)	Leucine	L	CTA CTG CTC CTT
<i>trnMe</i> (CAU)	Méthionine	M	ATG
<i>trnMf</i> (CAU)	Méthionine	M	ATG
<i>trnN</i> (GUU)	Asparagine	N	AAC AAT
<i>trnP</i> (UGG)	proline	P	CCA CCC CCG CCT
<i>trnQ</i> (UUG)	Glutamine	Q	CAA CAG
<i>trnR</i> (ACG)	Arginine	R	CGC CGT CGA
<i>trnR</i> (CCG)	Arginine	R	CGG
<i>trnR</i> (UCU)	Arginine	R	AGA AGG
<i>trnS</i> (GCU)	Sérine	S	AGC AGT
<i>trnS</i> (UGA)	Sérine	S	TCA TCC TCT TCG
<i>trnT</i> (UGU)	Thréonine	T	ACA ACG
<i>trnT</i> (GGU)	Thréonine	T	ACC ACT
<i>trnV</i> (UAC)	Valine	V	GTA GTC GTG GTT
<i>trnW</i> (CCA)	Tryptophane	W	TGG
<i>trnY</i> (GUA)	Tyrosine	Y	TAC TAT

codons de terminaison du code génétique IUPAC, seul deux d'entre eux, TAA et TAG sont utilisés. En grande majorité (93,7%), le codon de terminaison TAA est privilégié et seulement 6,3% des gènes se terminent par TAG.

III.4. Répartition des gènes le long de la structure quadripartite

Les 111 gènes de l'ADNcp de *P. signiensis* se répartissent dans les quatre compartiments de la séquence et en majorité dans la LSC (81 gènes).

Les figures 7, 8, 9 et 10 présentent les répartitions de ces gènes dans les différents ADNcp des Trebouxiophytes en comparaison de la place qu'ils occupent au sein de l'ADNcp de *M. viride*.

Il existe peu de changement entre la répartition génique de l'ADNcp de *P. signiensis* et celle de *M. viride*. Les IRs ont gagnés les gènes *psbA* et *trnS(gcu)* de la LSC, les gènes *trnT(ugu)* et *R(acg)* sont passés respectivement des IRs à la LSC et à la SSC tandis que les gènes *trnS(uga)*, *G(gcc)*, *D(guc)*, *Me(cau)* et *psbM* ont migrés de la LSC à la SSC. Ce dernier remaniement est aussi retrouvé dans l'ADNcp de *P. pseudoalveolaris*.

Bien que l'ADNcp de *C. vulgaris* ne présente pas d'IRs, les gènes inscrits dans les différents compartiments de l'ADNcp de *M. viride* restent regroupés dans cet ADNcp, lui permettant de garder une pseudo structure « quadripartite ». *L. terrestris*, ne possédant pas non plus d'IR, ne présente pas ce type d'arrangement génique. L'ordre de gènes a été remanié de telle façon que rien ne subsiste de la structure quadripartite de *M. viride*, contrairement aux autres ADNcp des Trebouxiophytes.

Les ADNcp d'Ulvophytes comptabilisent 12 gènes communs ayant migrés de la LSC à la SSC. Chez l'ADNcp de *P. signiensis* on note que sur les 5 gènes échangés entre LSC et SSC, 4 sont communs à ceux des ADNcp d'Ulvophytes : *psbM*, *trnS(gcu)*, *D(guc)*, *Me(cau)* et *G(gcc)*. Chez les ADNcp d'algues chlorophycéennes, seul les gènes de l'opéron *rrn* sont conservés. Ces derniers présentent la même organisation que ceux de l'ADNcp de *M. viride* contrairement à ceux de *P. signiensis*.

Figure 7 : Comparaison entre la répartition génique de l'ADNcp de *P. signiensis* et celui de *M. viride*. En bleu sont représentés les gènes appartenant à la LSC de l'ADNcp de *M. viride*, en rose ceux de la SSC et en jaune ceux des IRs. Les gènes absents de l'ADNcp de *M. viride* sont matérialisés en gris. *P. signiensis* est désignée par la même abbréviatiion que celle décrite dans la légende du tableau 1.



Figure 7

Figure 8 : Comparaison entre la répartition génique de l'ADNcp de *P. pseudoalveolaris* et celui de *M. viride*. En bleu sont représentés les gènes appartenant à la LSC de l'ADNcp de *M. viride*, en rose ceux de la SSC et en jaune ceux des IRs. Les gènes absents de l'ADNcp de *M. viride* sont matérialisés en gris. *P. pseudoalveolaris* est désignée par la même abréviation que celle décrite dans la légende du tableau 1.

Figure 9 : Comparaison entre la répartition génique de l'ADNcp de *C. vulgaris* et celui de *M. viride*. En bleu sont représentés les gènes appartenant à la LSC de l'ADNcp de *M. viride*, en rose ceux de la SSC et en jaune ceux des IRs. Les gènes absents de l'ADNcp de *M. viride* sont matérialisés en gris. *C. vulgaris* est désignée par la même abréviation que celle décrite dans la légende du tableau 1.

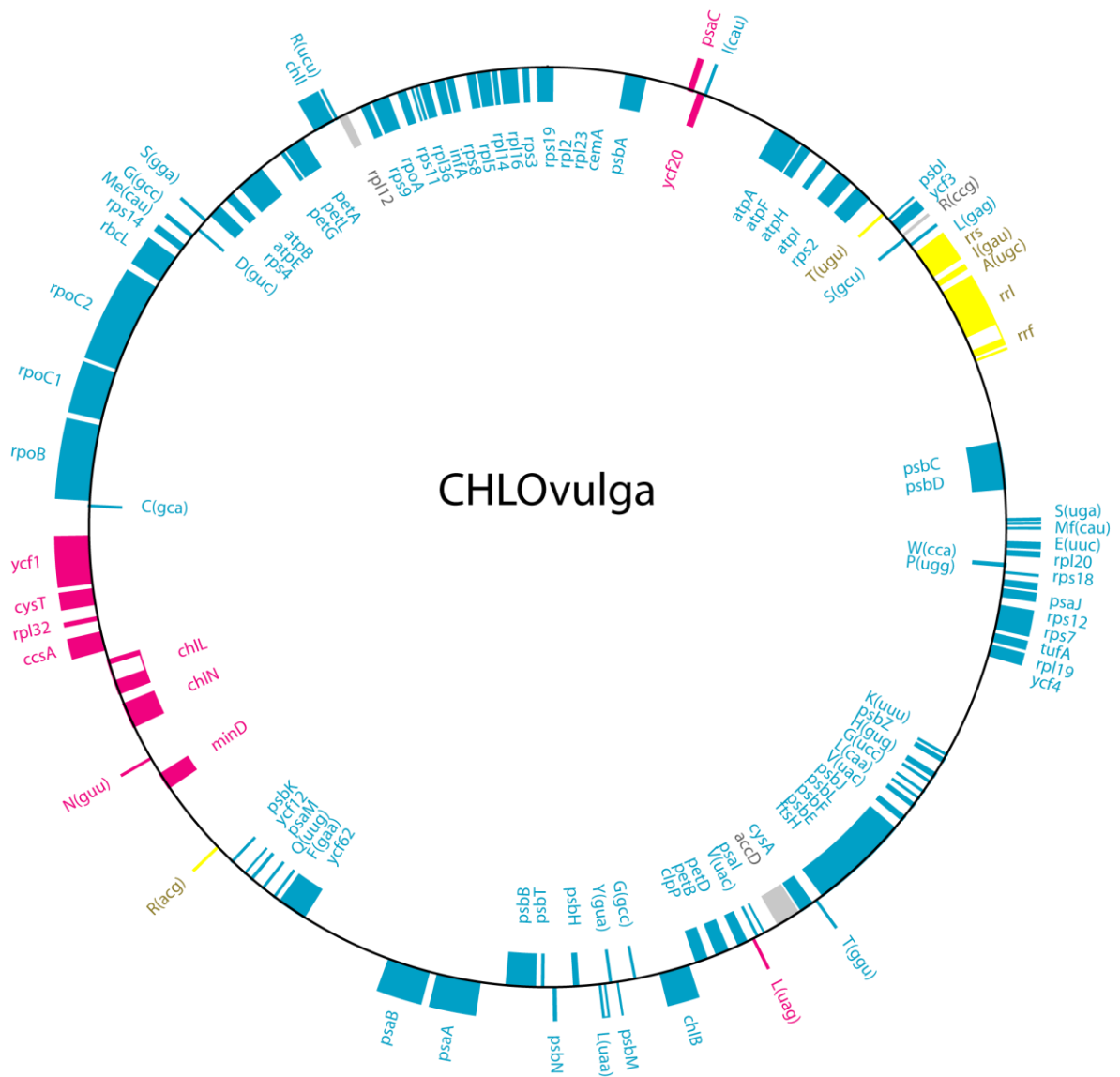


Figure 9

Figure 10 : Comparaison entre la répartition génique de l'ADNcp de *L. terrestris* et celui de *M. viride*. En bleu sont représentés les gènes appartenant à la LSC de l'ADNcp de *M. viride*, en rose ceux de la SSC et en jaune ceux des IRs. Les gènes absents de l'ADNcp de *M. viride* sont matérialisés en gris. *L. terrestris* est désignée par la même abréviation que celle décrite dans la légende du tableau 1.

III.5. Organisation génique et réarrangement de gène

On dénombre 23 regroupements de gènes identiques entre *P. signiensis* et *C. vulgaris* incluant 83 gènes, tandis qu'entre *P. signiensis* et *M. viride* on relève 19 regroupements de gènes incluant 71 gènes. Selon les résultats du tableau 4, l'organisation génique de l'ADNcp de *P. signiensis* est plus proche de celles de *M. viride* et de *N. olivacea* que des Trebouxiophytes *P. pseudoalveolaris* et *L. terrestris*.

Parmi les ADNcp des Trebouxiophytes, on dénombre 23 regroupements de gènes communs dont 83 gènes avec *C. vulgaris* mais aussi 15 regroupements communs incluant 67 gènes avec *P. pseudoalveolaris* et 17 regroupements comprenant seulement 54 gènes avec *L. terrestris*. Ces clusters conservés entre les ADNcp de Trebouxiophytes sont illustrés par la figure 11.

Les ADNcp de *M. viride* et de *N. olivacea* présentent 26 regroupements de gènes identiques. Sur ces 26 regroupements de gènes dits ancestraux, 13 sont retrouvés dans l'ADNcp de *P. signiensis* tandis que 12 sont présents chez *C. vulgaris*, 11 chez *P. pseudoalveolaris* et 10 chez *L. terrestris*.

En plus de ces regroupements ancestraux, tous les ADNcp des Trebouxiophytes présentent le regroupement de gènes : *rps18-rpl20*. Les regroupements *psbK-ycf12-psaM*, *trnL(uaa)-trnY(gua)* et *trnC(gca)-rpoB-rpoC1-rpoC2* ne sont présents que chez les ADNcp de *P. signiensis*, *C. vulgaris* et *P. pseudoalveolaris* tandis que le regroupement *trnR(acg)-minD* n'est retrouvé que dans les ADNcp de *P. signiensis*, de *P. pseudoalveolaris* et de *L. terrestris* (figure 11).

La figure 12 représente les différentes paires de gènes présents ou non dans les ADNcp de Chlorophyta. La classe des Trebouxiophytes est celle qui a retenu le plus de paires de gènes dits ancestraux, soit les paires de gènes conservés dans les ADNcp des deux Prasinophytes *M. viride* et *N. olivacea*. Avec 29 paires, la classe Trebouxiophyceae dépassent les 26 paires des Ulvophyceae et les 7 paires de gènes ancestraux des Chlorophyceae. Ils partagent aussi 3 paires de gènes avec les Ulvophytes en dehors des paires ancestrales que sont *rps19-rps3*, *rps9-pl12* et *rpl20-rps18*, et aucune avec les algues chlorophycéennes.

Tableau 4**Comparaison des différents contenus en regroupement de gènes entre les ADNcp des Trebouxiophytes et ceux des Prasinophytes**

		<i>PABIsigni</i>	<i>CHLOvulga</i>	<i>PARIpseud</i>	<i>LEPTterre</i>
<i>MESOvirid</i>	% gènes conservés	63	60	54	46
	nb regroupements	19	19	15	15
	nb gènes regroupés	71	68	59	50
<i>NEPHoliva</i>	% gènes conservés	58	57	54	47
	nb regroupements	16	16	14	15
	nb gènes regroupés	65	64	59	51
<i>PABIsigni</i>	% gènes conservés		74	60	48
	nb regroupements		23	15	17
	nb gènes regroupés		83	67	54

Les algues vertes sont désignées par les mêmes abréviations décrites dans la légende du tableau 1

Figure 11 : Comparaison de la conservation en ordre de gènes entre les différents ADNcp des Trebouxiophytes et celui de *P. signiensis*. En vert et rouge les regroupements de gènes communs. Les algues vertes sont désignées par les mêmes abréviations décrites dans la légende du tableau 1

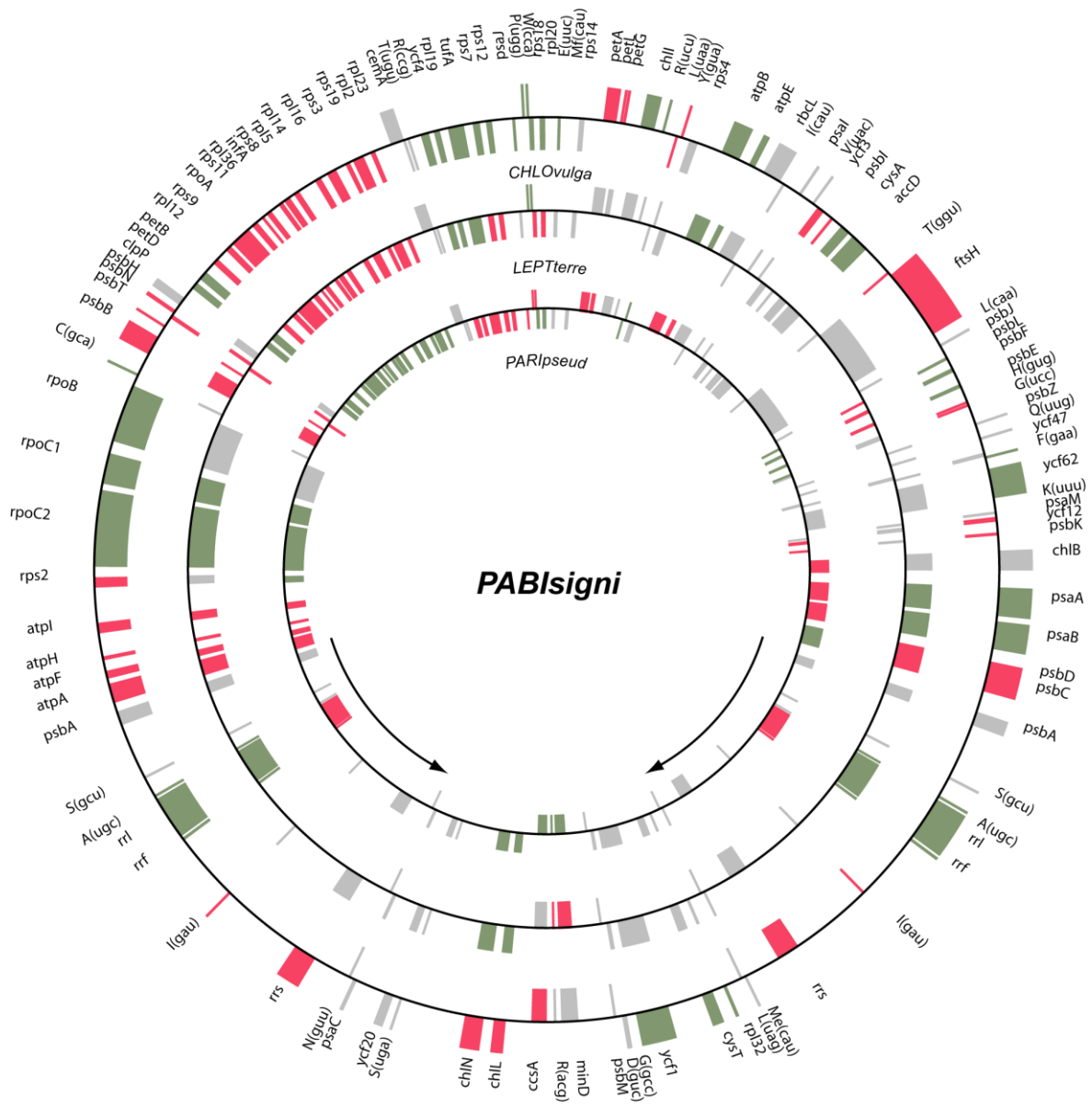


Figure 11

Figure 12 : Contenu en paires de gènes ancestrales ou dérivées entre les différents ADNcp des algues vertes du sous-règne des Chlorophyta. Les cases noires correspondent aux paires de gènes présentes dans le génome tandis que les cases blanches indiquent une perte de gènes. En gris sont matérialisés les gènes présents dans le génome mais non regroupés en de tel paires. Les algues vertes sont désignées par les mêmes abréviations décrites dans la légende du tableau 1

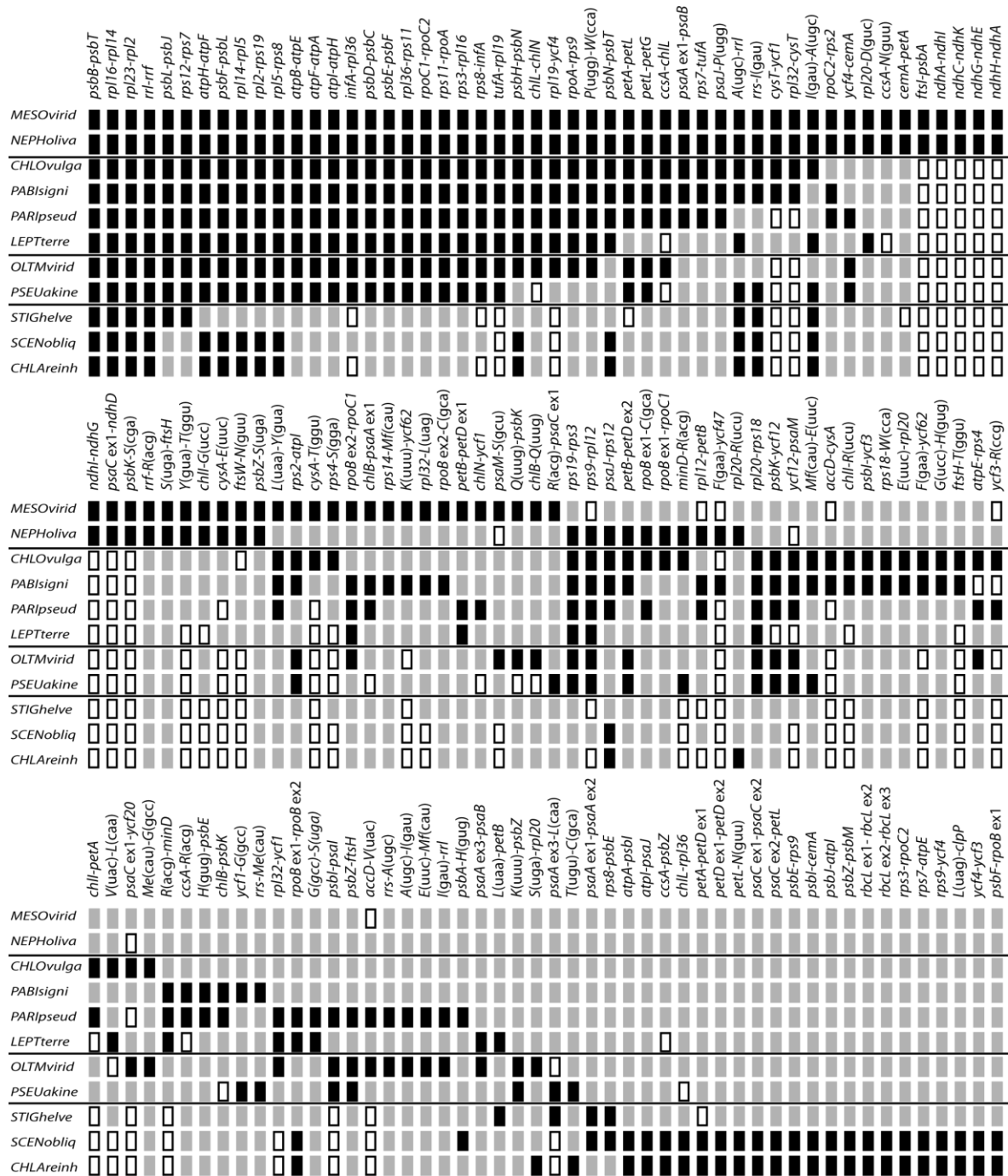


Figure 12

Cette proportion de paires de gènes communes entre les espèces et non plus entre les classes va aussi en diminuant au cours de l'évolution. *L. terrestris* qui est la Trebouxiophyte la plus divergente présente moins de paires de gènes communes (6) avec *C. vulgaris* (la Trebouxiophyte la plus ancestrale) que *P. pseudoalveolaris* (11). Toutefois au sein des Trebouxiophytes, bien que l'ADNcp de *P. signiensis* et de *C. vulgaris* comptent le même nombre de paires dites ancestrales (40), l'ADNcp de *P. signiensis* dénombre le double de paires de gènes communes (8) avec *M. viride* que l'ADNcp de *C. vulgaris* (4). En marge de ces paires de gènes ancestrales, les ADNcp des Trebouxiophytes ne présentent qu'une seule paire commune qui est *rpl20-rps18*. Celle-ci est retrouvée chez les deux Ulvophytes *P. akinetum* et *O. viridis*.

Le nombre d'inversions de gènes entre les génomes renseigne aussi sur le degré de remaniement des génomes de Chlorophyta. D'après les résultats de GRIMM, tableau 5, l'ADNcp de *P. signiensis* a subi autant d'inversion de gènes (39) que *N. olivacea* en comparaison de l'ADNcp de *M. viride*.

Au sein des Trebouxiophytes, cet ADNcp ne diffère de celui de *C. vulgaris* que par 35 inversions et par 38 de celui de *P. pseudoalveolaris*. Bien que *L. terrestris* soit aussi une Trebouxiophytes, on note plus d'inversions de gènes entre l'ADNcp de *P. signiensis* et ce dernier (50) qu'entre l'ADNcp de *P. signiensis* et ceux des deux Prasinophytes *M. viride* (39) et *N. olivacea* (43). Le nombre d'inversions de gènes avec l'ADNcp de *L. terrestris* est le même que celui obtenu avec les deux Ulvophytes (50).

En moyenne l'ADNcp de *P. signiensis* diffère par 41 inversions avec les ADNcp des Prasinophytes, 41 inversions avec les autres ADNcp de Trebouxiophytes, 50 inversions avec les ADNcp des Ulvophytes et 67 pour les ADNcp des algues chlorophycéennes.

III.6. Les séquences répétées

L'ADNcp de *P. signiensis* contient 54,1% de régions non codantes dont 30,5% sont composées de petites séquences répétées dispersées (SDRs) représentant 18,7% du génome, (tableau 6). Avec ces 39 089 pb de répétitions, l'ADNcp de *P. signiensis* est l'ADNcp des Chlorophyta le plus riche en séquences répétées si l'on se base sur la

Tableau 5**Nombre minimum d'inversions lors des réarrangements géniques entre les différents ADNcp de Chlorophyta**

ADNcp	Nombre de gènes inversés ^a									
	<i>NEPHoliva</i>	<i>CHLOvulga</i>	<i>PABIsigni</i>	<i>PARIpseud</i>	<i>LEPTterre</i>	<i>OLTMvirid</i>	<i>PSEUakine</i>	<i>STIGhelve</i>	<i>SCENobliq</i>	<i>CHLAreinh</i>
<i>MESOvirid</i>	39	42	39	46	50	49	49	70	64	65
<i>CHLAreinh</i>	65	64	65	67	64	67	67	73	51	
<i>SCENobliq</i>	66	66	66	66	67	70	67	71		
<i>STIGhelve</i>	69	69	70	72	69	70	69			
<i>PSEUakine</i>	53	48	50	54	55	50				
<i>OLTMvirid</i>	53	48	50	45	54					
<i>LEPTterre</i>	49	51	50	51						
<i>PARIpseud</i>	44	42	38							
<i>PABIsigni</i>	43	35								
<i>CHLOvulga</i>	43									

Les algues vertes sont désignées par les mêmes abréviations décrites dans la légende du tableau 1

^a les nombres de gènes inversés ont été dénombrés avec GRIMM

Tableau 6
Abondance des SDRs au sein des ADNcp des algues vertes du sous-règne des Chlorophyta

ADNcp	Taille maximale des répétitions (pb)	Nombre de répétitions ^a		Répétitions non-couvrantes ≥ 30 pb ^b		
		≥ 30 pb	≥ 45 pb	Taille totale (pb)	Fraction dans le génome (%)	Fraction dans les régions intergéniques (%)
<i>MESOvirid</i>	55	5	1	217	0.2	0.7
<i>NEPHoliva</i>	60	15	6	1,061	0.5	1.7
<i>CHLOvulga</i>	84	269	44	11,743	7.8	20.8
<i>PABIsigni</i>	357	2,235	828	39,089	18,7	30.5
<i>PARIpseud</i>	145	432	164	10,120	6.9	15.9
<i>LEPTerre</i>	194	407	50	9,558	4.9	8.8
<i>OLTMvirid</i>	172	1,205	161	18,033	11.9	30.1
<i>PSEUakine</i>	171	1,047	203	10,073	5.1	13.6
<i>STIGhelve</i>	625	2,856	640	39,941	17.8	40.3
<i>SCENobliq</i>	112	86	21	4,817	3.0	8.7
<i>CHLAreinh</i>	221	3,247	551	32,244	15.8	31.9

^a. Le nombre de répétitions de séquences identifiées a été estimé en utilisant REPuter.

^b. Les répétitions non-couvrantes ont été mises en évidence par RepeatMasker.

séquence entière et le troisième derrière *S. helveticum* et *C. reinhardtii* si l'on se base sur les portions non-codantes.

La majorité des unités de répétitions ont une longueur inférieure à 30 pb. Dans le tableau 7 sont répertoriés les unités les plus fréquemment retrouvés ayant une taille supérieure à 15 pb. Certaines répétitions forment des structures en épingle à cheveux par leur séquence palindromique. Très peu de répétitions en tandem ont été mis en évidence par ETANDEM, ne permettant pas de conclure quand à une présence significative de tels enchaînements dans l'ADNcp de *P. signiensis*. De très nombreuses unités de moins de 10 pb ont été fréquemment repérés. Les séquences adjacentes de ces séquences n'ont pas permis de mettre en avant de grandes familles d'unités de répétitions. Dans le cas du patron TAAAAC, l'ajout de séquences adjacentes a fortement diminué sa fréquence dans l'ADNcp. Exemple : AAGCGTTAAAACTTATGTTTA, qui contient le patron, n'est plus retrouvé que 28 fois alors que la séquence de 6 pb est présente près de 700 fois. Il en est de même avec les autres séquences issues du même patron :

- TTGGGTTAAAACTTTATTAAAA retrouvée 5 fois
- TAAAACATTTAGG retrouvée 33 fois
- TAAAACTTGA retrouvée 38 fois
- TAAAACAAA retrouvée 55 fois

Pourtant ces SDRs de moins de 10 pb sont très fréquentes dans l'ADNcp de *P. signiensis*, AAACAAA retrouvée 359 fois, TTTTTTAAAA 129 fois, TTGCCG 70 fois, CTTTTTTTC 52 fois ou CTCCCT 46 fois.

Bien que majoritairement présentes dans les séquences intergéniques, certaines unités de répétitions ont toutefois été mises en évidence dans les extensions nucléotidiques des gènes d'ARNr (ITS voir annexe) ou dans des gènes peu conservés tel *ftsH*.

La figure 13 présente les résultats de l'analyse PIPMAKER permettant de visualiser les répétitions selon leur emplacement dans le génome. Les unités se retrouvent tout au long de la séquence et principalement dans les IRs. Ces derniers comptabilisent 44% des répétitions de l'ADNcp de *P. signiensis*.

Tableau 7**Unités de répétitions retrouvées dans l'ADNcp de *P. signiensis***

Désignation	Taille (pb)	Séquence ^a	Nombre de copies ^b	
			100%	90%
A	18	<u>GTGTGTGC</u> -(W) ₂ -GCACACAC	28 (63)	57 (198)
B	18	<u>ATGTGTGC</u> -(W) ₂ -GCACACAT	24 (50)	54 (149)
C	25	<u>AGTATACAAAC</u> -(N) ₃ -GTTGTATACT	8 (16)	8 (19)
D	23	<u>ATTTAACAAC</u> -(N) ₃ -GTTGTAAAT	31 (77)	48 (159°)
E	27	<u>CCCTTGCCGT</u> -(N) ₃ -CGGCAGGGCAACGG	53	59

^a Les séquences soulignées correspondent demi branches des structures en épingles à cheveux pouvant être retrouvés seules dans le génome

^b Le nombre de copies de chaque répétition a été mis en évidence par le logiciel FINDPATTERNS avec une similitude de 100% et de 90%. Les nombres entre parenthèses correspondent aux copies des demi branches d'épingles à cheveux présents dans l'ADNcp de *P. signiensis*

Figure 13 : Analyse PIPMAKER de l'ADNcp de *P. signiensis*. La séquence du génome a été alignée contre elle-même. Les gènes sont représentés par des boîtes noires, lesquelles sont surmontés de flèches indiquant leur polarité. Chaque point correspond à une séquence répétée dans le génome pour laquelle l'homologie trouvée dépasse 50%. En lettres colorées sont représentées les séquences répétées directes de plus de 15 pb les plus fréquemment trouvées dans l'ADNcp de *P. signiensis*. Leur signification est donnée dans le tableau 7.

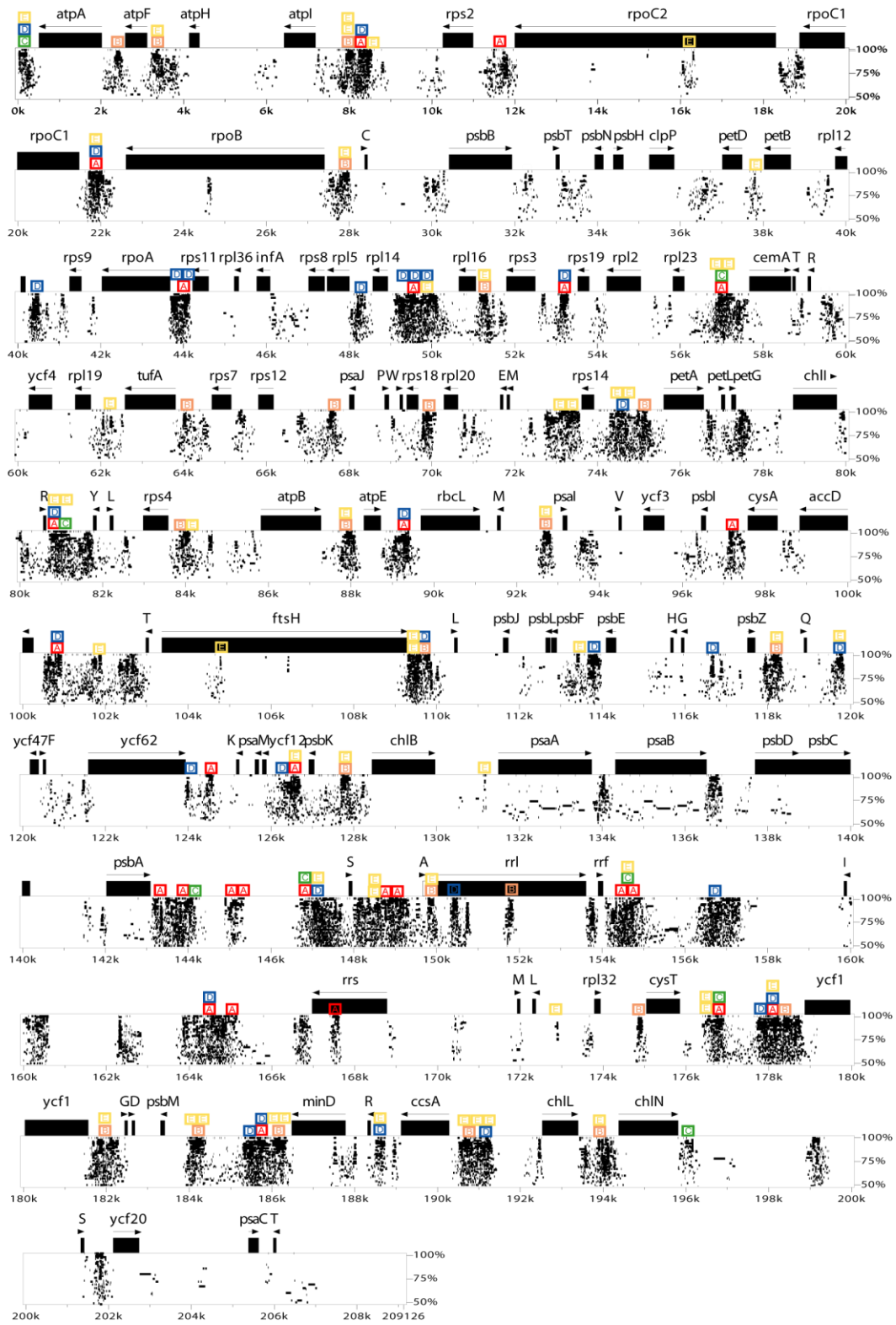


Figure 13

L'IR de l'ADNcp de *P. signiensis* étant très similaire à celui de *C. ellipsoidea* (voir paragraphe plus bas), ils ont été comparés dans le but de mettre en évidence leurs répétitions communes, (figure 14). L'IR de *C. ellipsoidea* possède 16% de répétitions, soit moitié moins que celui de *P. signiensis* qui comptabilise 31,9% pour cette portion de génome. Aucune des unités de répétitions répertoriées dans le tableau 7 n'ont été retrouvée dans l'IR de *C. ellipsoidea* contrairement aux petites unités de moins de 10 pb :

- TAAAAC présent 87 fois dans l'IR de *P. signiensis*, est dénombré 57 dans celui de *C. ellipsoidea*
- AAACAAA présent 47 fois dans l'IR de *P. signiensis*, est retrouvé 13 fois dans celui de *C. ellipsoidea*
- ou encore CTTTTTTC trouvé 13 fois dans l'IR de *P. signiensis* et présent 6 fois dans celui de *C. ellipsoidea*

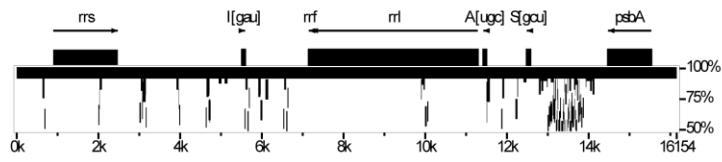
III.7. Les régions inversées répétées

Comme le montre le tableau 1, la présence et la taille des IRs varient entre les classes et les espèces, tout comme leur contenu génique.

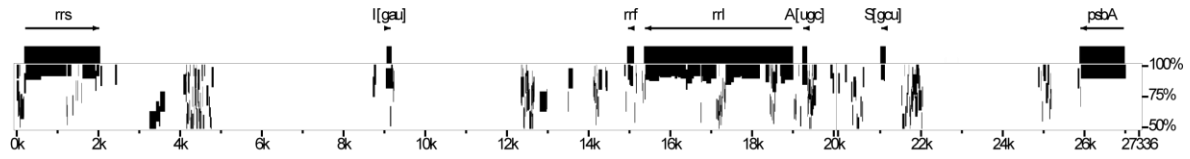
Les deux IRs de l'ADNcp de *P. signiensis* totalisent chacune 27 336 pb et contiennent l'ensemble des gènes composant l'opéron *rrn* (figure 6) soit les gènes *rrs*, *rrf*, *rpl* et les gènes d'ARNt *A(ugc)* et *I(gau)*. Aucun de ces gènes ne présentent d'intron contrairement à l'ADNcp de *C. vulgaris*. A ceux-ci s'ajoutent les gènes *psbA* et *trnS(gcu)*. Bien que la présence d'IRs dans l'ADNcp de *C. ellipsoidea* n'ait pas encore été démontrée, le contenu génique des IRs de *P. signiensis* est semblable à celui de l'opéron *rrn* de *C. ellipsoidea* et des gènes qui bordent ce dernier. *L. terrestris*, *C. ellipsoidea* et *P. signiensis* ont la particularité de présenter un opéron *rrn* brisé en deux fragments. Dans les IRs de *C. reinhardtii* et d'*O. viridis* est aussi présent le gène *psbA*, tout comme le gène *trnS(gcu)* l'est dans l'IR de *S. obliquus*.

Figure 14 : Analyse PIPMAKER des IRs de *P. signiensis* et de *C. ellipsoidea*. Dans un premier temps les IRs de chaque organisme a été comparés contre eux-même, puis l'un contre l'autre afin de mettre en évidence les séquences communes. Les algues vertes sont désignées par les mêmes abréviations décrivent dans la légende du tableau 1

CHLOellip/CHLOellip



CHLOellip/PABIsigni



PABIsigni/PABIsigni

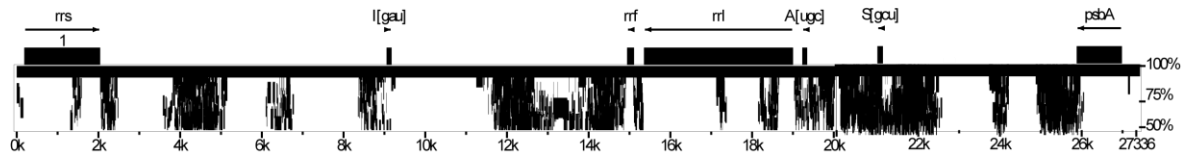


Figure 14

La grande taille des IRs, 27 336 pb pour l'ADNcp de *P. signiensis* contre 6 786 pb pour l'ADNcp de *P. pseudoalveolaris*, est due aux régions intergéniques riches en SDRs et à l'extension des gènes *rrs* et *rfl*. L'espace intergénique entre *rrs* et *rfl* chez *C. vulgaris* est de 1000 pb alors que celui de l'ADNcp de *P. signiensis* approche les 13,5 kb. Sans ces IRs, l'ADNcp de *P. signiensis* ne serait que le deuxième plus grand génome des Chlorophyta derrière *S. helveticum* qui n'en possède pas.

III.8. L'opéron *rrn*

L'opéron *rrn* ancestral contenu dans les ADNcp de *M. viride* et *N. olivacea*, présente la même répartition génique que celui de l'ADNcp de *C. vulgaris* (figure 15), soit *rrs-trnI(gau)-trnA(ugc)-rfl-rrf*.

Dans l'ADNcp de *P. signiensis*, tout comme dans celui de *C. ellipsoidea*, cet opéron est fragmenté en deux morceaux inversés l'un par rapport à l'autre. Le premier fragment se compose des gènes *rrs-trnI(gau)*, et le second de *rrf-rfl-trnA(ugc)* et sont tous les deux transcrits en sens opposé. La fracture se situe entre les gènes *trnI(gau)* et *trnA(ugc)* où l'on retrouve d'autre part une grande concentration d'unités de répétitions (figure 14).

L'opéron *rrn* de *P. pseudoalveolaris* présente une inversion entre les deux ARNt. Celui de *L. terrestris* est fragmenté en deux morceaux espacés l'un de l'autre par près de 60 000 pb. La fracture se situe entre les gènes *rrs* et *trnI(gau)*. Bien que *L. terrestris* ne possède pas d'IRs, on note toutefois la présence du gène *psbA* à proximité du fragment contenant le gène *rrs*.

Comme cité ci-dessus, les régions intergéniques et les gènes composant l'opéron *rrn* de *P. signiensis* sont plus grands que ceux retrouvés chez les ADNcp des autres Trebouxiophyceae. Dans le gène *rfl*, trois extensions sont présentes et nommées ITS $\alpha\beta$, ITS $\beta\gamma$ et ITS $\gamma\delta$ (annexe IV). Elles sont respectivement retrouvées entre les sous unités α/β , β/γ et γ/δ du gène *rfl*. Chez les autres ADNcp de Trebouxiophytes, *C. ellipsoidea* et *P. pseudoalveolaris* présentent aussi des extensions dans le gène *rfl*. Ces dernières ne sont présentes qu'entre les sous-unités γ/δ . Bien que la structure de ces ITS $\gamma\delta$ diffèrent (figure 16), ces extensions sont toutes riches en base A+T (plus de 55% chaque). Les ITS $\gamma\delta$ contenu dans les ADNcp de *P. signiensis* de *C. ellipsoidea* et

de *P. pseudoalveolaris* sont respectivement longues de 366 pb, 255 pb et 134 pb. Seul le gène *rrl* de *P. signiensis* présente les deux autres extensions (ITS $\alpha\beta$ et ITS $\beta\gamma$) longues de 268 pb et de 138 pb, et elles aussi riches en bases A+T (56,3% et 55,7%).

Autre particularité des gènes de l'opéron *rrn* : la taille du gène *rrs* (annexe II) qui présente une autre extension au niveau d'une zone variable en nucléotides. Cette extension est longue de 282 pb et riche de 58,5% en bases A+T.

Figure 15 : Comparaison des différents opérons *rrn* au sein des ADNcp des Trebouxiophytes. Lorsque ces opérons sont présents dans les IRs, l'ensemble du contenu génique des IRs ont été rajoutés. Il en est de même pour l'opéron de *C. ellipsoidea* bien que la présence d'IRs dans son ADNcp n'a pas été confirmée. L'opéron *rrn* de *C. vulgaris* est identique en composition et en répartition génique à ceux des ADNcp ancestraux de *M. viride* et *N. olivacea*. Les algues vertes sont désignées par les mêmes abréviations décritent dans la légende du tableau 1

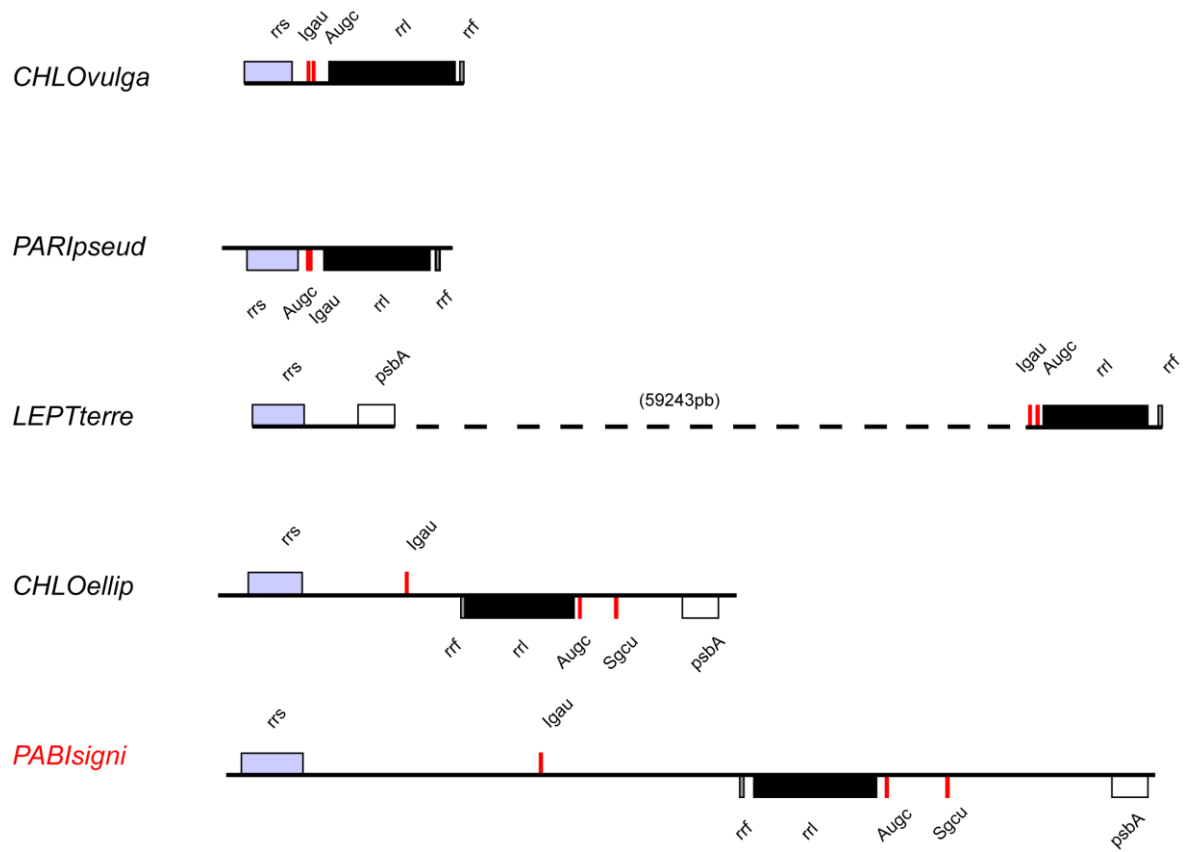


Figure 15

Figure 16 : Les différentes structures secondaires des ITS $\gamma\delta$ de l'ADNcp de *C. ellipsoidea*, *P. pseudoalveolaris* et de *P. signiensis*. Les algues vertes sont désignées par les mêmes abréviations décritent dans la légende du tableau 1

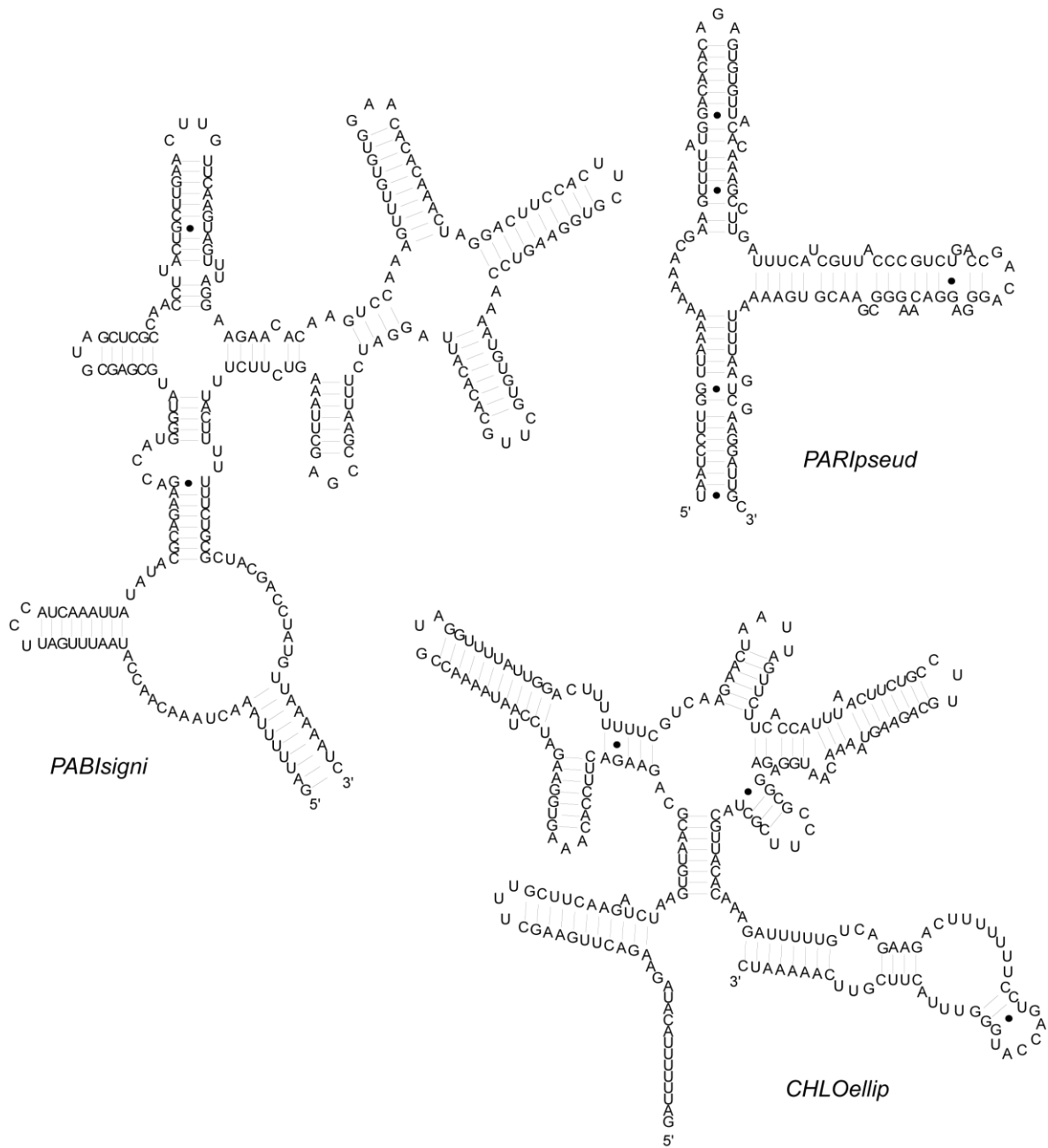


Figure 16

IV. Discussion

IV.1. L'ADNcp de *P. signiensis* : un génome très conservé

Comme la majorité des ADNcp d'algues vertes (Palmer 1985), l'ADNcp de *P. signiensis* possède une structure quadripartite répartie en deux IRs, en une LSC et en une SSC.

Dans cet ADNcp, les proportions entre les différentes régions ont été respectées contrairement à celles observées dans les ADNcp de *C. reinhardtii* (Maul et al. 2002) et *S. obliquus* (résultats non publiés) chez lesquelles LSC et SSC sont approximativement de même taille. Et contrairement aussi à l'ADNcp d'*O. viridis* (Pombert, Lemieux, and Turmel 2006) qui possède une SSC plus grande que sa LSC.

Cette structure quadripartite n'est retrouvée chez les Trebouxiophytes que chez *P. pseudoalveolaris* (résultats non publiés). Elle est toutefois présente dans l'ADNcp de *C. ellispodea* (Wakasugi et al. 1997) bien que la séquence entière n'ait pas encore été déterminée. Ces deux organismes présentent des IRs dont la taille est cependant plus restreinte que celles des IRs de *P. signiensis*.

Tout comme *P. pseudoalveolaris*, l'ADNcp de *P. signiensis* ne présente aucun intron. La présence de tels éléments dans les ADNcp fait généralement l'objet de caractère évolutif. Son absence appuie le degré de conservation de ce génome.

Autre caractère conservé partagé entre *P. signiensis* et *P. pseudoalveolaris* : la répartition des gènes le long de la séquence. L'ADNcp de *P. signiensis* a gardé une distribution génique proche de celle de *M. viride* (Lemieux, Otis, and Turmel 2000). En dehors des IRs qui sont connus pour leur variabilité en gènes (Simpson and Stern 2002) (gain ou perte de gènes issus des régions voisines (Yamada 1991; Sugiura, Hirose, and Sugita 1998)) et leur processus "Ebb and Flow" (Goulding et al. 1996), la LSC et la SSC ont été très peu remaniées. Ces régions contiennent la plupart des gènes retrouvés dans les LSC et SSC de *M. viride*.

Quelques gènes ont toutefois été échangés entre ces deux régions : *psbM*, *trnS(gcu)*, *D(guc)*, *Me(cau)* et *G(gcc)*. Ces gènes appartiennent aux mêmes changements de LSC et SSC observés chez *P. pseudoalveolaris*. Ces remaniements ont donc des origines communes. Dans les ADNcp de *O. viridis* (Pombert, Lemieux, and Turmel 2006) et de *P. akinetum* (Pombert et al. 2005), 12 gènes communs ont migrés de la LSC vers la SSC et parmi ces 12 gènes quatre sont communs à ceux de *P. signiensis*. Les remaniements géniques relevés chez les autres membres du clade UTC semblent avoir été initiés au sein de la classe des Trebouxiophytes.

En dehors de ces quelques mouvements de gènes, on ne relève que peu de critère de remaniement. 73% des regroupements de gènes de *P. signiensis* sont communs à ceux de la Trebouxiophytes la plus basale : *C. vulgaris* (Douglas and Penny 1999; Friedl and O'Kelly 2002) et ce chiffre atteint 63% avec la Prasinophyte *M. viride*. Parmi les 60 paires de gènes ancestrales, 40 sont retrouvées dans l'ADNcp de *P. signiensis*. Cet organisme est, parmi ceux du clade UTC, celui dont le génome est le moins réarrangé puisqu'il présente le plus de regroupements de gènes ancestraux (13/26) et les plus faibles taux d'inversions de gènes vis-à-vis des ADNcp de *M. viride* (39) et de *N. olivacea* (43).

Bien qu'il soit le seul ADNcp des Trebouxiophytes à contenir le gène *ycf47*, son contenu génique est équivalent à ceux des autres ADNcp de Trebouxiophytes (même nombre, même nature de gènes). L'ancêtre des Trebouxiophytes possède au moins 113 gènes. Le gène *ycf47* est un critère de conservation puisqu'il n'est retrouvé que chez *N. olivacea* (Turmel, Otis, and Lemieux 1999) pour les algues vertes, chez *Porphyra purpurea* (Reith and Munholland 1993) et *Guillardia theta* (Douglas and Penny 1999) pour les algues rouges et chez *Synechocystis* (Kaneko et al. 1996) pour les cyanobactéries.

Toutes ces caractéristiques permettent d'affirmer que l'ADNcp de *P. signiensis* est très conservé. Il diffère cependant des autres ADNcp par trois grandes particularités.

IV.2. Un génome de grande taille

P. signiensis possède le plus grand ADNcp séquencé parmi les Trebouxiophytes. Avec 236 463 pb, il dépasse la plupart des ADNcp des algues chlorophycéennes : *S. helveticum* (223 902 pb) (résultats non publiés), *C. reinhardtii* (203 827 pb) (Maul et al. 2002) et *N. olivacea* (200 799 pb) (Turmel, Otis, and Lemieux 1999). Il reste toutefois inférieur au plus grand ADNcp des algues vertes trouvé chez *C. moewusii* (292 kb) (Sugiura, Hirose, and Sugita 1998). Avec sa position basale supposée au sein du clade UTC, la taille de l'ADNcp de *P. signiensis* réfute l'hypothèse (de Cambiaire et al. 2006) selon laquelle la taille des génomes augmentent au cours de l'évolution de ces lignées.

Une des explications de cette grande taille réside dans la présence de deux grands IRs. En effet, tout comme les ADNcp de *C. reinhardtii* et de *N. olivacea*, l'ADNcp de *P. signiensis* possède deux IRs de plus de 20 000 pb. Chez les plantes vertes, la variation de taille des ADNcp est due aux variations de taille de leur IRs (Boudreau and Turmel 1995; Sugiura, Hirose, and Sugita 1998). Chez les algues vertes, la présence ou l'absence d'IRs ne semblent pas jouer sur la longueur de ces génomes. *N. olivacea* possède un ADNcp de grande taille pourvu des plus grands IRs (40 137 pb chacune), pourtant *S. helveticum* qui lui aussi présente un ADNcp de taille conséquente, ne présente pas d'IRs. Il n'existe pas non plus de corrélation entre la taille de ces IRs et la taille du génome puisque chez *P. pseudoalveolaris* (145 947 pb) et *O. viridis* (151 933 pb), les ADNcp sont approximativement de même taille alors que la longueur de leur IRs diffère (6 786 pb pour le premier contre 18 510 pb pour le second). Bien que la présence de ces grands IRs participe pour 23% dans la taille de l'ADNcp de *P. signiensis*, même en leur absence cet ADNcp reste imposant (209 020 pb sans deuxième IR).

Un deuxième facteur rentre dans la grandeur de l'ADNcp de *P. signiensis* : la prolifération de séquences répétées dans les régions intergéniques. *P. signiensis* présente que 54,1% de régions non codantes. Chez le genre *Chlamydomonas*, c'est l'extension de ces régions qui est à l'origine de leur grande taille (Boudreau, Otis, and Turmel 1994). Chez *P. signiensis*, elles contiennent 30.5% de séquences répétées dispersées (SDRs).

En dehors de l'ADNcp de *N. olivacea*, les grands ADNcp sont ceux qui présentent les taux les plus élevés en SDRs tels *S. helveticum* (40.8%) ou *C. reinhardtii* (31.9%). Il n'existe cependant pas de proportionnalité entre la taille de ces génomes et leur contenu en séquences répétées puisque *C. vulgaris* avec ses 150 613 pb présente 30.5% de SDRs et *N. olivacea* avec ses 200 799 pb n'atteint que 1.7%. Le gain de taille n'est pas forcément dû à une augmentation de SDRs bien que chez *P. signiensis* elles représentent un grand nombre de nucléotides.

IV.3. La prolifération de SDRs dans l'ADNcp de *P. signiensis*

Les ADNcp présentent généralement peu de SDRs (Palmer 1985) contrairement aux génomes nucléaires (Maul et al. 2002). Pourtant, l'ADNcp de *P. signiensis* a un pourcentage de séquences répétées qui atteint les 30.5 %. C'est l'ADNcp des Trebouxiophytes le plus envahi par les SDRs. Celles-ci se localisent tout le long de la séquence bien qu'elles se concentrent à 44% dans les IRs.

Ces SDRs sont majoritairement de courtes séquences de moins de 10 pb. Il est retrouvé quelques séquences répétées plus longues qui souvent se présentent sous forme de palindrome. Toutefois, seules les SDRs de petites tailles sont communes à celles retrouvés dans les IRs de *C. ellipsoidea*. Généralement le manque de similarités des SDRs chez les membres du clade UTC indique que celles-ci s'acquièrent de façon indépendante (Palmer 1985; Maul et al. 2002; Pombert et al. 2006). Dans le cas particulier des IRs de *P. signiensis* et de *C. ellipsoidea*, leur origine est sans doute commune.

Il semble que la présence de SDRs soit un facteur de réarrangement et de délétion de gènes (Pombert et al. 2006). Il existerait d'autre part une corrélation entre l'abondance des SDRs et le degré de remaniement (Pombert et al. 2005). Des ADNcp très réarrangés comme *O. viridis* ou *S. helveticum* présentent effectivement un pourcentage élevé de SDRs. Néanmoins, l'ADNcp bien conservé de *P. signiensis* présente un important taux de séquences répétées ce qui réfute les deux hypothèses précédentes.

Toutefois, 44% des SDRs de l'ADNcp de *P. signiensis* sont présents dans la partie la plus remaniée de ce génome : les IRs. Suspectées d'être à l'origine des inversions de gènes (Palmer 1991), on les retrouve au sein des IRs, concentrées aux extrémités des deux fragments inversés de l'opéron *rrn*. Cette disposition de SDRs est aussi retrouvée chez *C. ellipsoidea*, confirmant ainsi l'origine commune des remaniements de leurs IRs et l'hypothèse précédente quant à l'inversion des gènes. Malgré des remaniements géniques identiques, les IRs ne contiennent pourtant pas le même pourcentage de répétitions. Chez *C. vulgaris* 16% des IRs comprennent des séquences répétées tandis que les IRs de *P. signiensis* en contiennent le double : 31.9%. Ceci réfute une fois de plus l'hypothèse de la corrélation entre remaniement génique et abondance des SDRs mais confirme que les répétitions sont à l'origine de réarrangements géniques bien que ceux-ci ne soient pas forcément dus à la présence de SDRs (Milligan, Hampton, and Palmer 1989).

IV.4. Les IRs de l'ADNcp de *P. signiensis* : les régions les plus réarrangées

Les ADNcp d'algues vertes ne présentent pas tous d'IRs. Chez les Trebouxiophytes, cette présence est très aléatoire puisque *C. vulgaris* et *L. terrestris* en sont dépourvus contrairement à *P. pseudoalveolaris* et *P. signiensis*. Malgré leur origine commune et leur supposé présence dans l'ancêtre commun des algues vertes, leur perte serait multiple et indépendante des lignées (Turmel, Otis, and Lemieux 1999). Des études menées sur *C. vulgaris* et *C. ellipsoidea* ont clairement indiqué que les IRs subissaient des réarrangements et des délétions au cours de l'évolution des Trebouxiophytes (Yamada 1991; Wakasugi et al. 1997), de part la présence ou non de deux IRs, la variation du contenu génique de ces IRs, les réarrangements géniques de l'opéron *rrn* ou la prolifération de SDRs..

Comme de nombreux IRs, ceux de *P. signiensis* présente les gènes de l'opéron *rrn* (Palmer 1985). Ils contiennent aussi des gènes additionnels : *psbA* et *trnS(gcu)*, et de nombreux ORFs. Le contenu génique de ces régions est identique en tous points à celui trouvé chez *C. ellipsoidea*. La présence de gènes non ribosomiques dans les IRs est limitée aux plantes

vasculaires et aux algues vertes (Simpson and Stern 2002). La présence du gène *psbA* dans les IRs a été relevée chez les membres du genre *Chlamydomonas* : *C. reinhardtii* (Maul et al. 2002), *C. eugametos*, *C. moewusii* et *O. viridis* (Pombert, Lemieux, and Turmel 2006). Bien que *L. terrestris* ne présente pas de structure quadripartite, le gène *psbA* est aussi retrouvé à proximité d'un des fragments de l'opéron *rrn* (résultats non publiés). Ces variations de contenu génique dans les IRs sont souvent observées et découlent du phénomène « Ebb and Flow » (Goulding et al. 1996). C'est-à-dire : la possibilité pour les frontières IR/LSC et IR/SSC d'être sujet à un processus dynamique et aléatoire qui permet aux IRs de s'étendre ou de se contracter au profit des régions voisines. Par ce processus, les IRs acquièrent ou délètent des gènes supplémentaires à ceux de l'opéron *rrn*. Il est probable que l'ancêtre commun des membres du clade UTC présente le gène *psbA* à proximité de ses propres IRs, expliquant entre autre la présence de la paire *psbA-rrs* chez *L. terrestris*. Il en est de même pour le gène *trnS(gcu)* retrouvé dans les IRs de *S. obliquus* (de Cambiaire et al. 2006).

La grande taille des IRs de *P. signiensis* découle à la fois de ce phénomène d'« Ebb and Flow » mais également de l'extension des régions intergéniques. Les IRs de *C. ellipsoidea* (16 154 pb) qui possèdent le même contenu génique sont presque deux fois plus petites que celles de *P. signiensis* (27 336 pb). Comme cela est expliqué dans le paragraphe précédent, cette différence est due à la prolifération de SDRs. Ces SDRs s'accumulent au niveau des fragments de l'opéron *rrn*.

Cette fragmentation est une des particularités des ADNcp de *P. signiensis* et de *C. ellipsoidea*. L'opéron *rrn* est brisé en deux morceaux : *rrs-trnI(gau)* et *rrf-rrl-trnA(ugc)*, qui sont transcrits vers la LSC pour le premier et vers la SSC pour le second. *L. terrestris* possède également un opéron *rrn* fragmenté dont les morceaux sont séparés par 59 243 pb. Dans les deux cas, l'opéron *rrn* a perdu son processus de cotranscription.

Chez les Ulvophytes, *O. viridis* et *P. akinetum*, l'ensemble des IRs contenant l'opéron *rrn* sont transcrits dans le sens inverse de celui des algues les plus basales, soit vers la LSC. Il a été émis dans une étude précédente, que ce changement de sens de transcription aurait eu lieu chez l'ancêtre commun du clade UTC. Ce dernier présenterait un IR transcrit vers la LSC et l'autre vers la SSC. A défaut d'observer cette caractéristique dans l'ADNcp de *P. signiensis*, on note

tout de même un début d'inversion qui pourrait être à l'origine du changement de sens de transcription. Ce remaniement semble avoir pris sa source dans la classe des Trebouxiophytes.

Dernière particularité des IRs et plus précisément de l'opéron *rrn* : la présence d'ITS dans les gènes d'ARNr. Chez les ADNcp des algues vertes, les gènes de cet opéron présentent très souvent des introns, exemple chez *C. vulgaris* (Wakasugi et al. 1997). Bien que leur présence ne soit pas systématique et souvent aléatoire au sein des différentes classes, ils sont de plus en plus retrouvés chez les Chlorophyta et reste un critère de divergence (Lemieux, Otis, and Turmel 2000). L'ancêtre des algues vertes ne disposerait pas plus d'un intron dans son ADNcp. Au sein des gènes de l'opéron *rrn* de *P. signiensis*, aucun intron n'a été mis en évidence. En contre partie, les gènes *rrs* et *rrl* comptent respectivement un et trois ITS.

Les 3 ITS du gène *rrl*, si l'on se réfère à ceux trouvés dans les mêmes positions chez les membres du genre *Chlamydomonas* (Turmel et al. 1991; Turmel et al. 1993), permettent la formation de 4 ARNr appelés α , β , γ et δ . Les quatre structures secondaires de ces 4 ARNr assurent les mêmes fonctions que celle obtenue par l'ARNr 23S entier (Sugiura, Hirose, and Sugita 1998). Parmi les ADNcp des Trebouxiophytes séquencés, *P. pseudoalveolaris* et *C. ellispoidea* présente aussi l'ITS $\gamma\delta$ dans leur gène *rrl*. Tout comme les ITS des *Chlamydomonas*, ceux des Trebouxiophytes varient en taille et composition mais aussi en structures secondaires et sont localisés dans les régions instables du gène *rrl*.

Un quatrième ITS est retrouvé dans le gène *rrs*. C'est la première fois que ce type d'insertion est relevé. Cette extension longue de 282 pb est présente dans une région variable en base et ne présente aucune caractéristique intronique.

V. Conclusion

L'ADNcp de P. signiensis appuie-t-il la position basale des Trebouxiophytes au sein du clade UTC ?

Dans un premier temps, l'ADNcp de *P. signiensis* s'inscrit parfaitement dans la classe des Trebouxiophytes. D'un point de vue structural, sa structure quadripartite, sa répartition génique en fonction des régions (LSC, SSC et IRs) et son absence d'intron montrent sa ressemblance à l'ADNcp de *P. pseudoalveolaris*. D'un point de vue génique, son contenu, ses ordres de gènes et leurs regroupements le rapprochent de celui de *C. vulgaris*. Dans les deux cas, cet ADNcp a gardé les caractéristiques les plus ancestrales des ADNcp d'algues vertes, suggérant l'occupation d'une position basale au sein de Trebouxiophytes.

Cette hypothèse est renforcée par les ressemblances que cet ADNcp partage avec les membres de la classe des Prasinophytes. L'ADNcp de *P. signiensis* présente d'une part un pourcentage de gènes conservés plus élevé avec *M. viride* (63%) qu'avec d'autres membres de la classe des Trebouxiophytes : *L. terrestris* (48%) et *P. pseudoalveolaris* (60). Il en est de même pour le nombre d'inversions : *M. viride* (39) et *L. terrestris* (50). A ceci s'ajoute l'absence d'intron. D'autre part, l'ADNcp de *P. signiensis* est le seul parmi les Trebouxiophytes à posséder le gène *ycf47*. Chez les algues vertes, seul *N. olivacea* le contient (Turmel, Otis, and Lemieux 1999). Ce gène est également retrouvé chez *Porphyra purpura* (Reith and Munholland 1993) et *Guillardia theta* (Douglas and Penny 1999) pour les algues rouges ainsi que chez *Synechocystis* pour les cyanobactéries (Kaneko et al. 1996). La présence d'un tel gène appuie la position basale de *P. signiensis* et indique sa présence dans le contenu génique de l'ADNcp de l'ancêtre commun des Trebouxiophytes.

Au sein des Chlorophyta, l'ADNcp de *P. signiensis* présente, comme beaucoup de Trebouxiophytes, des caractéristiques intermédiaires à la classe des Prasinophytes et des Ulvophytes :

- Le contenu en gène : 111 gènes pour l'ADNcp de *P. signiensis* alors qu'en moyenne les ADNcp de Prasinophytes en possèdent 131 et ceux des Ulvophytes 105. Ce nombre va en diminuant au cours de l'évolution.
- La répartition génique selon celle de *M. viride*: sur les cinq gènes échangés entre LSC et SSC, quatre sont identiques aux 12 gènes communs relevés dans les génomes d'Ulvophytes (Pombert et al. 2005; Pombert, Lemieux, and Turmel 2006)
- Le nombre d'inversions de gène en comparaison avec l'ADNcp de *M. viride* : 39 pour *P. signiensis* tandis que les Ulvophytes s'en éloignent avec 49 inversions.
- Le nombre de paires de gènes ancestrales : *P. signiensis* en présente 40 tandis que *O. viridis* n'en compte que 32 et *P. akinetum* 29.

Les différentes particularités de l'ADNcp de *P. signiensis*, expliquées dans le chapitre précédent, illustrent d'autre part les remaniements observés chez les autres membres du clade UTC. Le changement de sens de transcription retrouvé chez les Ulvophytes aurait pris son origine au sein de la classe des Trebouxiophytes tout comme les échanges de gènes entre les régions. A noter aussi l'intrusion des ITS que l'on observe chez les algues chlorophycéennes.

Ces caractéristiques appuient l'hypothèse selon laquelle les Trebouxiophytes auraient divergé en premier dans le clade UTC.

Perspectives

Afin de conforter l'idée que *P. signiensis* occupe une position basale au sein des Trebouxiophytes et du clade UTC, il est nécessaire d'ajouter cet ADNcp aux analyses phylogénomiques en cours. D'autre part, l'étude de l'ADNmt de *P. signiensis* augmenterait la base de données mitochondriale et serait utile aux analyses comparatives tant structurales que phylogénomiques.

Une fois la position de *P. signiensis* définie dans la lignée des Chlorophyta, la compréhension de certains événements moléculaires pourra être précisée comme le changement de sens de transcription des IRs, la présence d'ITS dans les gènes de l'opéron *rrn* ou encore la perte aléatoire de la structure quadripaite au sein des Trebouxiophytes. Des études comparatives sur les SDRs de l'ADNcp et l'ADNmt renseigneraient sur la présence de telles structures dans ces génomes.

L'échantillonnage des Trebouxiophytes et des autres membres du clade UTC doit continuer afin d'avoir au moins les ADNcp et les ADNmt d'un représentant de chaque ordre et de pouvoir résoudre l'ordre d'embranchements au sein de la lignée des Chlorophyta.

Bibliographie :

- Bhattacharya, D., K. Weber, S. S. An, and W. Berning-Koch. 1998. Actin phylogeny identifies *Mesostigma viride* as a flagellate ancestor of the land plants. *J. Mol. Evol.* **47**:544-550.
- Boudreau, E., C. Otis, and M. Turmel. 1994. Conserved gene clusters in the highly rearranged chloroplast genomes of *Chlamydomonas moewusii* and *Chlamydomonas reinhardtii*. *Plant Mol. Biol.* **24**:585-602.
- Boudreau, E., and M. Turmel. 1995. Gene rearrangements in *Chlamydomonas* chloroplast DNAs are accounted for by inversions and by the expansion/contraction of the inverted repeat. *Plant. Mol. Biol.* **27**:351-364.
- de Cambiaire, J. C., C. Otis, C. Lemieux, and M. Turmel. 2006. The complete chloroplast genome sequence of the chlorophycean green alga *Scenedesmus obliquus* reveals a compact gene organization and a biased distribution of genes on the two DNA strands. *BMC Evol. Biol.* **6**:37.
- Delsuc, F., H. Brinkmann, and H. Philippe. 2005. Phylogenomics and the reconstruction of the tree of life. *Nat. Rev. Genet.* **6**:361-375.
- de Reviere, B. 2002. Biologie et phylogénie des algues. Belin, Paris.
- Delwiche, C. F., M. Kuhsel, and J. D. Palmer. 1995. Phylogenetic analysis of *tufA* sequences indicates a cyanobacterial origin of all plastids. *Mol. Phylogenet. Evol.* **4**:110-128.
- Douglas, S. E., and S. L. Penny. 1999. The plastid genome of the cryptophyte alga, *Guillardia theta*: complete sequence and conserved synteny groups confirm its common ancestry with red algae. *J. Mol. Evol.* **48**:236-244.
- Dixit, R., P. K. Trivedi, P. Nath, and P. V. Sane. 1999. Organization and post-transcriptional processing of the *psb B* operon from chloroplasts of *Populus deltoides*. *Curr Genet* **36**:165-172.
- Friedl, T., and C. J. O'Kelly. 2002. Phylogeny relationships of green algae assigned to the genus *Planophila* (Chlorophyta): evidence from 18S rDNA sequence data and ultrastructure. *Eur. J. Phycol.* **37**:373-384.
- Glockner, G., A. Rosenthal, and K. Valentin. 2000. The structure and gene repertoire of an ancient red algal plastid genome. *J. Mol. Evol.* **51**:382-390.
- Goulding, S. E., R. G. Olmstead, C. W. Morden, and K. H. Wolfe. 1996. Ebb and flow of the chloroplast inverted repeat. *Mol. Gen. Genet.* **252**:195-206.
- Jansen, R. K., L. A. Raubeson, J. L. Boore, C. W. dePamphilis, T. W. Chumley, R. C. Haberle, S. K. Wyman, A. J. Alverson, R. Peery, S. J. Herman, H. M. Fourcade, J. V. Kuehl, J. R. McNeal, J. Leebens-Mack, and L. Cui. 2005. Methods for obtaining and analyzing whole chloroplast genome sequences. *Methods Enzymol.* **395**:348-384.
- Kaneko, T., S. Sato, H. Kotani, A. Tanaka, E. Asamizu, Y. Nakamura, N. Miyajima, M. Hirose, M. Sugiura, S. Sasamoto, T. Kimura, T. Hosouchi, A. Matsuno, A. Muraki, N. Nakazaki, K. Naruo, S. Okumura, S. Shimpo, C. Takeuchi, T. Wada, A. Watanabe, M. Yamada, M. Yasuda, and S. Tabata. 1996. Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions (supplement). *DNA Res.* **3**:185-209.

- Kelchner, S. A., and J. F. Wendel. 1996. Hairpins create minute inversions in non-coding regions of chloroplast DNA. *Curr. Genet.* **30**:259-262.
- Kurtz, S., J. V. Choudhuri, E. Ohlebusch, C. Schleiermacher, J. Stoye, and R. Giegerich. 2001. REPuter: the manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Res.* **29**:4633-4642.
- Kusumi, J., and H. Tachida. 2005. Compositional properties of green-plant plastid genomes. *J. Mol. Evol.* **60**:417-425.
- Lemieux, C., C. Otis, and M. Turmel. 2000. Ancestral chloroplast genome in *Mesostigma viride* reveals an early branch of green plant evolution. *Nature* **403**:649-652.
- Lewis, L. A., and R. M. McCourt. 2004. Green algae and the origin of land plants. *J. Bot.* **91**:1535-1556.
- Martin, W., O. Deusch, N. Stawski, N. Grunheit, and V. Goremykin. 2005. Chloroplast genome phylogenetics: why we need independent approaches to plant molecular evolution. *Trends Plant Sci.* **10**:203-209.
- Martin, W., and R. G. Herrmann. 1998. Gene transfer from organelles to the nucleus: how much, what happens, and Why? *Plant. Physiol.* **118**:9-17.
- Mattox, K. R., and K. D. Stewart. 1985. Classification of the green algae: a concept based on comparative cytology. Pp. 29-72. *The systematics of the Green Algae.* Academic Press, London.
- Maul, J. E., J. W. Lilly, L. Cui, C. W. dePamphilis, W. Miller, E. H. Harris, and D. B. Stern. 2002. The *Chlamydomonas reinhardtii* plastid chromosome: islands of genes in a sea of repeats. *Plant. Cell.* **14**:2659-2679.
- Milligan, B. G., J. N. Hampton, and J. D. Palmer. 1989. Dispersed repeats and structural reorganization in subclover chloroplast DNA. *Mol. Biol. Evol.* **6**:355-368.
- Moret, B. M., and T. Warnow. 2005. Advances in phylogeny reconstruction from gene order and content data. *Methods Enzymol.* **395**:673-700.
- Nozaki, H., N. Ohta, M. Matsuzaki, O. Misumi, and T. Kuroiwa. 2003. Phylogeny of plastids based on cladistic analysis of gene loss inferred from complete plastid genome sequences. *J. Mol. Evol.* **57**:377-382.
- O'Kelly, C. J., and G. L. Floyd. 1984. Flagellar apparatus absolute orientatoin and the phylogeny of the green algae. *BioSystems* **16**:227-251.
- Palmer, J. D. 1985. Comparative organization of chloroplast genomes. *Annu. Rev. Genet.* **19**:325-354.
- Palmer, J. D. 1991. Plastids chromosomes: structure and evolution. Pp. 5-53 *in* L. Bogorad, and I. K. Vasil, eds. *Cell culture and somatic cell genetics of plants.* Academic press, San Diego.
- Pombert, J. F., P. Beauchamp, C. Otis, C. Lemieux, and M. Turmel. 2006. The complete mitochondrial DNA sequence of the green alga *Oltmannsiellopsis viridis*: evolutionary trends of the mitochondrial genome in the Ulvophyceae. *Curr. Genet.*
- Pombert, J. F., C. Lemieux, and M. Turmel. 2006. The complete chloroplast DNA sequence of the green alga *Oltmannsiellopsis viridis* reveals a distinctive quadripartite architecture in the chloroplast genome of early diverging ulvophytes. *BMC Biol.* **4**:3.
- Pombert, J. F., C. Otis, C. Lemieux, and M. Turmel. 2005. The chloroplast genome sequence of the green alga *Pseudendoclonium akinetum* (Ulvophyceae) reveals unusual structural features and new insights into the branching order of chlorophyte lineages. *Mol. Biol. Evol.* **22**:1903-1918.

- Race, H. L., R. G. Hermann, and W. Martin. 1999. Why have organelles retained genomes? *Trends Genet.* **15**:364-370.
- Raven, J. A., and J. F. Allen. 2003. Genomics and chloroplast evolution: what did cyanobacteria do for plants? *Genome biology* **4**.
- Reith, M., and J. Munholland. 1993. A High-Resolution Gene Map of the Chloroplast Genome of the Red Alga *Porphyra purpurea*. *Plant Cell* **5**:465-475.
- Schwartz, S., Z. Zhang, K. A. Frazer, A. Smit, C. Riemer, J. Bouck, R. Gibbs, R. Hardison, and W. Miller. 2000. PipMaker--a web server for aligning two genomic DNA sequences. *Genome Res.* **10**:577-586.
- Simpson, C. L., and D. B. Stern. 2002. The treasure trove of algal chloroplast genomes. Surprises in architecture and gene content, and their functional implications. *Plant. Physiol.* **129**:957-966.
- Stoebe, B., and K. V. Kowallik. 1999. Gene-cluster analysis in chloroplast genomics. *Trends Genet.* **15**:344-347.
- Sugiura, M. 1992. The chloroplast genome. *Plant. Mol. Biol.* **19**:149-168.
- Sugiura, M. 2003. History of chloroplast genomics. *Photosynth Res.* **76**:371-377.
- Sugiura, M., T. Hirose, and M. Sugita. 1998. Evolution and mechanism of translation in chloroplasts. *Annu. Rev. Genet.* **32**:437-459.
- Tesler, G. 2002. GRIMM: genome rearrangements web server. *Bioinformatics* **18**:492-493.
- Turmel, M., J. Boulanger, M. N. Schnare, M. W. Gray, and C. Lemieux. 1991. Six group I introns and three internal transcribed spacers in the chloroplast large subunit ribosomal RNA gene of the green alga *Chlamydomonas eugametos*. *J. Mol. Biol.* **218**:293-311.
- Turmel, M., R. R. Gutell, J. P. Mercier, C. Otis, and C. Lemieux. 1993. Analysis of the chloroplast large subunit ribosomal RNA gene from 17 *Chlamydomonas* taxa. Three internal transcribed spacers and 12 group I intron insertion sites. *J. Mol. Biol.* **232**:446-467.
- Turmel, M., C. Otis, and C. Lemieux. 1999. The complete chloroplast DNA sequence of the green alga *Nephroselmis olivacea*: insights into the architecture of ancestral chloroplast genomes. *Proc. Natl. Acad. Sci. U S A* **96**:10248-10253.
- Turmel, M., C. Otis, and C. Lemieux. 2003. The mitochondrial genome of *Chara vulgaris*: insights into the mitochondrial DNA architecture of the last common ancestor of green algae and land plants. *Plant Cell* **15**:1888-1903.
- Wakasugi, T., T. Nagai, M. Kapoor, M. Sugita, M. Ito, S. Ito, J. Tsudzuki, K. Nakashima, T. Tsudzuki, Y. Suzuki, A. Hamada, T. Ohta, A. Inamura, K. Yoshinaga, and M. Sugiura. 1997. Complete nucleotide sequence of the chloroplast genome from the green alga *Chlorella vulgaris*: the existence of genes possibly involved in chloroplast division. *Proc. Natl. Acad. Sci. U S A* **94**:5967-5972.
- Yamada, T. 1991. Repetitive sequence-mediated rearrangements in *Chlorella ellipsoidea* chloroplast DNA: completion of nucleotide sequence of the large inverted repeat. *Curr. Genet.* **19**:139-147.

Annexe I

Fichier GenBank de l'ADNcp de *P. signiensis*

LOCUS PABIsigni 236463 bp DNA circular PLN 13-JAN-2006
DEFINITION *Pabia signiensis* chloroplast, complete genome.
ACCESSION
VERSION
KEYWORDS .
SOURCE chloroplast *Pabia signensis*
ORGANISM *Pabia signensis*
Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae;
Prasiolales; Prasiolales incertae sedis; *Pabia*.
REFERENCE 1 (bases 1 to 236463)
AUTHORS de Herce,L., Lemieux,C., Otis,C. and Turmel,M.
TITLE Complete chloroplast genome sequence from the green alga *Pabia signiensis*
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 236463)
AUTHORS de Herce,L., Lemieux,C., Otis,C. and Turmel,M.
TITLE Direct Submission
JOURNAL Submitted (13-JAN-2006) Biochimie et Microbiologie, Universite Laval, Pavillon Charles-Eugene Marchand, Quebec, Quebec G1K 7P4, Canada
FEATURES Location/Qualifiers
source 1..236463
/organism="Pabia signensis"
/organelle="plastid:chloroplast"
/mol_type="genomic DNA"
/strain="SAG 7.90"
/db_xref="taxon:173497"
misc_feature 1..141652
/note="Region: large single-copy region (LSC)"
gene complement(505..2025)
/gene="atpA"
CDS complement(505..2025)
/gene="atpA"
/codon_start=1
/transl_table=11
/product="CF1 alpha subunit of ATP synthase"
/translation="MVKIRPDEISSIIRQQIEQYNSEVKVVNVGTQVQVGDGIARIYG
LEKVMAGELLEFEFGTGVGIALNLESKNVAVLMGEGLTVQEGTSVRATGKIAQIPVGD
GYLGRVNSLARPIDGKGEIPTTENRLIESAAPGIISSRSVHEPLQTGLVAVDAMIPI
GRGQRELIIGDRQTGKTAVAVDTILNQKGDVICVYVAIGQKASSIAQVVNVLQDRGA
LDYTIIVAATADSPATLQYLAPYTGAALAEYFMYKGRHTLVIYDDLSKQAQAYREMSL
LLRRPPGREAYPGDVFYLSRLLERAAKLSDKLGGGSMTALPIVETQEGDVSAYIPTN
VISITDQIFLSGDI FNAIRPAINVGVISVSRVGSAAQPKAMKQVAGKLLKLELAQFAE
LEAFSQQFASDLQATQNLARGQRLRELLKQPQTSPLSIEDQVASIYAGTNGYLDKIK
VDQVRSFVLVGLRQYLANSKPKYGEIVKSTNTFNDEAQSLLEKAIQEYTEEFLLAVSK"
gene complement(2586..3122)
/gene="atpF"

```

CDS                complement (2586..3122)
                   /gene="atpF"
                   /codon_start=1
                   /transl_table=11
                   /product="CF0 subunit I of ATP synthase"
/translation="MTFLTLFAHIPSGEGFGFNGNILETNLINLSVVIADVVSFGGDA
LRSLLDNRKQTILNNLQEQAKAKEAEEKLTKARAQLELAKKKAIEIREQGVITAEQE
KKQAIRQMVEDVSRLNEIKQETIRFQQQKAI SQVVSQVSLALNQVREKLSNRDLSTF
HSSVNNFNIVLFTNYKPK"
  gene              complement (4134..4382)
                   /gene="atpH"
CDS                complement (4134..4382)
                   /gene="atpH"
                   /codon_start=1
                   /transl_table=11
                   /product="CF0 subunit III of ATP synthase"
/translation="MNPLIGAASVLAAGLAVGLAAIGPGMGQGTAAAGYAVEGIARQPE
AEGKIRGALLLSFAFMESLTIYGLVVALALLFANPFAS"
  gene              5759..6061
                   /gene="orf100"
CDS                5759..6061
                   /gene="orf100"
                   /codon_start=1
                   /transl_table=11
                   /product="hypothetical protein"
/translation="MYFKNTNRTLFLNKHLINSLTSLNLSVVLTFGMQSVSEVR
PQRSLFPIAFGNGKGWKILILIESKILFWTRIISYANRISFTCSSSLGLLSKQK"
  gene              complement (6426..7187)
                   /gene="atpI"
CDS                complement (6426..7187)
                   /gene="atpI"
                   /codon_start=1
                   /transl_table=11
                   /product="CF0 subunit IV of ATP synthase"
/translation="MVIMNETLGTIFYQLNSKNPIFEMSEVSVGQHFYWNFGDYQVHGQ
VLLTTWVVLGLIILLSSLLGNRDLKKIPTGLQNLTEYITEFIRDLAKTQIGEEYLA
WV PFLGTIFLFI FVSNWSGALLPWRLIELPHGELAAPTNDINTTVALALLTSIAYFYAGL
RKKGLGYFKRYIEPAAFLLPINVLEDFTKPLSLSFRLFGNILADELVVGVLIALVPLV
VPIPLMLLGLFTSAIQALVFATLAGAYIGEALGH"
  gene              complement (10261..10995)
                   /gene="rps2"
CDS                complement (10261..10995)
                   /gene="rps2"
                   /codon_start=1
                   /transl_table=11
                   /product="ribosomal protein S2"
/translation="MINSIDEMVEVGMHFGHQARKWNPKMAYIYAERNGIHIIDLIQ
TYYHLKKSTQFLNDSAAQGKKFLFVGTGKQASKLIAKVALQCDSYYVNQRWLGMLTN
WKTIKSSIGKLNQLEAREKNGEFNNLPKKEAASCKKEKDRLQKYLGGGLKKMDSIPDVV
IIVGQLEEMNAVRECQKLGIRSIITILDTCNPALADLFIPANDDSVASLQILNSFLE
AIRQGQKLYVEKKLPNDFKTRQFRKN"
  gene              complement (11998..18306)
                   /gene="rpoC2"
CDS                complement (11998..18306)
                   /gene="rpoC2"
                   /codon_start=1
                   /transl_table=11

```

```

        /product="beta'"subunit of RNA polymerase"
/translation="MVSNLSFFNGCFDKGRLKSFISWSLLNCGEQFTIELVENLKNLG
FEYATQAGVLSLDDLKIPPTKSKLVSEAELQITSAHIQYQQGHLTAVEKFQQLIDTW
HRTSETLKQNVIQYFRATDILNPVYMMAFSGARGNVSQVRQLVGMRLMADPQGQIID
FPIRSNFRGLTLTEYVISCYGARKGLVDTALRTANSGLTRRLVDVVSQHVIVCQLDC
GTQRGIFLSDMMEGQKVLNLSLENRLIGRILAENIYFDFAQVPSQRSDDIDETILTGP
LISLSLNSNRKPNLNLISSNKKFLSTKNAGSPTNLIGSKDQEISATLAKKIAKLRKQV
LVRSPLTCEAKSSICQLCYGWSLAHGNLVSLSGEAVGILAAQSIGEPGTQTMRTFHTG
GVFSGDLMTEVRAPFDGIEFTELLQGILIRTPHGKIAFLTKVQGEFYIKQLKKSQQL
TSLGFETILKQNSQKEGILSRLEQKDKTGSKQKFKIPASTILFVRQGEFVLEKQLIAE
FSSISAQSNQRIQAKHNLNSEMEGQVFFENVVLAVKESKEGDITRVAKKLGSIWILSG
KIYQTMIPTSFFPKCDLVNSGVSVINQILVISPYTGFLTQNFYNSKASASWKAFPSKS
IFNTNATRKPISFSECLKRYENCQNLFISSDFKNIQISTLPVALPGTARVPKSTKFD
YPRDYSKDKGENLSLNPVLSFPIKNISSYYQMGYFLSFWSNKKLSFDLVKSFTKKT
FDYDFLSFVPSSTNQDLPETGSQLECGLLTTGFTYCSTGHSKHQNRRLRKKLFDK
TGSNLIQNAQKPSIKSRMPLAGLPNPVSTDLFFLSSSLKQELDNVNLPKLLYLQSF
PNSNYKTKTGGLVFYDSLYLENDCGQIFWVPQESYKVNPKTLFVNPNDTGFVLFPT
FEMTPNNLSYFVKKFAKKWVNKNFPLLSKYNSQGIIRNFSSKMSGWLEIKSTK
KLTGYC SIQPNGEILNSFPKSWFEVETCLKNFLSKQFINFKNRSFLTLVFI
LSQKNEVINSGLTSLIFNKA RKGQAKLSNLTTVGLLSIKNNCFRPNPYIL
VKQRHGETYKSVSYKKNYLNNINEARQISQFTINSIKLNWFLKFEEKYTC
NKVFLKYDKITRLTSRTPFINSQQSFITKNSVTSHFQVSDCTQLFNYSASKI
KPYISISLKKGEGFQIDSQTRPPSSVLPVNLVQENKKSSCDNCL EIGIKSG
WVYFPRNQTDILDYHKQLFKPGFSFIDNVEFDQHITYLECISIPFLSVDTY
TYQFSAFQVKEKNVYLGQLNFKELFKFPLSFLKLNNTQVPFSSYDSSNVEN
FIMNLDIKKSKVLFKMGSFNKIQNCNNLNLFLFKLDFKIFFKKSVDLAFALA
QGNEQNSKTA EFHPFTRKELNFIKSELPDFVKFGSNLTLCSDKIYYIKNKN
HRYFEFLQNKINVNKS FLPKLKNISSFYLVKSYQSVKKNVSLANFNKRLTP
KEQGCQAKLEYPIPKIFPNLN SFGFQSVLPEQFCNIAFSTKVDTLN
NNFSTPSFFILIRKVKEYSVFKSKHYKVLN NQKPNLIIQSQNDIGRLT
NLTFWNKQVITKLFSSFP SADFQFNSSFRFKNSKHMFFTKELNIVEFFI
LLNIPKFFPKTAKVGFVSKHTLLSTNSRFKHKVRSQEKFTNFAIT
KVASLRIAASNKSELGKQLNGTTNSSSNYRRNLGVSRAQQGNLSYITNYLE
FSLSRQIDNASFLIPELKNKTIQFTNTKLVLHPQNSIAAHNPLSLTDFFS
PYQGEITDIKTDFL GKQSFFLLTEKDQISFSTEQKTPLTFVGKLI
RYGEQIANNLSIANSQIIQVDKSKVILRKAQPILFSSKGVFYVH
HGDFVEKNSPLLTFLFYQRLKTGDIVQGIPKIEQLFEARQT
KEGEILPENLHDKLHNFFEKYKQKYSRDAARKSLEKTQQLLVDG
VQRVYQSQGVTIADKHLEIIVRQMTSKVRIVEGGQSGLLRGELI
DLDRIEIVNNGINSQKAEYEPVILGIT KAALETESFISAASFQET
TRILARAAIERKTDFLRGLKENVILGHLIPAGTGFSRSFD
PKNSAYSKKSEKEKLFKNFFLKVQKNEN"

```

```

    gene          complement (18966..21437)
                  /gene="rpoC1"

```

```

    CDS           complement (18966..21437)
                  /gene="rpoC1"
                  /codon_start=1
                  /transl_table=11
                  /product="beta' subunit of RNA polymerase"

```

```

/translation="MNIENTSKITKVD SIRVGLAPPERIKKWAERTLPNGKIVGQVLS
SQTVNYKTLKPEKDGLFCERIFGPVKDFECACGRKKNKTQQQFCPCDVEFTSARV
RRYRLGYIQLVSPVTHVWYLGKTPSYISIILEMRRKKEAITYCTETTNI
SGINFADLYSGITPSRFSNAKSNQTLPGI AVLNLLENLQEQRFEQTKWNSDFIQ
SKDTHKTQEIEQDSN SESKQRNYNVFTREPLIPGYDDLSSFSKEELHSSK
GKTLGERFNI SLEKKKEGRSLG PSFEETSLTKTPVEKSSQNKQVNNY
YTI FQTASWEVQDDL NKF SYMSASFEAEDTPI PIYFDRIKQFLL
LNIIYINPTHFDIPLTGAEALRSFLKNL DLQLLNQRIRVELFELN
EEIHEFENQNFLFLIEQRRLKFLLSLRAKCLRRLKLVHRFRRTKAQPE
WMILSILPVL PPDLRPIIQLDGDQVAVSDLNKLYQKVLFRNNRMKR
HKKSNC SNKSDEIKYAQRLLQE AVDALIENGKGGADPICASNDRPL
KLSLSDMLKGKKGRFRQNL LGKRV DYSGRSVIVG PKLKIHECGLPK
EMAIELFQPF LIRRLMTTQIVRTIVSAKRLIQRQDPVIWEILQOVL
KNHPILLNRAPT LHLRGLGIQAFQPKLV DGRAILLHPLVCPAFNAD
FDGDQMAVHVPLSF

```

QARAEAWKLMWSRNNLLSPATGQPILIPSQDMVLGCYYLTTTNSKALKGIGKYFIDLS
DSLKAFNQKKIALHAPIWVRWSGQFETGNLIEKPLEIRIDCFGNISNLYHKYQRTSDK
NLTQKSQFIRTTAGRILLNQILLNN"

gene complement (22607..27403)
/gene="rpoB"

CDS complement (22607..27403)
/gene="rpoB"
/codon_start=1
/transl_table=11
/product="beta subunit of RNA polymerase"

/translation="MSFSSVRISYFIPDFVEIQRKSFLDLLDKGLIQELYKRNPIITNT
KQNLLELRFYPEYYKLNPPDWTPEKECILKSKSYACRLYVPIHLINKATKESKLQWVLLG
NLPLMTRKRGHFIILNGSPRVIINQMVRSPGIYYQVHVHKNKRTYYADLISHRGAWLRL
ETDKKRKVVARMKKTpkvSILVFLQALGLSKEKIFQSINYSNFLKSSFLKEEHSFAHP
ISSGQALQALYFETHPKKEKSEITTDLGQKFLLRKFFNSRTYDLSRLGRIQLNKKLGL
SVPLTKSVLTEQDILFATDYLIKLEYGLGLTDDIDDLKNRRVRASGELIQNLGTFI
RLEKMIREKLLKPKRKLTVRSLITTKPVNGALREFFGSSQLSQFMDQTNPLAELTHKR
RLSSLGPGGISRETAGMAVRGIHPSHYGRICPIETPEGPNAGLVNSLTYYAKVNSEGF
IETPFYSVYKGGIQMSAGPIFFSAAQEEHVKIAPGDLKLSQLDFLPKSTIPVRIKEF
KRISRSEVEYISISPIQMISIATSLIPFLEHDDANRALMGSNMQRQAVPLIAPERPIV
GTGLEGRVAADSGHAIQSRTSGFVSVYVSGERITIKNFIEPTLSLVATQGNITKSLVSN
LVKKIKKKRVFSRNSPNSKFRSQAPVYNLLENSRSRFLKTGKVNShVQLALNNQKKS
RITQEKPFIFRSLKKSQNKLHSIELKMNLSITKTSPLKFKSILNSFLSSNFVLPVTN
KIGRDINKVKSCQNKINKEKLLFEFSHVQNLRSRQVLLKQQNFVNSEDPNTIFRNPS
FCTPRDCILEFCSSCLTLRIDKKQWKTQPVLGENCNLDKIYLNKNVEFQPQLTKKLN
EKYQILSFKKMICREICVSKKKRYFLENSHYTKKSFLIRKNILKLQFTSIFNWHFLF
QHLIDFQFLEDPGLREAQTLDFHSVEDLNTISDCILENNNFYNNTHSKKPWKGFHFKF
FDELQGWNRASKLSKFSFLNFPGISSESPFEVPLVLRQNNVENTGTVKVSFNIEFIRQLS
VLDYFSLNKWSSQSTKHNTYVEFNKQLTKPSRLLKIKTRLSNNFSSISRTDFLFKQR
SELATRLSVPKWKTTNNSINHQSANSNTFNIEKNFSLSSINFLDQYQRSNQDTCLT
QRPIVHEGEWVQKGDLLADGTASVGGELSLGKTIIVAYMPWEGYNFEDAILISERLIY
DEIYTSIHIERYEIEIRDTKFGVEQITNQIPEIEASQVSHLDNRGIAKIGSWIKEGDV
LVGKITPIQKKSLSPEKLLYDIVKKEIPTTRDTSLRVPKNVHGRVNNIQILETENIP
PEIDFEGPGRVHIYVAEKRKIQVGDKMAGRHNKGIISKILPRQDMPYLPDGTVPDMV
LNPLGVPSRMNVGQVFECLLGLAGKQLEQQFKIIPFDEIHGPEASRSLVYSKLFARM
KTGQKWLFDPNFAGKTKLFDGRTGECFDQAVTVGQAYMLKLVHLVDEKIHARSTGPYS
LVTQQPLRGRSKHGGQRLGEMEVWALEGFGAAYTLQELLTVKSDDMKGRHQVMDAILN
NRPISLGTPESEFKVLIRELQSLCLDVGVTIIESSGRRKQIDVMKLS"

gene 28374..28444
/gene="trnC (gca) "

tRNA 28374..28444
/gene="trnC (gca) "
/product="tRNA-Cys"
/note="codons recognized: UGY"
/anticodon=(pos:28406..28408,aa:Cys)

gene 30411..31937
/gene="psbB"

CDS 30411..31937
/gene="psbB"
/codon_start=1
/transl_table=11
/product="CP47 chlorophyll apoprotein of photosystem II"

/translation="MGLPWYRVHTVVLNDPGRLLISVHLMHTSLVAGWAGSMAFYELAV
FDPSDPVLNPMWRQGMFVLPFMTRLGITQSWGGWTISGETANNPGIWSYEGVAASHIL
LSGALFAASIWHVWYDLELFRDPRTGDPALDLPKIFGIHLFLSGLLCFGFGAFHITG
LFGPGIWSVDPYGITGVSQVPVAPSWGAEFGDPYNPGGIPAHHIAAGILGVLGAGLFHLC
VRPPQRLYNGLRMGNIETVLSSSIAAVFWAAFVAVAGTMWYGSAAPIELFGPTRYQWD


```

LGFFQQEIEKRVQSNLAEGKSLSQAWSEIPEKLAFYDYIGNNPAKGGLFRAGAMNSGD
GIAVGWLGHAVFKDKDGNELYVRRMPTFFETFPVLLIDKDGVRADVPFRRAESKYSI
EQVGVSVSIFYGGELDGVTFSDPATVKKYARRAQLGEVFEFDRATLQSDGVFRSSPRGW
FTFGHLCFALLFFFHGIWHGARTIFRDVVFAGIDVDLDEQVEFGAFQKLGDSSTRRQSV
"
    gene          32989..33084
                  /gene="psbT"
    CDS           32989..33084
                  /gene="psbT"
                  /codon_start=1
                  /transl_table=11
                  /product="T protein of photosystem II"
                  /translation="MEALVYTFLLVGTGLGIIFFAIFFREPPRIVK"
    gene          complement (33929..34138)
                  /gene="psbN"
    CDS           complement (33929..34138)
                  /gene="psbN"
                  /codon_start=1
                  /transl_table=11
                  /product="N protein of photosystem II"
/translacion="MATLAFGQGKGSRIKNYKSFNLKLVMESPAFFYTIFLWCLLLSI
TGYSIYVGFPPSKQLRDPFEEHED"
    gene          34380..34628
                  /gene="psbH"
    CDS           34380..34628
                  /gene="psbH"
                  /codon_start=1
                  /transl_table=11
                  /product="10 KDa phosphoprotein of photosystem II"
/translacion="MAIGTTSKTKQSPLDTGKVTALGTLKPLNSEYGKVAPGWGTTV
LMGVFMALFAVFLVIIIEIYNSSVLLDDLTVSWDSLAK"
    gene          35251..35856
                  /gene="clpP"
    CDS           35251..35856
                  /gene="clpP"
                  /codon_start=1
                  /transl_table=11
                  /product="proteolytic subunit 2 of clp protease"
/translacion="MPIGVKVPFRLPGEPTAQWVDLYNRLYRERVLFLCQDLDELA
NQLIGIMLYLNAEEKSKGLYIYINSPGGSVTCGIAVYDTMNYVKADVTTICVGTAAAS
ASFVLAGGDRGKRIALPHSRIMIHQPEGGSQGQASEVLSESEEVMRIRRVQVGKIYAER
TGQTLRSRISRDMDRDQFMSAREAKEYGIVDQVAVDVKWVSGS"
    gene          complement (37011..37493)
                  /gene="petD"
    CDS           complement (37011..37493)
                  /gene="petD"
                  /codon_start=1
                  /transl_table=11
                  /product="subunit IV of cytochrome b6/f complex"
/translacion="MAVTKKPDLSDPILRAKLAKGMGHNYGEPAWPNDDLYIFPVVI
LGTAFAGVIGLAVLDPAVIGEPANPFATPLEILPEWYFYPVFQVLRVTPNKLLGVLFMA
AVPAGLLTVPFIENINKFQNPFRFPVATTVFLVGTLVVAIWLIGIGATLPIDISLTLGLF
"
    gene          complement (38019..38666)
                  /gene="petB"
    CDS           complement (38019..38666)
                  /gene="petB"

```

```

/codon_start=1
/transl_table=11
/product="apocytochrome b6 of cytochrome b6/f complex"
/translation="MSKVYDWFEEERLEIQSIADDITSKYVPPHVNIIFYCLGGITFTCF
LVQVATGFAMTFYRPTVAEAFASVQYIMTDVNFGLIRSIHRWSASMMVLMMLHVF
RVYLTGGFKKPRELTVVTVGIMAVCTVSFGVTGYSLPWDQVGYWAVKIVTGVDPDAIPL
VGPTLVELLRGGVGVGQSTLTRFYSLHTFVLPVLLTAVFMLMHFLMIRKQGISGPL"
gene complement (39728..40120)
/gene="rpl12"
CDS complement (39728..40120)
/gene="rpl12"
/codon_start=1
/transl_table=11
/product="ribosomal protein L12"
/translation="MSTTTNEIIIEKLKSITLLEAAELVSQIEETFGVDASASVGAGFM
PASGEGAGGQEAEEKTSFDIILEDVPSDKRVAVLKIIRNLTSLGLKEAKEFTTSLPK
ALKESVSKEEADTAKQOLEEAGAKVKIQ"
gene complement (41242..41634)
/gene="rps9"
CDS complement (41242..41634)
/gene="rps9"
/codon_start=1
/transl_table=11
/product="ribosomal protein S9"
/translation="MNILAKGTGRRKSAIAQIKLVKGTGEFIINGKPGVLYMQENPAS
VLSIQAPLELLRLQKTYNTIVTVQGGGLIGQAEAIKLGARALCDIEKSYRPFLLKTKG
YLTRDSRCKERKKYGLKKARKAPQFSKR"
gene complement (42024..43676)
/gene="rpoA"
CDS complement (42024..43676)
/gene="rpoA"
/codon_start=1
/transl_table=11
/product="alpha subunit of RNA polymerase"
/translation="MPNIASHTLVASQKETLKVENLLLSIESRVENNRSFYGRFQL
GPFDLGQGLTVATALRRTLLSELTGLAITAVVIKASHEYSTLSGVRESVLDILLNLK
QVVLAPVDTELEFESFPDSSNKQILTEPQIGFLKVCVGPVIVKARDMKLPKSIQCVDPD
QYIATLTANGNLEMKFIISEGKNYIVQTIISPKLSSRLSKQNFSSSSLNLFVSSTDTSS
LLGLESHFSKNSKNSLKNELMSNEIINSRRLNSDKVSDLDLIDQSQVIENTNGTGGN
YPLINANSILNSDGLALPKGNNTNPLEQMSRRQNSSEQINFPNLNKNKSFKANSI
EHQINLRQDSIKKSQIGLLKTNKTKNVLLIDAVFMPVNRVNFLENDEETIKPKEKII
LEIWTNGSIHPRKAIHEAASALIQMIAPFQESKLLKAFTLNWSLKDSONYQESVRRRE
LYSNQKKVKFAAKKDNLAIKKNRASIDIGNLELSLRPYTCLKRANINTVSDLLQYSAD
ELLLLKNFGRSLEEVEATLAPLGFGLGKQNSKELLSKNIIR"
gene complement (44214..44606)
/gene="rps11"
CDS complement (44214..44606)
/gene="rps11"
/codon_start=1
/transl_table=11
/product="ribosomal protein S11"
/translation="MARQNRKSLPKKIKRKPVGVVHVQASFNNTIITITNLKGEVIS
WSSAGACGFKGARKGTPFAAKTAAETAAKQSLDQGMKQARVMVKGPGAGRETAIRGLQ
EAGLIITLIRDITSIPHNGCRPPKKRRI"
gene complement (45222..45335)
/gene="rpl36"
CDS complement (45222..45335)

```

```

        /gene="rpl36"
        /codon_start=1
        /transl_table=11
        /product="ribosomal protein L36"
        /translation="MKVRASVRKICDNCRLIRRKRKIMVICSNPCHKQRQG"
gene      complement(45763..46092)
        /gene="infA"
CDS      complement(45763..46092)
        /gene="infA"
        /codon_start=1
        /transl_table=11
        /product="translational initiation factor 1"
/translation="MVSVGCFFKRKSTLDKCLFAKKKIPIIKLNSEKQSLIEMQGVVTQ
CLSNAMEFRVKLENGFQVLAHVSGKIRRNIRILLGDRVTVELSPYDLTRGRIVYRLRQ
NEQKTET"
gene      complement(47018..47407)
        /gene="rps8"
CDS      complement(47018..47407)
        /gene="rps8"
        /codon_start=1
        /transl_table=11
        /product="ribosomal protein S8"
/translation="MVNDTISDILTRIRNANLVKSQTVSIPLTRISQKISVILEKEGF
IESFQIFPSDEIALRLKYKGREKKPCITNLKRISKPGLRIYANHKEIPKILGGMGIVI
LSTSQGIMTDREARFYGIGGEILCSIW"
gene      complement(47464..48006)
        /gene="rpl5"
CDS      complement(47464..48006)
        /gene="rpl5"
        /codon_start=1
        /transl_table=11
        /product="ribosomal protein L5"
/translation="MAQRLKKLYFQNIIVPKLLEQFQYKNTLQAPRIQKIVINRGLGDA
SQNAKILDSSLKELTIIAGQKGVITRSKKAIAAGFKLRQKMPVGVSVTLRGERMYGFLD
RLINLALPRIRDQGINPKSFDGHNYSGLGLEEQLMFPEIDYDKIDQIRGMDISIITT
SKNDQEGLLKKEFGMPFKN"
gene      complement(48566..48934)
        /gene="rpl14"
CDS      complement(48566..48934)
        /gene="rpl14"
        /codon_start=1
        /transl_table=11
        /product="ribosomal protein L14"
/translation="MIQPQSYLNVADNSGARKLMCIRVLGGNQRQSANIGDVIIAVVK
DAIPNMPLKKSIVRAVIVRTCKGLRRENGMTIRFDDNAAVVLNKENNPRGTRVFGPI
ARELRDRNFTKIVSLAPEVL"
gene      complement(50647..51060)
        /gene="rpl16"
CDS      complement(50647..51060)
        /gene="rpl16"
        /codon_start=1
        /transl_table=11
        /product="ribosomal protein L16"
/translation="MLSPKRTKYRKQHRGRMRGKANRGNTIAFGDFGLQALEPCWITS
RQIESGRRAMTRYARRGGKLWIRIFDPKPIITMRPAETRMGSGKGSPEFWVAVVKGPKI
LYEMKGISESVARISAMRIAAYKMPIRTQFIMKEAS"
gene      complement(51797..52495)

```

```

        /gene="rps3"
CDS      complement (51797..52495)
        /gene="rps3"
        /codon_start=1
        /transl_table=11
        /product="ribosomal protein S3"
/translation="MGQKIHPLGFRLGITKKHRSQWFAKTNQYPQLVVEDSFIRKLLL
EKFADAGITHIEIQRKLDQIKIEIRAARPGILVGRDGANLDILKKLLEQKLTYYRLKT
IFNFKSSNVVPSKIETVKHSIQVAIHVTKLANPDSEAAFITDFLVEQLEKRVAFRRAV
RQAIQRAQRAGVSGIKIQVSGRLNGAEIARSEWVREGRVPLQTLRADIDYCYKTAKTI
YGLLGIKIIVVFKNN"
        gene      complement (53519..53797)
        /gene="rps19"
CDS      complement (53519..53797)
        /gene="rps19"
        /codon_start=1
        /transl_table=11
        /product="ribosomal protein S19"
/translation="MSRSLKKGPFVASHLLKKIEKLNLDLGEKKVVLTSRRASTIVPIM
IGHTIAVHNGREHIPIFVTDQMVGHKLGFEFLTRTFRGIKSDKKSQR"
        gene      complement (54218..55045)
        /gene="rpl2"
CDS      complement (54218..55045)
        /gene="rpl2"
        /codon_start=1
        /transl_table=11
        /product="ribosomal protein L2"
/translation="MGIRLYRAYTPGTRNRSVSEFNEISQTKPEKSLTYSVHRSKGRN
NRGVITSRHRGGGHKKLYRQIDFKRNKIGIIAKVATVEYDPNRRNARIALLLHYQDGEKR
YILYPRGLKLGDNIIADIEAPIVIGNSLPLFNMPLGTEVHNVELHPGSGGQLVRAAGT
VAQLVAKEGKFVTLRLPSGEVRLISKECWASIGQVGNVDVSNLTLGKAGRTRWLGRRP
TVRGVVMNPVDHPPHGGGEGRAPIGRSHPVTPWGRPALGQRTRQSKKYSNKLILKRRK"
        gene      complement (55815..56096)
        /gene="rpl23"
CDS      complement (55815..56096)
        /gene="rpl23"
        /codon_start=1
        /transl_table=11
        /product="ribosomal protein L23"
/translation="MMIDLVKYPVLTEKSVRLIENNQYTFDVEPKLNKIQIKKLIIEGL
FEVKVIAVNTHRPPRKKKRLGISQGYRSRYKRVIIITIKSGESIPLPDPT"
        gene      57668..58675
        /gene="cemA"
CDS      57668..58675
        /gene="cemA"
        /codon_start=1
        /transl_table=11
        /product="chloroplast envelope membrane protein"
/translation="MAKQAIEQTGLIPRSIIRTFDRFKQLLPGAEKLVIQEFRISRF
QVLVSVKCLVSLTIIPLIINYLLKNFLLIPVTENFWNRHQNEIFLNSYQQDRAFTEIK
KFEEKIYFEFLINDLNNFIFESDKSDQENTTLGESNKKIKYSLNSNFKELSEKDLAKK
QTTEEIKELYFESTMQKRKKIQLIESKNVIQEDLQEKIIELAIHYNQQSIIALTNLFG
DFITFGIISFLVIWMKPQIIILKSFLTESIYSLSDTTKSFLILGTDLLVGFHSPRGW
EIFLELILRHFGFPENQDFIYLFVATFPVLLDTPVKYWIFRYLNQISPSTVATYHSMI
E"
        gene      complement (58709..58782)
        /gene="trnT (ugu) "

```

tRNA complement (58709..58782)
 /gene="trnT (ugu) "
 /product="tRNA-Thr"
 /note="codons recognized: ACR"
 /anticodon=(pos:complement (58746..58748), aa:Thr)

gene complement (59078..59151)
 /gene="trnR (ccg) "

tRNA complement (59078..59151)
 /gene="trnR (ccg) "
 /product="tRNA-Arg"
 /note="codon recognized: CGG"
 /anticodon=(pos:complement (59114..59116), aa:Arg)

gene complement (60258..60827)
 /gene="ycf4"

CDS complement (60258..60827)
 /gene="ycf4"
 /codon_start=1
 /transl_table=11
 /product="hypothetical chloroplast RF4"
 /translation="MVDLYETQESDLIRRYIVVGSRRFSNYWWASVIFLGASGFLLTG
 ISSYLQFNLLPFIKSKDILFFPQGLVMCFYGILGFLFSIFLWLTLFWSVGGGFNEFNK
 KDGIVRIFRWGFPGKNRRIDL LYSIKDIEAIRVELKEGLNPRRTIYLRVKGKRDIPLT
 RIGQPMTLEEIEKQAAELAKFLQVSLEES"

gene complement (61379..61759)
 /gene="rpl19"

CDS complement (61379..61759)
 /gene="rpl19"
 /codon_start=1
 /transl_table=11
 /product="ribosomal protein L19"
 /translation="MKIKELVRTIEQTQLKQGLPKICVGD TVKVGVL IQEGNKQRVQP
 YEGTVIALHRAGLSTTITVRRIFQGIGVERVFTIHSPSIQDIEIVRAKVRKAKLYYL
 RNRIGKGT RLKEKFDNRTTNTDN"

gene complement (62580..63809)
 /gene="tufA"

CDS complement (62580..63809)
 /gene="tufA"
 /codon_start=1
 /transl_table=11
 /product="translational elongation factor Tu"
 /translation="MARAKFERKKPHVNIGTIGHVDHGKTTLTAAITMALATSSGGKG
 RKYADIDSAPEEKARGITINTAHVEYETENRHYAHVDCPGHADYVKNMITGAAQMDGA
 ILVVSGADGPMPTKEHILLAKQVGVPNVVVFLNKEDQVDDEELLELEVELEVRETLDN
 YEFPGDEIPIVSGSALLALEALTENPTLTRGENKWVDKIFDLMDQVDSYIPTPERDTD
 KSFLMAVEDVFSITGRGT VATGRVERGSVKVGESIEIVGLRKRTRTTT VTGLEMFQKTL
 EESVAGDNVGVLLRGIQKVDI ERGMVLAKPGTITPHTKFESQVYVLTKEEGRHT PFF
 SGYRPQFYVRTT DVTGKIESFRADDDTATQMVM PGDRIKMNVELIQPIAIEKGMRF AI
 REGGR TVGAGVVSTIIE"

gene complement (64681..65151)
 /gene="rps7"

CDS complement (64681..65151)
 /gene="rps7"
 /codon_start=1
 /transl_table=11
 /product="ribosomal protein S7"
 /translation="MSRRRTPKKRFLSPDPVYD SHLVHMIVNHLMKKGGKSLAYRMFY
 ESMKQIGEVTKQDPVKIIEQAVRNATPLVEVKARRVGGSTYQVPLEVTPERGTALAIR

WILTSCRSRSGRNMISKLSNEFLDASKNMGNAIRKRDELHRMAEANKAFTKYRF"

gene complement (65802..66173)
/gene="rps12"

CDS complement (65802..66173)
/gene="rps12"
/codon_start=1
/transl_table=11
/product="ribosomal protein S12"

/translation="MPTIQQLVRSARKNISKKTKSPALQACPQRRGVCTRVYTMTPKK
PNSAFRKVARVRLSSGFVETAYIPGIGHNLQEHSVVLRGGRVKDLPGVRYHIVRGTL
DTSGVKDRVQGRSKYGVKRAK"

gene complement (67999..68124)
/gene="psaJ"

CDS complement (67999..68124)
/gene="psaJ"
/codon_start=1
/transl_table=11
/product="subunit IX of photosystem I"

/translation="MKYFTTYLSTAPVVALIWFTLTAGLLIEINRFFPDPLVFSF"

gene 68856..68929
/gene="trnP(ugg) "

tRNA 68856..68929
/gene="trnP(ugg) "
/product="tRNA-Pro"
/note="codons recognized: CCN"
/anticodon=(pos:68890..68892,aa:Pro)

gene 69223..69295
/gene="trnW(cca) "

tRNA 69223..69295
/gene="trnW(cca) "
/product="tRNA-Trp"
/note="codon recognized: UGG"
/anticodon=(pos:69256..69258,aa:Trp)

gene complement (69387..69671)
/gene="rps18"

CDS complement (69387..69671)
/gene="rps18"
/codon_start=1
/transl_table=11
/product="ribosomal protein S18"

/translation="MTFSNRISKNKTLQSQRIPPIIQRKSQSKLIPFQGTIDYKNIVL
LRKFITSEGKILPRRINGLTAKQQRHMAKAIKNARMIGLLPFINKSQPSQ"

gene complement (70287..70631)
/gene="rpl20"

CDS complement (70287..70631)
/gene="rpl20"
/codon_start=1
/transl_table=11
/product="ribosomal protein L20"

/translation="MTRVKRGNIAARKRRKVLKITEGFRGSSSILFRPANQQKIKALK
SSYRDRNQRKREFRRLWITRNMNAAVRLYGLNYSQFLHGLKSNMFLNRKMLAQLAIRD
HEAFRQLVQSLN"

gene complement (71648..71720)
/gene="trnE(uuc) "

tRNA complement (71648..71720)
/gene="trnE(uuc) "
/product="tRNA-Glu"

```

        /note="codons recognized: GAR"
        /anticodon=(pos:complement(71684..71686),aa:Glu)
gene      complement(71808..71881)
        /gene="trnMf(cau)"
tRNA     complement(71808..71881)
        /gene="trnMf(cau)"
        /product="tRNA-Met"
        /note="codon recognized: AUG"
        /anticodon=(pos:complement(71845..71847),aa:Met)
gene      complement(73613..73915)
        /gene="rps14"
CDS      complement(73613..73915)
        /gene="rps14"
        /codon_start=1
        /transl_table=11
        /product="ribosomal protein S14"
/transl="MAKKSMVEREKKRQQLVMKYSKKRLTLKAQMKQASFLKLTIQ
RQLQKLPNNSNPVRLHNRCLITGRPKGFFRDFGLSRHVLREMAHEGFLPGITKSSW"
gene      75599..76567
        /gene="petA"
CDS      75599..76567
        /gene="petA"
        /codon_start=1
        /transl_table=11
        /product="apocytochrome f of cytochrome b6/f complex"
/transl="MKKIFCYKIFYFLNVSKVILFSLALLAFDLTITNPAQAYPIFAQ
QNYPSPREANGRIVCANCHLAQKPVELEVPOQSVLPDPTVFEAVVKIPYDQOIKQVLGNG
KKGPLNVGAVLILPEGFELAPADRIPEEMKTKVGKLYFQPYSDDKNILVVGPLPGKK
YSEMVPILSPDPAKKNKNSYLKYPYIYLGNGRGRGQVYPDGSKSNNTVYNASAAGKIS
QIEPIEKGGYNITTIETNNGEQKIDKIPAGPELIVQVGVQSVQNDQPLTNNPNVGGFGQM
DTEIVLQNPTRVQGLIVFFGFGVIVAQIFLVLKQKQFEKVQLAEMNF"
gene      76988..77083
        /gene="petL"
CDS      76988..77083
        /gene="petL"
        /codon_start=1
        /transl_table=11
        /product="subunit VI of cytochrome b6/f complex"
/transl="MITIISYIGLLTGVLFFTLFVYVSLKIKLI"
gene      77231..77344
        /gene="petG"
CDS      77231..77344
        /gene="petG"
        /codon_start=1
        /transl_table=11
        /product="subunit V of cytochrome b6/f complex"
/transl="MVEPLLSGIVLGLVPVTLAGLFVTAYLQYRRGDQLNL"
gene      78726..79784
        /gene="chlI"
CDS      78726..79784
        /gene="chlI"
        /codon_start=1
        /transl_table=11
        /product="magnesium chelatase subunit of
protochlorophyllide reductase"
/transl="MSDIITLDQSRPIFPFTAIVGQEEMLSLILNVIDPKIGGVMIM
GDRGTGKSTTVRALVDLLPEIKVVLNDPFDNSDPEDPELMSEEVREKLKNEKIEVTNR

```

KILMVDLPLGATEDRVCGTIDIEKALTEGVKAFEPGLLAKANRGILYVDEVNLLDDHL
VDVLLDSAASGWNTVEREGISISHPARFILVSGNPEEGELRPQLLDRFGMHAQIGTV
KESNLRVQIVEQRASFDEAPLEFRKNYKDSQKELTEKLTQARKNLKNVEIDYDFRIKI
SQICSELNVDGLRGDIVTNRASKAIAAFENRTEVTPEDIFRVIPLCLRHLRKRDPLES
IDSGDKVRETFFKKVFGYQ"

```

gene      80563..80635
          /gene="trnR(ucu)"
tRNA     80563..80635
          /gene="trnR(ucu)"
          /product="tRNA-Arg"
          /note="codons recognized: AGR"
          /anticodon=(pos:80597..80599,aa:Arg)
gene     complement(81768..81850)
          /gene="trnY(gua)"
tRNA     complement(81768..81850)
          /gene="trnY(gua)"
          /product="tRNA-Tyr"
          /note="codons recognized: UAY"
          /anticodon=(pos:complement(81814..81816),aa:Tyr)
gene     82173..82256
          /gene="trnL(uaa)"
tRNA     82173..82256
          /gene="trnL(uaa)"
          /product="tRNA-Leu"
          /note="codon recognized: UUA"
          /anticodon=(pos:82207..82209,aa:Leu)
gene     complement(82976..83584)
          /gene="rps4"
CDS      complement(82976..83584)
          /gene="rps4"
          /codon_start=1
          /transl_table=11
          /product="ribosomal protein S4"
/transl="MSRYRGPRLRIIRKLGELPGFTSKSSARQTPPGQHGTSSVKKLSQ
YGIRLQEKQKLRFNYGITESQLIRYVRKARKIKGSTGEILLQLLEMRLDNIVFRLGIA
PTIAAARQLIGHGHIIVNNEKVTIPSYECKPKDTISVSKKKHSRNVVSNFFETSNTI
IPIHLSFNKETLIGTVNNIVNRQSVNLLNELLIVEYYSRKV"
gene     85822..87276
          /gene="atpB"
CDS      85822..87276
          /gene="atpB"
          /codon_start=1
          /transl_table=11
          /product="CF1 beta subunit of ATP synthase"
/transl="MSISVDTKNVGRITQIIIGPVLDAFPAGKMPNIYNAIIVRGKNE
AGDEISVTCEVQQLLDHCVRAVSMNATDGLMRGMDVVDTGKPLTVVPGPATLGRIFN
VLGEPVDNLGPVDAKEGLPIHRSAPLFDLDTKLSIFETGIKVVDDLAPYRRGGKIGL
FGGAGVGKTVLIMELINNIKAHGGVSVFGGVGERTREGNDLYMEMKESKVINEDNIS
ESKVALVYGQMNPPGARMRVGLTALTMAEYFRDINKQDVLLFIDNIFRFVQAGSEVS
ALLGRMPSAVGYQPTLATEMGGQERITSTKDGSITSIQAVYVPADDLTDPAATTFA
HLDATTVLSRGLASKGIYPAVDPLDSTSTMLQPWIVGEEHYECAQNVKETLQRYKELQ
DIIAILGLDELSEEDRLVVARARKIERFLSQPFFVAEVFTGSPGKYVSLAETIQGFNL
ILSGELDDLPEQAFYLVGTLDEAIIKAASLQEK"
gene     88316..88720
          /gene="atpE"
CDS      88316..88720
          /gene="atpE"

```



```

/codon_start=1
/transl_table=11
/product="CF1 epsilon subunit of ATP synthase"
/translation="MSLQVCIMTPDRIFWNDQAEIILPTNTGQMGVLSNHAPLITAL
DIGVMLIRANNDWISIALMGGFALVKQNQITVLVNEAESSKTIEAEEAENSFNKAKIK
LEEAEGQKQKVEANFVFKRARARYQVVKQSSN"
gene 89686..91113
      /gene="rbcL"
CDS 89686..91113
     /gene="rbcL"
     /codon_start=1
     /transl_table=11
     /product="large subunit of Rubisco"
/translation="MAPQTETKTAAGFKAGVKDYRLTYTTPDYQVKDTDILAAFRMTP
QPGVPVEECGAAVAESSTGTWTTVWTDGLTSLDRYKGRCYDIEPVPGEENQYIAYVA
YPLDLFEEGSVTNLFTSIVGNVFGFKALRALRLEDLRIPPSYSKTFQGPPIHQVERD
KLNKYGRGFLGCTIKPKLGLSAKNYGRAVYECLRGGLDFTKDDENVTSQPFRWRDRF
LFVAEAIYKSAETGEIKGHYLNCTAGTSEMLKRAQVAKELGVPIVMHDYLTGGFTS
NTSLSHYCRDNGLLLIHRAHHAVIDRQKNHGIHFRVLAKALRLSGGDHLHSGTVVVK
LEGEREVTLGFDLMRDDYIQKDRSRGIYFTQDWASLPGVMPVASGGIHWVHMPALVE
IFGDDACLQFGGGTLGHPWGNAPGAAANRVALEACTQARNEGRDLAREGGDVIRAACK
WSPELAAACEVWKEIKFEFETIDTL"
gene complement(91529..91612)
      /gene="trnI(cau)"
tRNA complement(91529..91612)
     /gene="trnI(cau)"
     /product="tRNA-Ile"
     /note="codon recognised: AUA; C in the first position of
the anticodon assumed to be post-transcriptionally
modified to lysidine, which A rather than G"
     /anticodon=(pos:complement(91576..91578),aa:Ile)
gene 93112..93222
      /gene="psaI"
CDS 93112..93222
     /gene="psaI"
     /codon_start=1
     /transl_table=11
     /product="subunit VIII of photosystem I"
     /translation="MAASYLPSILVPLVGLVFPVAVAMASFLYIEQEEIN"
gene 94462..94534
      /gene="trnV(uac)"
tRNA 94462..94534
     /gene="trnV(uac)"
     /product="tRNA-Val"
     /note="codons recognized: GUN"
     /anticodon=(pos:94495..94497,aa:Val)
gene complement(95061..95567)
      /gene="ycf3"
CDS complement(95061..95567)
     /gene="ycf3"
     /codon_start=1
     /transl_table=11
     /product="hypothetical chloroplast RF3"
/translation="MPRSQRNDNFIDKTFVADILLKVLPTSNREKQAFSYYRDGMS
AQSEGEYAEALQNYEAMRLETDAYDRSYILYNIIGLIHTSNGEHGRALEYYYQALERN
PSLPQALNNIAVIYHYRGEQAIENGQSEISKMLFDKAADYWKEAIRLAPTNYIEAFNW
LKMTGREV"

```

```

gene          complement (96453..96569)
              /gene="psbI"
CDS           complement (96453..96569)
              /gene="psbI"
              /codon_start=1
              /transl_table=11
              /product="I polypeptide of photosystem II"
/translation="MLTLKIFVYTVVSFFVSLFIFGFLSNDPGRNPGQKDLLD"
gene          complement (97580..98302)
              /gene="cysA"
CDS           complement (97580..98302)
              /gene="cysA"
              /codon_start=1
              /transl_table=11
              /product="probable transport protein"
/translation="MYKLIDYKKFMSILVEDLSKNFNGFQALDHVNLEIKTGSLVALV
GPSGSGKSTLLRRIIAGLDRPNKGRIWLAGRDSTFLSIQEREIGFVFQNYALFKHMTVY
ENIAFGLQVRKTSLSIINNRVKQLLQLIQLENFADRYPSQLSGGQRQRVALARALAI
PKVLLLLDEPFALDVKVRKELRNWLRRLHEEVPVTTIFVTHDHQEALEVANEIVVFQK
GRIEQLGNPQDIYHQLSINYLP"
gene          complement (98832..100262)
              /gene="accD"
CDS           complement (98832..100262)
              /gene="accD"
              /codon_start=1
              /transl_table=11
              /product="beta subunit of acetyl-CoA carboxylase
carboxytransferase"
/translation="MSILAWIEDQRKLLNAPKYNHPTSDGSQGLWTRCDHCGVILY
IKHLKENQRVCFGCGYHLQMNSQERIEHLIDSGTWRPLDETLSPCDPLEFRDQKAYTE
RLKEAQERTGLQDAVQTGTGMLDGIPIALGVMDNFNMGSGMSVVGEEKITRLIEYATQ
EGLTLILVCASGGARMQEGILSLMQMAKISAALQVYQSCANLLYISVLTSPTTGGVTA
SFAMLGDLIFAEPKALIGFAGRRVIEQTLQEQLPDDFQTAEYLLHHGLLDLIVPRSFL
KQALSETITLYKDAPFKRFGSIPHGIQQPLNFINEEKIRRRWSNWHNLRLSGPEKIIK
KETS LNQPEFVALETNLNQDET KFFYNNEHSYRQILTSFQNMFFCLQNKGFVSTQDKL
SNYGGSEFEPLNQPNWYSFEKTNSTFDNFINTSILLSTNKVIQWRALHNTKISSQKT
NFLKNLINKKNKNFVYRTISKSNKAR"
gene          102008..102358
              /gene="orf116"
CDS           102008..102358
              /gene="orf116"
              /codon_start=1
              /transl_table=11
              /product="hypothetical protein"
/translation="MTIEKKVNPENEVKRTSNVVSLLIEIKIKSIDIFIENRLIIDPS
SKDTLKNMYT DYVLSCDSCVPLGKQYRKQLEKAFQDRNAKVYFVSNTRGVRLEGVTL
KKEEVRGTTTSSFD"
gene          complement (102975..103047)
              /gene="trnT (ggu) "
tRNA         complement (102975..103047)
              /gene="trnT (ggu) "
              /product="tRNA-Thr"
              /note="codons recognized: ACY"
              /anticodon=(pos:complement (103012..103014),aa:Thr)
gene          103363..109299
              /gene="ftsH"
CDS           103363..109299

```

```

        /gene="ftsH"
        /codon_start=1
        /transl_table=11
        /product="cell division protein"
/translation="MEKFLPTRSQRLKTTNIIFNISIQLEVQLLKQDREFYNPNKFSSQ
IQKSYICNRRLELIGYLALYSVNSTIFYFFPLWFFMFPFYLFNSFKIWRSAWLKISVS
ISEYKLTWTKNSLTFTLKGAAANLQVVISFFPFIFGISYIIYSRYQNQVFSYFLHKNLPG
IYPSYPKLTWETFQHITSSQLKLDPTQKTCVLDYEISNKHFSRQSLTNQIEILKDFK
SVKYQLTDFDEGVAQEEDFRSKQNDKSQKNYTKSYFQLTKEARGQKHSGFQKQDFYS
DSFLIKLNSKWANKKNQGYVGI FPKKSFLGSFSVSPNFFKKPKISLGLALPQVNEQS
TEATPHNNFNSIGTTWHLVSPKDFLTFSSDGLFQTNQEKNVVKIENPILTQRTNLKFN
LNTNSSKANFFNNQPPSKIMKNYQWIFYFNLDELPKKLNSITTTNISALNLTGSNLNLK
KTPLLPIKKYSEERQKFLTNTLANTDSPVALPGTARAQAEYKFNQKSKKVLFSKSNF
LGNKIDPQINLTSQNKLSRILHGFSNQYQAQSRVLDNRANTINKVGLKSESIFFNKL
HLLQNELRQFFYNTQLKPNFLLLEPTLNPSSSQSLLNLVIGKDFNFDETKMPKNSDFS
KILEQSNNEILLTNPHLKQLTSSSKSEDEYLEQILQILQDFVSFKANLATPPRAMSGY
DFPDSTNTEVRSLILQFFYQKKLPIIKSLITTSQMKNKSNLFSPSFKVELSPSFTSYS
NYNYPQSKVPTLKLNYRPTFLEGVKKTLYQGGVLRRETFTHDFSLRNKEEIRRWIKPF
LSSDNPLTDRRAFFGHKLMQFDVPAFAPTGLTPTVALPGTAREDOGYQAKLDYQGH
DGFSLKKDSDSLAPLNSNTVRIELESIKKNI SPNKFLKLQKDI LKLKNVQTHFHLDPK
SKSIKLPVVAFTEEFQVPYLT KSKWHTILEKLRQAIAKQLEDKSEEEKSELEKKLQIN
VPVVRIRNPKQQA IQWPLNKL DYENLNDV FVSSPVSSF KHDNWT TAKDLTPSDQGYQA
KLDYQINSISNKL GDSQK NLL EKYL NTRY SNKLAALTNSSILFEDI IYEKLNQKENT
FLGLNRRVKPVTEKNNLIRINYHYVPSVQMVLSENATQKKIFGNVYKKAFTVYQNIIE
NRYESFNYSYKRKNYVKNTFLNNSNVFYQSWEPITSGSWMMITQLSFGLVVLQIFQDV
YRKYGKELISYVFDLLAALGIMDPSLKEELGIDDRRTGFRIIRKVSKRFRDIAGIDGI
LPELGEIVWFLRNSGRSFKVGNII PKGILLVGS PGTGKTL LVQAIAGEAEV PVLVQSG
SSLNDPEQEEGGAQQLTNLFEQARQMAPCIVFIDEIDALGEKRENVIONPMGTDEVIE
AIQESSESNPASLPSNTKGLRDLTSTPFI PKPKIKFNKTQINNLDGDQQNVLDVSVQE
EGAFRDFESYENSSLGIIQOSIDKQEAQEQRLRLLMQFLIEMDGLRARKGVIVIGATN
RPDVLDLALTRPGRFDQILHLGLPEKQKRIEILKLYSKNLGTENTISWDYLANRTIGF
SAADLAAVMNESSMKAILSETVHTIQTI EQGIESITSYSTDKPKLQIPNSIDPFLISR
LAYYQAGKAVVHTLLLQHPSVTVLHLWPRQKNARQTYISTIIQNKFLKISRKLELESR
IVGLYAGKAAEFLVLSNLSKSEIPKSFSSNKPWQSDIGIEDLNFATSLVQSMISQW
YFYSKDLVVRKSNPIFSQRNLQEIQELDTLELFNQLAIETETEIIQKTRSSGFNRDFQ
KWSIRPWWQTQITRQIGFLDPAYDDWYRIYLPDPEESERNDEWI PPDEFYHNNSNLTD
LAVNSVNSSVNWNDLYAMDRDYIYHGLLLSCFN TAFSILDQNRLELLDYFAAYLMRNDI
LRQHEINEI LYKFKPESSLSKNT ELIHKNKILNENKTNEQKLF LQKNKKTRVIEKAWG
QYSRRTIFRFFNFDSILSKL"
    gene          110428..110507
                  /gene="trnL(caa)"
    tRNA          110428..110507
                  /gene="trnL(caa)"
                  /product="tRNA-Leu"
                  /note="codon recognized: UUG"
                  /anticodon=(pos:110461..110463,aa:Leu)
    gene          complement(111605..111733)
                  /gene="psbJ"
    CDS           complement(111605..111733)
                  /gene="psbJ"
                  /codon_start=1
                  /transl_table=11
                  /product="J protein of photosystem II"
/translation="MSNTGTTGRIPLWLVGTVVGTAAALALLSIFFYGSYVGLGSSL"
    gene          complement(112634..112750)
                  /gene="psbL"
    CDS           complement(112634..112750)

```

```

        /gene="psbL"
        /codon_start=1
        /transl_table=11
        /product="L protein of photosystem II"
/translation="MAKPNNPKQSVELNRTSLYWGLLLIFVLAVLFSSYIFN"
    gene        complement(112773..112901)
                /gene="psbF"
    CDS         complement(112773..112901)
                /gene="psbF"
                /codon_start=1
                /transl_table=11
                /product="cytochrome b559 beta subunit of photosystem II"
/translation="MATKKSSYTYPIFTVRWLAVHALAVPTVFFLGSITAMQFIQR"
    gene        complement(114092..114337)
                /gene="psbE"
    CDS         complement(114092..114337)
                /gene="psbE"
                /codon_start=1
                /transl_table=11
                /product="cytochrome b559 alpha subunit of photosystem
II"
/translation="MSGATGERPFS DILTSIRYWVIHSITIPSLFIAGWLFVSTGLAY
DVFGSPRPNEYFTEDRQETPLITDRFNALEQVKQLAE"
    gene        complement(115651..115722)
                /gene="trnH(gug)"
    tRNA        complement(115651..115722)
                /gene="trnH(gug)"
                /product="tRNA-His"
                /note="codons recognized: CAY"
                /anticodon=(pos:complement(115688..115690),aa:His)
    gene        complement(115909..115981)
                /gene="trnG(ucc)"
    tRNA        complement(115909..115981)
                /gene="trnG(ucc)"
                /product="tRNA-Gly"
                /note="codons recognized: GGR"
                /anticodon=(pos:complement(115946..115948),aa:Gly)
    gene        117508..117696
                /gene="psbZ"
    CDS         117508..117696
                /gene="psbZ"
                /codon_start=1
                /transl_table=11
                /product="Z protein of photosystem II"
/translation="MTLIFQLTLFALIGVSFLLVVGVPVAFASPDGWTQNKSSIFSGL
GLWFFLVFAVGLNSFVI"
    gene        118874..118945
                /gene="trnQ(uug)"
    tRNA        118874..118945
                /gene="trnQ(uug)"
                /product="tRNA-Gln"
                /note="codons recognized: CAR"
                /anticodon=(pos:118906..118908,aa:Gln)
    gene        complement(119229..119534)
                /gene="orf101"
    CDS         complement(119229..119534)
                /gene="orf101"

```

```

/codon_start=1
/transl_table=11
/product="hypothetical protein"
/translation="MKNWKMLKIGGYLIFWLRGASNIPPAIPTQIREGCIIDKIMGVS
FDGEKKYKILINKSLFLNAIYNRDLVPTNLGGQTESYLKNFVRIPLRGEGNKLSLPS"
gene complement (120177..120389)
/gene="ycf47"
CDS complement (120177..120389)
/gene="ycf47"
/codon_start=1
/transl_table=11
/product="hypothetical chloroplast RF47"
/translation="MLNIIQLTFAILVIVLIVPQTPTENALLRIFYETRVFPNSNYGQ
TKKFLNLLTWVCIFIFLILFTTNVFS"
gene 120494..120566
/gene="trnF(gaa)"
tRNA 120494..120566
/gene="trnF(gaa)"
/product="tRNA-Phe"
/note="codons recognized: UUY"
/anticodon=(pos:120527..120529,aa:Phe)
gene 121585..123936
/gene="ycf62"
CDS 121585..123936
/gene="ycf62"
/codon_start=1
/transl_table=11
/product="hypothetical chloroplast RF62"
/translation="MKIQKNSNQIIDLVNKHVTITKQNLVIPTQFILIAISGGQDSI
CLFFILLQLKNQWKWDFGVLYCNHFQWQTDSEFYTSALIFRVGFIFSIPAYLNLPSENIF
TEQKSRNWRYYQFKRLSYFYNYDIIFTGHTATDRIETIILNLIRGTSTRGLSTLNWIR
PISRQSFYFSRSFCITKQKLANLTDNRNKVFKNQDLPLNDVLFVNFNHLFLKLQNN
LFTFYKQINNESLKRRNQFEQNIVQCKIQVVLQKTKNFNQSPFGLFLIAETFKRCV
GVQNKKRSLFSMNNI WVYDDFFFLTLEKFRIPNEVYELKFLKYYYLTDKECQKCFIRP
LYFDKSQV LVERELNISTIIILNYTNLNQSNFYRLKWNLWHKKNKNTKSGFINLVNSS
SKAKKYKFVTKKIKKIKLVTRAGLVKVPVKLCFPKIKSQILFKSYNAQNSSFII RQKV
SVFNFKPIFFSKKISKIIKFKVKS YWTQFKQNFLELEFLKLNPGAFNTALRNSLNLK
NTESFNLSAKTMNQNI SKSIIITLNLFDKQPQISNKFKLFLPEKENLKQKDFFNKKFFI
LEVITFPPNFSKDSVSVYVYQMN SFNQIRDLIVFLFFRQALVSNFIHIIRYNNQNL SFI
SPNTLLFGSCQLTNKQLCVFRPVLFLNRFQVQKQLCTFWNLPVYPDYTNNQLTYRNRV
RKQLPLLRFFFNPQIDKLFQFAEIANTEQIYLNSVSTRLQQEFQIKKTKSFELNMS
LFHFLPLAIRRLLKQFLDDSLTKKVNFFQIDFLLNLINRKKK"
gene complement (125162..125233)
/gene="trnK(uuu)"
tRNA complement (125162..125233)
/gene="trnK(uuu)"
/product="tRNA-Lys"
/note="codons recognized: AAR"
/anticodon=(pos:complement (125199..125201),aa:Lys)
gene complement (125612..125707)
/gene="psaM"
CDS complement (125612..125707)
/gene="psaM"
/codon_start=1
/transl_table=11
/product="M polypeptide of photosystem I"
/translation="MILSDSQIFIALFTALVTGILALRLGIALYK"

```

```

gene      complement (125790..125891)
          /gene="ycf12"
CDS       complement (125790..125891)
          /gene="ycf12"
          /codon_start=1
          /transl_table=11
          /product="hypothetical chloroplast RF12"
          /translation="MNLEIIFQLASLVFILAAGPIVIILLASRGGNL"
gene      complement (126916..127047)
          /gene="psbK"
CDS       complement (126916..127047)
          /gene="psbK"
          /codon_start=1
          /transl_table=11
          /product="K protein of photosystem II"
/translacion="MTFLLAKLPEAYAPFDPIVDVLPPIIPVLFILLAFVWQASVSFR"
gene      128440..129966
          /gene="chlB"
CDS       128440..129966
          /gene="chlB"
          /codon_start=1
          /transl_table=11
          /product="ChlB subunit of protochlorophyllide reductase"
/translacion="MKLAYWMYAGPAHIGTLRVASSFKNVHAIMHAPLGDDYFNVMS
MLERERDFTPVASIVDRHVLARGSQEKVVENITRKDKKEERPDLIVLTPCTCTSSILQE
DLQNFVDRAAIESESDVILADVNHYRVNELQAADRTLEQVVRFYIEKARRSNTLELEK
TEKPSANIIGIFTLGFHNQHDSRELKRLNLDLGVINEVIPEGGLVTNLRNLTKAWFN
LIPYREVGLMTATYLEKEFNIPYVTTTTPMGIVDTAECIQEIKEILKKNYYKNFDEYI
EQQTKFVSAAWFSRSIDCQNLDTGKKAVVFGDATHASSMTKILAREMGIRVSCAGTYC
KHDADWFREQVQGFCEVLITDDHTQVGDMIARIEPSAIFGTQMERHIGKRLDIPCGV
ISAPVHIQNFPLGYRPFLGYEGTNQIADLVYNSFTLGMEDHLLLEIFGGHDTKEVITKS
LSTDSILNWAPDGLAELNKIPGFVRGKIKRNTEKFAREKNINEITIEVMYAAKESVGA
"
gene      131492..133747
          /gene="psaA"
CDS       131492..133747
          /gene="psaA"
          /codon_start=1
          /transl_table=11
          /product="P700 apoprotein A1 of photosystem I"
/translacion="MTISPPEREAKKVKIVDRNPVETSFEKWAKPGHFSRTLSKGPT
TTTWIWNLHADAHDFDSHTSDLEEISRKVFSAHFGQLGIILIWLSGMYFHGARFSNYE
AWLTDPTHIKPSAQVWVPIVQGEILNGDVGQGGFQGIQITSGFFQLWRASGITSELQLY
STAIGGLVLASAMFFAGWFHYHKAAPKLEWFQNVESMLNHHLAGLLGLGSLAWAGHQI
HVSPLPINKLLDAGVDPKEIPLPHEFLLNRDLIAQLYPSFQKGLAPFFFTINWAEYSDFL
TFKGGGLNPVTGGLWLTDTAHHHLAIAVLFLVAGHMYRTNWGIGHSMKEILEAHKGPFT
GEGHKGLYEILTTSWHAQLAINLALFGSLSILVAHHMYSMPYPYLATDYATQLSIFT
HHMWIGGFCIVGAGAHAAIFMVRDYDPTNNYNNLLDRVIRHRDAIISHLNWVCIFLGF
HSFGLYIHNDTMSALGRPQDMFSDTAIQLQPVFAQWIQNTHFLAPELTAPNALAATSL
TWGGDVVAVGGKVAMMPSISLGTADFLVHHIHAFTIHVTVLILLKGVLYARSSRLIPDK
ANLGRFRPCDGPGRGGTCQVSAWDHVFLGLFWMYNSLSIVIFHFSWKMQSDVWGTVNA
SSVSHITGGNFAQSANTINGWLRDFLWAQSSQVIQSYGSALSAYGLMFLGAHFVWAFS
LMFLFSGRGYWQELIESIIWAHNKLVAPAIQPRALSITQGRAVGVVAHYLLGGIATTW
SFFLARIIAVG"
gene      134318..136522
          /gene="psaB"
CDS       134318..136522

```

```

        /gene="psaB"
        /codon_start=1
        /transl_table=11
        /product="P700 apoprotein A2 of photosystem I"
/translation="MATKFPKFSQGLAQDPTTRRIWYGIATAHDFESHDMTEENLYQ
KIFASHFGQLAIIFLWTAGNLFHVAWQGNFEQWGRDPLHVRPIAHAIWDPHFGQPAVE
AFTRGGASGPVNIISTSGVYQWWYITIGLRSNQDLYNGSIFLILGAGLFLFAGWLHLQPK
FQPSSLWFKNAEARLNHHLSSGLFGVSSSLAWTGHLVHVAIPEESRQHVRRWNNFLTTLPH
PQGLGPFPSGNWAAAYAQNPDSSNHIFNTAEGSGNAILTFLGGFHPQTQSLWLTDIAHH
HLAIAVVFILAGHMYRTNFGIGHSMREILDAHTPPAGGLGAGHKGLFDTVNNSLHFQL
GLALAAVGTVCSLVAQHIYSLPPYAFLAQDFTAQASLYTHHQYIAGFILCGAFAHGAI
FFIRDYDPEQNKGNVRLARILDHKEAIIISHLSSVSLFLGFHTLGLYVHNDVVQAFGTPE
KQILIEPVFAQWIQAAGKTLYGFDLLSSSASPFTASQSLWLPGWLDAINNNTNSL
FLTIGPGDFLVHHAIALGLHTTTLILVKGALDARGSKLMPDKKDFGYSFPCDGPGRGG
TCDISAWDAFYLAVFWMLNTIGWVTFYWHWKHLGIWQGNVNFNESSTYLMGWLRDYL
WLNSSQLINGYNPFGMNSLSVWAWMFLFGHLIYATGFMFLISWRGYWQELIETLAWAH
ERTPLANLVRWRDKPVALSIVQARLVGLAHFSAGYVLTAAFLIASTSGKFG"
    gene          137695..138753
        /gene="psbD"
    CDS          137695..138753
        /gene="psbD"
        /codon_start=1
        /transl_table=11
        /product="D2 reaction center protein of photosystem II"
/translation="MTIAIGKTEEKRGLFDLVDDWLRDRFVFGWSGLLLFPTAYFA
LGGWLTGTTTFVTSWYTHGLASSYLEGCNFLTAAVSTPANSMGHSLLFLWGPEAQGDFT
RWCQLGGLWTFVALHGSFALIGFMLRQFEIARSVKLRPYNAIAFSAPIAVFVSVFLIY
PLGQSGWFFAPSPFGVAIIFRIFLFFQGFHNWTLNPFHMMGVAGVLGAALLCAIHGATV
ENTLFEDGDGANTFRAFNPQAEETYSMVTANRFWSQIFGVAFSNKRWLHFFMLFVPV
TGLWMSAIGVVGLALNLRAYDFVSQEIIRAEDPEFETFYTKNILLNEGIRAWMAAQDQ
    PHEKLVFPPEEVLPRGNAL"
    gene          138659..140122
        /gene="psbC"
    CDS          138659..140122
        /gene="psbC"
        /codon_start=1
        /transl_table=11
        /product="CP43 chlorophyll apoprotein of photosystem II"
/translation="MKVVFVHGWLKISLMKNLYSLRRFYHVETLFGNTLTVGGRDQES
TGFAWWAGNARLINLSGKLLGAHVAHAGLIVFWAGAMNLFVAHFVPEKPMYEQGLLIL
LPHLATLGYGVGPGGEVIDTFPYFVSGVLHLISSAVLGFGGVYHSLIGPETLEESFPF
FGYVWKDKNKMTTILGIHLVILGIGAWLLVWKALYFGGVYDTWAPGGGDVRLITNPTI
SPAVIFGYILKSPFGGDGWIVSVDNMEDIIGGHIWIGTLNILGGVWHILTKPWAARR
AFVWSGEAYLSYSLGAIISLMGFIACCMSWFNNTAYPSEFYGPTGPEASQSQAFTFLVR
DQRLGANVASAQGPTGLGKYLMSPTGEIIFGGETMRFWDFRGPWLEPLRGPNGLDLN
KLKNDIQPWQERRAAEYMTAPLGLSLNSVGGVATEINAVNFVSPRSWLACSHFCLGFF
FFVGHVHAGRARAAAAGFEKGIDRDNEPVLMSRPLD"
    repeat_region 141653..168988
        /note="copy B of inverted repeat region (IRB)"
        /rpt_type=inverted
    gene          142027..143088
        /gene="psbA"
    CDS          142027..143088
        /gene="psbA"
        /codon_start=1
        /transl_table=11
        /product="D1 reaction center protein of photosystem II"

```

/translation="MTAILERRESASLWARFCEWVTSTENRLYIGWFGVLMIPTLLTA
 TSVFIIAFIAAPPVDDIGIREPVSGLLYGNNIIISGAVVPTSNAIGLHFYPIWEAASL
 DEWLYNGGPPYQLIVCHFFLGICCYMGREWELSFRLGMRPWIAVAYSAPVAAATAVFII
 YPIGQGSFSDGMPLGISGTFNFMIVFQAEHNILMHPFHMLGVAGVFGGSLFSAMHGSL
 VTSSLIRETTENESANAGYKFGQEEETYNIVAAHGYFGRLIFQYASFNNSRSLHFFLA
 AWPVVGIIWFTALGISTMAFNLNGFNFNQSVVDSQGRVINTWADIINRANLGMEVMHER
 NAHNFLDLASVEAPSINA"

gene complement(144242..144859)
 /gene="orf205"
 CDS complement(144242..144859)
 /gene="orf205"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein"

/translation="MLYVYIRLLKEIKKMIQISDFVLEFRDFILLSWPSLKKLGDDDR
 PYYCKENCIEAIWELLLEPLLATKLQTTVYIQYWLSGAEIYFRSPRAFFPDVKKSTH
 EIFCLPKNNEFFFEFYKKKIFLNKPGALDPGLRSSRCLTLTQGENPEQNYFQFSHF
 VTLSENESMPDLEKPPFDYAVCNFENKEVIFKVDSCFEFYIKKIED"

gene complement(145341..146639)
 /gene="orf432"
 CDS complement(145341..146639)
 /gene="orf432"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein"

/translation="MKNFSNVQKIKSMENFSEDELQIFIFRYYFEDSNHFYKNFCHKK
 KNKLLSDFVFNFSIFIGLGLVLLTQSKSAYAMQOPDPIWATGAEASSPNQADYHSHLS
 EFSGAPTSIGIPPMEPTTETPLSCFNWMTCCFPSPPIAPPDQPPTHASPFMSGPF
 PRRSSFASLAENPLSSSSPRARSVGAAATSVQYTRRGESNVLSHSSRSETPRAQRMLS
 LSTQAVQFKLKKSTLNPVKVLLNELQYHGEIPTLAVSSQDTLVNFQIILTSSDDRELV
 AFIRDADLLAFRNTTTNGKLDGKLGVTICKLRKDFANFETADKLLAKDQKERKKNYNDEY
 FNTHSEECWDITEVRKKIFRFNETRKRLEPKQFELHERAVGYENSTDKGKVTGETQL
 VQTLRQLFSPKTVTLYENSGTNLKSETSIYKKGNPAPIAPK"

gene 147878..147964
 /gene="trnS(gcu)"
 tRNA 147878..147964
 /gene="trnS(gcu)"
 /product="tRNA-Ser"
 /note="codons recognized: AGY"
 /anticodon=(pos:147912..147914,aa:Ser)

gene 149663..149735
 /gene="trnA(ugc)"
 tRNA 149663..149735
 /gene="trnA(ugc)"
 /product="tRNA-Ala"
 /note="codons recognized: GCN"
 /anticodon=(pos:149696..149698,aa:Ala)

gene 150007..153611
 /gene="rrl"
 rRNA 150007..153611
 /gene="rrl"
 /product="23S ribosomal RNA"

gene 153902..154022
 /gene="rrf"
 rRNA 153902..154022
 /gene="rrf"
 /product="5S ribosomal RNA"


```

gene          155480..156253
              /gene="orf257"
CDS           155480..156253
              /gene="orf257"
              /codon_start=1
              /transl_table=11
              /product="hypothetical protein"
/translation="MTKRNLFKINSSENSNKSRSSETAEDTIAQHKVPKPNISGDPTP
QEAPGNSNEFRPSEKTEQAIAETEVRITPGLGEPKIQGGSRNSAKSRSAQEAHEITIVE
AQSPISSAFVEFNTQGGIVNPELGQQRYPSPPEVKPNDTRYGNVLTFFVSLLSLISLVG
VAIYFHLTRHNDSKKTKKRKEVLELTVLFPFEKNTPILMCKLHPKIKLGQIGKVKKK
LSANLASVEFQKLEFENSPIKDLIGLALSSKGEGLVFL"
gene          157807..158883
              /gene="orf358"
CDS           157807..158883
              /gene="orf358"
              /codon_start=1
              /transl_table=11
              /product="hypothetical protein"
/translation="MYTLSSSLNAAIPLHRPDMPLRRPYEFGNLIKTKGQIKKIPKPIW
RAIPTLAKDQQSNPFNQLAALSSRPTWPPASPYPKFQAEQWYTWVENVPGRSINYLDLDR
LILTVGNFLPSLDFQYDFPAWFNRIFEHCLETKHILDIEICNLPNIYKSTTFYNTRT
SQLTEKGEKLSLFTLFFMVFI FNEIYENPITETNLSTNYDNFQNLTANDLDFIAYSK
NNNLKTPQILTARFRHDVHREISVDKPLYGGSFTLIYTI CLFNELWWIEPVTDFVKGG
TNQSNKNRYKQVGSWNQNSYQELRKILEILQKFYFNFDVLSKCTYLSEQQKKEVTNTL
YKVFNGKDPFNDDTKAILTSLDM"
gene          158907..159725
              /gene="orf272"
CDS           158907..159725
              /gene="orf272"
              /codon_start=1
              /transl_table=11
              /product="hypothetical protein"
/translation="MKKQTENLGYESKSKNNKAKENTSIAFEIYSRERLISPGDYVA
LRTDLIDKKTLSCEFFIGQIGRVGPKVKLQKQESETYFINFYDELNKLKNVNSYDAFAT
PILHLDEILKLHFKQIQIANDYPVNSVKKALEEGTTLLEEQEDLDLINKEENIIINHN
FMSSRFDPGNLVTIVYDFAIKQENENIVIFSGQVGMIKKVLNNNVSVVFRQLGLSQL
LDKETNLEEIQDISALKKEENWVIETTHKSFLPLYYTKLQLTDKGNQLLNA"
gene          complement (159838..159910)
              /gene="trnI (gau) "
tRNA          complement (159838..159910)
              /gene="trnI (gau) "
              /product="tRNA-Ile"
              /note="codons recognized: AUU"
              /anticodon=(pos:complement (159874..159876),aa:Ile)
gene          complement (160647..160979)
              /gene="orf110"
CDS           complement (160647..160979)
              /gene="orf110"
              /codon_start=1
              /transl_table=11
              /product="hypothetical protein"
/translation="MKIEDVVEKLSKHTNGIYDKRRQDYMFGIYSTAILFKLLDFDHQ
LFLPEDGTAFMDELSNEELVDEQGMKAQIQNSRTFLQNAVKNPSGSTVFMVLVVFLARN
RWAFHFL"
gene          complement (161157..161708)
              /gene="orf183"

```

```

CDS complement (161157..161708)
    /gene="orf183"
    /codon_start=1
    /transl_table=11
    /product="hypothetical protein"
/translating="MQNLNNPTGSSSSSESLPINTDDATRSARLTAYLNEAQVKIIPP
AAESSLEKMIEAYFIRFLELKKKYFPNYTTFHDFIAVNILDPLFTYNFRHWATIKN
RAVIQSLHSRRLQLLTNETWRELLDVLKNNGLSLPDDDFNSSGNLNRRRRGGTAVAPP
RRIPTPWSTPYTDFKISDSDNEN"
    gene complement (161743..162273)
    /gene="orf176"
CDS complement (161743..162273)
    /gene="orf176"
    /codon_start=1
    /transl_table=11
    /product="hypothetical protein"
/translating="MNQIISNVLLTVSYQIYLRIKVNLELPPASKSYSNFLINIIA
PSYFKNPFYLYLGVKLNLFYCRNFKNISYYRELIISFMQDEYVDGIIISNSKNYFE
RKPIHLFPKLPKRNIEILKQNPQLQKQIAQDIHYITQVEYTVQRLENLEAYLLWCM
AYWFGCAFCLNSKAFF"
    gene complement (162988..163806)
    /gene="orf272"
CDS complement (162988..163806)
    /gene="orf272"
    /codon_start=1
    /transl_table=11
    /product="hypothetical protein"
/translating="MSIFNEDFLPDFLKYLNKFTKNLGSIFYKKNKNQEKKKQKAI
LQDLKDKQENNDNCNCKLSCNYCRTKFTKDLQSVFRPILNIMWELQKETLHKARQT
LQLKLVTAEDYSTVSKEIKKQYYKEHGFSENYKLSQEQTSDFDQFCVDRLEKEYEKSDN
AKIARFHRCARYKYKEDSRLDPVGNLKLAYTLLNPAYVLQGNNEKILIDFFMSGAFQ
LFAPDDFSSAQLFNKHFERLGHYLTGNKATQSTAPSNQPSTPNQPSTFWSLPET"
    gene complement (165228..166007)
    /gene="orf259"
CDS complement (165228..166007)
    /gene="orf259"
    /codon_start=1
    /transl_table=11
    /product="hypothetical protein"
/translating="MTNETNALWGVISEILGLFEKEFGPNCPYHWTFSLDFAYYAFDE
ASSNLNLEQLFNQWNQLPSDESFDQDEDEILDVLKQSPAWEWIPDYIFGSALFFIYDN
LETTGKLYPNISKLLYSLMFIDTGKQIKLLNVIERNWFFELYPEFAINGAKNFFLRQ
ERLPDLFFQYIQAPRQTLTYILKLLQKEAQIQLLVNETFKTRSFSKIINFLESYNLP
TPNLFNITELEVLTKFGLPEEKRYLLTNISQNLNNSGKN"
    gene complement (166052..166387)
    /gene="orf156"
CDS complement (166052..166387)
    /gene="orf156"
    /codon_start=1
    /transl_table=11
    /product="hypothetical protein"
/translating="MESISRTDQTRLEKNNMIKALKALRLKDKKLVDFAVTDRIVLEPA
CTERLSDLYMSFFKYSGNNPEIGRRAFKEILIHIIIVSKFEIEGVVEVKSSKGVLIQGI
GLLDNPKYM"
    gene complement (166970..168779)
    /gene="rrs"
rRNA complement (166970..168779)

```

```

        /gene="rrs"
        /product="16S ribosomal RNA"
misc_feature 168989..209127
              /note="Region: small single-copy region (SSC)"
gene         complement(169313..169621)
              /gene="orf102"
CDS         complement(169313..169621)
              /gene="orf102"
              /codon_start=1
              /transl_table=11
              /product="hypothetical protein"
/translation="MMMENDIHEMFIKRRQIKKGWSLPAYHAFKVKHGILADVLALEK
VRIKHQLLSYQYVDEFCLNFEQQIQFLDRATNIYKTNVKLSNEFKRCRQEHTLFLNN
"
gene         complement(170084..170389)
              /gene="orf101"
CDS         complement(170084..170389)
              /gene="orf101"
              /codon_start=1
              /transl_table=11
              /product="hypothetical protein"
/translation="MKFFNYFQFFRKFWARNTASETISSSLEQPSLLTSQKSGVESGL
LTFSESKIQPNMKDLVNYQNI IISKQSIRKSKAVYDSLQDLSLIQNFTLHQSEYIP"
gene         170775..171503
              /gene="orf242"
CDS         170775..171503
              /gene="orf242"
              /codon_start=1
              /transl_table=11
              /product="hypothetical protein"
/translation="MNLNPKQIYSFSKREIYSNYPFTDEFLLFTLVNDPYVSVSKSLEE
FLSFHLIDIDSEININCGVFNLDRLDKEDLEIDLEQVLKSPNLIYDLENVYKINGLIY
YYNAKSQVYYNLTNINFSDLVSSSFNYFYFIQTNKDVNNQNRWKIITEYLFTKKFRKK
NFWYLKKFQILEDPTIKKQFVRNVSSRCRQYLSPLASSSTARYASTMLFDVLFNFIGKN
KMKSLLELVSNFKTTANPKPILRYI"
gene         171924..171996
              /gene="trnMe (cau) "
tRNA        171924..171996
              /gene="trnMe (cau) "
              /product="tRNA-Met"
              /note="codon recognized: AUG"
              /anticodon=(pos:171957..171959,aa:Met)
gene         complement(172293..172374)
              /gene="trnL (uag) "
tRNA        complement(172293..172374)
              /gene="trnL (uag) "
              /product="tRNA-Leu"
              /note="codons recognized: CUN"
              /anticodon=(pos:complement(172338..172340),aa:Leu)
gene         173788..173937
              /gene="rpl32"
CDS         173788..173937
              /gene="rpl32"
              /codon_start=1
              /transl_table=11
              /product="ribosomal protein L32"
/translation="MAVPKCRASKTKTKTRRTIWKKKAQANKAFSLAKSVLKNLLN

```

```

EAGNK"
gene      175044..175853
          /gene="cysT"
CDS      175044..175853
          /gene="cysT"
          /codon_start=1
          /transl_table=11
          /product="probable transport protein"
/translation="MTLFKNSWPWLFLLYLLFMLILPITSLTASQTLVIRFWEIA
TEPVAVSAYFVTISMALCACFFNAIFGFILAWVLVRYNFPGKLLDAVDLPPALPTS
VAGLTLATVYSDQGWIGSILNDFGIQIVFTRLGITIAMIFVSPFPVVRTLQPVLOEME
IELEEAASWLSGASPWQTFIKIVFPPLVPALLTGVTLAFSRAIGEYGSIVIVSSNFPLK
DLIASVLIQFQLEQYDYIGATVIGTVVVISLFLLLISINLFQSWNQNYKMN"
gene      178917..181547
          /gene="ycf1"
CDS      178917..181547
          /gene="ycf1"
          /codon_start=1
          /transl_table=11
          /product="hypothetical chloroplast RF1"
/translation="MSFVTAIKDYIDVLNHVYDSVGNVTIQQIFQQTFFVYIFSSIKFL
AGYLVSFQWLRDLAYLPILVPEIESTTILKGNVFLDAPLSNFFTLLETPTYENNKFLIG
FLNSFFLCLPLSTTHFISIRLLVQGIPAGIVSNIGTILGYTFFIFSVLFGRLRVIIP
WLAYEPFSYIIGVLFLLTIVYDMTHERTIRKFDWSQKSTLIQIFFINFLLTWTEQSCI
FSYFGNLTLPQPTIFDNFSSLTKSGSILAHSTYLIGILLGSVFFTSLFTLICLQLSN
LSLKL SAVPYSRWLRRTNFVLLTTIIAFTFASIPFYSLDYLFGGPLGFISQDKAFDKT
ILSSKNVEDSTRMLGENSNFKSLGTDVTPFDRGSYLLPDSNENFEDLNYQGEYAWTAR
KDHRAVYGTESRKLFSNFFKKAQNLSQKQDEIQNPISQNRYSAPKEKNSPLADRS
SEQKSFGLKKNKEKLNQLDNSDLIEKSDSEKIVQNEVNTNEFPDNLSDFETDDLYGT
SQKGI PSSYFYKLNHRINTDFKINQSLQPLLDTSFSPQFFGETLNTLDPELEKTLKQR
YYSNPVYKTL LSADIDLFLRRQPSSYLLSPKEEKNLFQKRLMLSNNYDSLRYQKMPY
TKEFQNFSGSKSYADRVYNQQFKGTLKVVRRLFSITFNEEENLPEKITLKFQPLYK
KLKQNNNLIGHEELVVKKNKPFIELVNPPIPFYTGWDEQLRKLVI TNRLLPSTAGYN
MQFDSKKELT DYPTLSELLKSTKKVDFTTWPLTKTNLEKPKEDLKIPYNVLYEQISDP
QNKLLAESLKDLEGSPWTFETVPPNLKKIDDPKINDVLPPTRGGFVWSGHSSLKINFK
KIFKEGINKFYTFKMPFSQS"
gene      182443..182514
          /gene="trnG (gcc) "
tRNA     182443..182514
          /gene="trnG (gcc) "
          /product="tRNA-Gly"
          /note="codons recognized: GGY"
          /anticodon=(pos:182475..182477,aa:Gly)
gene      182618..182691
          /gene="trnD (guc) "
tRNA     182618..182691
          /gene="trnD (guc) "
          /product="tRNA-Asp"
          /note="codons recognized: GAY"
          /anticodon=(pos:182652..182654,aa:Asp)
gene      complement(183307..183411)
          /gene="psbM"
CDS      complement(183307..183411)
          /gene="psbM"
          /codon_start=1
          /transl_table=11
          /product="M protein of photosystem II"

```

```

        /translation="MEVNILGLIATALFILIPSTFLLILYVKTASQEE"
gene      complement (186480..187778)
        /gene="minD"
CDS      complement (186480..187778)
        /gene="minD"
        /codon_start=1
        /transl_table=11
        /product="septum site-determining protein"
/translation="MINSKFKWSSRLLFPCSIASITESSKISELYSIYPCFAANQPAR
QRVKQCIYGGLKKNYPSYSALSQIITPRNKMFSSKFSYAFKSSFTNKVCKTFHVQSGI
GVSLSKYKSSSITNFSFSRSIFKISYRESQQGEKYPQLKKKFLTKISSGNEISISENS
NTVESRIIVITSGKGGVGVKTTATANLGM SIARLGYKVALIDADIGLRNLDLLLGLLENR
ILYTAMDILEGQCRLDQALIRDKRWNLSLLSISKNRQRYNVTRKNMENLVNSISSLG
YHFILIDCPAGIDVGFINAISPAHEALIVTTP EITAIRDGDRVAGLLEANGIYNVKLL
VNRVRPDMIQRNDMMMSVRDVQEMLGIPLLGAI PEDNHV IISTNRGEPLVLK KKLTLTSG
IAFENAARRLIGKQDYFIDLTPYKGVFQKVQEFFLGGNG"
gene      complement (188313..188387)
        /gene="trnR (acg) "
trnA     complement (188313..188387)
        /gene="trnR (acg) "
        /product="tRNA-Arg"
        /note="codons recognized: CGH; A in the first position of
the anticodon assumed to be post-transcriptionally
modified to I (inosine)"
        /anticodon=(pos:complement (188350..188352),aa:Arg)
gene      complement (189118..190281)
        /gene="ccsA"
CDS      complement (189118..190281)
        /gene="ccsA"
        /codon_start=1
        /transl_table=11
        /product="heme attachment to plastid cytochrome c"
/translation="MTHLNFEIILGNFSFILVFGIMIFYWLEVSFSTFSKHFVAFSA
TILANFSLLLLLITRWFNYGHFPLSNMYESLIFLSWSLLLLQLVLEKLSKNIFLGAI
SPIALFTNSFASFLPKEMQKASVLVPALQSNWLMHVTVMMLSYAALLGGSLLAIAF
LIITFGKNIDLTGNSLGRDPGVAGTQRVKFQGFYQTKLDTRLVSTNSINDINSKANLS
NLDFSFKELQYNTGNTKQFLAAVSSNAADTSSKDEYVIENSSSSLEKNIKKYFINHEG
SEYKLAKILDNLSYRTLIGIGFPLLTIGILSGAVWANEAWGSYWSWDPKETWALITWLV
FAIYLHTRITKGWQKKPAIIASAGFVIVWICYLGVNLLGKGLHSYGWVHLNF"
gene      192529..193395
        /gene="chlL"
CDS      192529..193395
        /gene="chlL"
        /codon_start=1
        /transl_table=11
        /product="ATP-binding subunit of protochlorophyllide
reductase"
/translation="MKLAVYGGKGGIGKSTTSCNISIALARRGKQVQLIGCDPKHDSTF
TLTGFLIPTIIDTLQSKDYHYEDVWPEDVIYQGYGGVECVEAGGPPAGAGCGGYVVG
TVKLLKELNAFYEYDIILFDVLDVCGGFAAPLNYADYCIITDNGFDALFAANRIV
ASVREKARTHPLRLAGLIGNRTAKRDLIDKYVESCMPVLEVLPLIEDIRVSRVKGKT
LFEMA ESETALTYVCFYLN IADQLLSHPEGVVPNELPDRELFSLLSDFYLN PQENIT
DAHNDSLDFMVV"
gene      194375..195811
        /gene="chlN"
CDS      194375..195811
        /gene="chlN"

```

```

/codon_start=1
/transl_table=11
/product="ChlN subunit of protochlorophyllide reductase"
/translation="MSETLSETLTFFECETGNYHTFCPI SCVSWLYQKIEDSFFLVVGT
KTCGYFLQNALGVMIFAEPYAMAELEEGDISAQLNDYKELKRLCIQIKQDRNPSVIV
WIGTCTTEIIKTDLEGMARLEAEIGIPIVVARANGLDHAFTQGEDTVLAAMAHRCPT
QDSKILVGTISDSENNDQWPLNKVLNSKIVNNLFQFNKKTQPN DSTNYYNHPPLVLF
SLPSTVTTQLNLELKRQGIKIVSGWLPSQRYIDLPTLGDGVYVCGVNPFLSRTATTL
MR
RRKCKLIGAPFPIGPDGTRAWIEKICSVFGIIPKGLKTRENQIWEGLDQYIDL
VKGKS
VFFMGDNLLEVSLARFLIRCGMIVYEVGIPYMDKRFQAAELALLEDTCNEMNVPL
PRI
VEKPDNYNQIQRIKELQPDLAITGMAHANPLEARGISTKWSVEFTFAQIHGFTNARDI
LELITRPLRRNQSLEGLGWTKLVKSATS"
gene 196873..197229
/gene="orf118"
CDS 196873..197229
/gene="orf118"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/translation="MKEKTAGPLLLALSAYEARCSNSSTFDPAFTLKKKG YQKINENL
QLGIQATVQQPEDLLAFNSFKKYLIQKKMLSRIFFFKLLFKKKQIRYFYDLSNEI
PVT
QLSAVGVWTIPVVLLD"
gene 197904..198548
/gene="orf214"
CDS 197904..198548
/gene="orf214"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/translation="MLIFRHRSSQTTLLDDLRE STYEKAKAMRIFSTLNQTRAIYGAQE
RKEDLKN TDLLTNRWEAFWQTL DHDKFIYLLVSNCEGNENVVKELLKRYSELYI
SCYL
PNVILWLLKYQVPKELLLAAYECCVGITRLLTPIPNLESCWELFTKDKDRTEQAD
FMI
NSIHNCIVYKNILVQLCKKYFPTQYPVFKTLFEDLITSSINSLEESKKKMYSKF"
gene 200324..200713
/gene="orf129"
CDS 200324..200713
/gene="orf129"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/translation="MIDFIKDNPHYQRVETTKVIFAENVLKHLRTSIGKKIEVCS
DLC
FLNGRLNFKTQELSDYTPNVYCFNFKNMYYLKTEVLF DSTIKFITNL CNKNR
LYINCL
RIWLRNCLIRYISDEYGLYLFVYSSHN"
gene 201325..201411
/gene="trnS (uga) "
tRNA 201325..201411
/gene="trnS (uga) "
/product="tRNA-Ser"
/note="codons recognized: UCN"
/anticodon=(pos:201359..201361,aa:Ser)
gene 202109..202735
/gene="ycf20"
CDS 202109..202735
/gene="ycf20"
/codon_start=1
/transl_table=11
/product="hypothetical chloroplast RF20"

```

```

/translation="MDSNTRIFKKVIQFTNIGKNKLFSAKNNLPIKIFLLFSGFLIGS
LFGTFLPSLPEQINSYSVIIILITITAITAIEIVNYLVYSSKKRQFFFSGSFFGILLQFFLSI
KTKLVSFFFNRDPTLQSYEAQKNNKVSSELQIQTMVPQFEHGSITPKGEGYQAQLDY
PLILHSKEDSSFRVKQWQNTNKNFYKNLNSFKIGIMLGFFIDAFKVGVS"
    gene                203418..203945
                        /gene="orf189"
    CDS                 203418..203945
                        /gene="orf189"
                        /codon_start=1
                        /transl_table=11
                        /product="hypothetical protein"
/translation="MNFVKLNNLNRTNDSRESFIAFLNELLFQARKNDLCLKTHIQK
YQVPQKHQLISKNFTIRQSIQVGMHLQVAHNLEFFIYSPTYKQKLNAQHLALTKEFK
GLLEQFIEVNKNTLYFNAEIAKKDLKEFNKQKPDQNFELEIICNTDLDLGNSYGRNGE
ILDLRIQTAINLAKI"
    gene                204090..204455
                        /gene="orf121"
    CDS                 204090..204455
                        /gene="orf121"
                        /codon_start=1
                        /transl_table=11
                        /product="hypothetical protein"
/translation="MKPSAQTDLGLITTLSDDEWMTEIVDRLNKQSQNDKSGDKNSDK
FIKNFPESEDKGRYFENEIDLPTVFIQDLDPDYKPRRRSQKEELPRVPRRETPEERLQ
HLQEDKLPQVPNTPEEQ"
    gene                205379..205624
                        /gene="psaC"
    CDS                 205379..205624
                        /gene="psaC"
                        /codon_start=1
                        /transl_table=11
                        /product="subunit VII of photosystem I (Fe-Spolypeptide)"
/translation="MSHAVKIYDTCIGCTQCVRACPTDVLQVMPWDGCKASQIASAPR
TEDCVGCKRCESACPTDFLSVRVYLGSETTRSMGLAY"
    gene                complement (205977..206048)
                        /gene="trnN (guu) "
    tRNA                complement (205977..206048)
                        /gene="trnN (guu) "
                        /product="tRNA-Asn"
                        /note="codons recognized: AAY"
                        /anticodon=(pos:complement (206014..206016), aa:Asn)
    gene                206336..208378
                        /gene="orf680"
    CDS                 206336..208378
                        /gene="orf680"
                        /codon_start=1
                        /transl_table=11
                        /product="hypothetical protein"
/translation="MKDKIQNSFRIITQLFETEFNLCKFWWPFSLMAEYVFTAGSS
CTSFEQILFIWKFTRQELLEKQNWLNENEGNEKGNEDSKLEYIDNFVNQIRNSEVYNW
PYNKLLCSVIFVFHKKLCSGGIDSVSQCQLITILTSVNSAIRNEIYDLIQKNGYFH
LNPELAINSARDYFNFKETTSFFFEFKDSNQKQKLTFLVLRLLQIQESQIQLLVSNAF
KNKSFNSIIIEFLKNYSFEKPKLLTFKQLQVFLNQIDLSSSEKKEIIQDVSTELSMKS
IDLLSTWDDLTLSSNVAIFEFYNELPQRQKLLNKEIIAIVTTLIALAKFVEINTPK
IILYPLQIIQNRLSSISSENLANLYELFVLDENVLDTSSKSQLKEGLLIRQVILCLL
FGTLITTSVLGTRSAFAAQLPQGQAGTVAVARPEKNSQMMAKIVTKNKTEALPASLRS
GEKKVKRLTVTRGRYGLDVFKEYDLPEERKHETYDIKYEDSYEIIIMDGIASFQVSFQTE

```

IDKKTNKPRKLLSIYVTEAPYGHLLQNVIKQFPLLFGPNVSNREHCQLDHLCTVWLQK
AFGIPASRAGVITVLTTPHNPSRTGGDDGLNDLGRSSSITYRVAIQDGLHPVEGLAPGI
LKSFHRVVKNPTEIPYEQENIEKVLSQVKPAISDITGTLTHSLYFLMLSHKAMVLW"
gene 208372..209034
/gene="orf220"
CDS 208372..209034
/gene="orf220"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/translation="MVMEACGSADTLGRATFMLDMMKQAPEKFPSTFFINQGVIEQI
QRNYQFVRGVQESVKLAGLSTETSDYLERRLNRRATQVLSTIPSEVRAHRFRKSRIFEN
EKLEISENEKAFFELFNKLDQAINSIGNPYIEQLEEFSTRKLAIDKHAGPRGCPLASGF
IEPQDWGVTQSRLLDQYGNPFGDPLPVRRLVDYQGNPLSSYETTVIRNRVEPNDKSI
YL"
repeat_region complement(209128..236463)
/note="copy A of inverted repeat region (IRA)"
/rpt_type=inverted
gene 209338..211146
/gene="rrs"
rRNA 209338..211146
/gene="rrs"
/product="16S ribosomal RNA"
gene 211777..212064
/gene="orf156"
CDS 211777..212064
/gene="orf156"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/translation="MIKALKALRLKDKKLVDAFVTDRIVLEPACTERLSDLYMSFFKYK
GNNPEIGRRAFKEILIHIIIVSKFEIEGVVEVKSSKGVLIQIGILLDNPKYM"
gene 212109..212888
/gene="orf259"
CDS 212109..212888
/gene="orf259"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/translation="MTNETNALWGVISEILGLFEKEFGPNCYPHWTFSLDFAYYAFDE
ASSNLNLEQLFNQWNQLPSDESFDQDEDEILDVLKQSPAWEWIPDYIFGSALFFILYDN
LETTGKLYPNISKLLYSLMFIDTQKQIKLLNVIERNWFFELYPEFAINGAKNFFLRQ
ERLPDLFFQYIQAPRQTLTYILKLLQLKEAQIQLLVNETFKTRSFSKIINFLESYNLP
TPNLFNITELEVLTKFGLPEEKRYLLTNISQNLNNSGKN"
gene 214310..215128
/gene="orf272"
CDS 214310..215128
/gene="orf272"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/translation="MSIFNEFDLPDFLKYKANFTKNLKGKSIIFYKNKNKQEKKKQKAI
LQDLKDKQENNDNCNSCNKLSNYCRFTKFDKLDQSVFRPILNIMWELQKETLHKARQT
LQLKLVTAEDYSTVSKEIKKQYYKEHGFSENYKLSQEQTSDFDQFCVDRLKEYEKSDN
AKIARFHRCARYKYKEDSRLLDPVGNLKLAYTLLNPAYVLQGNNEKILIDFFMSGAFQ
LFAPDDFSSAQLFNKHFERLGHYLTGNKATQSTAPSNQPSTPNQPSTFWSLPET"
gene 215843..216373


```

                /gene="orf176"
CDS             215843..216373
                /gene="orf176"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein"
/translation="MNQIISNVLLTVSYQIYLRIKVNLELPPFPASKSYSNFLINIIA
PSYFKNPFYLYLGVKLNLFYCRNFLKNISYYRELIISFMQDEYVDGIIISNSKNYFE
RKPIHLFPKLPKRNIIEILKQNPQLQNKQIAQDIHYITQVEYTVQRLENLEAYLLWCM
AYWFGCAFCLNSKAFF"
    gene         216408..216959
                /gene="orf183"
CDS             216408..216959
                /gene="orf183"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein"
/translation="MQNLNNPTGSSSSSESLPINTDDATRSARLTAYLNEAQVKIIP
AAESSLEKMIEAYFIRFLELKKKYFPNYTTFHDFIAVNILDPLFTYNFRHWATIKN
RAVIQSLHSRRLQLLTNETWRELLDVLKNNGLSLPDDDFNSSGNLNRRRRGGTAVAPP
RRIPTPWSTPYTDFKISDS DNEN"
    gene         217137..217469
                /gene="orf110"
CDS             217137..217469
                /gene="orf110"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein"
/translation="MKIEDVVEKLSKHTNGIYDKRRQDYMFGIYSTAILFKLLDFDHQ
LFLPEDGTAFMDELSNEELVDEQGMKAQIQNSRTFLQNAVKNPSGSTVFMLVVFLARN
RWAFAHFL"
    gene         218206..218279
                /gene="trnI (gau) "
tRNA           218206..218279
                /gene="trnI (gau) "
                /product="tRNA-Ile"
                /note="codons recognized: AUH"
                /anticodon=(pos:218240..218242,aa:Ile)
    gene         complement (218391..219209)
                /gene="orf272"
CDS             complement (218391..219209)
                /gene="orf272"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein"
/translation="MKKQ TENLGYESKSKNNKAKENTSIAFEIYSRERLISPGDYVA
LRTDLIDKKTLS CFFIGQIGRVGPKVKLDKQESETYFINFYDELNKLKNVNSYDAFAT
PILHLDEILKLFKQQIANDYPVNSVKKALEEGTTLLEEQEDLDLINKEENIIINHN
FMSSRFDPGNLVTIVYDFAIKQENENIVIFSGQVGMIKKVL SNNNVSVVFRQLGLSQL
LDKETNLEEIQDISALKKEENWVIETT IHSFLLPLYTKLQLTDKGKNQLLNA"
    gene         complement (219233..220309)
                /gene="orf358"
CDS             complement (219233..220309)
                /gene="orf358"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein"

```

```

/translation="MYTLSSSLNAAIPLHRPDMPLRRPYEFGNILKTGQIKKIPKPIW
RAIPTLAKDQQSNPFNQLAALSSRPTWPPASPYPKFQAEQWYTWVENVPGRSINYLDR
LILTVGNFLPSLDFQYDFPAWFNRI FEHCLETKHILDIEICNLPNIYKSTTFYNTRT
SQLTEKGEKLSLFTLFFMVFI FNEIYENPITETNLSTNYDNFQNL TANDLDFIAYSK
NNNLKTPQILTARFRHDVHREI SVDKKLYGGSFTLIY TICLFNELLWWIEPV TDFVKGG
TNQSNKNRYKQVGSWNQNSYQELRKILEILQKFYFNFDVLSKCTYLSEQQKKEVTNTL
YKVFNGKDPFNDTDFKAILTSLDM"
    gene          complement (221863..222636)
                  /gene="orf257"
    CDS           complement (221863..222636)
                  /gene="orf257"
                  /codon_start=1
                  /transl_table=11
                  /product="hypothetical protein"
/translation="MTKRNLFAKINSENSNKSRSSETAEDTIAQHKVPKPNISGDPTP
QEAPGNSNEFRPSEKTEQAI AETEVRITPGLGEPKIQGGSRNSAKSRSAQEA EHTIVE
AQSPISSAFVEFNTQGGIVNPELGQQRYPSPPEVKPNDTRYGNVLT FVSLLSLISLVG
VAIYFHLTRHNSKKT KKRKTEVLELTVLFPFEKNTPI LMCKLHPKIKLGQIGKVKKK
LSANLASVEFQKLEFENSPI SKDLIGLALSSKGEGLVFL"
    gene          complement (224094..224214)
                  /gene="rrf"
    rRNA          complement (224094..224214)
                  /gene="rrf"
                  /product="5S ribosomal RNA"
    gene          complement (224505..228109)
                  /gene="rrl"
    rRNA          complement (224505..228109)
                  /gene="rrl"
                  /product="23S ribosomal RNA"
    gene          complement (228381..228451)
                  /gene="trnA (ugc) "
    tRNA          complement (228381..228451)
                  /gene="trnA (ugc) "
                  /product="tRNA-Ala"
                  /note="codons recognized: GCN"
                  /anticodon=(pos:complement (228418..228420), aa:Ala)
    gene          complement (230146..230238)
                  /gene="trnS (gcu) "
    tRNA          complement (230152..230238)
                  /gene="trnS (gcu) "
                  /product="tRNA-Ser"
                  /note="codons recognized: AGY"
                  /anticodon=(pos:complement (230202..230204), aa:Ser)
    gene          231477..232775
                  /gene="orf432"
    CDS           231477..232775
                  /gene="orf432"
                  /codon_start=1
                  /transl_table=11
                  /product="hypothetical protein"
/translation="MKNFSNVQKIKSMENFSEDELQIFIFRYYFEDSNHFYKNFCHKK
KNKLLSDFVFNFSIFIGLGLVLLTQSKSAYAMQQPDPIWATGAEASSPNQADYHSHLS
EFSGAPTS GIPPMEPTTETTPLSCFNWMTCCFPSP IAPPDQPPTHASPF SMSGPF
PPRSSFASSLAENPLSSSSPRARSVGA AATSVQTRRGESNVLSHSSRSETPRAQRMLS
LSTQAVQFKLKKSTLNPVKV LNELQYHGEIPTLAVSSQDTLVNFAIQILTSSDDREL V
AFIRDADLLAFRNTTTNGKLDGKLGVTICKLRKDANFETADKLLAKDQKERK NYNDEY
FNTHSEECWDITEVRKKI FRFNETRKRLEPKQFELHERAVGYENSTDKGKVT PGETQL

```

```

VQTLRQLFSPKTVTLYENSGTNLKSETSIYKKGNPAAIAPK"
  gene                233257..233874
                      /gene="orf205"
  CDS                 233257..233874
                      /gene="orf205"
                      /codon_start=1
                      /transl_table=11
                      /product="hypothetical protein"
/translation="MLYVYIRLLKEIKKMIQISDFVLEFRDFILLSWPSLKKLGDDDR
PYYCKENCIEAIWELLLLEPLLATKLQTTVYIQYWLSGAEIYFRSPRAFFPDVKKSTH
EIFCLPKNNEFFFEFYIKKKIFLNKPGALDPGLRSSRCLTLTQGENPEQNYFQFSHF
VTLSENESMPDLEKPPFDYAVCNFENKEVIFKVDSCEFYIKKIED"
  gene                complement(235028..236089)
                      /gene="psbA"
  CDS                 complement(235028..236089)
                      /gene="psbA"
                      /codon_start=1
                      /transl_table=11
                      /product="D1 reaction-center protein of photosystem II"
/translation="MTAILERRESASLWARFCEWVTSTENRLYIGWFGVLMIPTLLTA
TSVFIIAFIAAPPVDDIGIREPVSGLLYGNIIISGAVVPTSNAIGLHFYPIWEAASL
DEWLYNGGPYQLIVCHFFLGICCYMGREWELSFRLGMRPWIAVAYSAPVAAATAVFII
YPIGQGSFSDGMPLGISGTFNFMIVFQAEHNILMHPFHMLGVAGVFGGSLFSAMHGSL
VTSSLIRETTENESANAGYKFGQEEETYNIVAAHGYFGRLIFQYASFNNSRSLHFFLA
AWPVVGIWFTALGISTMAFNLNGFNFNQSVVDSQGRVINTWADIINRANLGMEVMHER
NAHNFPDLASVEAPSINA"
BASE COUNT          78176 a  39473 c  39445 g  79369 t
ORIGIN
   1 tgttttagatt acccgggagc cctagattcc tcattaggta ttgctaata gcaatctagg
  61 gctacggggtc gaaatttaac aacaaagttg ttaaataaag ttttccgcgt cgggctttac
 121 gcgaagcgta aaactttatg ttttaccggt gccctgccgt caacggcaag ggtctaagat
 181 aaaacattta ggaaagcctt tagggaagtc caaagtatac aaacaaagtt tgtatactag
 241 gctttgcca tagagcaaag ctttatggtt tctcagagat aaaacattta gatattaacg
 301 cctagcttta agtcttttgt attgttaaat ttagccttcg cctattgggc aaagcccaat
 361 aggacaacgg gttcaaatta aagccaggct acattacggt agtaatttaa gacaagattg
 421 tactttaagt ttttttatac cacctagata aatgtttcag tttacattta aactaagaag
 481 atatttttac tataaattaa aaaattatct agaaactgct aaaaattctt ccgtgtattc
 541 ttgaattgct tcttttaaaa gtgattgagc ttcattcatta aaagtattag tagatttaac
 601 aatctcacca tattttgggt tactatttgc taaatactga cgtaatccaa ctaaaaacga
 661 acgaacttga tcaaccttaa ttttgtctaa ataaccattt gttcccgcgt aaatgcttgc
 721 aacttggtct tcaatagaaa gaggtgacgt ttgaggttgt ttaataaatt cacgtaaacg
 781 ttgaccacga gctaattggt tttgagtagc ttgatctaaa tctgaagcaa actgagagaa
 841 agcttcaagt tctgcaaact gagctaattc tagttttaat ttgccagcaa cttgtttcat
 901 agcttttggt tgagcagcag aaccaactcg agatactgaa atacctacgt taattgctgg
 961 acgaataccc gcgttaaaaa tatcaccaga taagaagatt tgcccatcag taatagaaat
1021 tacgttggtt gggatataag cagaaacatc accttcttga gtttctacaa ttggaagggc
1081 agtcatactg cctccaccta acttatcact tagctttgct gcacgttcta ataaacgtga
1141 atgtaaataa aatacatcac ctggataagc ttctcgccct ggaggacgct ttaataataa
1201 cgacatttca cgataagctt gagcttggtt tgataaatcg tcataaataa ctaaagtatg
1261 tcgaccttta tacataaagt attctgctaa tgcagacca gtataaggag caagatattg
1321 taaagtagct ggagaatcag cagtagcagc aacaataata gtataatcta aagcccctcg
1381 atcttgtaaa acgttaacaa cttgagcaat tgaagaagct ttttgaccaa tagctacata
1441 gacacaaata acatcttttc ctttttgatt taaaatggta tctacggcta ctgcagtctt
1501 accagtttgt ctatcccaaa taattaattc acgttgtcca cgtccaatag gaatcattgc
1561 atctactgct actagtccag tttgtagggg ttcattgaacg gaacgtcgag aaataattcc
1621 cgggtgcagct gactcgatta atctattttc agtagttgga atttcacctt taccatcaat
1681 agggcgagct aatgaattta caactcgacc taaataacca tcacctactg gaatttgagc

```

1741 aatTTTTccg gtagctctaa cagacgttcc ttcttgtact gttaagcett ctcccattaa
1801 tactgcacca acgtTTTTtag attcaagggt tagagcaatt cctactgtac catcttcgaa
1861 ttctaaaagc tccccagcca ttactTTTTc aagtcctgaa attcttgcaa taccatctcc
1921 tacttgghaaa acagtaccca catttactac tttaacttca ctattatact gttcaatttg
1981 ttgtcgaatg atgctactaa tttcatcagg tcgaatttta accattggta aacagactcc
2041 ctaaaattat aaatataatc taagtgtttt ctacttataa attaaaactt attttaaaaa
2101 aattttatatt taaattttta ccttaaacta aaaactataa taaatttcat acaaagttaa
2161 aatttcagct ttaaaagcaa aaactcctaa atttttaatt gtttaaaatt aggaataaaa
2221 cccccagttt taatattcca ttaaaaccag gtcgtaaagt ctctaagtca taaacaaagt
2281 gttaatcagt aaattcaatt ttactaagcc gttgtctaag gttaaaagct ctgcttttaa
2341 acatagaaga ttttaagcga agcttaaaga tcctaagtgt tgcaagcaca cattttggac
2401 ttccacaaaa tggaagatcc taatagaact ctgttctatt ttggactttt tttttacgct
2461 ccgcaaaagt tttctaaatg ttattcaata tttagaaaag cctatcggat acggcaaaaga
2521 tacaataaat aaaaaagta caaattattt gaaactcaac gatttatttc aaagcgtatt
2581 gaaatttatt taggtttata attggtaaaa agcacaatat taaaattatt tacagaagaa
2641 tgaaatgtcg aatctaactg attacttagt ttttctcgca cttgatttaa agcaagagaa
2701 acaacttggt gagaaacttg actaatcgtt ttttgctggt gaaaacgaat agtttcttgt
2761 ttaatttctg ttagacgtga tacatcttct tgcatttgtc gaatagcttg tttttttct
2821 tgctcagcag ttataacacc ctgctcacga atttcaatag ctttcttctt agccagctct
2881 aattgagctc gagcttttag taatttttct tcagcttctt tagctttttg atcagcttct
2941 tgaagattat ttaaaattgt ctgtttacgg ttatcaagta aagatcgtaa agcatctcct
3001 ccaaagtaa caacaaccgc aattacaact gataaattta taagatttgt ttctaaaata
3061 ttaccattaa acccaaaacc ttctccagat ggaatatgag caaagagggg taaaaaagtc
3121 ataaaaacct ccgtcttttag aatgctctat agtttatattc gaaactactg aattttctcc
3181 aactttctaa gctaacaacac gaataaccaa gtttttctag attacaaaca cagaaaaaaa
3241 gtccaaaaata gaacagagtt ctattaggat cttttatfff atggacgtcc aaaatgtgtg
3301 cttgcacaca ttaggatctt taagcgtcgc ccgttgccct gccgggacag gcaaggggta
3361 aagtcttctg cgtttgtaat aaagttttac aagcgtagct tgtaaagccc gtagggagaa
3421 ttcaaagttt gtgtccttac cccttggtct tgaccatta gctgttacta atgagtaata
3481 gccaaagggg aaggacacaa accaggcttt gcccataggg cttacatact taccatgtct
3541 aataagataa tggcaaagcg taaaaattta tgtttgatcc ccattgacct tggcaatggc
3601 aaggatgat caaacattta ggcaaccctt tatgttttat cggaaataaa acatttagat
3661 cttccacccc ttggctcctta ttcagtagcg gttgttaatg agcaatagac aaggggtaag
3721 gactaattaa acatagttta attagtagaa aatacttctt ttattggctt tggtaataaa
3781 ggaaggatct ttcagttttg gaagtctttt aactttcatc ttaaaattat aagacctaaa
3841 tagaagtttt gccaaaatga ttctagtttg aaaaaatcgc taaaaacagt tgtttgctat
3901 tcagctcttg tttttaggat tgaaaaatct aacttttttc tttgaagaga aaaaaatctg
3961 attgaggctt ttagtctgag tttgagaaat ttagtttata atatacttca gtagtttcca
4021 ttttatgtaa ttttaagttt taacaagtag tttgctaaaa cttaaaatta cataagaatt
4081 taattaatac taaataaatt cttatgtaat tttctaaagt aactgatagt tttttatgaa
4141 gcaaaaggat tagcaataa taaagctaaa gctacaacta aaccataaat agtttagagat
4201 tccataaaaag caaaacttaa taataatgca ccacgaattt taccttccgc ttcaggttgt
4261 cttgcaattc cttcaacagc ataaccagca gctgttcctt gaccattcc aggtccaata
4321 gcagcaagac ctacagctaa tccagccgca agaacagacg cagcgccaat aagtgggttc
4381 atgatgattt cctccaattt ttatatttaa aaaattttta ctatatgttt aaacatttat
4441 ttaagtgtat agtaacactt aaatatttaa gtttatcaaa taaaaagtaa attaaaataa
4501 ttataactta aattgtaaat aacaaaaaca ttttaaccgaa tagaattttt attcatattt
4561 tgttgaaatg aataattttt ggttactaca cttaatattt tagtctaatt attttaacaa
4621 caaggttgtc ttctttttta tttaaaaaag ccgaaattat aaatttttga ttttgactaa
4681 aagaaaaata ttaacgctag taagcttagc ttctctagtt taattgctta aaaataaaaa
4741 gatattgtat tagtttagtt taaaagctaa atcttttaag taagctttgc tcagatgctg
4801 attatcttta aattaatttt taaataaact aagctaataa ttttgctaac catttaaaaa
4861 tgacatctag gttataaact agatttaaaa atgcaaaaaa attgacatat tttatttgtt
4921 tttcaaaaag ctagattttta aaatcagggt tgtactaaca ctaaaagtta aagcctattt
4981 attattttata cttttcaata ttattaataa aaattcaagc taggtctata atttagccaa
5041 ataatatatt ttttcatagt gtttagacttt aaaataaaac tcaagtcatg aataatttac
5101 tttttttaat acaattgaaa tatttacata atttaaaata aagcttctta attcttttaa

5161 aactctacca gtatatgcct actaagactt cattatattga aacaaataaa gactaaaaaa
5221 ccaaactag cgctttctta gattatntaa aaataaagtg ttgaaaaacc aagtaaagtt
5281 tagtcaaatt acgattctat tgaattctta atttggctct atttcagaat atgttattaa
5341 aataaacttt aaattataga gattttaaac taaattacct aaaaaaaaaat aaaactttta
5401 ataaaagaga ctaaaaaatt tagttgcctc caatattttt ttaagaaaaa gttttttaca
5461 aaaattctat taaagtacac atataatcct gatacttagc tcactctaaa actagccatt
5521 ttaatagttt attttttagtg taatacttaa ttgttcacag cttatagctt ttccgacaaa
5581 gaattaaaaat ttgccaaaac cagtataaac cgttaaaaat atcagaaaaac gtcatttttc
5641 tagtatcttt cagtttacca tttctataaa gtaaagatac aaataacaga caaactctca
5701 aatattcaaaa tttactaaat tacgtgttag tatttttaatt ttaatagaaa actaagcgat
5761 gtatttttaaa aataactaata gaactctatt tttaaacaaa catttgatta atagcttaac
5821 ttcactaaat ttaaatctaa attttaagtgt tgttttaacc tttgggatgc aatctgtaag
5881 cgaagtctgc cctcaaagac tttcattttc gattgccttt ggtaatggca aggggtggaa
5941 gatcttaatc ttaatagaat caaagatttt attttggact cgtattataa gctacgctaa
6001 tagaataagc tttacttgct caagtagttt aggtctttta tccaagcaaa aacttaaata
6061 aaatctgata ataaaaatac acctatctaa gagattgttt tacctgaaaa taagaatctc
6121 gtaactcaaa agatataata cctaaactta ggttttaca aaaaagtttg tgaaacttga
6181 cgtagcacag aaaaaagtt ttctgtgttt aaaaatgtag ttttttaatt tataaaatta
6241 tggcaatttt tgttttagtct taaattgaaa gctagagatg tacaataaag ctttataaaa
6301 cttcggtaaa taagtcgtat ttcatactac tatagttttt aattaggatt gttttcaaaag
6361 aactccctat gcgaaagtct ttttgtctcc ccaaattccc aatgtaatt tggttttgaa
6421 aatgtttaat gtccttctaa ggcttctccg atatatgcac cggctaattg agcaaaaaca
6481 agtgcttgaa tagcactggt aaataagcct aaaagcatta gtgggattgg aacaacaaga
6541 ggaactagag caattaaaac acctacaact aattcatctg ctaaaatatt tccaaaaagt
6601 cgaaaactaa gtgataaagg ttttgtaaaa tcttcaagaa cattaatgg taataaaaat
6661 gcagctggtt caatataacg tttaaaatag ctagtcctt tttttcttag accagcataa
6721 aaataagcta tagatgttaa taaagctaaa gcaactgtag tattaatctc attagtagga
6781 gctgctaatt caccatgagg taattcaatt aatcgccatg gtaataaagc accagaccaa
6841 tttgaaacaa aaatgaaaag aaaaattgta cctaagaaag gtacctatgc taaatattct
6901 tcctcaccaa tttgagtttt agctaaatct cggataaatt ctgtaataa ttcagttaaa
6961 ttttgtaagc cagtaggaat tttttttaa tcacgatttc ctaatagaga taataaaaata
7021 attaagccaa gaactacca cgtagttaat aaaacttgac catgcacctg ataactcca
7081 aaattccaat aaaaatgttg accaactgat acttcagaca tttcaaata aggattttta
7141 gaattaagct gataaaaagt acctaaagt tcatcatta taaccatatt aattaaattt
7201 tttttctaaa taagtttgat taatttttaa tatttgaccg ttgccctgcc gggcacggca
7261 agggttacta gaataaagtc gaatatcgaa tattaatatt ttaaattaaa attttatctt
7321 gaaaaattca acctcacatt ctaattttga tgagagtaaa ttctatagta cctcaacta
7381 gagattttct ctttcaagat aaagcttacc cgggaacctt aggttcccgg gctacggtaa
7441 ggtattttgt tataaaactt atacaaacgt aacggagtgc acctctagat agaaggtttt
7501 gatctgcaaa caaaactttg ttttaagttat atttttactc gtggcaaatg cctaaactat
7561 ttaaataagt aaagcttggt taagtaacat agtccacca gaatagaatc tttcatccct
7621 tgacattgcc aagggcaatg gcaagtagaa atcacctaga taaggctaac aaggatgcca
7681 aaggctaaaa caaagcttaa acttaaataat ccagagcgt taattttagg tttttataaa
7741 gctttctaac aatttagata gatataaata tttaaaaaga aatttttgac taaaaacttc
7801 aatcattaaa aattttgttg taaatttctt caatttttaa acagaaaaaa ttgtgcctca
7861 atttatacgt aaaacagatt ttgttttaca acgaaatggg ttttccctac gggctttaca
7921 agcgtagctt ataaaacttg attacaaacg cagaacactt taacccttgc cgtgcccggc
7981 agggcaacgg gcgaagcttc aagatcctaa tgtgtgcaag cacacatttt ggactccac
8041 aaagtggaaag atcctaatag aactctgttc tattttggac ttttttcta tgctgcgcaa
8101 agtttgtaat ttgaaaaaac ttggtttaaca aacaaagttt gttagcttag agaacttaaa
8161 aagtttgctc ttatcgtgtc caaagttaaa aaattttatt tattaacaaa cccgggagcc
8221 ctaaactacc ggaggtagta aggttgacaa gcttcgcttg tttaccctag ccttccacta
8281 cgtggaagac cttaatacaa tctcggattg tattttgccc ttgcccgtgc cggcagggca
8341 acgggggaag aacacaagtc taaaatacaa tcagagattg tattaatca ctatgtgaaa
8401 gtccaaagtg tgtgcttgca cacactaggc tttgcccgtg gggcaaacct ttatttaaca
8461 acgaagtgtg taaatttaga cctaactctaa acaactatg tttgttttag atctttcctt
8521 taggaaagtc ttcttcctag gttaccgggc tacaggtcta aagtctttat attggatttc

8581 aaaaaacgtc agcttgaacg gtttatttta tagtctacaa aaaaataaaa aagcagaaaa
8641 tcttttcgcc ccactatata ctatagttaa aggttatagt cataaataca taaaagtttt
8701 tgtactctta actgattggg caaacctcat taaaaattgt tctcaaatca aacatagctt
8761 attaaatttt tatccaaggc tttgccctat gagcaaaact tgatgtttta tcaaagataa
8821 aatatttagt aaacacccaa atttccactc aaagtaatat ttttaaaatc agtctttcaa
8881 aaaaaacttt cagttaaaaa ggttgtttaa actttacttg aacacaaatg tagataaact
8941 cattgacaca tgactcaaac tcggtactcg gtattattat cccttgacca ttgctttgaa
9001 gtaatgcaaa gaaaggttta aaaacccaag tgggtgaagat caaactgact actcaaacaa
9061 tcttcactta gcaaaactaa ttaggacttt cgaacttcag ttttagattta gtttagacc
9121 tgacaaaagc aaaacttaaa tttagaacgc aactgatcac tatagtatgg aaagggtttt
9181 cactttatgt aaaacgtagt tacgaacttt ctctgcatt taaattcttt atagatttat
9241 ttgaataaaa cattaaaaact taaaaattha ctttaaaaat cgtctactct ttaattataa
9301 aattgattaa gaaattttta agtttagaat aactttcaaa ataagttttc taataaagtc
9361 ctgcaaaaaa aattgtagta atctttcata ttgtttgctt taagagtaaa tctactaaaa
9421 aagcaaaaaca ggatttattt taaagaaaga acaaaactata cgaaactaag cttttttaa
9481 aatgccttcc ccataaaatc ataaacgttt aaaatcaatt tgacaataat tttttaactt
9541 ttacttaagt aacaaagtcc taatttttta actaccgtag ttagttttaa tgtaaactat
9601 gatattttaag taaatgcttt tataaagtgt ttgataaaaa ctaaatttta gaaaaaatt
9661 gtcgagtact tggtaaaatt tttagaaca aaaatttaga ctagagaaaa tttatagttt
9721 caagtgtatg caactttttt taaattttta taatataaaa aactagaaaa aagcgccgtg
9781 ttgagtttta aatatcaatt ttttagtcaa aataaataat aaaaaagggt cttccagct
9841 aaaaaatagt gataactaat aattttagtt aagaagtgtc tagtaagata aatttatttg
9901 ataacctcag catctaagac gagaagctct ttaaatttat aacttacttg aagtttatat
9961 taggattttc cactccttag tctttaccct ttggtaaaga ccaaagggcc agggctaagt
10021 agtggaaagt ttcttcatag acgaaactata atgtgttatt gaactaaaaa aagtagctca
10081 agtaagaaac ccgggaaccc tagattcccg ggctacgcta aattattact tttggttatc
10141 tagtatttta atttattaac gttactaaaa actttatctt tgacataaat tttattcgag
10201 ctgttttatt tttgttacia tattgaagaa gaaatcattt tacagaatag aatgtctact
10261 ttaatttttt ctaaactgtc ttgttttgaa gtcatttggg agcttttttt caacgtataa
10321 cttttgtcct tgacgtattg ctcttaaaaa actattttaa attaattgta aagaagcaac
10381 tgaatcgtca tttgctggga taaaaagatc tgcataagct gggttacaat cagtactcaa
10441 gatagtaata ctccgtatac ctaatttttg acattccctt acagcattca tctctctaa
10501 ttgtccaact ataattacaa catctggtat agaatccatt ttttttaaac cgcccaaata
10561 tttttgtaaa cgatcttttt cttttttaca tgaagcagct tctttttttg gaagattatt
10621 aaattctcca tttttttcac gagcttccaa ttgatttaat tttccaattg aactttttat
10681 ggtttttcaa tttgttaaca ttccacctaa ccatcgttga ttaacgtagt aagaatcaca
10741 ttgtaatgca acttttagcaa ttagtttact tgcttgcttt tttgttccaa caaataagaa
10801 tttttttcct tgtgccgccc aatcatttaa aaattgagtt gattttttta agtgataata
10861 agtttgaatt aaatcaataa tatgaattcc atttctttcc gcataaatat aggggtgccat
10921 tttaggattc cattttcttg cttagatgtc aaaatgcatg cctacttcaa ccatttcatc
10981 gatgctatth atcataatcc gaaattattc ctagtthaaa ttttttgta gcctacgttc
11041 taaattatct agttttaaat ttgaaaactt aagtattgca aagcattgct aaagttaaca
11101 aacgagattt ggtactcgag cgttgcgctg ttttattgtc ttactactta acttaacata
11161 agataaatca agaggcaaaa caataaaaca acggtctaaa gtttaaattha taagtataac
11221 ccaagttact ccataacttc actcggtttg tctaaattta agaatgaagt tccaaggtat
11281 gtgcttttcta caaagcagaa tcttccactt tatgaaaatc ttaaataaaa gtttaacaag
11341 cttaaagatc taaatatttt atctctgaga aaatataaag ggttgcccgt agggcaagc
11401 ctgactaaga aacaaggttt cttaattttg gaaataattg tagaaaaaaa tcttgtagct
11461 ctgctttaag tccaaagtgt gtgctgtcac acactaggat ctccctaaa gggagagctt
11521 cgaactaaca aattatttaa aaagtttttt ttatatttcc tatataaata atatgtaaat
11581 aatagttttt tataattata aaaaatttga acctaaaatc ttaagtataa aggtttatatac
11641 ttaagggccc ggggtattcgt tttgctgtaa gcaaaacgtc aactttacga ccttcggaca
11701 ttgctcatta gtccaaagt tgtgtaaaac acaaaactagg atcttccata aagtggaggt
11761 ctttaccctt tgggtctttac tcattagccg ttgctaataga gcaacggcca aagggcaatg
11821 gccgaaggag tagcttaggc tgcaaccttt aaaaacttat ttaataaaat agaaattttt
11881 ttactaattt atacgtctca atgcaactaa tagtttagca aaaaaaaga agtgagcgaa
11941 ataaatttct tttgttttgt taaaaaatct taaattactt ttaaaggtac aaaaaattta

12001 attttccattt ttctgtacct ttaaaaaaaaa atttttgaaa agtttttctt tttcactttt
12061 tttggaataa gctgaatttt ttggatcaaa cgagcggagaa aaacctgttc cagcaggaat
12121 taaatgccct aaaataacat tttcttttag tccccgaaga aaatcagttt ttcgttcaat
12181 agctgctcta gctaataatac gtgtagtttc ctgaaaacta gctgctgaaa taaagctttc
12241 agtttctaag gccgcttttg taattccaag tattacaggt tcgtattcag ctttttgtga
12301 gttaataccg ttattaacaa tttcaatcct atctaaatca ataagttctc ctcgtaataa
12361 accactttgt ccaccttcga caattctgac ttttgaagtc atttgtcgaa caatgatttc
12421 aagatgctta tccgcaatag taacaccttg agattgataa actctttgaa cgccatctac
12481 aagaagttgt tgagtttttt cttaaactttt tcgagcggca tctcaggagc tatatttttg
12541 tttatacttc tcaaaaaaat tatgtagctt gtcatgtaaa ttttcgggaa gaatttcacc
12601 ctcttttggtt tgtctagctt caaaaagttg ttcaattttt ggtataacct gaacgatgct
12661 accagttttt aaacgttggg aaaagagcgt aagaagagga gaattttttt caacaaaatc
12721 accatgatga acataaaaaa caccctttga agaaaaaaga attgggtgtg ccttacgtaa
12781 aataactttc gatttatcaa cttgaatgat ttgaccagaa ttagctattg aaagattatt
12841 agcgatttgt tctccatata gaattaattt gccacaaaag gttaaagggtg tcttttgttc
12901 tgtactaaaa cttatttgat ctttctctgt tagaagaaaa aaactttgtt tgcccagaaa
12961 atcagttttt atatccgtaa tttctcctcg atatggagaa aaaaaatctg ttaaactcaa
13021 ggggttatga gcagcaatag aattttgggg gtgtaatact aattttgtgt tcgtaaattg
13081 aatagttttg ttttttagtt caggatcag gaaactagcg ttatcaattt gccgagataa
13141 tgaaaattct aaataattgg ttatataact caggttgcct tgctgagctc tgaccgaacc
13201 taaatttcga cgataattgg aactagagtt tgttgtacca ttttaattgtt ttccaagttc
13261 tgaacttttg ttactagcag ctattcttaa actcgcaacc ttagtgatag caaagttggg
13321 aaacttttct tgtgaacccc taactttatg cttaaatcta ctattagtag acaacaaagt
13381 gtgttttgaa acaaatccta cttttgctgt tttaaaagga aatttttttg gaatatttaa
13441 taagataaaa aattcaacaa tattaagttc ttttgtaaaa aacatgtgct ttgaattttt
13501 gaaccgaaat gaagagttga actgaaaatc agcagatggg aaagaagaaa aaagttttgt
13561 tctacttgtt ttattaaacc aagttaaatt tgttagtctt cctatgtcat tttgggattg
13621 aataattaaa ttatttggtt tcccttgggt attattttaa agtacctttt tataatgttt
13681 agatttataa acagaatact cttttacttt tcgtatttaa ataaaaaagc taggtgtact
13741 aaaattgtta ttcaatgtat caacctttgt agagaaagct atgttacaaa actgctcagg
13801 caacacactt tgaaatccaa atgaatttaa attgaaagga aaaattttaa gaattgggta
13861 ttcaagtttc gcttgacaac cttgctcctt cggagtggg cgatttttaa aatttgctaa
13921 gctaactttt tttttaacac tttgataaga cttgactaag taaaacgaac taatattttt
13981 tagttttttg ggaagaaaag acttattttac gttgatgttt ttattttgta aaaattcaaa
14041 ataacgatga tttttatttt ttatataata aattttgtct gaacataatg ttaggtttga
14101 accaaattta acaaaatcag gtaattcaga ttttataaaa ttaagttctt tccgtgtaaa
14161 aggggtggaac tctgccgtct tagagttttg ctcgttgcct tgtgctaacg caaaggctaa
14221 atcaactgat ttttttttaa aaaaaatgaa tttatcta ataaataaaa ataaatttaa
14281 attgttacia tttttgtatt tattaanaact ccccatttta aaaagtacct ttgacttttt
14341 gatgtccaaa ttcatataa aattctctac attagaagaa tcatagctcg agaatggaac
14401 ttgagttgtg ttcagtttta aaaatgacaa cggaaattct ttaaaaagtt ctttaaaatt
14461 aagttgaagg cctaaatata cttttttttc ttttacttga aaagcgctaa actgggtatg
14521 ataagtgctc actgacaaaa atggaattga aatacattct aaataagtaa tgtgttgatc
14581 aaattcaaca ttatcaataa atgaaaatcc aggtttaaag agttgtttat gatagtctaa
14641 aatatcagtt tgatttcggg gaaaatatac ccaaccagat ttgattccta tttccaaaca
14701 gttgtcacac gaagattttt tgttttcttg aactaaatta acaggcaaaa cagaagatgg
14761 tgggcgagtc tgtgaatcaa tctggaacc ctccactttt tttaaagata tagaataagg
14821 tttgatcttt gaagctgagt aattaaaaag ttgggtacaa tcagaaactt gaaaatgaga
14881 tttaaccgag tttttagtaa taaaagattg ttgagaattt ataaaagggt tctcgtatgt
14941 taatcttgta attttgtcat attttaaaaa aactttgtta catgtgtatt tttcaactt
15001 taaaaaccaa ttttaattga tagaattaat tgtaaattgt gaaatttgtc gagcttcatt
15061 aatgttattt agatagtttt ttttataaga tacactttta taagtttcac catgtctttg
15121 ttttaactaaa atataagggt ttgggcgaaa acaattattt ttaattgaca atagccctac
15181 ggtagtaagg ttgctaagct tcgcttgttt accccttgcc ttattgaaaa tcaatgaggt
15241 aaggaaaccc gagttgatca cttcattttt accttgggat aaaataaaaa ccaacgtaag
15301 gaatagagat ctatttttaa aatttataaa ttgcttggaa aggaaatttt tcaagcaagt
15361 tacttcccag aaccaagact ttggaaaaga atttaatatt tcgccgttag gttgaatgga

15421 acaataacca gttaatthct tagtagactt aatthctaac cagccactca tcttagacga
15481 aaagthctta atthctcctt gtgaattata thtagataaa agcgggaaat ththattaac
15541 ccactthctt gcaaatthct taacaaaata agataaattg thtgaggatca thtcaaatgt
15601 cggaaaataaa acaaatcctg tatcattagg acaaaaaaga gththctggat taactthata
15661 actthcttgc ggaaccctaaa aaatthctcc acaatcattt tctaaataaa gagagtcata
15721 aaaaactaat cctcctgttt tagththata atthctatgga aagctthgca gatataataa
15781 ththtgtaaa thcactgaaat tatcgagtht thgththaaa gagctcgata aaaaaataa
15841 atcagtggaat actgggthtg gtaaaccagc gagaggcatc ctgaththta tactaggctt
15901 thgagcattt tgaattaaat tggaaacctg gtctthaaac aactththct thagthctct
15961 atththgatgt thactatgtc cggtagaaca ataggtgaaat ccagtagtaa gcaaacca
16021 thctagthga ththcagtht ctggaagatc thgththgtg ththactaaa thcaagctta atthcttatt
16081 atcataatca aaagththth thgtaaagct ththactaaa thcaagctta atthcttatt
16141 gctccaaaaa thtaaaaaat aacctatthg ataataagaa atathththta taggaaaaga
16201 caatacagga taatthtagag ataaatthct cctcttatcc thactatagt ccttaggata
16261 acgatcaaac ththctgact thggtaccct tgccgtgcc ggcaggcaa cgggcaaatg
16321 agaaatthgt atgthththaa aatctgagga tataaataag ththggcaat thtcataacg
16381 ththththaa thcactaaatg aaatthgtht thctagthgca thagtathaa aaatthgatt
16441 ggaaggaaaa gctthctcatg aagcagatth tgaattataa aactgathth tagthaaaaa
16501 accagtataa ggactaataa ccaaatthg atththact gaacctgaaat thactaaatc
16561 acaagththt ggaaaaaaag aagthggtat catgththga tatactthct cagataagat
16621 ccaaatagac ctagththth tagcaacagc ggttatgtcc cctccttac thctthaac
16681 agctaaaaca acaththcaa aaaatacctg acctctcatt thtgaathta aattatgtht
16741 tgcttgaatt cgthgattac ththgagctga aatthgatgag aatthcagaa taagthgtht
16801 thctgagaaca aactcacctt ggcgtacaaa thaaatagtg gaagctggta ththaaatth
16861 thgththctgaa ccagthththt thththgtht tagccgagac aaatthctct thththgca
16921 atththgtht aatathgtht caaaccaag gctthgthaa thgthggat thththaatg
16981 ththaatataa aatthcacct gaactthtagt aagaaaagct atththctcat gaggtgthct
17041 aatthaaatth cctthgthaa atthctgthaa thcaataata ccatcaaatg thgthcagaa
17101 thcagthctt aaatcacag aaaaaactc acctagthga aaagthctca tagthtagtht
17161 thgtthctggt thcacaattht atthggctgc gagaatacca actgctthct ctaatgatac
17221 taagththca thgagctaaag thcaccata acataatthg caaatagaa thththgctt
17281 acaagththt ggagatcgaa thaaagctth ththctthaa thtagcaatth thththgctaa
17341 agththcagat atthctthgat ththactgct gattaagthta thctggcagc thgcatthth
17401 agththgataa aatthththt tactagaaat thaaatthaa ththgththt thththgatt
17461 thaaacttht gagathtagag thththctcag thaatathgt thctctatat thctcagaa
17521 thggctggga acctgagcaa aatcgaaata aataththct thtagaathc thctatthaa
17581 thgaththct aaagaaagta atactththg gcctthctat atactactta agaaatthct
17641 gctthgctgth ccacaatcaa gthgacaaac gataacatgt thgagaaacat thactaatct
17701 acgtgthtaag taaccagaaat tagctgthct thaatgagth thctaccagth thctaccagth
17761 cccataacac gaaataacat atthcggthca agthtagcct thctctaaaat thactthgaa
17821 thgaaaaatca ataatthgct ththgggagc agccatthaa cctcgcathc caactaathg
17881 acgaaacttht gaaacathth ctctgthctc thgaaaacgc atcatgthaa cgggaththaa
17941 aatathcagth gctctaaaat atthgaaatc gththgthth aaagththct thgthctatg
18001 ccagthtca ataatthgtht ggaactthct aacagcagth aaatgthct ththgthattg
18061 aatathggct gaagthaatth gthaatthcag thctgaaact aatththgath tagthaggag
18121 ththththaa thcatcaagac thagthgaaac gcctgctthg ththgctact caaagctag
18181 atthththaa thththcaact atthcaatggt aaatththca ccacaathaa thaaagacca
18241 thgatataaaa gathththaat thctththct aaaaagcca thaaaaaaag thaaatthca
18301 aaccathact thgtctctth thgagththth thctathcat thtagactct accgtctth
18361 atgthctacc gctgthataag ththaaagact thagththcaag ththgthgct thgcaaaacc
18421 aggatcttht actacgthga agthctthaa aaaaatthct thgthgthgth agacctaccg
18481 aacacaataa gathtagcaag ththctctcat atacctgth agagathact aatctthgth
18541 thcgggthaat gththtagaaa caagccaaac ththththtag ththththgth ththaaataa
18601 thgctththaa thththgaaat thggthththaa agctagthgth ththaaactthg ataacgthg
18661 gatcaatgth aggthactth agaaataact aatthththg thgaaatct thcaatththg
18721 ththgataatg thagctctg agthththcaag gthththact thctcaaaacc aggactaat
18781 thaaacaaac gthagththgth thagthctth cacctthg thththactca thagththgth

18841 ttaatgagca atagccaagg ggcaaggact aattaaatta tgtttaatta atagaaaata
18901 cttcttttat gggctttgcc aataaaaaaa agatcttcca taaactggaa gtctatctaa
18961 acaaattaat tatttaaaag aatttgattc agtaaaatcc tacctgcagt tgttcgaata
19021 aattgacttt tttgagttaa atttttatct gaagtacggt gatattttatg ataaagatta
19081 cttatatattc caaaacaatc aatttctaatt tctagtgggt tttcaattaa attacctggt
19141 tcaaactgac ctgaccatcg taccagatt ggcgcatgta atgcaatfff tttctgattg
19201 aaagctttta aactatcact taaatcaata aagtatttcc caatcccttt taatgctttt
19261 gaattttag ttgttaagta ataacaacct agcaccatat cttgacttgg aataagtatt
19321 ggttgcccag ttgctggtga caataagttg tttcgagacc acattaatff ccaagcttca
19381 gctcgcgctt gaaatgataa aggaacatgc acagccatff ggtcaccatc aaaatcagca
19441 ttaaaagctg gacaaaactaa tggatgtaaa agtatagctc taccatcaac taactttgggt
19501 tgaaatgctt gtataccaag gcggtgaagt gttggagctc tatttaataa aattggatgg
19561 tttttcaaaa cttgttgtaa tatttcccaa attacaggat cttgtcgttg aatcaatcgt
19621 tttgcactta caattgtccg aacaatttga gttgtcatta gtcgtctaaf taaaaagggt
19681 tgaaaaagtt caattgccat ttcttttggc aatccacact catgaatfff taatfffggc
19741 ccaacaacaa ttacggaacg accagaataa tctacacggt tgcctagtaa atfffgtcta
19801 aagcgcctt tttcccttt taacatatct gataaagatt ttaaagggcg atcattagaa
19861 gcacaaattg gatctgctcc ccctttacca ttttctatta atgcatcaac tgcttcttgt
19921 aataatcttt gtgcgtatff aatttcatca gatttattag aacaattact tttttatga
19981 cgtttcattc tattatffct aatagaact tttgataga gtttatttag gtcagatact
20041 gctacttgat caccatctaa ttgaataatt ggtcgtaaaf caggaggaag aactggaaga
20101 attgataaaa tcatccattc tggttgtgct ttagttcgac gaaaatgacg aacgagcttc
20161 aatcgtctta atffftttagc tcgtaaagaa agtaaaaatt ttaatcgacg ttgttcaatt
20221 aagaataaaa aatfffgatt ttcaaattca tgaatffctt catttaattc gaaaagttct
20281 actcgaatff gtctgtffaa aagctgcaaa tctaaatfff taagaaaaga gcgaagtgtc
20341 tctgcacctg tcaaaggfat atcaaaatgt gttggattaa tgaataatf atffaaaagc
20401 aaaaattgat cttttattcg atcaaaaata attggfatag gtgtatcctc agcttcaaat
20461 gaagctgaca tataatagga aaatffatff agatcgtctt gaacttccca agaagcgggt
20521 tgaaaaatag tataataatt atffacttgc ttgtfftgac tagatffttc aacgggaggt
20581 ttagttaaac ttgtctcttc aaaactaggf ccaaggctac gaccttctt tttttttct
20641 agactgatat taaatctttc acccaatggt ttcccttttag aagaatgacg ttcttctttt
20701 gaaaaagaag acaacaatc atcgtagccc gggattaagg gctcccgggt aaaaacatta
20761 taatffcttt gcttcgactc tgaatffgaa tcttgttcta tttcttgagt tttgtgggta
20821 tcttttagact gaataaaatc agagttccac ttagtttgtt caaaacggtg ctctgtaaa
20881 ttttccaaat ttaatactgc gatacctgga agtgtttgat tactctttgc attgctgaaa
20941 cgcgaggggg taataccact atataaatct gcaaaattaa ttctgaaaf atffgttgtt
21001 tctgtacaaf aagtaattgc ttctactfff tttcgtcgca tttctaaaaf aatagaaata
21061 taactgggtg tccctfftaa ataccaaaca tgagtaacag gtgaaaccaa ttgaatatat
21121 ccaaggcgat atctgcgaac acgcgctgaa gtaaattcaa catcacaatc tggacaaaaf
21181 tgttgttgcg tcttatfftt ttttcgacca caagcacatt caaaatcttt tacgggacca
21241 aaaatccgct cacaaaataa gccatctfff tcaggttffa atgttffata gttaacagtc
21301 tgtgaactta gaacttgtcc tacaatffta ccattcggta aagttcgttc tgcccattff
21361 ttaatacgtt caggagggtgc taacccaact ctaatagagt ctactfftagt aatfftactt
21421 gtatffttcta tattcataaa aagcgtcaa tttcaatfff tgcagtaagg tttffttata
21481 atgaatattc attataaaaa atataaagct tgtffaaagc gttffatffta ctgattatgt
21541 caaaaaatca aaccttaag ccattctfff taatattgtt ctttaacgtg tactaaatat
21601 tttatctaag tffaagttff gctffttgff agctgtaaac ataaaccga gaatffttat
21661 tcaggggatc cagggagccc taactaccg tgaactaccg gaggtagtaa ggttgacaag cttcgcttgt
21721 ttaccctatg cttccactac gtggaagacc ttaatacaaf ctcggtatgt atffttccct
21781 tgccgtgccc ggcagggcaa cgggggaaga acacaagtct aaaatacaaf ccgagattgt
21841 attaaatcac tatgtgaaag tccaaagtgt gtgcttgac acactaggct ttgcccgtag
21901 ggcaaacctt tatttaacaa cgaagtgtt aaatffagac ctaatctaaa caactatgt
21961 ttgtfftaga tcttcccttt agggaagtct tcttcttagg ttcttaata ttgttgcagt
22021 tggcaacaaf ataaagttff aacaagctac gcttgttatt aggttccgct actffaatcc
22081 gtaccggtat ccagtaggct ttaccttacg ggtaaaaatt tattcaagac tttcaccctt
22141 tgccattaca agggcgctgg aagggtgaag atactggttt gtgctffgca caaagcttgc
22201 aacttcgctt ttgaatffac acagcggtaa gtttaaatag tttaggcaag gttataaaa

22261 acagttccga agtttgtgct cggcctcttg gtgtttacca aaaggtaaag acacaaacta
22321 aggttttatt gttaaaacgt ctaccaaggg gtaaataagc atttaaaaag ctaaaattct
22381 taagttacac ccctcgctaa atataaaact aaaagtttat atacagtggt aaccttctaa
22441 tgaattttgg ctatcttcgt tatgtttttt tgcaaaacat aacgaaaatg ttattaaana
22501 tgaatttcg accagaaaca atatagacaa actacacttg ggtaataccc ttagtgtaaa
22561 actggaccgt atgggtaatg tttcgagatg aaaaatatta aaaaaattaa gaaagtttca
22621 taacatctat ttgcttacgt cgcccagatg attcaatagt atagactcca acatctaaac
22681 acaaagattg aagctcacgg attaaaactt taaaagattc tgggtgtccc aatgaaattg
22741 gtcgattatt taagatagca tccattactt gatggcgctc tttcatatca tcggatttta
22801 cagttaaaag ttcttgaagt gtataagctg caccaaatec ttcaagtgcc cacacttcca
22861 tttcaccaag tcggtgcccc ccatgcttag atcttctctg tagaggctgt tgagtaacta
22921 acgaatatgg tccagttgaa cgagcgtgaa ttttctctgc aactaagtga actaatttta
22981 acatgtaggc ttgaccaaca gttacagctt gatcaaagca ttctccagta cgaccatcaa
23041 aaagttttgt ttttctgca aaattgggat caaacaacca tttttgacct gttttcattc
23101 ttgcttcaaa aagttttgaa taaactagac ttctagacgc ttcaggaccg tgaatttcat
23161 cgaaaggaat aatttttaaat tgctgctcta attgtttacc tgctaaacca agtaaacact
23221 caaaaacttg acctacattc attctagatg gaacacctaa tggatttaat accatatcaa
23281 cgggggtgcc atcaggtaaa tatggcatac ctggcgcggg taaaatttta gaaataattc
23341 ctttattacc atgacgacca gccattttat ctccaacctg tttttccgt ttttcggcaa
23401 cgtaaataatg tacacgacca ggtccttcaa aatcaatttc aggtggaata ttttctgttt
23461 caagaatttg aatattaact actcgaccat gaacattttt tgggactcgc aaagatgtat
23521 cccgtgttgt aggaatttct tttttaacta tgtcatataa taattttctg tgagggtgaaa
23581 gtgatttttt ctgaattggg gttatttttc caacaagaac atctccttct tttatccaac
23641 ttccaatttt agcaattcca cgattatcta gatgtgaaac ttgagaagct tcaatttccg
23701 ggatttgatt tgttattttg tcaacaccaa attttgtatc acgaatttca atttcatatc
23761 gttcaatatg aattgacgta tagatctcat cataaataag tctttcacta attaaaatag
23821 cgtcttcaaa attatatccc tcccacggca tataagctac caaaatagtc tttcctaattg
23881 ataattcacc cccaacgctc gctgtccccat cagctaataa atctcctttc tgaaccact
23941 ccccttcattg gacaattggc cgttgagtca aacaagtatc ttgatttgac cgttgatatt
24001 gatcaagaaa ataattgata gaagacaaac taaaattttt ttctatatta aacgtatttag
24061 aatttgcact ttggtggttt attgaagagt tagttgtttt ccatttctgct ggaacagaca
24121 atctagttgc tagttccgat ctttgtttaa acaaaaaatc agttcgagag atagaagaaa
24181 aattatttga taaacgtggt tttatttcta ataacctact aggttttagtt aattgtttat
24241 taaactctac gtacgtattg tgcttagtgc tttgaaaact cgaccacttc aaattactaa
24301 aataatctaa aactgatagc tgctaataa attcataatt aaaagatacc ttagttccag
24361 tgttttctac gttgttttgc cttaacacaa ggggcacttc aaaaggagat tcagagatac
24421 cagggaaatt caaaaaactg aattttgata atttagaagc tctattccaa ccctgcagtt
24481 catcaaattt gaacttaaaa tgacctttcc acgggttttt agaatgagtg ttattataaa
24541 aattattggt ttctaaaata caatcagaga ttgtattcaa gtcttccacg gagtggaaatg
24601 ccaagggttg tgcttcgctc aaaccaggat ctccagaaa ctggaagtct attaggtggt
24661 gaaaaagaaa atggtaccaa ttaaaaatac tagtaactg gagttttaaa atattttttc
24721 ttattaaana accttttttc gtataatgtg aattttctaa aaaatatcgt ttcttttttg
24781 aaacacaaat ttcacgacag atcatttttt taaatgatag tatttgatat ttctcttcat
24841 ttaatttttt tgtaagttga ggttgaaatt ctacattatt tttcaaataa attttatcca
24901 agttgcaatt ttccccaaagc accggttgtg ttttccattg ctttttatcg atacgaaggg
24961 taagacaaga ggaacaaaac tctaaaatac aatccctggg tgtacaaaaa ctgggattac
25021 ggaagattgt attagatct tctgaattca ccacaaaatt ttgctgcttt aataaaacct
25081 gacgagataa attttgaacg tgtgaaaatt caaataattt tttttcttta tttttttgt
25141 tttgtgtaca ggttttact ttatttatgt caccaccaat tttattagta accgggtaaaa
25201 caaaattaga agacaaaaaa ctgtttagaa ttgaaaattt ttttaattgt ggactagttt
25261 ttatagtcga cagattcatt ttttaactcta tggaaatgtaa tttattttga ctttttttaa
25321 gacttcgaaa aataaaagggt ttttcttgag taattcgatt tgattttttt tggttattta
25381 aagctaactg tacatgagaa ttgacttttc ctgtttttaa aaaagatcga gaactgtttt
25441 ctaataaatt atagaccgga gcctgagaac gaaattttga attgggtgaa tttcgagaaa
25501 aaactctttt tttcttaatt tttttaacta agtttgaac taaagatttt gttatattgc
25561 cttgtgttgc cacaagggat aaagttgggt ctataaaatt ttttatagta attcgttcac
25621 cactcacata agacacaaaa ccacttgttc ttgattgaat agcatgtcct gaatcagcgg

25681 ctaccgctcc ttctaatecg gttcctacta taggtcgttc gggcgcaatt aaaggaacag
25741 cttgacgttg catatgtgag cccatcagag ctctatgtgc atcgtcatgt tctaaaaatg
25801 gaatcaacga agttgcaatt gaaatcattt ggatagggga tattgaaata tattcaacct
25861 cgcttcgact tattctttta aattctttta taatcctaac tggaatagtg ctttttggtg
25921 aaaaatctaa ttggcttagt ttcaaatcac ctggagctat tttaacgtgt tcttcttgag
25981 cggcagaaaa aaaaataggt ccagcagaca tctgaatttg gcctttataa acactataaa
26041 atggagtttc aataaatcct tcagaattca cttttgcata tgcgtttaat gaattaacca
26101 gtcctgcatt agggccttca ggagtttcaa taggacatat tcggccataa tgactaggat
26161 gaatacctct tactgccata ccagctgtct cagcactgat accacctgga cctaaggaac
26221 ttaaactgctc tttatgagta agttcagcta aaggatttgt ttgatccata aattgtgaaa
26281 gttggcttga cccaaaaaat tctcttaaag ctccattaac tggtttagtt gttattaaac
26341 ttcgaacagt taattttttt cgaggctttt tcaatttttc tcgaatcatt ttttctagtc
26401 gaataaatcc agtacctaatt tgattttgaa ttaattcacc agatgcgaga actcgtcgat
26461 tttttaaatc atcaatatca tctagtgtgc caagcccgta ttccagttta attaaataat
26521 cggtagcaaaa taaaatatct tgttctgtta atactgactt agtaagtggg actgataaac
26581 ctaatttttt atttaattga atacgacctt atcgacttaa atcatatgta cgagaattaa
26641 aaaattttct aagtaaaaaat ttttgacctt gatctgtagt aatttctgat ttttcttct
26701 ttggatgtgt ttcgaaataa agagcttgta atgcttgacc acttgagatt ggggtgagcga
26761 atgaatgctc ctctttttaa aaagaagatt ttaaaaaatt ggagtaatta atagattgaa
26821 aaattttttc tttacttaac cctaattgctt gaagaaacac taaaatagaa acttttggtg
26881 tttttttcat tcgagcccaa acttttctgt ttttatctgt ttccaatctc aaccaagcac
26941 cacgatgaga tattaaatct gcataataag tacgtttttt atttttatga actacttggt
27001 gatagtaaat tcctggactt ctaaccattt gattgataat aaccctaggt gaaccattta
27061 agataaaatg acctcttttt gtcattaaag gcaaattacc aagaagtacc cattgtagtt
27121 tgctttcttt agttgcttta ttaattaaat ggataggaac atataaccga caagcataac
27181 ttttagattt taaaatacat tcctttggtg tccagtcggg agggtttaat ttgtaatatt
27241 cgggataaaa acgaagctct aaattttgtt tagtattagt aattggattt cgtttatata
27301 attcttgaat taagcctttg tctaaaagat ctaaaaagct ttttcgttga atctctacaa
27361 aatcaggaat aaaataagaa attcttacgg aagaaaagga cattttttgt cagctccaat
27421 tgaatttata ataaatttga cttggaacaa ggctccaaga ccttagttaa caaacacaga
27481 agacttctac tatgtagaag atcctaatac aatctctgat atgaaatata ttttcttaaa
27541 ttaccgggaa gccttagatt cccgggctac gccctagtaa ttttaaggaag gattgtactt
27601 tggacttttt tccacagaag actccactga aggcgaagac cttgggttaac aaacaaagtt
27661 tgttaacttt tgacttctac gtagtagagg atcctggttt gttctttgca caaaccttgg
27721 acttgtgttg tgtgcccttt ggtctttact cattagctat tgctaattgag caatagccta
27781 aactaacttg ggtagcggag cgcagaaaaa aaagtccaaa atagaacaga gttctattag
27841 gatcttccac tttgtggaag tcaaaaatgt gtgcttgcac acattagaat ctttaagctt
27901 cgcccggtgc cctgccgggc acggcaaggg ttaaagtctt ctgctgaagg ttggcaagct
27961 tcgcttgctt agcaaagggg aaagactgct ctgtgctacg cgaagtttat taactttgaa
28021 tttttgttac ttctactctt tagcaagtaa agcttaggtt tagatctttt gatgtttctc
28081 ctaaggccat tactcgggtg tttttgtaa gaggcttcac tctaggggtg aaatttgcac
28141 aagcttgggt tgttttagga aagacaaaaa ggttttgctt aagctacgga aggtagctta
28201 ggcaaagaaa aagacaacga tgaagtggaa ggctttgaca aagtgaattg ttttaatatt
28261 ttattaaaac ttgaacgcta agaacaaagt ttttaactt aaacaaaagc tcaagcttct
28321 atatattaat tagtagatct aattttagaa atatgttata attgtataaa aaagggggcg
28381 tggccaagtg gtaaggcatc ggattgcaa tccgcgacct ccagttcgaa tctgggtgcc
28441 gcctaactt tttatctttt taagatgtaa acctttcaaa ttaaattggtc tttgtcaaga
28501 ttacaaactg actaataaaa tttgtatgag tttcaatata tttataattg tttactttt
28561 atttatattg taataaagct aatctaattt atataattag gcttttaatt tttttgata
28621 tggtaactat aaaaacatta aattgtaaaa gttgttacct tagtaataaa tatgttttca
28681 aaaacaaacg ggtaatcttc aaactaatat catagagact acttaacaaa tttcttaaaa
28741 aaaattttagc ggcgttttat ttagactacg tctgtaaaat ttataacgac ttctctttca
28801 gagaaaacct tagttaataa agtttgtttg ttaacttttg agattttaag aatttttgca
28861 tataaataaa ttttattttt tgtcttatga gtgattgaag cctttgtttt taaagaaata
28921 cagaaaatth ttaaatataa tttgattaat tttattttat ccattgtgga ggccttatat
28981 ctgagactac ggctattttt taacaagatc tttgctcgaa actttagaat taatthaaat
29041 tgcctgatta actgcacgct ttagattcat aagttgataa aaaattgtta ccctcttctt

29101 gtattatctc tagatgatca ccacccggta aatcccgtac aaaagaagga taggaagttc
29161 aaagtttctc cgaaaattca aattttagaaa aaaaaattta ttttttagag atataaggtt
29221 tagaagaatt tttaattttt gggtaagtta gtaaagctcg cgaattttac ctaaaagaat
29281 tgtcttccgt gttaagctag taaaataagc tttgtttggt aaaatagggt aggtgtagtc
29341 tccgggtggtt gcgcacggct tgatatttac ttttgctgaa gattatcttt gaaattggaa
29401 tactctatgc tccaaaaatg aaacaattag aacaagattt tccgatttcta aaaatcaaat
29461 attcaaataa tgaacaggca atttttctaa atcttttcga tgaagcctcg ataacaaatt
29521 ttttagattga acttaaaaaat taaattatca acacttagta agtctagcaa attgcatatg
29581 tgaatctaata caaaaattgt tataagtgtt taattttagc ttttagcgaa ggtccattag
29641 ccctggctat ctaaaataac ctttaatggt gctactttta cgtcctgaca tgattccagt
29701 ttttaagatta cagggcaatg gacttaataa tctatcctag caatttattt tgtaaaaaatc
29761 aagttgattc caatttccca aagtttggtgc aaagcacaaa ccaggctact aattaacat
29821 agtttaattc gtagaaaata cttcttttat tggctttgct aataaaggaa aaatcttgca
29881 gtttatggaa gtcttctata tttcaaatta aaaatttgat ttatttggtg gaactttagt
29941 taaaaacccg cgagtttcac tcttattgac ccgttgctaa atgtagtata aaagatcaaa
30001 cataaagttt tacacttcga cccgggaccc ctagggtccc gggctacgat aaagcctagg
30061 gggaaagcct atggttattt gagcgaaact cgacaacctt tccgagaaaa ctttttttct
30121 acgttacgct accgaaggta gtttaggcaa agtctttttg taaagaccac tccaatttta
30181 gcagcgcatt gggagtaagg agtaaaaggc actttatgtc taaattttta aacacagcaa
30241 taggacctaa tctaaacaaa ctatgtttgt ttaggaatac gccttgctct taaactttta
30301 aaattaaatt ataaaatggt gaaaaaaaaa attctatatt tatggtataa tttaaaataa
30361 ttatataaaa gaaagaaagt aaaacagttt ttattttata agaaatattt atgggattac
30421 cttggtatcg tgtccataca gttgtattaa atgaccaggc tccgttaatt tctgtacact
30481 taatgcatac atcacttgta gctggttggg ctggctctat ggctttttat gaacttgccg
30541 tttttgacct atctgatcca gttttaaatc caatgtggcg tcaagggatg tttgttctac
30601 cttttatgac tccgttagga ataacacaat cttgggggtg ttggactatt agtggtgaaa
30661 ctgctaataa tccaggaatt tggagttatg aaggtgtagc tgcttctcat attctactgt
30721 cgggtgctgt atttgctgca tctatttggc actgggttta ttgggatctt gaactttttc
30781 gcgacccaag aactggtgat ccagcattag atcttccaaa aattttcggc atccatctat
30841 tcctttctgg acttttatgt ttcggatttg gtgcatttca tattactggt ttattttggtc
30901 ctggaatttg ggtttctgat ccatatggga tctactggaag tgttcaacct gtagctccat
30961 catggggagc agaaggtttt gatccatata acccaggagg aattcctgct catcatattg
31021 ctgcaggat tctaggtggt ttagctggac tttttcatct ttgtgttcca cctcctcaaa
31081 gactttacaa tgggtcttaga atggggaaca ttgagactgt cctttcaagt agtatagcag
31141 cagttttttg ggcagcattt gttgttgctg gaacaatgtg gtatggatca gctgcaactc
31201 ctattgaatt attcggacct actcgttctc aatgggattt gggttttttc cagcaggaaa
31261 ttgaaaaacg agttcaatct aatttagcag aaggaaaatc actttctcaa gcttggtcag
31321 aaattcctga aaaattggcc ttttatgatt acattggtaa taatcccgca aaaggtggat
31381 tattccgagc tgggtgcatg aatagtggtg atggatcgc tgtaggctgg ctaggacatg
31441 ccgtattcaa agataaagat ggaaatgaat tatatgttcg tccgatgcca acgttctttg
31501 aacattccc agttttatta attgataaag atgggtgtgt acgtgcagat gtaccattcc
31561 gacgagctga atcaaaaatac agtattgagc aagttggggg gtctgttagt ttttacgggtg
31621 gagaattaga tggggttact tttagtgatc cagcaactgt taaaaaatat gctagacgag
31681 cacaattagg tgaagttttt gaatttgatc gtgcaacttt acaatctgac ggtgtattta
31741 gaagtagccc acgtggttgg tttacttttg gtcatttatg ttttgcttta ttattcttct
31801 ttggtcatat ttggcatggt gcaagaacca ttttccgtga tgtttttgca ggtattgatg
31861 ttgacttaga tgaacaagta gaatttggag cgttccaaaa attaggtgac tcttctactc
31921 gccgtcaatc agtataaaca tttttaaatc tccaaagtta acaaaacaac tttgtaact
31981 aaggtgttgt cttccagaca agtcgtagaa caaaatttaa tttaatagct ttgaaagtta
32041 gcgaacagat attagatatt ttctaactct tgttttaaaa attcatcgcg atagaacttt
32101 ataaagtaaa aaataaagggt taaaaccaag ggccattctt tcatttaatt ttgtcaaaat
32161 ccaatcttcg attccacttc gtggaattaa agtttctcta cgtagtggat catagcccgg
32221 ctattgctca ttagcaatgc ctaatgagga acctagggct cccttctat aagtaaaaac
32281 ttgaccttac ccccttgccg tgcccgggaa cctagggctc cggggttgc cctattgctc
32341 aacgaagttg ggcgtagggg tgagggtaat agttagagaa gcttccgattg tattaggatg
32401 tttcacgaag tgaaaatcag tttttgaaca caaaatgttt aaatagtgca aacttattta
32461 aaatatgcgc tctaaaattt ggaacacctt actcaaat ttactttggt ttacgtttta

32521 tgtaaaacca ggcataatgtg acacacgtag tataaattac gtaggaagag ttttttatgg
32581 ttataaagtt tacatattaga ttacaaataa aatattgata cgagttcatg agtgaaataa
32641 aacataatgt tcgcttatag aagctatttc aaaaccttta tattccgtaa aaaagcaaaa
32701 taatatgtga agtaaataat ttcgtaaaca tggttgtttac aagctttatt ttttgaacat
32761 tacgtatggt tacattcaaa ttcaaataag atttatgcaa agttttcact aagagtgaaa
32821 tcttatataa aaatcttacg aaaacctgta cttttaaag gtaattggaa agcttttttc
32881 ttagtcacac tttagattta ctattgaatt aatatgtaag cattttgtgt ccagactgaa
32941 aaaactttct tttaatattg ttgccaatc aataaaaact ttatttttat ggaagcttta
33001 gtttatacat ttttattagt aggaacttta ggaattatct tttttgcaat tttctttaga
33061 gaacctcctc gtattgttaa ataaaagtaa aaaaactttt cttaaaaaat ttttttttaa
33121 gacttccagt ttctggaaga tctttccttt attggctttg ccaataaaaag aagtattttc
33181 tactaattaa actatgttta attagtggga gatcctgggt tatgctttgc ataaactttg
33241 gaaaacttaa tgactttaca acttggggaa tttaaaaatt attattaatc tttttattag
33301 caatatttta agaacgatat gaggtttttag aacattaaat aaaaatttat aaaattaaga
33361 ggctttgaaa agattaagac ttcttttagtt taaattttta gtgaaaacct tggagaaatt
33421 ttaccaaaca tggtaaaaat aaaaaacctt agtttgcaaa aatttgggac ttctgcaacg
33481 tgaatctcc taatacaatc tcctattcca ctttatggaa ttaaaatttc tctactacgt
33541 aaattccctt ataagggaaa acttactatt acctttcggg agtcaagcgc agcttgacaa
33601 tccctagccc gggaaacctg ggctcccggg ttgacctat taccacga agttgggctg
33661 aggggtgagg gtaatattta gagaagcttc gattgtattt tagattcgtt ttacgagtga
33721 cactttatgt tttatttacg ataaaatatt tagaaacgca acttttaaac ttttattcta
33781 aaagctgtta ttttttattt attacactta gtctttacct tttggtaaag accaaaagga
33841 ggataaaacg tagattttga ctttaatttt gttcagttta cagcttatga agatagggtc
33901 gtaaaactgg atatgaaacc aagtaaattt aatcttcgtg ttcctcaaaa gggtcacgaa
33961 gttgttttga tgggtggtcca aatcctacat aaattgaata tccagtaata cttaaaagaa
34021 gacaccataa aaagattgtg taaaaaaaaag caggactttc cataacgagt tttaaattaa
34081 aagatttata attttttatt cttgagcctt tgccctggcc aaaggctaag gtggccataa
34141 agcaaaaagc tagattttta ctcttttgtt aaaacaaagg taagtaataa attttttgc
34201 ttagtataa atttattaac agtttaagtt taacaaaaaa aagcgtttta taatttagtt
34261 ttttttaat attttttagc aataaacta atttaaaagtt ataaattttt tatttaacct
34321 tccccctttt aataatatta cagaaagtaa aaaaatctaa gaattttttt gaacatttta
34381 tggcaatagg aactacctct aaaactaac aaagtccatt agatactggg aaagtaactg
34441 ctttaggaac tttactaaaa ccattaaact cagagtatgg aaaagtagca cctgggtggg
34501 gaacaactgt tcttatgggt gtttttatgg ctctttttgc agtgtttcta gtaattattt
34561 tagaaattta taatagctct gttcttttag atgatttaac tgttagctgg gattcattag
34621 caaaataaat ctagtaaaaa ttgtatttta gtttggctta aatttttaag ccaaactaat
34681 ttattttaat ataaataata ttcaataact atgtattaat atcattaatt accgacaaat
34741 ttaaaacaat tatatagcta ttaattactg aaaataaac ttgtaacaac gtaagaaaca
34801 gattaaaatt aaattagttt aaacttttta aaaaagtttt tataagtata aggactttgt
34861 tcttaaactt agacatcatt ttaaaatata acttatccaa ttttctataa aataatttat
34921 aagcctcca tggagtggaa taccctagta aataaatttc gtttgtttac tttggatttc
34981 tgtgaaatag acaatcttaa ttaagatagt ttttgaatta taaacgacgt ttataattca
35041 attgtactat aaaaacttat aaagtctatg taaaaacctt tagaaaatat gtagecgttag
35101 aatatttgaa attaggtgtc caaagaactt tatagccaaa cttttttctt acaaaaaata
35161 gtgtattttt gagtaacagt cccatcaatt ttcaccataa ctagataatt taaaaataa
35221 tctatataca agtttaagga gatagtattt atgccgattg gagttcctaa agtaccattt
35281 cgtttaccag gagaaccaac agctcagtgg gtcgatttat acaatcggct ttatagagag
35341 agagttttat ttttgtgtca agacttagat gatgagttag caaaccaatt aattgggatt
35401 atgctatatc tcaatgcaga agaaaaatca aaagggttat atatctatat taactgccta
35461 ggaggggtccg ttacttgtgg aattgcagta tacgatacaa tgaattatgt aaaagcagac
35521 gtaacaacca tttgctgtgg aacagcagct tctatggctt cttttgtatt agctgggtggg
35581 gatcgaggaa aaagaattgc tttgccacat tcgcgaatta tgattacca acctgaagga
35641 ggaagtcaag gccaaagctc agaagtactt tcagaatcag aagaagttat gcgaattcgt
35701 cgtcaagttg gcaaaattta cgctgaaaga actggtcaaa ctttaagtag aatttcacga
35761 gatatggacc gagatcagtt tatgtctgct cgagaggcta aagagtatgg aattgtagat
35821 caagttgctg ttgatgtaaa atgggtctgg tcttaatttt gaagaatttt ctttttcaa
35881 aataaaaatt caccctttg gtaaaggctt taacaaaaaa agcttgtcct tagtttgtgc

35941 tctacctctt ggtcaagatc aagaggtaaa gacacaaact ttagaaccgc ggtcgccaac
36001 cttaccccct cgagctgaag ttttaaacca gggcaaagac aggagcctaa cctcgaaaga
36061 gcaaagtgat tatgcttagt tactttatat ttttaagcta cacattgtct catgtacgat
36121 cttttactag ctcttggtaa aaaccaaag gtaataaaag cctaaatfff ggttggaag
36181 gaatgaaaat atttatffff aatacaagca aaggatgatg ttttagagta acatgggatt
36241 ttaatfttagg ttcagtagca caatftftta taaaataact tcttataaat tftaatcagc
36301 atatatftgat aaaaaatatt aaacaaacca tftaatfttag gatgftgctt aaaaaagftt
36361 gagcgggtga taaatfttatt ftataatata atftctftct aaaaaaaaa acgftctfta
36421 ataagftftft tacaatftgtt tftftaatga ataataact cftaatfttat catftftftcg
36481 gtatfttaagt tftftaattca caaatftaatt aaaaagactt ccaatftftct gaagatcctt
36541 cctfttagtgg cftftgccaat aaaagaagta tftfttacta attaaacata cftftaattgg
36601 tagaaaaact tggftftgtgc gaagcacaaa ccttgaaata cfttagftftft ftattatcac
36661 tftftatctc aaaaaatftt tftactcaaa ccaatftaac atftftcgtft tftaataaaa
36721 aaactfttaaa aaaatagatt tgactftftta taactcgtct aaaaftftaaa gaagftftta
36781 aatftftacag taagtaactc gggagcccta gattcccggg ctacgcacca gtctaaaatt
36841 taatftaatt taacaaacta agctgagata gcgaagtagt tftctcgata tctftaatta
36901 agatfttatt tgctaatcaa taagtactat atfttaataat tctgftftftg cftaatgtca
36961 atfttaacaaa gacagaatta tftaatatag taaataagft aaaaaaaaa tftagaataaa
37021 ccaagagfta atgaaatct aatftggtaat gftgaccaa tftcftaacca aatagcaaca
37081 agtftaccta cftaaaaatc agtctgtagca actggctgac gaaaaggatt tftgaaatfta
37141 tftaatatftt caatgaaagg aacagftaatt aaacctgctg gtacagcagc catgaaaagt
37201 actcctaata atfttattagg tactgttctt aaaacttggg aaacaggata aaaaaccat
37261 tcaggtaaaa tfttctaatgg agtagcgaac ggatftgtag gfttctcctat tacagctgga
37321 tftaaaaactg cftaatccaat aacaccagca aaagtaccta agataacaac tgggaaaata
37381 tataataaat cgtfttgcca agcaggtctt ccataataat tgtgacctat tctftftgct
37441 aatftcgtct tftaaaatagg atcggataga tctggftftft tftgtaactgc catggaataa
37501 aatctccttc aaaafttaag tftagftaacg gacaaaftfta cataatftggg tattcaagct
37561 tctgcttgaca accttgcctc tftcggagtga agccgttctt aaaaacagca aaatfttgca
37621 tftgtcttaca acggtaaggg tftcaagftt cctggaagag gtgaggtftt gaaactftft taaaacggga
37681 agcaaggctt gactaccttg cctggaagag gtgaggtftt gaaactftft taaaacggga
37741 aaacatcaag aggtctftt ccttggcgtg cccggcaggg caacgggaaa acaagftatc
37801 tfttgacacc tctfttgcatt tgctgggggtg ctacaattca aagftftacct aatftfttct
37861 ctatgacctt tftataaaatt taaaagtata gctftftaaat taaacctcac cctactgaca
37921 aagcctacga gataggatat tgagcttctg tctgctgctt gftaaagtgc acttagctta
37981 aacatcagta taaaactgaa cftftaaatct gtatagattt aaagtggctc agaaatacct
38041 tfttttacgaa tcatftaagaa atgcatftagc atgaaaactg cagfttaataa tggftaataca
38101 aaagtatgta aactataaaa acgagftftaa gftagattggc ctacaccaac acctccacgt
38161 aacagctcaa caagagfttg ccaacaaga ggaatftgcat cgggtacctt agfttacaatt
38221 tfttacagccc agftaacctac tftggftccaa ggtftaaagaat atcccgtftt accaaacgat
38281 actgtacaaa cagccataat tactccagft acccatgfta atftcacagg tftftftaaag
38341 ccaccagfta aataaactct aaaaacatgt aaaaatcatca tftaaaacct catacttgc
38401 gaccaacgat gaatftgaac gattaaccaa ccaaaatfta catctgtcat aatgtattgt
38461 actgaagcaa atgcttctgc aactgfttgg cgataataaa aggtcatagc gaaaccagtc
38521 gctactftgta caagaaaaca cgtgaaatgta ataccaccta acaataaaa aatattftaca
38581 tgaggtggaa catatfttact tfttaaatctg tctgcaatcg atftgaatftc aagacgttcc
38641 tcaaaccaat catatactftt actcftaataa aagagctcct tatagftatt taaatcftaa
38701 tftaggctaaa tftftatftt tfttacftaa taaaagftatt gftftatftta tcatagaatt
38761 tftftfttata cftaatfttatt ftatagaaaa atftgfttagt tftaataatftt tftcaatcaa
38821 tftftatftc aacttctgtt gftatcaact tctagfttctt ggtftaaagata gagatftgca
38881 agctaaaatg ggtcatcttc cftatattat aaactftgtc atftagcttgt cacaaaacga
38941 cgtgatctaa ctggctaaft aaagctfttct ataactctft ggftaaacaaa gftftatftat
39001 tfttactcacc cgtagftgaat tfttftacct gtctacgcta ccgaagatag taaggtftgat
39061 cftfttactft tagcaaaaac aataccagft gctfttgtgat aacgcaagcg cftaaatgftt
39121 tftgctfttgt afaaatcgaa tftgggctaa aaaggcccta aaaaagataa tftgfttaaaa
39181 gftfttagftaa tftatcaaaa aatatftftaga ctgcatataa cfttctfttata afaataagat
39241 caaggtftftt taaaatcctt tagagfttctt aaaaacagftt tftaaagtgc tftgftaaatt
39301 aatgaagta tftaaaatatt agtftftftaa aatfttattgt cftaattgaga gatagftcca

39361 atgcttttgct ctttggacgg taacacaagc tttgtttata tcagactaga atttctcagt
39421 ctattcagag tccccacttc tctcaacaaa ttcgttacta tttaaatagt ttaggtgaat
39481 tttcttgatt aattaagaaa agcttaacag tctctaaatt taaaagctaa gcttttaaat
39541 aaagttttac aaactttggt tgtaaaacct ccctatcacc caacttggat aggtgatagg
39601 ggggaaggtct ttcgtcatta gttggcgagc tccgctcgca taccaagaaa taaagactta
39661 ttagcaatgg ctaattaaaa agttaaataag tgagtacgca gagcttttgt tccagagcta
39721 tgaactttta ttgaattttt acttttagcac ctgcttcttc taattgttgt tttgcagtat
39781 cagcttcttc tttagatact gattctttta atgcttttgg taatgacgtt gtaaattctt
39841 tcgcttcttt taatcctaaa gaagttaaata ttcgaattat ttttaaaaca gctactcttt
39901 tatcacttgg tacatcttct aaaataatat caaaagaagt tttttcttca gcagcttctt
39961 gaccaccagc tccttcaccg ctagecggca taaaaccagc tccaactgat gcagaagcat
40021 ctactccaaa tgtttcttca atttgagaca ctagtctggc agcttctagc agagtaattg
40081 attttcaatt ttcaataatt tcattagtgt tagtagacat aattttttaa agataataat
40141 taataaattt actatattta tagttttatt tgttcacatg ttggcaatga tctttcaatt
40201 tcagccatag tctaagttac gaactttagg tagtcgttac ttagtgcgac agaccaaagg
40261 gtaatcaagc ttcacttaac aagctgacct taaaacaagc tacgcttgtt aaagtgccac
40321 ccttcggtaa taaagcttca cttgataacc ttgcccctct tggcttttac tcattagcaa
40381 ttgctaatta gcaattgcta atgagtaaag accaaaggct tccctaaat gttttatcgt
40441 agataaaaca taaagtttta cgcttggcgt aaagccttct gggaaaactt tatttaacaa
40501 cgaagttggt aaatthagac aacggcaacg gccaaaggat agcaaggata aggatgaagc
40561 cttcggtaat caagccccgg cttaaaaatt cgtaacaaaa gttttttttt aaaagccaag
40621 cgctcgcagc gcaacggagc acagaaaaaa gtcttctgtg atgattaagt tcaatttttg
40681 tccaggcttt acaaacgtag cttgtaaaac tttattataa acaaagttta taatthagaa
40741 gtcaaatata aaaaagcaat aagattaatg actttctttt ataaaacat aagtttatgc
40801 ttaaaaaacta aatattaaaa gtaaagttaa acaaaaacat actttattag ttaagtcata
40861 aaatactctt caacagagat caataatcaa ctttatacac cataaggttt tgctttttta
40921 gacaaaacga ggtctaacc cttggaaaac gaaacattgg ttcgcagaat ttagtatggt
40981 aaggttttaga tatagaaacg aagtttttgg tttatttcaa tttgttacga acaaattggt
41041 gatttctaaa tttaaaagct acgcccattg gtttgtgcat tcgcacaaac ttttaaatca
41101 agttttacaa actttgtttg taaaactgac gtgcttgaat ctttagtctt tctttcaact
41161 aaaaatagta aaaaaaagc ttagactttt agaacataaa ttttgagttt tttacttgat
41221 aactaaactt atgtggctaa attaacgttt cgagaattga ggagcttttc gagctttttt
41281 taaaccatat tttttacgct ctttacaacg agaatctctg gttagatagc ctttagtttt
41341 aagaaaaggt cgataagatt tttcaatata gcaaagagct cgagcaattc ctaatttaat
41401 tgcttcagct tggccaataa gccccacc ttgaaactgt acaatagtat tataatgttt
41461 ttgtaatctt aataattcta aaggagcttg aatagataat acagaagctg gattttcttg
41521 catatataac actcctgggt ttcattaat aataaactca ccagttcctt taacaagttt
41581 tttttgtgcg atagctgatt ttcgacgtcc agtaccttta gctaaaatat taatagacaa
41641 tgtatgtcct cctcaataaa attaatagt actataaata aaaaagctca aaaattcaag
41701 atgacaaata tttacttaat tattatctaa gtttaagaac tttattttta acaaaagctt
41761 acccgacacc gtaaaaataa aatgattttt tcatgcaccg tagcccggct attgctcatt
41821 agcaatgcct aatgaagaat ctagggctcc cggttttttt taagatcaaa aaaccaagtt
41881 ttatttttta gataaaacac ctaaaatttg aatactttgt gtttttttgy gtttaagaat
41941 agtccgtttt gttttacata caacacttta tgtgttatgt gaagtctagt tcgataatc
42001 agggctccca taaaaaata taattaacga attatatttt ttgaaagtaa ttcttttgaa
42061 ttttgtttac ctaattttaa cccaagtgga gccaaagtag cttctacttc ttctaaagag
42121 cgtttaccga aatttttaag taaaagtaat tcatcagcag aatattgtaa aaggtcgtt
42181 actgtattaa ttttgctcg ttttaacaa gtatacggtc ttaaagaag ttctaatttt
42241 ccgatatcaa tcgaagctct atttttttta atggctaaat tatctttttt tegtcaaat
42301 tttacttttt tttggtttga atataattct cgacggcgta cactttcttg ataattttgt
42361 gaatctttga ggctccaatt taaagtaaag gcttttaata acttactttc ttgaaaagga
42421 gcaatcattt gaataagtgc tgaagctgcc tcatgaattg cttttcgtgg atgaatactc
42481 ccattagtcc aaatttctaa aattattttt tcttttgggt ttattgtttc tcatcattc
42541 tctagaagga aatttactcg attaacaggc ataaaaaccg catcaattaa aagaacattt
42601 ttagttttat ttgtttttta tagacctatt tggctttttt ttatagaatc ttgtcttaaa
42661 tttatttgat gttctaaaga attcgctttc gaaaaagatt ttttttgtt tagatttggg
42721 aaatttattt gttcggatga attttgacct ctagacattt gttctaaagg gttatttght

42781 tcgttacctt ttggcaacgc aagggaaacca tcagaattta aaatactggt agcattttatc
42841 aatggataat ttccaccagt tccattcgtat ttttctataa cttgagattg atctatcaag
42901 tccgaatcaa ctttatctga atttaatcgt ctagagttga taatctcgtt gctcattaat
42961 ccttcatttt ttaaggaatt tttttttgag ttctttgaaa aatgagattc aagtcctaata
43021 aaactacttg tatctgtaga tgacacgaaa ttaagagagc tactcgaaaa attttgtttg
43081 gataatcgtg aagatagctt tgatggagat atagtttgaa caatataaatt tttaccttca
43141 gagataataa acttcatctc taaattaccg ttagccgtaa gagttgctat gtattgatca
43201 ggatcaacac attgaattga cttaggtaat ttcatatcac gggctttgac aattccagga
43261 ccacaaaactt ttaaaaaacc aatctgtggt tcagttagaa tttgtttggt tgatgagtca
43321 ggaaaagatt cgaattccaa ttctgtatcc acaggagcta atacaacctg ttttaagattc
43381 agaagtatat ctaatactga ttccgcaact ccacttagag tagaatactc atgactttgctg
43441 cctttaatga ccacagcggg aatcgcctaatt cccgttaatt cagataataa agttcttctt
43501 aaagctggtg caactgttaa acctgtcctt aaatcaaagg gccctaattg aaatcgccca
43561 taaaaactac gattattttc aactctcgat tcgatacaag aaagcaataa attttccact
43621 tttaatgtct ccttctgaga agcagatact aaagtatgag aagcaatggt tggcatagcc
43681 ccatgccttc cactacgtgg aagaccttaa ccgttgccct attgggtaag atatttatga
43741 ctaaataatct aaccacaatag gcaaaggcta aatttaacaa cttcgtttgtt aaataaagtt
43801 ttccccgaag gctttacgcc aagcgtaaaa ctttatgttt tatctacgat aaaacattta
43861 gggaaagcct gccgggctaa gatgtttaca agtaaaccatc taagcaaggg tacaatcgaa
43921 gattgtatct tgccgttgcc gtccccggaa cggcaacggg ggaagaacac aagtctaaaa
43981 tacaatcaga aattgtatta aatcattatg tgaaagtcca aagtgtgtgc ttgcacacac
44041 taggctttgc ccgtagggca aacctttatt taacaacgaa gttgttaaat ttagacctaa
44101 tctaaacaaa ctatgtttgt ttaggatatt ccctttaggg aagtcttctt cctaggttcc
44161 cgggctacgg ggcaagggtg acgaacgcag ttcgtcaacc gagctgctta gtattaaatt
44221 cgtcgttttt taggtggtcg acagccatta tgaggaatag aagtgatata acgaattaat
44281 gtaattatta agcctgcttc ttgtaaacct ctaatagctg tttctctccc cgcacctggg
44341 cctttttacca taactctage ttgtttcatt ccttgatcca aagattgttt agcagcagtt
44401 tcagcagcag ttttagccgc aaatggagtc cctttacgag ctccctttaa accacaagcc
44461 ccagcagatg accatgaaat aacttcacct tttaaattag ttatagtaat aattgtattg
44521 ttaaaacttg cttgtacatg tacaacacct ttccggaact ttcgttttat tttttttggg
44581 agggattttc gattttgtct tgccatagta acaattcctt aattttatgat tttttttata
44641 tattactaaa tcctttaatt gtatccaggg tcttaaatgt aaatcataca aatatataac
44701 ccctaccctg gattaccgat aataagctga ttttcagata atactacacc cgggaaccta
44761 ggttccccgg ctacggtaca ataaaaat tttccaata aaaaaaagct tgttctagac
44821 aaaaatctta gagcaagagt ttttaagaaa gatcttaact gttatgtttt atcgtagata
44881 aaacctttta attaaaaatt cgtacaaact ccaccttcc agagtaaagt tagttggaga
44941 aaattttctt acgtcttgta aaataaaatc ttgttttaca aactttgttt gcaaaacaat
45001 cattgctttg tctaaactac ttttaacaagc aaaatttatt aaagtttcag ctgtagagtg
45061 aagccttgct ttaccttaaa ataaggagca atcactttcg caaagcaaat aatcctgatg
45121 gagtcgaaga ctctatcttg gcaaagaaca taattcaaag attaaattaa ggctttctaa
45181 aaaatagaaa ttgtaactta gtaattttaa aatatatata attaacctg tctttgttta
45241 tgtttaggat ttgaacaaat gaccattatt tttctttttc gacgtattaa tcgacaatta
45301 tcacaaaatt tgccaacaga tgcacgaact ttcatataat tagtttttta aaatttaata
45361 atacaaatct agaacaaata cactttttat caaaatttcc gattttatta tatgaactat
45421 taaccactca cttaaaaaatc aaatttggaat catttttaat aaagtatttt gtttaaaaac
45481 ctagtgtctt aggcctttaa tttacgtgtc taataaaaact ctattttcac taaaagtaaa
45541 aattacatta ccgctttaa atcctaatt tttaggattaa ggttagtagt cgcagctttg
45601 cactttgagc aaaattttag aagttatact tttattcacc cccttgtaaa gattaagggg
45661 tccacaatta taaacgaagt ttattataac aagctttggt ttaatttaat gcgaaatttt
45721 acaataaaga ttcttttttt ttagtcttaa gcttatttta ttttaagttt ctgttttttg
45781 ttcgttttgt cgtaatcgat agacaattcg ccctcttggt aaatcgtagt gacttaattc
45841 tacagtaacg cgatccccta ataataatag tatataattc cgtcgtat tcccagagac
45901 atgagcaagc acttgaaatc cattctctaa tttcactcga aacattgcat ttgagagaca
45961 ttgggtcaca acaccttgca tttcaataag actttgtttt tcactattca atttgattat
46021 tggatatttt ttttttgaga agaggcactt gtctaagggt ctttttcggt taaaacaacc
46081 tacgcttacc attgtctcaa ttcaagactt ccaccttctg ccattgcaat ggcaaggggt
46141 ggaagtccaa accttatgag tagcataagg taggatcttc tattatag aagtcttgaa

46201 taaagtttta cccgtaaggt aaagcctact gaacacggta aggggtgcaag tattcccttt
46261 ttaaaaaataa aagtatatcg aatttatttt taataaattt aaaagagggtg gacaaaagtt
46321 tatcctaatt tagcagttta aactaaaac tacaaattgc ttgtttaagc aatcactaaa
46381 gagtaaaagc caaaggttaa ctcaaggaat acactaagtt tagacaaaaa cttaagtcca
46441 aatagaatc ttctattcta gtaagatttg cgacttcgtg aaaatTTTTTg aagaagggtg
46501 gttttcttcg tagggtaaca cttaacactt agacaagaat ccagatattt aggcgaaatt
46561 tgacttccac tttttggaag ttctgagtac aatctcctat tgtattctga cgttgtttaa
46621 gtttaagaac ttcgttctta aacaaagtggt cacttgtagg agaaaacctt ccaggcacia
46681 taagattggg gaaactttgc tcgcctatcc atagatcgag acttgctatt accttgcgta
46741 taatgtttag aaaacaacca aagtttgttt tagcaaaagg aaaagcaatg gagctacgct
46801 ttgttcttaa acttgtacaa cgccttcacc acctttggga caaggtagtc gagcaaagct
46861 tgattatcga aatataatga gctttaacaa agcgataagg tctttgcctg taatagataa
46921 agttttttta aaacacttta tgtttttatg agaaacaatt agcaacaagc aaccggggag
46981 ccttaggttc ccgggctacg gtgcaaatta aaaaatTTTta ccaaattgag cataaaattt
47041 ctccgccaat tccgtaaaaa cgggcttctc gatcagtcac aattccttga gaagtagata
47101 aaattacaat acccatacca cctaaaattt taggaatttc tttatgattt gcataaatcc
47161 taagtccagg cttaactaatt cgttttaaat tcgtaataca aggttttttt tcccgccctt
47221 tatattttta acgtaacgct atttcatctg acggaaatat ttggaaagat tcaataaatc
47281 cttctttttc taaaattact gaaatTTTTt gactgattcg agttaaaggt attgatcacg
47341 tttgactttt tactaaattt gcatttcgaa tccgagtcaa gatataccta atagtatcat
47401 tgaccatttg ttttttattt tccttaaaat aataaatatt tataaatatt tttgagaacc
47461 tgattagttt ttaaaaggca ttccaaattc ttttaaaaga gctaactcctt cttgatcatt
47521 ttttgaagtt gtaataatag agatatccat tccacgaatt tgatcgattt tatcataatc
47581 aatctctgga aacataagtt gttcttctaa ccctaaacta taattacatc gtccatcaaa
47641 actctttgga ttaattccct gaaaatccct aattcgagggt aatgctaaat taattaatcg
47701 atcaagaaat ccatacatc gttcgcctct aagagtaact gataccctaa caggcatttt
47761 ttgacggagt ttgaatcctg caattgcttt ttttgaacga gtgataactc ctttttgtcc
47821 tgcgataatt gttaattctt ttaaagaaga atctagtatt tttgcatttt gtgatgcac
47881 tcctaaaccg cgggttaataa caattTTTTt aattcgtgga gcttggagtg tatttttata
47941 ttgaaatagt tctaataatt ttggaacaat attttgaaaa taaagTTTTt tcaacttttg
48001 tgccataata gttttatttt aatTTTTaat gtttatcctt tgcattactt aaactagcga
48061 aggtaataag gttggcgaat ttcgctcgga tacaccttac ttcacagtaa ggcttcagcc
48121 tacgggtgaa actttctggt ttatcaaagc ttttccaagc tatgcttgat aaaatTTTt
48181 attgtccgaa ggatttgctg gtagagcaaa actttattta acaactttgt tgttaaat
48241 agacaacgta aagataaaac atttaggttt tataaataaa atttgtaaaa cttgatttaa
48301 aagtttggtc gaatgcacia accaatgggc gtagctttta aatthagaaa gagcaatgta
48361 aaaaacgaag gacctaactt aaacaaacta tgtttgttta gggtttagat ttttaattg
48421 taaacctcta tgaagtaaga tagaaaatat atctttgatg gagaaatata tacgtttaca
48481 attaaaatac aggttgatg tttagaagta tggtttataa cgtctatttt taatacaatt
48541 gaaaaatggt ttttagtttg taaaattaaa gaacttctgg agctaatagaa acaatTTTTt
48601 taaaattccg atcacgaagt tcacgtgcta taggaccaa aactcgcgtt cctctaggat
48661 tattttcttt attgagaact acagctgcat tatcatcaaa acgaatggtc attccgtttt
48721 cacgtcttaa tcctttacaa gttcttacia taacagctcg aacaatatca gattttttca
48781 aaggcatatt tggaaatggca tctttaacia cagcgataat aacatcgccg atattagcag
48841 attgtctttg atttccacc aaaactcgaa tacacattaa ttttcgcgca ccgctattgt
48901 cagctacgtt gaggtatgat tgaggttgaa tcatagcatt atatttataa aaaaaatgaa
48961 aacattattt aattgatagt aaaatattta actggaaagg tataaataaa gtttataact
49021 tcaatatttc ttaataggct ttggtcccgt tgctaattgt gttagcaaa gtttagcata
49081 aggtttaaag caaagccctg acccctcacc tgggagccct caattaccg aggtagtaag
49141 gttgacaagc ttcgcttggt taccatgcc ttccattaca tggaaagacct tcaccgttgc
49201 cctcttgggt aagatattta tgaataaata tctaaccctaa taggcaagg ctaaatttaa
49261 caacgaagtt gttaaataaa gttttccccg aaggctttac gccaaagcgtt aaactttatg
49321 ttttatcgta gataaaacat ttagggaag cctgcccggg acggcaaggg tacaatcgaa
49381 gattgtattt tgccgttggc gtccccggaa cggcaacggg ggaagaacac aagtctaaaa
49441 tacaatcaga gattgtatta aatcactatg taaaagtcca aagtgtgtgc ttgcacacac
49501 taggctttgc ccgtagggca aacctttatt taacaacgaa gttgttaaat ttagacctaa
49561 tctaacaata ctatgtttgt ttaggatctt ccctttaggg aagtcttctt cttaggttgc

49621 taaatattgt tgtcaacggc aacaatataa agttttacca agcgtagttt gttaaagcca
49681 ggctacgcta cttgataccg gtgtccggta aaatgtgctg gtagagcaaa actttattta
49741 acaacttcgt tgttaaattt agcctttgcg tgttgggta gatatttatt cataaatatt
49801 ttacccaacg ggatttcccc tttgggtaaa actttattaa aaaacttcgt ttttaattt
49861 agacaacggc ggtagacttt tcccgtaggc tttaccatag ctcggaacc tagggctccc
49921 ggggcgaagc gtaaaacttt atgttttacc gttgccctgc cgtcaacggc aagggctcaa
49981 gataaaacat ttaggaccta atctgaacaa actatgtttg tttagggaaa actttattta
50041 acaacgaagt tgttaatttg ttaaatttag cttttgccta ttgggttaga tatttattca
50101 taaatatcta acccaatagg gcaacgggccc aggtttatca agcgaagctt gattaccggy
50161 aagccctaaa taccggacta aggtcaagcc ttgcgataac gtcaaagaca aacaaacttt
50221 aatcctataa acacatatat ggtttatcta aattgtacaa acgaagcttg tacaacttga
50281 tttaaattta aaaattataa atttagactt tttgtattg ttatttattt cttgccttaa
50341 atgtgtgata taacatcaaa ttttacacaa aacgtaaaaa ccaaagatga gcaaggaaaa
50401 ttagaagaaa acttagtgaa aacttttggg tttgtgttac gggtagcttt atcatttaga
50461 aaacacttag agatcttcta tatggcccta gaaaaaattt aaattattct agagacttgg
50521 agccttgttc aaacatccaa gtttaatttt tgaactgtgt ttttttctaa actgaaaaat
50581 ctttttctca cactgcattt taaaaatttt agctttatta actctaattt ttttaactca
50641 ttttaagctat gaagcctctt tcataataaa ttgagtctgt attggcattt tgtatgcagc
50701 aattcgcata gcagatcttg ctacgctttc agaaatacct ttcatttctg ataaaaattt
50761 cccaggctta actacagcaa cccaaaattc tggagaacct ttaccggaac ccattcgcgt
50821 ttctgctggt cgcattgtaa ttggtttctc tggaaaaata cgaatccaaa gtttaccacc
50881 tcgctcgagca taacgtgtca ttgcacgacg acctgattca atttgtcgag aagtaatcca
50941 acaaggttct aaggcctgta acccgaaatc tccaaaagct attgtattac ctcgattagc
51001 ttttctcgcg atacgtccac gatgttgctt acgatatttt gttctttttg ggcttaacat
51061 aatgatcttt ttaataattg taatcgtttt cctaattgggt gaacacaaaag tgtaacagtt
51121 taggtaaaac tttgctttat agtaaggtaa aaaagatttt ctaaattaca aactttgcyg
51181 agcgcagaaa aaaagtccaa aatagaacag agttctatta ggatcttcca ctttgtggaa
51241 gtccaaaatg tgtgcttgca cacattagga tctttaagcc tcgcccgttg ccctgcccgg
51301 cacggcaagg gttaaagtct tctgcttttg taatcaagtt ttacaagcta cgcttgtaaa
51361 gccttatggg gaaaatttgg tgaacaaacg aagtttgttc acttaaacac ggtaaaaatt
51421 taaaactttt taggtttgaa tttatctaaa agtaagactt taccctattt gcccaacgaa
51481 gttgggcgta ggggtaagg ttaataagct cgcttgatta gccaaagcgt gtttagacct
51541 tgtaaatata ttttttagt tgttatgaaa cccgtgtgtc caaataataa aaggatagga
51601 ttttctctac gagaaaaact tggttaacaa acacaaacac aacgtagcac aaaaaagg
51661 ccaaaataca ctacagaggt gtattaggat cttctactat gtagaagtct tctgtttttg
51721 ttaacttaaa aaacaaaatt taattttatt ttaagttcaa ctttatgaga aacttttatt
51781 ttaggaacaa attattttta ttatttttaa aaacccaaat ttttatacct aataatccgt
51841 aatagttttt ggcggtttta taacaataat ctatatctgc tcgtaaagtc tgtaaaggca
51901 cacgtccttc tctaaccac tcaactacgc caatttcagc accatttaaa cgacctgata
51961 cttgaatttt aataccacta acccctgctc gttgtgcacg ttgaatagct tgacgaacag
52021 ctcgctgaaa ggcaaccctt ttttcgagtt gttctactaa aaaatcagtt ataaaagctg
52081 cttctgaatc cgggttagct aatttagtta catgaattgc aacttgata gaatgtttaa
52141 cagtttcgat tttcgaagg actacattag atgatttaaa attaaaaatc gtttttaaac
52201 gatagggtgg taatttttgt tctaataatt tttttagaat gtctaaatta gcccatctc
52261 gacctactaa aatccctggc cgtgctgcac gaatttcaat ttttatttga tcgagctttc
52321 tttgaatttc gatatgagtt attccagcgt cagcaaattt ttcaagaagc agttttccta
52381 taaaactatc ttcaactaca agttgagggg actgatttgt ttttgcaaac cattgagatc
52441 tatgcttttt cgtaatacca agacgaaacc caagaggatg aattttttgt cccataaggt
52501 tataattttt tataactaatt aatttgttga aactgtaaat ttggtttttg accaaaaaat
52561 atgtttcttt aaaaaatcc aataaaatc attttttaga gtttttagtt tttaccgctg
52621 tctattacc taaaggattg cttttttagt ttttctaag ttttacaac aaaatttcta
52681 aaacttgatt taaaagtttg tgcgaatgca caaaccaatg ggcgtagctt ttaaatttag
52741 aaatttgttt tctaagattt tcaactaaaag caaaattggt cccttttgaa tcttaatacc
52801 ggataataga ggtggcaatg aggtaatgat ttgcctacat ttcaatttaa aattgaagtt
52861 taggtcgaac ttaaaaaata ttacggattt tatataata gagatatact tttccaaagc
52921 caaactacac tcggttaaga tcttctagct ctactacaag ttcaaattta aaagcgtagt
52981 tttaaaagtt gttttttttt aatattttaa tataaagttg tcaacaatg actttcacga

53041 gaagatttaa aaaagatttt agtttttcta taccttgaat tttggtttac cggggagctt
53101 cacgggaaga acagaagtct aaaatacaat cagagattgt attaaatcac tatgtgaaag
53161 tccaaagtgt gtgcttgac aactaggct ttgccgtag ggcaaaccct tatttaacaa
53221 cgaagttggt aaatthagac ctaatctaaa caaactatgt ttgttttagga tattcccttt
53281 agggaagtct tcttcctagg tccccgggct acatcctttg cggatttaga aggccaacag
53341 taacccaaaa atttctaaaa ttatctgtgt ttagtgagaa agtagaaaaa aatttttaat
53401 aagcgaagct tattttaaag aagaaaaact taaaataaac tattaatfff aacaaagttt
53461 ataaaaaaat gtttacactg caattaccaa gttaattaaa aacctcataa ttttacattt
53521 agcgttttga ttttttatca ctttttatat gacctcgaaa agttcgtggt agtgaaaatt
53581 ctccataatt gtggcctacc atttgatcgg taacaaaaat gggaaatagt tcacgtccat
53641 tatgaacagc aattgtatgc ccaatcataa taggtacgat ttagaagctt aatttttttt aaaagatgac
53701 ttaatacaaa ttttttttct cggagatcgt ttaacttttc aatttttttt ctttatttaa
53761 ttgctacaaa agggcctttt tttaacgaac gtgacatgag cgtttttttt ctttatttaa
53821 aatgaaaccc ttcattacaa aataatgtgc ttaaagggtt tttactttac taaaatgtta
53881 caaagcgtag tagaacatac cccactgcg tgggagcatt attataattc tacttcagat
53941 tctatttcgg aaaacgtggg ttcgcaaaaa actttttttt tgcgtttgta ataaagtttt
54001 acaagctaca cttataaagc cctatgataa aaactcgggt aacagacgaa gttggttaac
54061 ttagagagtt aaccttggat ttagattctt acgctcacc tttattcata aaaaaattca
54121 gaacttaaat tttgtataaa aaataaattt tttaaaatta tttgtaggat aaaatagaaa
54181 aagagctaaa actctttgaa ttttataaaa acttatttta tttgcgtcgt ttaagaatta
54241 atttattact atatttttta gactgtcgg ttcgctgacc taatgctggt cgtccccaag
54301 gtgttacagg atgtgatcgt ccaattgggt cgcgcccttc tccccctcca tgtgggtgat
54361 ctactgggtt cataacaaca cctcgtaccg ttggctgacg accaagccat ctagtacgcc
54421 ctgctttacc taatgttaag ttactaacat caacattacc gacttgacca atgcttgccc
54481 agcattcttt tgaaattaaa cgaacttcac cagaaggtaa tctaagagtt acaaatttac
54541 cttcttttagc gacaagctgt gctacagttc ctgctgctct tactagttgt ccaccgatc
54601 cgggatgcag ttctacatta tgaacttcag ttcttagtgg catattaaaa agaggtagag
54661 agtttccaat aactatagga gtttcaaat cagcaataat attgtcacct aattttaatc
54721 ctcgaggata aagaatatat cgtttttcac catcttggta gtgcagtaaa gcaatcgtg
54781 catttctatt tggatcatat tcaacagctc ctacttttgc aataatacca atttatttcc
54841 gtttaaaatc tatttggcga tataattttt tatgaccgcc accacgatga cgactagtaa
54901 taactccacg attatttcga ccttttgatc tatgtaccga ataagttaaa gatttttcgg
54961 gtttcgtttg tgaaatttca ttaaattccg atacagaacg atttcgtgta ccagggtgat
55021 aagcacgata gagacgaata cccataatat ttaattacta ctaaaaaaaa agctatcctt
55081 aagtttttta ataaaagggt gtgagaaata tttgatagaa ttttctaaaa atttcgttaa
55141 gctgggtgatt tcgacattct tcaaaaacta attaaaagtt cacaattaag gtggataaag
55201 attaccaata gtttcctaaa tcaagcttta cttattttgt caaccagtag ggtactatca
55261 agttcactct ccaagtatcc gaacgaagtt cttaaatctt atcctgtacg ataaactagc
55321 gacaaaatat aatttgaaat ttattcaaaa cttttatcct tcccatgccc ggaagaatft
55381 acccgtaggc tttacgcgaa gcgtaaaatt ttatcgaaga taaaacattt aggcgaaact
55441 ttattcaaga acaagttctt aaatttcgat aatggccaag gtaaagatac aagactataa
55501 tacaatcttt aatcaccctc catcgtcaca aaataatcta aggtgggttt taaatcgtc
55561 gcactttgaa ataaaagtct aaaatgccga atttataaac gaaacaagac ataagaatga
55621 ttgagtaaca aggttgatga acaaaattca tagaattaaa agacttgttt caacaacaaa
55681 attttgttag taatthaggt aaatctgaaa taacaaaaat gtagtgaaga cttgaagcga
55741 agcttcaaga tcctaataga ctctttgagt ctattttgga caagatttgt tttgacaaat
55801 tatttagtat ttatttaagt gtctggaaac aatggaatag attcacctga tttaatgtta
55861 ataataactc gtttataacg acttcgataa ccttgagaaa taccaagtct ttttttttta
55921 cgtgggtggtc gatgagtatt gaccgcaata acttttactt caaataaacc ttcaattaat
55981 ttcttaattt gtattttatt aagtttaggc tcaacatcaa aagtatattg attattttca
56041 ataatcttta cagatttttc tgttaaaaca ggatatttta ctaaatcaat catcataaaa
56101 ttttttgttt tatecttaac aaaacagata tttataaatt tgtattattt tcgaacgaac
56161 taagtttttg gtcttttagct cttggtaaag accaaaatga gctttgtctg ttataaattt
56221 tgttgatatt tactatttat aaaaaccttt gatttacaca aagttttaca aacatagatc
56281 gaaaaacggt ttaaatagaa cttgaaaact aagataaagc tttcaaaata atatttttta
56341 aaaacaaaag ttttttttaa aaatcaattt tctagctacc ttttcaaaat taacaaactt
56401 tgtttgttaa ctaaggtggt ccattttatg aaagtcttac tatcgtaaat tactttaaca

56461 ttatTTTTagT ttaatttctt ttttaaagcg atatagtctt attttatttt taaaaggTtg
56521 attttttagtg ataaaatact attaaacata gcttttttaa aacaaaaaaaa atagattttta
56581 gttttcgcct aacgtttttt aggtttgtgc tctacacaaa ccaagattct gtgtttggta
56641 attcaaactt tccagcttca ctccaagtca ttatagagtt ttgtttaaca atatatttag
56701 cttataacgt cttacgttgg aagacctagg aacaaaagag ttaacaatta aaagtttata
56761 actgtagtgc ataaaactgt gattacaaac aaaactaaaa atttaaaaaa caatattttt
56821 aacttcccaa acgcagtttg ctaaccaagt tttttctaaa tgttttaccg ttgccctgcc
56881 ggttacggca agggtcgcag aagactttaa gactttaagc gaagcttaa gatcctagtG
56941 tgtgcaagca cacactttgg acttccctaa aggccttctt aaatgtttta tcttagaccC
57001 ttgccgttga cggcagggca acggtaaaac ataaagtttt acacttcgcg taaagcccga
57061 ctgggaaaaac tttatttaac aactttgttg ttcaatttag acccgtagcc cggccattgc
57121 tcattagcaa tgcctaataG aactctagtg gctcccgggt aatctaaaca aacgcagaag
57181 acttaaagca tagctttaag atcctagtat acaaacaaag tttgtatact ttggactttt
57241 tttctgcgct tgcctatggt tgtttaggaa acatcttaat agagcgtagc tctattatgg
57301 actttttttc tgtactatgc tgcgcttcgc tacgggttat aaacctctct gaatgaggca
57361 taaagttttta cgcttcgcgt aaagcccgtA gggaaaatcc aaggttttgtg tccttaccC
57421 ttggctattg ctcattagca atagctaata cgtaaggact tccaccattc ggtggaagat
57481 cctggttacc aaactttggt tggtaacctt gtaccaagag gcgaagcaca aaccaggatc
57541 ttctactaat taaaatata ttaattagta gaaaatactt cttttattgg caaagccaat
57601 aaaagccaat aaaagaagga tcttccaaaa attggaagtc tttgttaatt aaaattaaaa
57661 aattttctatg gcaaaacaag caattgaaca aaccggactc attccacgat ctattatccg
57721 gacctttgac cgttttagaa aacaactttt gcccgcgct gaaaagttag ttattcaaga
57781 atttcgtatt tcaagattcc aagttttagt ctctgtaaaa tgttttagtta gtttaacaat
57841 tataccctta attataaatt atttattaaa aaatttttta ttaataccag taacagaaaa
57901 tttttggaat agacaccaa atgaaatttt tctaaattct tatcaacaag accgcgcttt
57961 tactgagatt aaaaaatttg aagagaaaaat ctattttgaa tttttgatta atgatctaga
58021 taataatttt atttttgagt ctgacaaaag cgatcaggag aatacaactt taggagagtc
58081 taataaaatt aaatattctt tgaattcaaa ttttaaagaa ttatctgaaa aggatcttgc
58141 aaaaaaacia accactgaag aaattaaaga gctttatttc gaaagtacca tgcagaaaag
58201 aaaaaagatt caattaattg aatcaaaaaa tgttatacaa gaagatctcc aagaaaaaat
58261 aattgaatta gcaattcact ataactcagca aagtattata gccttaacta atttatttgg
58321 tgattttatt acttttggaa taattagttt tttagttatt tggatgaaac cacaaattat
58381 aattttaaaa tcatttttaa ctgaatcaat ttatagttta agtgacacaa caaaatcttt
58441 tttgttaatt ttaggaactg atttgcttgt tggttttcat tcaccacgtg ggtgggaaat
58501 ttttttagaa ttaatattaa gacattttgg atttccggaa aaccaagatt ttatttactt
58561 atttgttgcg acctttcccg tattattaga tacagttttt aaatattgga ttttccgtta
58621 tttaaaccaa atttcccat caacagttgc aacttatcat agtatgattg aataaactca
58681 aaatttttac taattatgaa atatggaaag cctgctatcg gagtcgaacc gataactttc
58741 ggtttacaaa accgatactc tgccaattga gttaagcagg cttaaagctc tatccatagt
58801 ttacttaaaa tttaaaaatt tttcaatttt ttttattctg gaattctaaa tttctattaa
58861 ttaagacttt tactttggga cctaattaat aaattctggt tatgcacttt ggatttaca
58921 acaaagtttg taaaacttga tttaaaagct acgcttttaa atttataaag ttaattttct
58981 tttagagcta gatttgtcga cctggaaatt tgtaatttcc agaactattt tttttaaag
59041 gacagaaaaa tatcgactta taaaaacaaa ttattttgag taaggaggga atcgaacctt
59101 cgacaccgcg cttcggaaac gcgagctcta atccactgag ctacaaactc atacaaaaat
59161 tcgtaatatg tcaattttaa cattttctaa attagtactt tcacccttac catatctaata
59221 agaatttaata tttttgtggt gtaaaattta ctttgcctgag atgacgtttt cgactctatc
59281 aggttacttt gtagaagtca ttgctttttg tatcatagta cacttcacac tacggatgaa
59341 cttgaaatca gctttgctta ttttcgagga tgtcagtgaa gttccctaac ctcacccaaa
59401 agacttccac gaggtgaaag atcttaatac aatcgaagct tctctaataa ttacccttac
59461 ccctacgccc aacttcgttg ggcaataggg gcaaccggg agccctaggt tcccgagcta
59521 cgggttgcca agctgcgctt gattaccgaa gggtaatatc aaagttttcc ttataggaaa
59581 cccgggagcc ctaggttcct cattaggtat tgccctttgg tctttactcc ttagccattg
59641 ctaaggagca atggcctaca ctacctttgg tagcgtagggc aaagcttaca aaatcttctt
59701 ttaatatttc cctcaactt aaatttatgt ttcacagaag accttaattt gtgctttgct
59761 aaatgtttta tcttcaataa aacaggaagt ttcacgcaaa gcgtaaagcc tcttgggtgtt
59821 tctcttcttt cacagaaatt taaaaaccta aaaggctttc aaaactatac ttagaaaaact

59881 tgaatacgc tgcgtttat tagtatccac acaaacttta gattttttaa aaaaaagcct
59941 gaccatttta ctttccatta ttgcaaagct ccacagaaaa ctttttctcg tgctatgttt
60001 cacttgataa ccttagaata cttttcatag ttatttagac ttctaaaaca aaactttggt
60061 accttttgta gcgtcttggt ttgctatttc attagaattc gcactaaact ttgttcatta
60121 aagcaaagtt agaggaatga ttcaatttat ttaaaatata tcgaaattaa acattgatct
60181 tttactttta taaggctttc acttcgtgaa caaccttagt ttatgcttcg cataaacttt
60241 ggaaacatac atgaaaatta actttcttct aaagatactt gtaaaaaact tgctaattca
60301 gcagcttggt tctcaatctc ttctaaagtc atagggtggc ctattcgagt taatgggatg
60361 tcacgctttc cttttacacg aaggtaaata gtacgacgtg gatttagccc ttcttttaat
60421 tcaactctaa tagcttcaat atcttttaata gagtataata agtcaatacg tcgatttttt
60481 cctggaaaaa cccatctaaa aattcttaca ataccgtctt ttttattaaa ttctgttaaaa
60541 ccgcctccaa cactccaaaa taaagttaac cataaaaaaa tactaaaaaa aaaacctaaa
60601 attccataaa agcacattac taatccttgt ggaaaaaata gaatatcttt tgatttaata
60661 aagggaagca aattaaattg taaatagctt gaaataccag ttaataggaa acctgaagca
60721 cctaaaaaaa taacagaggc ccaccaataa ttactaaatc tacgtgaacc tactacaata
60781 taacgacgaa ttaagtcaga ctcttgagtt tcgtataaat caaccatata atataaattt
60841 ttttaaaagc aacttttata tgttattcct aaactatcaa ctttattaaa tcagaaggct
60901 agggggcgta gtttacatac cccttgaaa taaaacaaag ttttatataa atattttttc
60961 tgcgaaactat tttagtcatt cacttagagt caatttataa aaaaacaacg gctatgtgtc
61021 cacccttact ctaagcagag tttctagctt agacgggaac ggtagattt ctagctttgt
61081 ttgattacct ctcccttttg atatttaata gttgacagtt aacactatta cagaaagctc
61141 ttcgtacaga gttcgagagt acactttatt taaaattgat ttctattata tctaaatggt
61201 taaaccaaat tttatttgga tttccttgaa aaggaaaact ttgtttaaga cttccatcct
61261 taccatgtcc gataaggaca atggcaaagt ggaagtctta aacttagaaa acaaaacttt
61321 gtttattttta taaaatctcg atttcattta taatcataaa aaatcgaggc atcgccaatt
61381 aattatcagt attagtgggt ctgttaaaat caaatttttc ttaagacga gtaccttttc
61441 cgatacgggt tctcaataaa tatagttttg cttttctaac ttttgctcgt cttacaattt
61501 caatatcttg aatcgacgga gaatgaatag taaaaactcg ttcaactcca atgccttgaa
61561 aaatacgtct gaccgtaatt gttgtgctta acccagcacg atgcaaagct ataactgttc
61621 cttcataagg ttgtactctt tgttattttc ctcttgaat taagacacct actttaaccg
61681 tatcacctac acaaattttt ggtaatccct gctttaattg agtttgttca atagtacgta
61741 caagttcttt tttttcata agtttgttcc ctattttgaa aaaataatct ttttaagtgt
61801 gtttaaagat aaaattctaa ctataaaaaa ttaaacacaa tgtttttttag gtttaggcta
61861 gctagtaccc taagaactaa gaagaacata aaattttgc taaatattta cttaaaaaga
61921 aagtaaattt ttctctaagt tatgagtttc atcttttaac aaagtacacg cttttatcat
61981 aaaacatcgt tcacattaac ataagtaaag ctaactgttt aaccattat atgagcgaag
62041 ttttgacttt taacgtaact cgaaaaaac agagaacact tggacccttg ccgtgcccg
62101 cagggcaacg ggcaaagttt caagtctttg cactactaaa tagatgaagg gttatgttct
62161 acaaaatgta aaatttgatt acaagcaaag tttgtaagtt aaaaatctta cgccgttggt
62221 ttaatgcggt agtagaagtt gaatttctat ttctttacta aatgttttat ctttcacct
62281 tgccgtgcc agcagggcaa cggtaaaaca caaagtttga cctgggagcc ctaggttccc
62341 aggctacgcc ctacagatat tcgagcgaag ctcataaacc ttactactta aacaagcaaa
62401 gcttatttaa gtagtttagg taaaataaat agtaaagtaa agaacttag aaagacctaa
62461 ttacggagct ataagtctca acctctactt ggatttgaac aacgacaaag attaaaaatt
62521 aaataagttt tacttatcgt agcccgggaa tctagggctc cggggttttt tcagcaaat
62581 tattcaatga ttgtagaaac aacacctgca ccaactgttc gaccgccttc ccgaatagca
62641 aatctcattc ctttttcaat cgcaattggg tgaatcaatt caacattcat cttaatacga
62701 tctccaggca ttaccatttg agttgcagta tcatcatctg cacgaaatga ttcaattttt
62761 ccagttacat cagttgtacg cacatagaat tgaggtcgat atcctgaaaa aaacgggtgta
62821 tggcggccac cttcttcttt tgtaagaaca taaacttgag attcaaattt agtatgtgga
62881 gtaattgttc caggttttgc tagcaccatt cctctttcaa tatccacttt ttgaattcca
62941 cgtagtaaaa ctctacatt atctctgca acactttcct ctaatgtttt ttgaaacatt
63001 tctaaaccag ttacggttgg agtgcgagtt tttcgtaaac caacaatttc aattgattca
63061 cctactttta cactgcctct ttcaactcga ccagttgcta cagtaccacg tccagtaatt
63121 gaaaatacat cttccactgc cattaaaaaa gatttatcgg tatctctttc tggagtagga
63181 atataactat caacttgatc cattaaatca aaaattttgt caaccactt gttttcacca
63241 cgtgttaatg ttggattttc agttaaagct tctaaagcaa gtagtgctga accagaaca

63301 ataggaatth catcaccagg aaattcataa ttatctaaag tttcacgaac ttctaactca
63361 actaattcta acagttcttc atcgteccact tgatcttctt tatttaaaaa aacaacaacg
63421 ttaggtactc caacttgtht tgctaataaa atatgttctt tagthttgchg cattggacca
63481 tctgcacctg aaactacaag aatagctcca tccatctgtg ctgcacctgt aatcatgtht
63541 ttaacataat ctgcatgacc tggacaatct acatgtgcat aatgacgatt ttcagthtca
63601 tactctacat gagcagtatt aattgtaata cctcttgctt tttcttcagg tgcagaatca
63661 atatcagcat atthttcgthc tttcccacca ctactagthg ctaatgccat tgtaatagct
63721 gcagthaaag tagthttthc gtgatcaaca tgaccaatthg taccaatgtht tacatgchggt
63781 thtttacgtht caaaththtg gcgtgccaata thtttaaththt gthccttaath gthtaaaatht
63841 thtaataaaa ththtgatht tcaththaaga thaatcgthc aatcacttht tctagththt
63901 thgtatagtht ththttataat cattaththtg gaactthgaa thaaataatht cgaaaaaat
63961 agcaaatthg ctgaththaga thttatctgta tgataaagtht caathcagtht tccaagtht
64021 ataaatacag aagactthaa gcgaagctta aagatcctaa tgtgtgcaag cacacatht
64081 ggactthctac tatgtagaag atcctaatac actctctgag tataththtg tttcgaaaag
64141 atcctaataa aaththctgat atgaaatata cthttctthaa thactaaggt aaththaagga
64201 gggathatat ththggacttht ththctgthg acgthaaagtht aathhaactag gthcctthcaa
64261 aaathgaaag thcgthagccc gggaaatctag gthcctccagg ththagthatht atththtagaa
64321 atthtagaca aatataatht cacaaththt atggthththt atcathththt agthtaagtht
64381 gacgaactthc gatcgthagcc thcaccctac gcttaataaa gthgggchgta gagggaaggt
64441 aataaagctth cgctthgatha thgaaggtth caccctcacc cthgctactc cthtagccatht
64501 gccatctthg gthctthactc atthagctatc gctgatgagtht aaagaccaag atgggcaag
64561 thatcaagcg aagcgthgath accgaggagtht gaaactthgaa caagthgagc thgththtagg
64621 gthagththgt caagctthcg cthgathgaccg aathththtgc aaaatataca atctacgtht
64681 thaaaaacga ththththgth aagctththt thgctthctgccc atthctgthgaa gthcgtcacg
64741 ththacgaathc gcatththcca thththththgaga agcatcththaa aaththctthc thththththg
64801 aatcataththt cggccagathc gthctthctgca cgaagththaa atccatcgaa thgcaagggc
64861 thgthaccagtht thctgggthca cthcththaaag aactthgthaa gthcgaaccac caacacgthc
64921 thgctththact thcaacththg gggthgcttht thctthactgct thththcaataa thththactgg
64981 atctthgththt gthaaacththc caaththgtht catagaththc thaaacatht thththgcthth
65041 thgathththt cthththththc thaggtgatht thacaaththc thgacththaa gacththc
65101 acagggthct ggagathhthaa aacgthththt aggggthcga cgacgagaca thagathhth
65161 actccaththt thththgatht gththgthhth caacththct gacaaacthth thththththg
65221 atththgcttht thctththaaag thaaacacga agcgtagtht thaccctthg thththththg
65281 aagactthcgt ccgthaggagc cthththacagc aaagthththgaa aactththgagc thgthcagaa
65341 aaagthcttht thgthgthhthaa aggtthcathhth thacaathctt cthththhthaa cththgcthth
65401 gccgggthaat cthgggctthc cgggthaththt aagaaagtht atththcththc agagaththgth
65461 thaggaththt gthcaagcttht cthgthgththt aagthththgthc gaathgthcathhth accaathggg
65521 atagctththt aaththththaa cththgththaa gththhthaaagtht thagacththaa cththcththc
65581 gthgathgathc cthgthgathhth cghhthhthgaa atgcatthct thgththththt thctgaththt
65641 thgathththc ththththgagth ththhthgctth cagththgtht thhthcththgath thththcaagth
65701 thththgthgac thhthgththt aathctththt thththgthhth thththgthhth thhthththt ththgththhth
65761 agathhthaaag thhthacagtht cthcththhth aacgcththaa cthththththg cthgctthhthaa
65821 ccatathththg aacgactthg aactcghctt thhthacactg atgthgthcthth ththctctgth
65881 acathththgath atcghhthccc aggtathhtht thhthactctt cthccththgath thhthhthhth
65941 gaathgthctt gthhathththg gcththctctt gghaththhth cthgthctthc aathctgthhth
66001 cthhthhthcghth cagghgcthac thththththhth gthcagaththg gthththththgg gthcththgth
66061 thhthctgthg thcagacact thcagctthgth gthcagactt thhthhthgth gaathgthgath cathththt
66121 ththththgaa thththththct gththththct thhthhthgath thhthhthhth thhthhthhth
66181 actththhthaa thhththgaa thhthhthhth thhthhthhth thhthhthhth thhthhthhth
66241 thhthhththg thhthhthhth ththgththgath aaththgatht thcghhthhth atthctthhth
66301 atacghhthaa cathhthhthhth cthctcghthgath cththgacththg thhthhthhth thhthhthhth
66361 atcgththhth thhthgctthcath thgththcghth cghthhthhth thctcghthgth actcththth
66421 aghthhthgath caathhthgth hthhthctthhth thhthhthhth hthhthhthhth aathcththt
66481 thththgghth hthcghacthga cththcththcth caathhthgth cthgagcghth aagactcghhth
66541 cghthhthhthcath aathgathgath gathhthhthhth cghthhthhthhth hthhthhthhth gggghthhth
66601 aagghthhthhth cghthhthgath aghthhthhth thhthhthhthhth thhthhthhthhth hthhthhthhth
66661 thhthhthhthhth athhthgththt thctthththgath athhthhthhth hthhthhthhth hthhthhthhth

66721 acttcgaaac cccggctaaa attactttat aaataaactc ttagtTTTTT cggcctcacc
66781 ccttcgagcc aaggttctca agcgaagctt gattaccgaa ggggtatagc aacaaagttc
66841 tttatctggg agccctaaac taccggaggt agtaagggtg acaagcttcg cttgtttact
66901 catgccttcc attacatgga agaccttgct ccttgcctaa attgaacaac aaagttgttc
66961 aatntagga agggaagc ttggttaagc tgccttgact atcctagtgg tagtttcaaa
67021 actttgttac aaaatTTTct ataaaaatTT ctaatTTtatt aaaaaagtaa aagcgggtatt
67081 ttattactcg cttaaaatta gcagtaaaat tcagTTTTaa ttttaaatta aaaactagtt
67141 tctatTTgtg ataaaaatct caacaaaatt tattagataa agactcattt taatTTgttg
67201 ataaaagctt acaaaaatata aatTTTTTTa actaattatg tcgtgattat accatacata
67261 agcgaagcagg ccaaattTaa atTTTTTaaa tatgggaaat aagttgcatt taactttact
67321 tttaaaaaaa agtattTTTc cagttactag taatTTTtaga acaaattTaa aagagttcac
67381 taaaagttag aatcaaatTT aaaatTTTgta taagagttag actgTTTtag ttttaagaaa
67441 ccgagtgaaa aataagagta gctgTTTata ttaacctTta aacaaaactg tacacctggg
67501 atTTTcccta cgggctTTac aagctacgct tgtaaaactt gattacaac gcagaaaatt
67561 ttaagcgaag cttaaagacc ctaatgtgtg caagcacaca ttttgactt ccacaaagtg
67621 gaagatccta atagaactct gttctatTTT ggactTTTTt tctgcgctgc gcaaagtttg
67681 taatTTtagg aaacttggtt aacaaacaaa gtttgTTaac ttagaaaatc ttatTTtgaa
67741 tttaatTTTT aagattgtat acaaaacgaa gctatacttt actTTaacct gatcctgctc
67801 gaaaaatTTc agTTTTgtt aaatgTTTTa tcttagataa aacatttagc aacgggtaaa
67861 gcttcaaaaa aaattatgag ttagcttagt ttttactgt tagacaaaac ttagagaaat
67921 tattcaactt ctggaaggtt accaaagaaa attggataac caaggTTTTa actTTtagtg
67981 aaaacaaata aaaatTTtTT aaaatgaaaa aactaatgga tcgggaaaaa atctattaat
68041 ttcaatcagt aaacctgctg ttaaagttaa ccaaattaaa gccacaacag gtgcagttga
68101 aagatatgta gtaaaatatt tcattgaact ggtccttatt ttaaattatt gataatatat
68161 tgtacaaaca atcagctTaa aatataggat aattgTTtagc cgttaaattt ctctctaata
68221 aaaatgaaat gttcgcgTaa attaaataga aaatcttagt ttttagtaat gagttaacca
68281 aggtTTTaaa ataaactTTa tttattctaa agtgtatTTT tttaaatgTT ttctctaagt
68341 ttaaaccgaa ggtTTTaaaca aagtgtcacc tcgctgccag caaagTTTtac ccagctgatc
68401 ttaagcggta ggtgTTacat ttataagtga acctTTaatt catactTaa cagtaaaa gcTTTcacac
68461 atacagctga tgaacactTc gccattTtaga acaagatTTT acagtaaaa cTTTtcacac
68521 aaaacagata aatTTcgttT gtatTTTctT tgaaggtaag cTTTTTaaa aaataatTTg
68581 taaattattc atTTTaaaca atgTTTTTac tTTaagctag tTTTTTTca atTTTgagtg
68641 cTTTTaatat tctccgTaaa aaaacagtaa gTTTTaatta gagagacgta gcccggaac
68701 ccagggtctac cgggaaaact ttctcctaaa aaatTTTtaga gctattgtac aatTTattat
68761 aaaactaaaa agTTaaagaa gctatTTTTt taaatatatt aaaaatcact tgcataacta
68821 agTTTgaaaa tGtaaattTt taatagataa attatcggga tgtagcgcag tttggtagcg
68881 ctctgtTTTt gggaacagga ggtcgcaggt tcgaaatcctg tcatcccgat tctaaattaa
68941 tcttactTTT tatcttcatt tacttctatt tTaaaagTTa cgcgTTTaaa taaaatgatt
69001 tgcgTTTaaa atctgaggtc cTaaatataa acgagcccaa agatacaaat cttatTTcgg
69061 taatcaagct tcgcttgata accttgcccc gaaggctTca ccctacgggt gaaactTTaa
69121 caagcgtagc ttgTTTTagg atgagggctg tggTcctTcc tccgaacact tttgTTatgg
69181 gctacgtTac tacaataaac tTTTTTaaa gaaaaagacg caacgcctt agttcagttg
69241 gtagaacgta ggtTTTccaaa acctaatgcc gtgggtTcaa gTcctacagg gcgTgTTTTt
69301 ttgaaactag ttggtTataa tTTTctctaa ataaagTaaa acaaattTTT aatattaaaa
69361 atTTgTTTTa cTTtatTTtag agaaaattac tgagatgggt gactTTTtatt tatgaaaggt
69421 aatagaccga tcattcagac atTTTTTata gctTTTgcca tgtgtcgttg ctgTTtagct
69481 gTTaaaccat taatacgtct aggtaatatT tTcctTccg aagtaataaa tTTTctaat
69541 aaaacaatat tTTtataatc aatTgtTcct tGaaatggaa tcagTTTtga ttgtgattta
69601 cgTTgaataa taggaattat tcgTgactt tgaagggtct tattTTTtga tattcggTTt
69661 gaaaaagtca taagctgcaa aattatataa tataaataat agTTaataaa atTTTtatt
69721 taatatgaaa aattgatgtg taattgattt gTaaaaatTT actgattTTT aattacaata
69781 tctTtgcgta gcctTgtctt attagtcatt acctTTTggc cgtTgctcat tagcaacggc
69841 taatgagtaa ataccaaagg gTaaagaaaa gactTTaagc gaagctTaaa gatctTaatg
69901 tgtgcaagca cacatTTTgg actTccacaa agtggaaagat cTaatagaa ctctgtTcta
69961 tTTTggactt tTTTctgca ctatgctacc aaaggtagtt taggtcattg ctcttatca
70021 atggcTaaag agTaaaaacc aaagcgaat acctaatgag gaacctaggg ctccgggtt
70081 gccctattg cccaacgaag ttgggcgtag gagcgaagggt aatattTtaga gaagctTcga

70141 ttgtatthtta gaaaactthta tcaaaaccca agatthtagat ttaaaaaaaa cttaatthtga
70201 aaactthaggt ttggcccttht tataaatcta tcaactagca aagcgtthgct tgacaacttht
70261 aatthttgcac tcaaaagtht tacatathtaa thtaaagatt gaactaattg acgaaaagct
70321 tctgtgatctc gaattgctaa ttgtgctaac atthtttcgat ttaaaaaacat attggattht
70381 ttcaatccat gtaaaaaattg gctatagtht aagccataaa gtcgtacagc agcattcatt
70441 cttgtaatcc aaaggcgtct aaattctcgt thtcgtthgat ttcgatctct ataagaagat
70501 thtaacgctt thattthtttg ctgattagcc ggtctgaata aaatcgaaga tgagccacga
70561 aaccctthctg taatctthtag tactthtttht cgtcgtthtac gagcaatatt ccctcgttht
70621 actcagagtca taataattat cthttaaatt taaaataatt tcgaattact thattgatta
70681 gctaaaaacc gctgattatg ctagtgcaac thattaacct tatatcaatt tggthtactag
70741 gthtagactth atggcaactt tgaaththta thctgaacca tghtccaaat thaaatccca
70801 thtctaaatt taaaagctat actthtaaat gaagthtttac aaacaaagtt tgtaaaaaaa
70861 cttctattht acctaaacta cttgaatacg cgaagthtcgc atatacaagt gaagaaaaaa
70921 cthttcttht tagatctcc thtacattgc tattgacctt ttggtaaagc aaggatthgc
70981 tcagagatca aaactthaat ttgaatacct tgtatthaaa thtaagthaa accaatgggt
71041 aagcaagcga agctthgccga ccttacgcag aagaccttht thtctgcgct acgctacaaa
71101 aggtagthtt agattccgct gthtcaaatta aagthttgccc tatatgthct cthtctggac
71161 cthtagataat gaaagaaagt thcattthgaa cattaaaaata aggctthtcta cgaaatagaa
71221 atcgcaactt ataactthagt aaggcaatt thttatcaga thaaatattt attatthgat
71281 atagthtttag taaaagtht gattthattat agctthgaag ctctaattth taaaaacgt
71341 gatataaact thattthctaa aaaatttat acgtcccatt thaaataaaa thaaacttht
71401 aaataatatt aagtaaaagt thatggthct atctaaacaa aaacgtattg gtaatatatt
71461 agtagthttaa tctaaatthaa ataaaaaatc aaaatthaat thtaaaagcc tghtthttat
71521 thaaatthag aaataaaaaa thcaagaaac actctactth thttaaagtat tacaatthaa
71581 ggtctaatat agactaaagt gctthatttht aaagatatct acaaaaaata atatattaga
71641 taatthaac cccagggga agtcgaatcc ccgthtctc cgtgaaaggg agatgtcctg
71701 ggcctctaga cgatgggggc gaacatctta cthtgtatct taatthattt tathtaagat
71761 gtcaataatt taaatattat actctaaacc aaatthtttht aataagthtag cgggagcagg
71821 attcgaactt ggcacctcaa gattatgagc cthgcgagct accaggctgc tctatcccgc
71881 gctcatataa thataagtht accagaatat aatagthtct acaagattht ggtthtttat
71941 agtccacgaa actctaacca cactgattth gthtattaga attatactta attaaaaaa
72001 thtccagtac gctthgcccag tctgatctac thtagaacct tcagthtgtg ggttatatat
72061 thttthtcag thttcataat aaagthtgat tathttcagc thttthtttag agcaagthtt
72121 attactctth thtatattata gccaatthtt aatthattth tactthattg thctatcctg
72181 ggttcataaa aaatccacat aactthattg acctthggag cgagaactgt atthgaaaga
72241 aaattatatt thatcatcgt ataaaagtag atataacgta atthtcgacaa thtathtttht
72301 tctaattagc taagthgtht aattgtccga acgagtactc tgattthgat cattgccaat
72361 thttthtcgaa tgggcacaaa thtagcaaat tcacaaaatt thatgthttc taactccaaa
72421 agcgtagctt thaaatcaag thttacaac aaagthtgta aaacctthaat aaattathtt
72481 cthcagagaag caaactcagt thaatactct taatctthac tcgttaagca thgctaaaac
72541 accctacaag taaagcctct tgatgthttat thtctagagg tgatcggggg gaggaaagga
72601 aagattcacg ggcaatgthc aacatgthaa gacaaaaaa tcaaaaaaag ctaataacgg
72661 actthattacc thttgaaagt aagthaagthc actaaaaaga aaaaatthta thggthcaata
72721 aatthaaagg gtaaaaaaa aatthagaaa aattacctca cccggaaggg gtatacgaac
72781 aaagthtctt atctagcagc cctatggaaa gaacacaagt ctaaaatata atcagagatt
72841 gtathaaagt cthccacata gtagaagthc aaagthgthc cctaactthaa acaaacatag
72901 thtggthtag atctthcctt tagggaaagc thctthgctaa gthtctaaat atthgtgccc
72961 thggcaacaa tataaagctt taccagcgt taccagcgt agctthgthaa agccagccta cgtactthga
73021 taccggtatc cggtaggatt tgctggthaa gcaaaactth atthaacac ggcaagthta
73081 atthggccct cttaccgthc cttaccgthc ccagtaggat thgctctacc ccthgccgth
73141 cccggcaggg caacgggagc aaaacttht thaaataact cgtthgthaaa thtagcctth
73201 gcctgthggg cgtagcctgg gaatctaggg ctcccaggth agatathttat tcataaatat
73261 thtaccacac gggatthccc cthtgggthaa aactthatta aaaaactthc thttthtaatt
73321 tagacaacgg cggcaggctt thcccgtagg cthtaccgth gctcgggaaac ctagggctcc
73381 cggggcgaag cgtaaaactt tatgthtttac cgtthgccct cgtcaacgg caagggthcta
73441 agataaaaca thtaggacct aatctgaaca aactatgtht gthtagggaa aactthattt
73501 aacaacgaag thgtthattt gthaaattht gcctthgcct atthgggthag atathttatc

73561	ataaatatct	tacccaagag	ggcaacgggc	aagattatca	agcaaagctt	gattaccaac
73621	tcgattttgt	aattcctggt	aaaaaccctt	catgagccat	ttcacgtaat	acatggcggtg
73681	ataacccaaa	atctcgaaaa	aaccctttag	gtcttccagt	tattaaacat	cgattatgaa
73741	gtcttacggg	attactatta	cgaggaagtt	tttgtaattg	tctttgaatt	gtagtttttt
73801	cttctaaaaa	agaagcctgt	ttcattttgtg	cttttaaagt	taaccgtttc	tttgaatatt
73861	tcattactaa	ttgtttggcgt	tttttttcgc	gctcaacat	acttttttta	gccataaatt
73921	caatttttat	acaataaata	atztataaat	acttttctaaa	ttttatttta	aacttttattt
73981	taaatttttg	ctagcttggt	taaaaagtca	taattaaana	tatgtttatt	ctactatttg
74041	tatattttcg	aagctaatat	cctaagtatt	ctcgtttttt	cagagctagc	ttaagtaaaa
74101	aacttttactt	tgaaccttag	aaaatattat	tcacatttta	ataaaaaatt	aaaagcttac
74161	ataactaaaa	accttttctag	gttgataatc	aggttgtttc	tattttttata	aaaatcgggc
74221	taaataacct	ccgttttattt	accattttat	aaaaaactta	ggatttttact	atgaataaat
74281	atataaagga	taaaagctag	atztataaat	ttataatatt	atgacttgaa	gcttcgccta
74341	aactacttaa	tcaagcttta	cttgattaag	tagcgtagct	cgggaatcta	gggctcccgg
74401	gtaagggtga	caaacttcgt	tcgcatagcc	attgggtttg	gcattctcaa	aaacttcaag
74461	atctaaatgt	tttattaaag	ctttccctaa	gtaaaaaacat	aaagttttta	cgcagaagac
74521	tttaaccctt	gccgtgcccg	gcagggcaac	gggcgaagct	taaagatcct	agtgtgcaag
74581	cacacacttt	ggactttcct	aaaggctttc	ctaaatgttt	tatcttagac	ccttgccggt
74641	gacggcaggg	caacggtaaa	acataaagtt	ttacgcttcg	cgtaaagcct	gacgcggaaa
74701	actttatttta	acaactttgt	tgttaaattt	agagctaatc	taaacaaact	atgtttgggt
74761	aaattagctc	tattatggac	tttttttctg	tgctacgctg	tgcttcgcaa	gggttgacct
74821	ttcagggtcaa	gcaagctatg	cttggtaaaa	ctttatgttg	tctgtccgcc	cttactaata
74881	agcaagggct	aatgagacaa	cgtaaacata	aaacataaag	cttcgcttaa	atgtttgatc
74941	agcccttgcc	attgtcaagg	gcaatgaaga	tcaaacatga	agttttacgt	tttgccccgg
75001	gagccctaga	ttcctcatta	ggcattgcta	atgagcaatg	gccgggctac	ggtaaagcct
75061	acgggtatcc	gttttgcttc	tagcaaaacg	taacccttcc	gcagaagacc	atggtaagaa
75121	tacaagtcca	aagtttgtgt	gaaacacaaa	ctaggatcct	ccacttcgtg	gaagtccaaa
75181	atgtgtgctt	gcacacatta	ggatctttta	gcgaagctta	aagtcttctt	actttttttc
75241	tgcgtctacg	taccttcggc	cgttgttcat	taacaatggc	taaagagttag	tttaggcaaa
75301	gtctagttaa	caaacaaagt	ttagtaactt	tggagctttg	ctcctaaaaa	aagtattctc
75361	taaatataaa	tcaaagaaac	aaacttaggt	ttaaagcata	atatgatatt	taaatccact
75421	tttccaaaat	acaattgaaa	attagagtcg	actattctac	taactgacat	tcttaatatc
75481	ttataaaaag	aaagccttcg	aacaaaactt	tacaaaaaaa	cttctttctt	tatcttaaat
75541	ttgacgataa	tatagaaatg	agaagaaaaat	ttaaaaatgt	tagaataaat	tagtaattat
75601	gaaaaaaaatt	ttttgttata	aattttactt	tctaaatgta	tcaaaagtaa	tattattttag
75661	tttaagtgtc	cttttagcct	ttgatttaac	tattacaaat	cctgcacaag	cttaccctat
75721	ttttgctcag	caaaattacc	ctagtctctg	tgaagctaac	ggtcgaattg	tatgtgcaaa
75781	ctgtcattta	gctcaaaaac	ccgttgaact	tgaagtacca	caatctgttt	tacctgatac
75841	tgtatttgaa	gcagtagtta	aaatacctta	tgatcaacaa	atataacaag	ttttaggaaa
75901	tggtaaaaaa	ggacctttta	acgtcggggc	tgtttttaatt	cttcctgaag	gatttgagtt
75961	agcacctgca	gatcgaattc	ctgaagaaat	gaaaactaaa	gtaggaaaac	tttattttca
76021	accttatagt	gatgataaaa	aaaatatttt	agtagttggg	cctttaccag	gtaaaaagta
76081	tagcgaaatg	gtttttccaa	ttctttctcc	tgatccagct	aaaaataaaa	atgtttctta
76141	tttaaaaatac	ccaatctatt	tagggggaaa	tagaggctcg	gggcaagttt	atccggatgg
76201	atctaaaagt	aataatactg	tatataatgc	ttcagctgca	ggaaaaatta	gtcaaattga
76261	accaatcgaa	aaaggtggct	acaatattac	aattgaaaca	aataatggtg	agcaaaaaat
76321	tgataaaaatt	cctgctggac	cagaattgat	tgttcaagtt	gggcaaagtg	ttcaaaatga
76381	tcaaccatta	acaacaacc	ctaacttagg	tggatttggg	caaatggata	ctgaaattgt
76441	cttccaaaac	cctactcgtg	ttcaaggttt	aattgtattt	tttgggtttg	taattgtttg
76501	ccaaattttc	ttagttttta	agaaaaagca	atgtgaaaaa	gttcaattag	cagaaatgaa
76561	tttttaaaact	attactaaaa	tctgacgtgg	tctaatcaat	ctttatgttt	ctaaaaataca
76621	gtcttcaatt	gtattatgat	tttttactca	ttctccattg	ctaaaaacaag	caaagcttgt
76681	tcaagtttct	aaatttaaaa	gcgtagcttt	taaatcaagt	tttacaattt	ttatgtgtaa
76741	aacaccctat	gagtgaaacc	ttttgggttt	tactcattag	ccgttgctaa	tgagcaacgg
76801	ctaaagggta	gtcaagcgaa	gcttgacaac	cttgccttga	aggcttcacc	ctacgggtga
76861	ggctttaaca	agcgtagctt	attttagggg	gaggcaatgg	tgaagtggaa	ttgattattt
76921	gaacacaaac	ctgggagctc	tagatcccag	gttctagtaa	atttttttaa	atatctattt

76981 taatcttatg ataacaatta ttagctatat cggtttactt actggtgtat tgttttttac
77041 tttatgtgtt tatgtcagtt tattaanaaat taaattaatt taatttttta taattatatt
77101 tttttgcttt ttactacaat taaaagcaac atttattttg aataaatttt aaaaaattca
77161 ttttttgttt aattaattag tttattaatt aaaagtttta aaaaaatact tatttatttt
77221 aaacgttttc atggttgaac ctcttttatc tgggaattgta ttaggattag tacctgttac
77281 ttttagcagg ctatgtgta ctgcatatct acaatcga cgtggagatc aattaaattt
77341 ataaatattt aattgatagt aaaattattt ttcaaagtta acaacaacg tttgttaact
77401 aggctttacc taaactatct cttagccatt gtcttttggc tttaccctt tggcttttac
77461 cctttggtct ttactcatta gctcttgcta atgagcaata gccagaggct ttacggaag
77521 cgtaaaactt tatgtttttat ctttgataaa acattttaggt aaagacttcc actttgtgga
77581 agatcctagt ttgtgtttta cacaaacttt ggactaataa gcaatggccg aaggtaggga
77641 ggttggtgag ctttgctcac ataccgtag gctttaccgt agcccggaaa cctaaggctc
77701 ccggggcaaa gcgtaagct tcttggtcgg taccctcttt caaacagatt aaaaaccaa
77761 gaggctgttc caattaaaaa cctagaatac ttccaccctt accgtgttta gaaaggtttg
77821 cctcaattta agaagaaagt tttagataag ggcggaatgg aagatttctg taaactttaa
77881 aactttcttt aaaaatgtag attttaaat taaaccttag ttactaaact tcgtttgta
77941 atattgtatt tggaaactga ctctaagta cgaacttctc ttagtgaatt agttaaaaat
78001 gtctactcct tacttaaaat tagcttaact tttttaagct ttacctaaac ccgggagcct
78061 taggtttccg ggctaggggt aaagatacaa aatttacta tagtcttcaa ttctatattg
78121 acctatccct tgcattaacc tttggttttt cattattcgt tagcgagctt cgcttaaagc
78181 gtctaacaaa gagtaaagat ccaaaagctt tactcgtagg gccagagggt aaagctttgc
78241 gtcgtccaaa gaccatggct aaaaaagtac aagacattca gcaattgctc aggtaaagag
78301 taatccagga ggttctacaa acaaagcttg taaaacttgg cgtagcacag aaaaaagtc
78361 ttctgtgttt aaaaatgtaa ctttcaaact tagaaattca aaaatccaac gctttataag
78421 ctctgcttat aaaactttgt ataaatagct ttcagctatt tatacttaga gaacaccctt
78481 aagcaatata ttttaaatat cgttttagtaa attttggttt gtctccgaag ttttctcata
78541 gatagatttt tcaactgctag attttttctt ttctgcaaaa taaattttta gtactaaatt
78601 tatcaattta ttaaaaatta tattgttttt ttcatttttt tcttttagta taagaagaag
78661 aaatactatt ttaaatagta tttttaaata agaaaatctc tttttttaca agttaatgag
78721 aatttatgct agatataata actcttgatc aaagtctgctc catttttccg tttactgcta
78781 ttgtagggca agaagaaatg aaattatccc ttatcttgaa tghtaattgac cctaaaattg
78841 ggggggtaat gattatgggt gatcggggaa ctggaaaatc aactactggt cgcgctttgg
78901 tagattttatt gccagaaata aaagtgtgat taaatgatcc attttaattca gatcctgaag
78961 atcccgaatt aatgagcgaa gaagtctgtg aaaaactgaa aaaaaatgaa aaaattgaag
79021 ttactaatcg aaaaatttta atggttgatt taccactggg agcaaccgaa gatcgcggtt
79081 gtggaacaat cgatattgag aaagcgttaa cagaagggtg taaggctttt gagccagggt
79141 tattagccaa agcaaactga ggaattttat acgtagacga agtaaattta ttagatgatc
79201 atttagtaga tgttttatta gactcagcgg catctggctg gaatacagtt gaaagagaag
79261 gtatttcaat tagtcacca gcacgtttta ttttagtagg cttaggtaac cctgaagaag
79321 gagaattacg gcctcaattg cttgatagat ttggtatgca tgctcaaatc ggaaccgtaa
79381 aagagtctaa ttttaagagt caaattgttg aacagcgagc aagttttgat gaagctcctt
79441 tagaatttag aaaaaattac aaggattcac aaaaagaatt aactgaaaaa cttacgcaag
79501 cccgtaaaaa tttaaaaaat gttgaaattg attatgattt ccgtattaaa atttcgcaaa
79561 tttgttcaga attaaatgta gatggattac gaggtgatat cgtgacaaat cgtgcaagta
79621 aagcaattgc cgcttttgag aaccgaactg aagtaacgcc tgaagatatt tttagagtaa
79681 ttctctatg tttacgacat cgcttacgaa aagaccctct tgagtctatt gattcagggtg
79741 ataaagttcg agaaactttt aaaaaagtat ttggatatca gtaaaacaat ctaagaaact
79801 cagcctaata aatcaaagct tgggttttcaa acacaaacta ttttaagctt cagggttaat
79861 tctttcttaa agcaagactt gaagcttctgc ctaaaactact taattaagct taccgtgatt
79921 aagtggcgta gcccggaat ctagggctcc cgggtaagggt tggcgaactt cgttcgcata
79981 cccattgggt tgtgcattcc cacaaacttt aagatcctaa tacactcttc gagtgtattt
80041 tggaaatctta gtctattgt tttacttgtt gttaaaacga gtaaagctta ttcgatatat
80101 attactcgta gggtaatacc cagtttcacc ctacgggtat acgagcgaag ctgcaccaatc
80161 ttactactga aggtagttta gccaatgcga aaaaataaag cagctaaaac tagctttctt
80221 attccttgtt tttactctaa attaaaactt tatattgact ttgaatttta tcttaaatat
80281 ttttcgttat tcataaattt ttttttcgta aaacaacatt ttacttaata aaactattaa
80341 gtaagattga tttaaaagcg tcatttttag ctttactcgt atatatcttt ggtaaaataa

80401 cttcatatca agttattact taaattttaa acaacattaa caagcaaaac tgacataaat
80461 aaaattttca caataaaata aaatgaaaaa aaattgtatg aatattgaag tcttagatta
80521 tttttgatat attaatatct aatagaaatg agaaaagtat tagggcttat cgtctaaggg
80581 attaggacag aaatcttcta agtttcta atcctactaa gctcatat
80641 ttagcactaa ataaaaaaaa ttaagacttg aattcaagtc ttaatTTTTT atcaaagctt
80701 tccccagta aaaacataaa gtttttacgc cgaagacttt aacccttgcc gtgcccggca
80761 gggcaacggg cgaagcttaa agatcctagt gtgtgcaagc acacactttg gacttctcta
80821 aagactttcc taaatgtttt atcttagacc ctgcccgttg acggcagggc aacggtaaaa
80881 cataaagttt tacgcttcgc gtaaagcccg acgcggaaaa ctttatttaa caactttggt
80941 gttaaatttc gaccgctagc ccggccattg ctctattagca atgcctaag aggaatctag
81001 ggctcccggg taatctaaac aaacgcagaa gacttaaagc atagctttaa gatcctagta
81061 tacaacaaa gtttgtatac tttggacttt ttttctgccc ttcgctatgt ttgtttaggg
81121 aagactttta tagagtgtgg ctctattatg aactTTTTTT ctgtgctatg ttgcccctcg
81181 caaggtttga ctttcaggt caaacaagct atgcttgata aaactttatg ttatccgtcc
81241 acccttacta atgagcaagg gctaatagaga caatgtaaac ataaaacaca aagcttcgcc
81301 taaatgtttg atcagccctt gccattgcca agagcaatgg agatccaaca taaagtttcc
81361 cctaaatggt gttggtacaa caacatcaag ttttgctaa atatttgatc aaagatcaaa
81421 cataaagttt tacgctttgc cgcgagagcc ccagggttcc gggctacagt aaagtatacg
81481 agaaaagctt acgggtatgt gagcaaagct caccaacctt tcttcggcca ttgctcatta
81541 gtccaaagtt tgtgtaaaat acaaaactagg atcttccata aagtggaagt ctttacctaa
81601 atgttttatc aaagataaaa cataaagttt tacgcttcgc gtaaagcctc tggctattac
81661 tcattagcaa cagctaataga gtaaagacca aagggcactg ggttaaggagt agtttaggta
81721 aagcctagaa actttgtttg ttaatTTTTG aagttaatta tttttttgg gccgagctgg
81781 atttgaacca gcgtagatat aaatcagcgg atttacaatc cgccccatt aaccactcgg
81841 gcatcgacc tttttataat tttaaaagtt tagtgaaaaa ataaaatggt tatttaaagt
81901 ttactaaatt tagattta attgcaatggt ttatctaaaa gataaatttt tgtaaaggcg
81961 taaaaatatt gacgtatcgt tgtttaactt taaaaacata gcattgctcc attgctttga
82021 tctaaaattt gaacactttc tgtttaaatt aaagttttcc ttatttaata tatttaaagt
82081 tataagtgt cctgtaatag aaaataatta aaaataattg ttttaaactg atttttatg
82141 atatacttat taaaataaaa aaaaactagt taggggatat ggcggaattg gcagacgcta
82201 cggacttaaa atccgttggg tgaaaaaccg tgagggttca agtccctcta tcccaata
82261 cagctctttc aaagaataaaa acaaatTTTT tcttatgcat tctttaattt ttttactgg
82321 gtattgaatc ttttctttta ctgggttagt ttgttttaaac tgcacttttc acttcagtca
82381 tttgcccagt gtgacattaa cggaaactaga acgtcaattc tatctgtctc tttgttacac
82441 ttagacatag ataatactaa atcaataaga tgggtcaaatc agaactttaa aaaggctaga
82501 tttctccgta ggctttaccg tagcccggga acctaggggt cccggggcga agcgtaaaac
82561 tttatgtttg atcaaagatc aaacatttag agaaaactta aacaagcttt cttattttta
82621 gaaaataata attatcatta tcaaaaagct ttctataatt attaattag gcattgtcaa
82681 aaaattttaa ataataattg atcttatctc aatataaaat tttagtttta tattgagata
82741 agatacaata tttaaactca tgggagcctg gaaaacttta cagtctttgt ttttttctt
82801 ttgatagaat tatagaaaga aagttttagt ctgacttat aaagtgatgc ttaagtattt
82861 ctagttaa ataaagctttaa tcaatcgtta aaacctataa aaagccaaa atatattggt
82921 ctagtaaa atgttctctaat ttattttaga aataaaaata attacttagt aattattaaa
82981 ctttacgcga ataatactca acaattaata attcattaag gttaaagatt actgattgtc
83041 tattcactat attattaact gtttctataa gggtttcttt attaaaagaa agatgaatcg
83101 gaatgggtatt tatatttgaa gtttcaaaa agttacttac tacatttcga gaatgctttt
83161 ttttagaaac tgatatggta tcttttgggt tacactcata gcttggaaata gtcacttttt
83221 cttatttaac aataatatga ccatgacctt ttaattgcct agctgcagct atagtaggtg
83281 ctattcccaa tcgaaataca atggtatcaa ggcgcatttc taataattga agtaaaattt
83341 caccagtaga acctttaatt tttctggctt ttcgtacata acgtattaat tgagattcag
83401 taattccata gttaaaacgt aacttttggt tttcttgtaa acggatacca tactgtgata
83461 attttttaac actagtacca tgttgaccag gaggagtttg tcgagctgaa gatttactag
83521 taaatccagg taattctcca agtttgcgta taattcttaa gcgtggacca cgatagcgag
83581 acataatatt ttagtagtat tttttattga aaaaatttaa atgtgtatc aatttttagt
83641 aaatcgagtt tttaaagatt tctgaaaatt aaaaacttaa atttgaccac agagtgttca
83701 aacaagagta agactatggt tatcttcaat tcttattttg ttggttttct aagtgaacaa
83761 acttcgttt ttaaccaagt ttttgtcgaa gggaaacct aagatgtttt gttcacgttg

83821 acctttggtc tttactcctc ctactacgta ggagtgtagg tccttagcca ttgctaagga
83881 gcaatgggat aaactacctt tggtagcgta gcgcagaaaa aagtccaaaa tagaacagag
83941 ttctattagg atcttccact ttgtggaagt ccaaaatgtg tgcttgcaca cattaggatc
84001 ttgaagcttc gcccgttgcc ctgccgggca cggcaagggg taaagtcttc tgcgtaaggt
84061 tggcaagctt cgcttgctta ccaaagggta aagactaaat gttttctctt ttacttacac
84121 tgttatgcag tagcagaata caacataaag ttttacgctt tgcgtaaaat cttgctggcaa
84181 cgcaaaggat gaataacact caaacttaga aatcaatfff taaaacaaat tttttgtgtt
84241 ttaacttttg ttttaaaaa actagaagac ttaaataaatt gatttcaact tagaatagaa
84301 ggataaaaa atttttaaaa tgaggtttta aaaaattata gccatcaatt acaaactaat
84361 aatcaaaaata taaatatgat aatattaaaa atattaacat atgtcaggat aaaaatacaa
84421 tctgtttcat ttgtactaat aatatagtca attaagaaca aattttttgt acgtcgtagc
84481 ctggcaagaa gcttcagtgt ttccaatffa gaagcttcgc ttgttcaagt tttccccgta
84541 ggctttaccg aagctcggga acctagggct cccggggcga agcgtaaaaat tttatgtttt
84601 atcgaagata aaacatttag ggaaagcaac ctttaggata aatttacaca aaattttagt
84661 cgtagaacaa agcctaacat acggcaacaa taaataaatt tgcttgagta aaagaattta
84721 ttatggtaaa ctatattctt tgaagagcct tattaaaact tcaaaattaa acgacgattt
84781 ttaagcttag aaacatffff ttaaaggaaa tacttcacct taaaaggctg aagctgtgtt
84841 ttacaaataa aacttaacta aaaactatat tttcaagctt attttttatg actttattaa
84901 aggctttctt tttttataa aaaaaaccaa tcattaaacc ataaattaa ccagtatatt
84961 aacgttttgt gttgaatfff tgtttttcga caaatffffa gatttttgct aaaaaacgct
85021 attctattaa ataccatfff taatagagat cccatfffga tttacttac aggaaagctt
85081 tataagcttt atgtataaaa gagcattggt ttagccaaaa cgtagctgat ctacactaat
85141 aactacataa catffffaat ttaaaatffa aagtcttcta cgtttgact aaagttttat
85201 aagctacgct tgtaaagccc gtagggaaaa tcaaataagc tttatftatt caaatftaga
85261 aatgataaag ctagaggtag ggtccgcgta aaaactagaa taatftcatt attacaaata
85321 acacaatgta aaataagctg gaatfttagaa agtaaaacta aatfttaacc tcaaacfttt
85381 tttaggatat ttgctftttac tfttaaaataa aagtaacatt ttcacatgta tataactaca
85441 tctggcaata aagttccaaa tggaaaccga gggctfttaa tgtaagcatt cacactaagg
85501 atgtgtftaa ggttaatfff tgtftttaca actfttaaca aactftftaa atgacagact
85561 ttgcctaagc tactcattag tccaaagftt gtgtfttaca caaacfttggt actaagtaga
85621 attatggccg aaggtagcaa ggttgatggt ttgctagaag taaaacgaat atctgttgct
85681 aaatgtfttc tcaaaaagaa aacgtfttaat gaagactgaa gcctcgctgct cggctgtgtt
85741 tttactatcg atftftagfta gagftftatgg cttcttgctt aaaactcgct acaataaata
85801 atatftfta aatftftaac tatgagtatt tctgtagaca caaagaatgt aggacgtatt
85861 acacaaatta ttggaccctg tfttagacgtt gctftttccag caggaaaaat gccaaatatt
85921 tataatgcaa ttattgttctg tggaaaaaac gaagcaggag atgaaatatt ggtaactftgt
85981 gaagttcagc aattactftg tgatcactgt gtccgtgcag tftctatgaa tgctacagat
86041 ggattaatgc gtgggatgga tgtagtggat acaggaaaac cctgacagat tccagfttgg
86101 ccagctacac taggccgtat tfttaatgta cttggagaa cggfttgataa tctaggtcca
86161 gttgatgcta aagaaggatt acctatftcat agatcagctc cftftatftgt ggatfttagac
86221 actaagctft caatftfttga aactggaatt aaagftgftg actftattagc gccttaccgt
86281 cgtggaggaa aaattggatt atftggtgga gcgggagtag gtaagactgt attaatattg
86341 gaattaatta ataatatcgc aaaagcacat ggaggagftt ctgftftftg gggtgtagga
86401 gaacgtacaa gagaaggtaa tgacctftat atggaaatga aagaatctaa agtaattaac
86461 gaagataata tftctgaatc aaaagftgct ttagtatatg gtcaaatgaa tgaaccacca
86521 ggagcgcgta tgcgtgtagg tctgactgca ctgacaatgg cagaatattt tagagatatt
86581 aacaacaag atgtftftgtt atftatftgat aacatftfta gattftgttca agcagggtca
86641 gaagtgtcag cfttactagg tctgatgctt tcagctgtag gttatcaacc aacattagca
86701 acagaaatgg gtggattaca ggagcgaatt acatctacaa aggatgggtc tatcatatca
86761 attcaagctg tftacgttcc tgctgatgat ttaacagatc cagcaccagc tactactftt
86821 gctcatttag atgctactac tgtactfttca cgaggattag cfttctaaagg aatfttatcct
86881 gcagtagatc cattagattc tacatcaaca atgttacaac cfttggattgt tggagaagag
86941 cattatgaat gtgcacaaaa tftaaaagaa acattacaaa gatataaaga attacaagat
87001 attattgcaa tftftaggftt agatgagfta tctgaagaag atagattagt tgttgcgca
87061 gctcgtaaaa ttgaacgtft cfttatcgcag ccattcttctg ttgctgaagt atftactggc
87121 tctcctggaa aatatgtaag tfttagcagaa acaatftcaag gattftaattt aatcctftct
87181 ggtgaattag atgatttacc agaacaagca tftftatftag taggaactct tgatgaagca

87241 attgctaaag ctgcttcatt acaagagaaa aactaaatth ggatccaaaa ttaccaaaaca
87301 aaatthggta actcggatcc aaatthaaaa gcacaactth taaaaaaagt thtactgth
87361 tgaacagtht ctgthtaaac aatggaaatt taaatctaag cththtgcca ttataaatag
87421 aggthtaaaa agcataagaa ctataaatag thththtgagt thththtagaa atthcttaca
87481 thgccccgtg ththtatccct tacattgacc thtaatctgt cthtattagt tgatgagctt
87541 cgctcgcata caaaaagata cagagtaaac agcaatggcc tggctctctac atcaacctcg
87601 thgaacaata tgggcaaggg tgatgagcac agctccaaaa thtatgcaga gtacaaatta
87661 taatthtaac aagthtcgct tggtaaagcg tctatcaaaa gataaagacc atcaggttha
87721 atctataaag taaaacttha thcaagaatt ccactthtgth aaagatcctg cththtgctt
87781 tgtacaaagc thggaacttc gthcttgaaat thaatattga ccccttggha aacaagcgaa
87841 gcttgccaag cttacgcaga agactthaac ccttgccgth cccggcaggg caacgggcca
87901 agctthaaaga thctaathgth tgcaagcaca cthththggac thccacaaag tggaaatcc
87961 taatagaact ctgthctatt thggacttht thctgctgct acgctacca aggthaattta
88021 ggctattgct cattagcaat agctaathgag caaagaccaa ggggacaacg gtaagcaaac
88081 cthgththtag caatgatcaa agthaaatth thgththgaaa caacgththa cthtactatgth
88141 aatthcgttha thccaatgca atatthaaag thagatcatt thaccgthgth ctatgcttha
88201 tatgththta thgacagatt aaactctatc agatthgagth gctgactggc acgghaaaga
88261 thcgcgctt thgtatthta actcaaggaa taacacatat thggagghga acagaaatgag
88321 thtacagghth thcattatga cgcctgacg thththththgg aatgatcaag cagaagaaat
88381 thththtacc actaatcacg thcaaatggg agthctatca aatcatgcac cthtaattac
88441 agcattagac atthggagtha thgthaattcg ggcaataat gattgghatt caatagcghth
88501 aatggghggh ththgctthtag thaaacaaaa thcaattacc gtactagtha atgagghcaga
88561 atctthcaaaa acaathgagth cagaagaagc agaaaatthct thtaataaac ccaaatcaa
88621 attagaagaa gctgaagghc aaaaaacaaa agthagaagct aatththgtht thaaaagagc
88681 aagagcaaga thacagghth thaaaacaatc aagthaactaa thththaaat thththaaat
88741 thgccagththc thagaagathc thgaththgtha thcacacaaa cththataaaa thgaaatth
88801 thaaaatghthc aatataagth ataaaaatta cthgththtagt aaaaaaaagt thctththaaa
88861 thaatghthta ctatthaatta agataattht ththagattha aaaaaattht thaatataaaa
88921 atagaaagth thgatgcaaaa thgthaaagct gaghctththgc thaaacaagc thtactgct
88981 aaagthtccat cthccggthag thcaagctthc cththgacaacc ththgcccghth cthaaattha
89041 aaaaagaaat thththtaata agthththgctc thacagcaaaa thctaccgha thaccgghatc
89101 aagthcagtha gcctggctth aacaagctac gctthgthaaa actththatt gthgcccghth
89161 acaacaatth thtaggaact agcaagaaga cthccctaaa gghaagatcc thaaacaaaca
89221 thgththgtht agatthaagthc thaaatthaac aactthcghth thaaataaac gththcccghth
89281 agghcaaacg cthagthgthg caagcacaca cththggact ccactatgth gaagacttha
89341 atacaatctc thgaththatt thtagactghth gthctthcca thggactctt agghaaggha
89401 cththgthcgh atacctththc ggagthgagat thaaagcctha ththththaaag cthcghctct
89461 aagthacctth thgghagaag gththgacgag thaaatthth ththcaaacct atthctctga
89521 atatccathc thaaaaatth gathththgag thaatgthct agctthacta gctthcaata
89581 aatththattth ccaaaatthc caaataaaat thgghaacta aatctththtag atthagaaac
89641 gghcagatha thagatcgtha aaaatththta thaaatththg thaaatggc thccacaaact
89701 gaaactaaaa cagccgctgh cththaaagca gghthaaaag atthacgath aactthactat
89761 actccagath accaagthaa agathactgath atththagcag cththcghat gactctca
89821 cctghgththc ccghthgaaga gthgghthgca gcagththcag cagaatctc aactghthact
89881 thggactactg ththggactga thgththaaact agthctagatc gatacaaacg acghthghat
89941 gacatcgagc cagthcctgh thgaagaaat caataththg catacghagc atatctthta
90001 gathctthththg aagaagghat thgthacaat thattcactc thaththgag thaaagththc
90061 gghththcaag cghthcghthc atthcghctc gaagatctc gththaccact atcathctca
90121 aaaaactthcc aagghactcc thcatghtht caagthagaac gthgataaac thaaacaaat
90181 gghcghgghth thcttagghth thactatthaaa cththaaatth gthctthcagc thaaaactat
90241 gghcghgag thththagghth thctacghghth gghththagath thactaaaga thgathaaaac
90301 gthaaactcac aaccatthth gcghthggaga gaccghththc ththcghthagc agaaagcaat
90361 thataaatctc aagcagaaac thghthgaaat aaagghthcact actthaaactg thactgaggha
90421 actthctgaag aatgththaa acghthcacaa gthagththaa aathagghthg acctathghth
90481 atgcacgath atththactgh thghatthact thcaataacta gthththctca cthactgthcgh
90541 gathaatghth thattactthc caththcagag gctathgcatg cthgthaaatga ccghthcaaaa
90601 aatcatghgha thcactthccg ththththtagc aaagctctc gthththcag thghthgathc

90661 cttcattcag gtactgtagt aggtaaactt gaaggtgaac gtgaagtaac tttaggcttc
90721 gttgatctta tgagagatga ttacattcaa aaagatcgta gccgtggtat ttactttact
90781 caagattggg cttcacttcc aggagtaatg cctgtagctt cagggtggtat tcacgtatgg
90841 cacatgccag ctttagtaga aatTTTTGGT gatgatgctt gtttacaatt cgggtgggtgt
90901 actttaggac acccttgggg taatgcacca ggtgCGGCTG caaacCGTGT tgcgttagaa
90961 gcttgactc aagctcgtaa cgaaggtcgt gatcttgctc gtgaaggtgg agacgttatt
91021 cgtgcagctt gcaaatggag tcctgaacta gcagctgctt gtgaagttt gaaagaaatt
91081 aagtttgaat ttgaaactat tgatacttta taagaaatat ttagtTTAAA ttttctgga
91141 gtttcgaaaa tttttataaa aaataaactt aagctcgtta ttctaagtta cactatacaa
91201 agtagaaggt ctttgccggt tctaccttgt tttacctaactactTTAAC aagcattaat
91261 gcttgTTAAAG gtagattagc ccagaacaca aatcTTACCC cttggTTTTT accaaggggt
91321 aaaattaacg agctccactt gcatactcga aaggTTAAAG ttgatgTTTT atctaacttt
91381 aatTTTTTGG tttaaataac gctTTTAAAG ggtcTTTCTA ttataagtaa atcTAAACA
91441 tgtaagtttc gaatTTTAAAC acaaatTTTT ttactcaggt ttgatacctt ctagtgttga
91501 cctTTTTTAAA aagaaatat aaaaaatatg cattcagttg ggttTGAACC aacgacgtct
91561 tttcagaacc gcattatgag tgcggtgcaa tCGACCCTC tgccatgaat gcttTAAATA
91621 ttttaatttt accttagatt taggactttt atgctTTAAA gctaagTTAA tgaagctagg
91681 tgatgctaaa gaataaatta ttttaatttt gctttcattc actactcaaa tagtTTTAAA
91741 atgtatttta aattatgata taataataac tacatttcgc agtgAAAAAA aactgatatt
91801 tgttatgacc ttatcttatt gtgaaagatg ctaatttagc ataagTTTTT tcgcttagaa
91861 aactaaaatt ttattaaact cactttccta aaaaagaac ttcaataaat taatagtctt
91921 tcaggaaaat tcgaacacat aagatTTTAA acaatgtagg aaaaattttc ttttTAAAG
91981 ttttctaaag ctaaaattag gttataaatg aaatTTGCGA tttggTAAAT gttgacttga
92041 attgaaactgc ttgaagtttt acaatTTTTT taaaacgaca ctataatttt gatactacat
92101 ccaaggttta aacttcgcta tTAAACCAA attagTAAAT cagttaggat ttttTAAAA
92161 atgagttctt gtgttttagt taaactaata actTAACTAT gTAAAGTGT tttactaaat
92221 aaattactac aaaaataatc ttatagaatt ttactaattc taaaagttct tgttttTGTG
92281 catcattact caaacaagca aggttttatta aaatgttcca tgctatattt ttgttttagt
92341 ctttaagtca aaaagtcaaa cccaataaaa atcTTTACAT aaaaagTAT gcgacggctc
92401 ggagcaaagc tccaagatct cCGCCTTGCg gggagagtaa atactctagt tactcaattt
92461 taattggtaa ctttggaaaca gctttcgcatt ttaagtttca cttttgTTTT atttTAAAC
92521 aaacctgaat acgacactct taatTTAACA gcatagcttt aaaatcagtt tTAACTAGT
92581 tgaacacaaa gtgttcaaac aatggcttaa attacccttg gtagcgtagc gcagAAAAAA
92641 agtccaaaat agaacagagt tctattagga tcttccactt tgtggaagtc caaaatgtgt
92701 gcttgcacac attaggtctt ttaagcttcg cCGTtGccc tgcgggcac ggcaaggggt
92761 aaagtcttct gcgtaaggtt agcaagcttc gcttgcatat ccattggtat gcaaacgaaa
92821 ttcgctactt cggtagttta agccattaaa tTCCcattgc ttttccgcta gattttTAAc
92881 ttctctaagt gtcaactcaa aggttacaca aacttttTGT cgtagaataa actcttctag
92941 cagaaaaaag atccgacggc taaaaaccCG agcaaagata aactTTAAAT aaagtgtcac
93001 cctctgagta aaaccgttgg ttttacttta gaaaagtcta ataccaattc ttcatgtggg
93061 tatataataa gactcactta aaaaaattaa tttaaaaaag gaatatttat tatggctgct
93121 tcttattttac cttctatttt agttccctta gtaggtcttg tatttccctgc tgttgcaatg
93181 gcgtctttat ttttatatat tgagcaagaa gaaattaact aaaataattc aaccatttaa
93241 aagctacgtt tataaaatgt caaatagaaa ttttatttga tgcctttatt ttactagaaa
93301 aactttgtta ccaacataat cttggtagac caccaaaacc gctcaactta tctcctgtat
93361 caataaattt gaatcctagt tagtaaactt tatttgctaa ctttggactt tacgtggtta
93421 tttaggttat actttatttc ttgattttta ctcaaaccct tcatctagaa cacaagtcta
93481 aaatagacta tttccctaaa acaagctttg tttgCGTtG ccctctcgat ctttacttat
93541 tagctattgc taatgagcaa tagcctaaac tactttcgggt agcgtacgct agggAAAAAA
93601 tcttctgCGT aaggttggca agctttgcta gcttaccAAG gggTAAAAc taaattcaag
93661 aacgaagttc caagctttgt gcaaaccaca aaccaggatc tTCCactttc cattgtcctt
93721 gcaatggcaa ggagtggaa gTCCAAaccct atgcgtagca taaggtagga tcttctatta
93781 tatagaagtc ttgaataaag ttttaccCGT aaggTAAAGc ctactggaca cgcTAAAGga
93841 gtcaagtttc accctcacc cgaaggagtg agggTGAAGc cttcaggttt taataagcat
93901 agcttattaa aactttatat tgttatcGGA acaatatcta gggagTATag gtcgTAAatt
93961 ctattacgat gttctacccc gtacaagTct tttatgaaag gttggtgagT tTcactcaca
94021 tatttataga atgaaattca atattttatc taagatcaaa aagactcaat ctgcttacgt

94081 ggtatgacaa aaacttttgc tagtttttgg aaaaaaagt aaagatgaga tgtcttatat
94141 tatctttttt atttaattgt ttttactgct aaagttgata tttgtattaa aactttaatt
94201 gaaactagag cctagattaa cactttcctt atatactata ttttaataaag tgccaacctt
94261 tataaagtct ttgactgtag tttgaagaaa ccaattataa atctaataatga tagcaaaaaa
94321 atgaaactag cttcatttga ttattttctc tttttaaaag gatattttaaa gcaatattaa
94381 aacaagttat tttcttaaaa aatcaaat acttaatcaa ttgctctcta ttaaaaaatg
94441 tattataatt tactatataa agggcctata actcagttgg tagagtgcaa ctcttacaaa
94501 gttgaggtcg ttggttcgag tccgactggg cccaaaaatt gactttttta tttctttcta
94561 gagcctttca tagtttctat tttttgacaa aatgaataa ttaataaaaag atgtaattgt
94621 tttcagctaa caatgtaggt aagtcgtttt ttctcgaaaa cactaagaaa ttgatttaat
94681 ttaaaaagtc tactctacaa gtagatttta cttatctatt ttaaacacta tgggatcaaa
94741 tctcgtaaag gtggaagtca aagctgtagt tttagactac atcacaaaga cttccactag
94801 gtgaaagatc ctaataaaat cttcgctgaa aacttgaatt ttttgcatag gctttataaa
94861 tataacttgt aaacctttat tgttgggaat aaaattcaaa gcttagtttt tctatttcgc
94921 agaaattttt atattttgatt tttaaaaatt ttatcataag gtggaaccta tattttttaa
94981 aataaagctt aaaacaaaat taaataacta aagaaataag acttgaagtt atgaaaatgt
95041 ccacaaaaga ctttaaaaat ttagacttct cgacctgtca tttttaacca attaaaagct
95101 tcgatataat ttgttggagc tagacgaatt gcttctttcc aataatctgc tgctttgtca
95161 aagagcattt ttgatatttc tgattgacca ttttcaattg cttgttcacc gcgataatgg
95221 taaataactg caatgttatt taatgcttgt ggtaaagacg gattacgctc aagagcttga
95281 taataatatt ctaatgctcg accatgttca ccattactag tatgaattaa accaatatta
95341 taaagaatat aacttctatc ataagcatcc gtttctagac gcatcgcttc ataatagttt
95401 tgtagtgcct cagcatattc tccttcagat tgagcagaca taccatctct ataataacta
95461 aaggcttgct tttcacgatt agatgtcggg aacactttta gcaaaatatac agcaacaaca
95521 gtaaaggttt tatcaataaa attatcattt cgctgtgatc gtggcatatt aatttttttt
95581 tatagtttta ttagagattc tcgattgtat taaataaatc atatttaaaa ttttaattatt
95641 gtctaaactt tattaagttt ctattaattt agacaatttc tttttttaat ttaaacctga
95701 aatacttaaa taattttatt ttcacgcaa cttttataca aaaataaaaa acagctgctt
95761 tcttttgtat ttattatacc tgtattttta aaaaaattaa aaaataaaaa atttacagta
95821 cttactgtat cgaagtgact caaattttgt tgttcaccaa attttttatt atggaaaaat
95881 cctttttacc gcaaggggca agtttttaac gttcgattat aaaattgatt gttttacaaa
95941 gtgaaacgag ttaatgaaag ttaaagaaga aaagattaaa tactttgact ggagagtctg
96001 aatctaagtg tgaagtgtat atcaagttta catcgaagtt caagttttgt tcgtcaaaca
96061 aaacctacgt taaaaataaa cctggtagct tgaaactaca gtagatagta agtttaacga
96121 gctttactca tatacttctg ttaaaaaaag tctaaaattt atatctttat taaaaaagg
96181 gaagcttatt caagtgttcc aagcgttaacc tgggaagcct cttgggtttt taaaactctc
96241 tgggaagaggg taaaaaccaa gaagcttcat ccgaagggta atacatcatg caaatcatag
96301 tcccgtaaac taggacctaa tctaaacaaa ctatgtttgt ttaggactct tggttttaa
96361 tttacattta acaaagcaca aactaattat gaaagttttc ttttttgggt accgggttac
96421 ggtttatact tcgtaaatac caagcttcgg tattaatcta aatctttttg accaggattt
96481 cgacctgat cattagaaag gaatccaaaa ataaataaag aaacaaaaaa actaacaaca
96541 gtataaacia aaatttttaa agttaaataa aatttaaaat tataaaaagtt ttttataggc
96601 tacccttatg tattattata ccataaaaaat ttttttttctg tataagttta atatcttcaa
96661 aacgaaattc ctcttagatt taaatataag aactaaaaaa gcttttgaca atgaaaccga
96721 acaattatag atctgctccc gtttcaggtc aaaataaaaa gccgtataac agtaaaataa
96781 aacttgtctc tctagttgtt taattttaag aacgtaattt caaggtttag cgaagcacia
96841 acccggtttt ttcacaaatt agaagttcaa agttaagaac caaggtttct taactaagta
96901 agccatgctt aaagtcttaa ataaaatttt attcaaaaaga ctaattttta ttagatagac
96961 taaagctttt tcttaccat taagtctatc tgaaggtcgg ggcactgctt aagggcaatt
97021 ttagacttaa gttgttgact tcatattatc ttattttata actaatttta tattaagac
97081 ttgaagcttc gcgtcaagtc caaagtgtgt gcttgcacac actaggcttt gcccgtaggg
97141 caaaccttta tttacaacg aagttgttaa atttagacct aatctaaaca aactatgttt
97201 gtttaggatc ttccctaaag ggaagtcttc tatgttttaa agtttgtgca gaacacaaac
97261 caactcttgt ctgaacattt tatgttcaaa tgggtatgag aacgaagttc gccaacctta
97321 cccgggagcc ctagattccc gagctacgct acttaatcaa gtaaagcttg attaagtagt
97381 ttaggcgtag cttttaaatt taggtatgtg agcaaagctc accaaccttc ctacctggg
97441 tagttaggc aaagcctagt taaaaaacia agtttgttaa ctttggaag taagtgacta

97501 tcaataaaat tagcacttgg gttcgtaaac cacctttgta aaaaatthtc ccgattggat
97561 acaaagtthta aaagcaaatt tatggcaaat aatthtataga tagctgatga taaatatctt
97621 gtggattacc aagctgttca attcgcactt tttgaaaaac aacaatthca ttagctactt
97681 ctaatgcttc ttgggtggca tgtgtaacaa atatagttgt cacaggaact tcttcgtgta
97741 gacgtcgtaa ccaatthctc aactctthtc gaactthtac atctaaagca ccaaatgggt
97801 catctaacaa taatacttht ggttcaatag caagagctcg tgctaaagca acgcgttgct
97861 tttggcctcc tgaaagctga cttgggtatc gatctgcaaa attctctaatt gaataagtt
97921 gtagaagttg tthtactcta ttattaatta tcgataaaga tgtthttcga acttgtaaac
97981 caaaagcaat attctcataa actgtcatat gcttaacaa agcataatth tgaaaaacaa
98041 aaccaatthtc tcgctcttga attgacaaaa atgtagaatc tctthccagct aaccaaattc
98101 thccttthgtt aggtctatct aaaccagcaa taatthcgtaa taatgtagat ttaccagatc
98161 cagatgggccc tacaagagca accaaagatc ctgthtttat thcaagattg acatgatcaa
98221 gggcttgaaa accatthaaa thttthgata gatctthcaac thaaatactc ataatthttt
98281 tatagtcaat aagctthtatac atthttgtatg gacaggtggt ctgattthatt cgaagagcaa
98341 aaattgattt ctacgaagta aaaaactthtc acatattgaa attcgcagat ttgactcttg
98401 cattgctaaa acaagcaaag cttgttcaag thtcacccta cggctgaagc ctcaaggatt
98461 catttcaaaa atgacacgtt atgttgatc thtctacgca gtgacactat aggtaaaaga
98521 tgaaatattt agtctthtacc cthttggacat tgctccttag ccatggctaa ggacctacac
98581 thctacgttag taggaggagt aaaggtcaaa gggcaaccgg tathctatth gcttatagga
98641 aatctthaat cthgtctactc thggtagctt aggcaaagta atggccgaaa aagggthttat
98701 ctgtagaaca aagthttata aaacaaact taagcttaga tataagcctt atataaatat
98761 aatathttac atthtaagccc ctgagttcgt tatatagaaa atctgtggct thaatthttga
98821 aataataaaa atthaacgagc thtathttgat thtagaaattg thcgataaac aaththtttg
98881 thttthtttat thaatthaaatt thtcaaaaaa thtgthtttt gcgaactaat thtagtattg
98941 tgtaaagctc gccattgaaat tactthtattt gtcgataata aatagatgt atthataaaa
99001 thtcaaaaag thgaattcgt cthttcaaaa gaataccaat thgggtthttg gththaaagg
99061 thcaaattctg atccaccata atthggataat thtctthtgag thgaaacgaa thctthtatt
99121 tgtaaacaaa aaaacatatt thgaaacgat gthtaatatt gtcgatacga atgctcatta
99181 thgtaaaaaa actthtgthtc atctthgatc aaththgtct caagagccac aaththcaggc
99241 tgatthtaag atgtthtctt thttataatt thttcgggtc cacttagctg thaatthgtgc
99301 caattactcc atctthctacg aaththttct thcattaataa aaththagggg ctgctgaaat
99361 ccatgtggta tactthcaaa acgtthtaaat ggagcatctt thataaagat aatagthttc
99421 gataaagctt gthtaagaaa agaacgtggg acaataaggc caagtaatcc atgggtgaagt
99481 aatathttctg cggthttgaaa gtcacacagg aathgttctt gthaaagthttg thcaataact
99541 cgtctactctg caaatccgat aagtgcttht ggthcagcaa atatgagatc acctaacata
99601 gcaaaaactg ctgtaacacc gccagthgtt ggagaagtht atacggaaat atacaataaa
99661 thtagcacatg atthgataaac thgtaatgca gcagaaatth thgcccattg cattaaactc
99721 aaaatacctt cctgcatccg agccccact gatgcacaaa thaaaataag agthaaacct
99781 thcttgagtag catathcaat thagcaggtt atthttthcac thactaccga acctatactt
99841 ccaccataaa aaththaaatc cattacacct aaagcaattg gaathccatc aagcatacct
99901 gtaccagtht gaacagcatc thgaagacca gthctthctt gtgctthctt thaatcgttca
99961 gtataagctt thtgatctct aaththcgagc ggatcacaag gagaaagagt thcatctaaa
100021 ggtcgcctatg thccagaatc aathtagatgt thcaattcgtt cthtgctatt catttgcaaa
100081 tgataaccac aacaaaaaca tactcgttga thttctthca aatgthtaat atataaaatt
100141 actccacaat gatcacaacg agtccagagc cthtggtctac catctgatgt aggatgatta
100201 ththttggag cattaagaag ththtaattht cgtthgatct caatccaagc thaaatagac
100261 ataaaaaaa atctccatat ththatttht thataaaaatt aaaagthcaa ththtataag
100321 thttataaaa aaaagcggtc tgaacgcagc gcatgctagc ththaggtct ththttthgg
100381 thtacccttc thtgtgcatt thttgtcgac cthtagctac cataaactth acctaaagag
100441 tggcatttht ththttgttc agcगतgcac tggctataga gggaaagat aaacctgta
100501 gacaaaactg ggctaaacga taccaaaagt thctagattc agggthaatca agcgaaactt
100561 gataaacctt gcccgthtgc thaatthaaa aacgaagtht thtaataaag ththgtctcat
100621 agagcaaagc thctthgggtat agatathttat gaataaatat thaaaccaag aggcgaaggc
100681 thaatthtaac aacgaagtht thaaataaag ththgtctct ccagthaaatc thaccggaca
100741 thcagtatcaa gtagcgtagc thggthttthaa caagcgcagg thgtthaaac ththattthgt
100801 gggcaacaat atthtagaac thtaggaagaa gactthctta aaggtththaa gcgaagcgtat
100861 aagcctacgg gcaaacctat gthgtgtgcaa gcacacactt thgactthca cgtagthgaa

100921 ggcattgggta aacaagcgaa gcttatcaac cttactacct cggtagttt agggctccca
100981 ggtggagggtg agtctattaa tattctagtg aaactcgcct agtagtattt atggctcttt
101041 gggcattgggt ttacaagaat ggaaaagaac ttttaaaact caagaggtaa agattaactt
101101 cactatggtg gaccattgct aaaacaagca aagcttattc aagggtcacc ctatgggtga
101161 agccttacgg tctaagtaaa aacttattgt taagggttac taaggaattht aggcctttgg
101221 tttttactaa gtctaaactg aaagtttact ctaactattht tttgtatcag aaaacattta
101281 aaaaaacaaa acttaaaaga taaactaagt ccagttgcct attagttttat gacgtttttaa
101341 aaaattttctc taaaagatac ctttttaaca acatgtagtt ttggtcacia ataaaaaaaa
101401 tatcacacag ttttttagag aaagccataa aaaaacttta aaagatattt ttttaggtca
101461 gaactcatat ttttttattt ttcaggaaaa attttaggggt ttgttttata aaaagctaac
101521 tctaataaat tttataaagc ttattcataa taagttttaa cttaaaagta aagtttttac
101581 cttaatatat aatcttaact taaagattta ggtagttcaa attttttcta aacaatgtht
101641 tagtcaatta aatttttaaa aaagttttta aagacttgaa gggagccttc aagactaaa
101701 tgttttatca aagttttccc caagtaaaaa cataaagttt ttacgcagaa gactttaacc
101761 cttgccgtgc ccggcagggc aacgggcgaa gcttgaacat cctagtgtat gcaagcacac
101821 acgtctgtht tgccttatcg tttaaactta gggttattaa acccgggagc cctaaactac
101881 cggaggtagt aaggttgaca agcttcgctt gtttaccat gccttccact acgtggaagt
101941 cttcttctca gattcccggg ctatgctaaa tgatagttaa attataataa agaaagaggt
102001 aaaaaaatg actatagaaa agaaagttaa tccagaaaat gaagtaaac gcacgagcaa
102061 tgtggtagt ttactagaga ttgaaaaaat taaaagtatt gacatattta ttgaaaatcg
102121 tcttataatt gatccatcaa gcaaggacac cttgaaaaat atgtatactg actacgttht
102181 aagttgtgat tcttgtgttc cgctaggaaa acaaaaatat cgtaaacaac ttgaaaaggc
102241 ctttcaagat agaaatgcca aagtgtactt tgtgtctaac acgcgcgggg tacgacttga
102301 aggtgttaca ttaaaaaaag agggagtaag aggtacaaca acatcaagtt ttgactaaat
102361 atttgatcaa agatcaaaca taaagttttt cgctttgccc cgggagcctt aggttcccgg
102421 gctacgataa agcccacggg ggaagcctcc gggtagtgga gcaaagctca ccaatctgtc
102481 tacctttggc cattgtctcat tagtccaaag tttgtataaa acacaaaacta ggatttgccc
102541 gtagggcaaa actttactta acaacgaagt tghtaagttt agacctaatc taacaaaacg
102601 tagcgaagcg cagaaaaaaa gtcttctgcy tttgtttagg atcttccaca aagtggaggt
102661 ctttacctaa atgttttatc aaagataaaa cataaagttt tacgcttccg gtaaaagcctc
102721 tggctattgc tcattagcac agctaagtga taaagaccaa agggcaatgg gtaaggcgta
102781 gtttaggtaa agcttagtta accaacaata tttgttaact ttaaaaaaca ttccgtttta
102841 gagccaatat tctaatatca gtatgattaa aagatttttag ttaaaacaata aaattaattht
102901 tagattatat tcttaatttg aaattaattht gtaaaatttg ttaaaattat tatataaatc
102961 taaaattaaa actaagcctc ttgcccggatt tgaaccggcg actttttgct taccatgcaa
103021 atactctacc tactgagtta aaaaggcttc tattttattht agtctatcat tttcaaaaga
103081 acctgtctac taaagggcag tagaaaaact aatthttttt ctgtgtgccc gaagcctcgc
103141 gcaagaactt ggcgcataat gaagtttttag aaatggtaaa ttccagattt tatcgaagct
103201 actthttttta agcatggtht aagtttgaag ttttactttg aaagtaatag ctatttataa
103261 caaataaaaa aagtttccgt tgtaaaaata tcaacgtht tttataattht aatagtht
103321 gtagagtaaa gaaataaaca ataaataaag gactaattgc atatggaaaa atthttacct
103381 actcgaagtc aacgattaaa aacaaccaat attatthttt atagthtca attagaagtt
103441 caattactaa agcaagatcg agaattttat aaccacaata aatthttctc acaaatacaa
103501 aatcatata tctgtaatag aagattagag ctaataggct atctagcttht atattcagta
103561 aattctacta ttttttattt ttttccgcta tggthtttca tghttcttht ttatttattt
103621 tcaaattttt aaatttggag atcagcttgg ttgaaaatta gthgttccat atccgagtht
103681 aaactttgga cgaaaaattc gttgaccttht actttaaaag gagctgcaa tcttcaagtt
103741 gtaattagtht tttttccatt tatttttggc atthtttata ttatttattc tcgthtcaa
103801 aaccaagtht tttcttactt ttttcataaa aatthaccag gaatctacct ttcttacc
103861 aaacttacct gggaaacct tcaacatatt acatcctcgc aatthaaact agatcctacc
103921 caaaagactt gthttctaga ctacgagatt tcaaaataac atthttctcc aagacaatcc
103981 ttaactaatc aaatcgaaat actgaaagat tttaaatcag tgaaatatca attactgac
104041 tttgatgaag gthgtgctca agaagaggat tttcgttcaa agcaaatga taaatctcag
104101 aaaaattata ctaaatcgta ttttcaatta actaaggaag ctcgaggcca gaagcactca
104161 gthttttatta acaactaga tttttatagtht gattctthtt taataaaact aaattcaag
104221 tgggcaata aaaaaacca aggctatgtht ggaatthttc cgaaaaaatc ttttttagga
104281 tctgtttcag ttagtcccaa tttttttaa aagcctaaga tctccttgg gthcttggc

104341 ttaccgcaag tcaacgagca aagcactgag gctacgccc ataataatth ctcaattggt
104401 actacttggc atttagtgtc accaaaagat ttccttactt tttcttctga tgggtctttt
104461 caaacttggga atcaagaaaa aaatgttggt aaaatagaaa atccgatcct cacccaacgg
104521 acaaatttga aatttaatth aaatacaaat tcaagcaaaa aagcaaatth ttttaataat
104581 cagcctccat ctaaaattat gaaaaattat tggcaaatth tttatttcaa tttagatgaa
104641 ttgccaaaaa aactaaatth aataaccaca aacatctcag ctttgaatth aacagggagt
104701 aattttaaata aaaaaactcc actttttacca ataaaaaaat attccgagga acgtcaaaaa
104761 tttttaacaa atactttggc caacacggac tcttttcccg ttgccctgcc gggcacggca
104821 agggctcaag cggaatataa atttaaccaa aaatctaaaa aagttctatt tagtaaactt
104881 aattttttgg gaaataaaat tgatccaca ataaatthaa cctcacaaaa taaattatca
104941 agtcgaatac tacacggttt ttcgaatcaa tatcaagccc aatctagagt tttagataac
105001 cgagctaata cgattaacaa agttggttta aagtcagaat caatthtctt taataaatta
105061 catttactcc aaaatgaatt aagacagttt tttataata ctcaatthaa accaaactth
105121 ttactattag aacctactt aaatccttct agttctcaaa gtctgtthaa cctagtaatt
105181 ggtaaagatt ttaatttctga tgaactaag atgcctaaaa attcagactt tagtaaaatt
105241 ttggaacaat ccaataatga aatattattg acaaacctc attthaaagca acttactagt
105301 tcaagtaaga gtgaggatga atatttagaa caaatattac aaatthtaca ggatttctgt
105361 tcattcaagg caaatthggc aaccccacca cgtgcaatgt caggatacga ctttccagat
105421 agtactaata cagaagtccg atcactgatt ttacaattht tttatcaaaa gaaattacca
105481 ataattaaaa gcttaataac gacaagtcaa atgaaaaaca aatctaatth gttttacca
105541 tcttttaaaag tagaattatc tccaagttth acttcttatt ctaattataa ttaccctcag
105601 tctaaagttc caactttgaa attaaactat cgacctactt ttttagaggg tgtgaaaaaa
105661 actctctatc aaggaccggg ggttttacgt gagactthta cacatgattt cagtctaaga
105721 aacaaagaag aaattcgtag atggattaag ccattthtatt cgtccgataa cccgcttact
105781 gatcgacgca aagctthtth cggccataaa ttaatgcaat tcgatgttcc tgcctttgct
105841 cctactaaag gtctcactcc gaccgttgcc ttaccgggca cggcaagggg ggatcaaggt
105901 tatcaagcga agcttgatta ccaagggcac gatgggttht ctttaaaaaa agactccgat
105961 ttaagccttg cccctthtct gaacacggta aggatagaat tagaaagcat taaaaaaat
106021 atatcgccaa ataaattht aaagctacaa aaagatattt taaagctaaa aaacgttcaa
106081 acacactthc acttgatcc ttaagtccaag tcgataaaat tacctgtagt tgcctthtact
106141 gaagactthc aagtacccta ttgactaaa tcgaaatggc atacaattht agaaaaactg
106201 agagctcaga tagcaaaaca attggaagat aaatcagaag aagaaaaatc agaattagaa
106261 aaaaaacttc agataaatgt tccagtggtt cggattcгаа atcctaaaca acaggcaatt
106321 caatggcctt taaataaact cgattatгаа aatthaaatg atthtgttht tagttctccc
106381 gttagtthct tcaaacacga caactggacc acggccaaag acctcactcc ttcggacca
106441 ggttatcaag cgaagcttga ttaccaata aactcaataa gtaataaact aggagattcc
106501 caacaaaaga atthattgga aaaatatttg aacacaaggt attcaataa attagcagct
106561 ttaacaaact cttcaattct atthgaagat atcatttatg agaaactthaa ccaaaaagaa
106621 aatactthtct taggtthtaa tcgacgagta aaaccagtha ctgaaaaaaa taacttaatt
106681 agaattaat accattatgt tccttcagtt caaatggat taagtgaaaa tgctactcaa
106741 aagaaaatth ttggtaatgt ctataaaaa gctthtaccag tttatcaaaa cattatagaa
106801 aataggatag aatctthttaa tagctataaa cgaaaaaatt acgttaaaaa cactthtcta
106861 aataatthta gcaatgtatt ttaccaatct tgggagccta ttacatcagg gtcttggatg
106921 atgatcacac aattatctth tggtttagtt gtgttcaaaa tattccagga tgtttatcgt
106981 aaatattggt aagaattgat ttcttacgta ttgatcttc tagccgcatt aggtatcatg
107041 gacccgagct taaaagaaga acttgggatt gatgatagac gtacaggatt tcgaataata
107101 agaaaagtat ctaaactgatt tcgagatatt gcaggtattg atggaattht accagaatta
107161 gccgaaattg tctggttcct tagaaattht ggacggtcgt ttaaagttgg aatattatth
107221 ccaaaaggaa tctthttagt tgggtcacct ggtactggaa aaactctact agtacaagca
107281 attgcggtg aagcggaggt acctgtthta gttcaatcag gaagttcatt aatgatcct
107341 gaacaagaag aaggtggagc acaacaacta acaaatthtatt ttgaacaagc acgtcaaatg
107401 gctccatgta ttgtattht agatgaaat gatgctcttg gtgaaaaacg agaaaatgth
107461 attcaaaatc caatgggaac agatgaagtc attgaagcaa tacaagaatc tagtgagtct
107521 aacccggctt ctcttcccag caatacaaaa ggtctgcccgg attthaacgag tactccatth
107581 atacctaac caaaaataaa gtttaataaa acccaatthata aactthtaga tggagacca
107641 caaatgtat tagatthtct cgttcaagaa gaaggtgcgt ttagagatth tgagtcatat
107701 gagaattcaa gtcttggat tattcaacaa agtattgata aacaagaagc aaaaacaagaa

107761 caattaagac ttttaatgca atttttaatt gaaatggatg gtcttcgagc tcgaaaaggg
107821 gttattgtta ttgggtgcaac aaatagacca gatgttcttg atttagcttt aacgcgacct
107881 ggaagatttg atcaaattct tcaccttggg ttacctgaaa aacaaaaacg aatagaaatt
107941 ttaaaacttt atagtaaaaa tctaggaact gaaaatacta tttcctggga ttacttagca
108001 aaccgaacta ttggtttttag tgcagcagat ttagcggctg taatgaatga atcatcaatg
108061 aaagcaatth taagcgaaac agttcatact attcaaacta ttgaacaggg tattgaatca
108121 attacaagtt atagcaccga caaacctaaa ctcaaattc caaattctat tgatccattt
108181 ctgattagta gatttagctta ttatcaggct gggaaagctg tagttcatac attattatta
108241 caacatcctt cagtaactgt actacactta tggccacggc aaaaaaatgc aagacaaact
108301 tatattagta ccattattca aaataaattt ctaaaaatta gtcgaaaatt agaattagaa
108361 tctcggattg ttggattata tgctggaaaa gctgcagagt ttttagtttt atcaagtaat
108421 ttgagttcta agtcagaaat accaaaaatcc ttttagttcaa ataagccatg gcaatctgat
108481 atcgggtatt aagacttaaa ttttgcaact tctttagtac aatcaatgat tagtcaatgg
108541 tttttttatt caaaagattt agttgttaga aagtctaata ctattttttc tcaacgaaat
108601 ttacaagaaa ttcaagagtt agatactctt gaattattta atcaattagc aattgagact
108661 gaaactgaaa ttatccaaaa aactcgttca tctggattca atcgcgattt tcaaaaatgg
108721 tcaatacgtc catggtggca aacacagatt acacgacaaa tcgggttttt agatccagct
108781 tacgatgatt ggtatagaat ttatcttctt gatcctgagg aaagcgaacg gaatgatgaa
108841 tggataccac cagacgagtt ttatcataat aatagtaatt taacagattt agctgttaat
108901 tcagtttaatt cgtctgttaa ttggaatgat ctatatgcta tggatcgtga ttatatttat
108961 catggttttac ttttaagctg cttaataact gcatttagta ttttagatca aaatagagaa
109021 ctttttagatt attttgctgc atatttaatg cgcaacgata ttcttagaca acatgaaatt
109081 aacgaaatth tatacaaatt taaacccgaa tctagcttat ctaaaaacac tgaacttatt
109141 cataaaaaata aaatthtgaa tgaaaacaag acgaatgaac aaaaattatt tctacaaaaa
109201 aataaaaaaaa caagagtgat tgaaaaagca tgggggtcaat attctagaag gacaatattt
109261 cgtttttttta attttgattc tatattatct aagttgtaat ctagtattaa aaaccttaat
109321 ttaaaccggg gagccctaaa ctaccggagg tagtaagggt gacaagcttc gcttgtttac
109381 ccatgccttc cactacgtgg aagaccttaa tacaatctcc gattgtactt tgcccttgcc
109441 gtgcccggca gggcaacggg ggaagaacac aagtctaaaa tacaatcaga gattgtatta
109501 aagtcattcca catagtggaa gtccaaagtg tgtgcttgca tacactaggc tttgcccgta
109561 gggcaaacct ttatttaaca acgaagtgtt taaatttaga cctaactctaa acaaacatg
109621 tttgtttagg atcttccctt tagggaagtc ttcttcttag gttaccgggc tacgttcttt
109681 taaaggthtt acaaaccttg cgcagcgcag cgtagcacag cgtagcacag aaaaaagtc
109741 caaaatagaa cagagttcta ttaggatctt ccactttgtg gaagtccaaa atgtgtgctt
109801 gcacacatta ggatctthaa gcttcgcccg ttgcccgtcc gggcacggca agggthaaag
109861 tcttctgctt ttataaaact tgattthaaa gcgtagcttt taaatttctga aagaattthaa
109921 aaaaattctc cttttaccce tgagctcctc gcctaactaa gcttgacttc thaaacattc
109981 tacaacatac tatgtggcgt tttttgaata gtatcataac cctattttca aggtttatgc
110041 ttcggacaaa ccaggatctt ctactacgta gaagtcttht tccattgtht tghtgtthta
110101 aagcaaggta aagaccaaag caaagatttg acaaagcaaa ccttgactth taaaagtcaa
110161 gatttgccca taggcttgaa aaactaacct tagcttgtht aaagthtatc ttttctacca
110221 aagaaaaggt ttaggtaaaag ccttacgacc atataagggc caaacaaaac thtattthca
110281 ataataaggt gctaaaaata agthtttatat aaattatact agaaaaacta aagatcttat
110341 aaaaatagtt actthatttc atactthttt aataaaaaat ttagaaattt tactaatatt
110401 ttcattthtac gtataatatt agaaaatgcc ctgggtggtg aatggttagac acgagtgatt
110461 caaaatcagt tgcgtaatgc gtggaggthc aagtcctctt cagggcaatt tcttaataaa
110521 ctattthttat taagaaaaaa gatataaata taattthtaga gacgthttth acctthttta
110581 aataaaaaata caaaatttht gttagact caaaatttht tgctataaat cctaatttat
110641 taaaaataat tactaaaaat cthttcgata gthgtctata agctthttaac cggatatcta
110701 gatttctcaa gthgtthtagt aaagagcaat gtaacatcat atctctaagc gcatcaacta
110761 gtcttatgth taaaaaaata ataatataa gcgaatgcaa aataattthaa gtatcgtcat
110821 tatagtthta ataaactaga ctgagtcggt tattatttht tattagattc thtatattga
110881 agaaaacagt tgaaataaat thtgatctta tatttcttga aaagaaaaag aaatagthta
110941 ataataacct aatgctthatt atthtatctaa taatggaagg ccataagcct tagtattatt
111001 aggagtgatt gccaaaaact agaatacagc aaatataagt thgtctthtt thaaagattt
111061 agtaaatatc thattthttct taatgatcgc thctgcttht accctacat gthgtattth
111121 ttgctagact thtaaaaggt caacaagtht tctctaataa gthtgthaa gthgacaaaa

111181	ttagtttaga	aatatTTTTct	tatatatcaa	gttgggctta	aacttaaatt	ataatTTtata
111241	aaaaatgata	atTTTgaaata	cgatcgcttt	agtagaaaag	ttaaaaatac	ggatTTTcaat
111301	aactcaaatt	ttgggctTTtg	cattgtataa	actactacta	tcccgcactt	aatgCGggggc
111361	aaggTccgaa	ggTtaaAacc	aacagTtaaa	ggTgaatgag	cttCGgtcat	ataccCatag
111421	gtTaaaactt	gatgTTTTat	ttTggaccta	tactatcttc	cctatttatat	Tgcgtcgacg
111481	caatataata	gggaagatag	tataggTcct	ctagcaacgg	gcaaagctcc	aacaatgTtg
111541	ttcgTaaaac	aagattTtaga	TTTTTgtTTa	aattaaataa	aactgagatg	aattaaataa
111601	aaaactatag	Tgaagaacct	aaaccaacat	aagagccata	aaaaaatata	ctTaaTaaag
111661	caagtGctgc	agtacctaca	actgtaccca	caagccataa	agggatacgt	ccagtagTtc
111721	cagtgtTtaga	cattgatatt	tttCctaaag	TatcaaAcgt	Ttaataattg	gagcaataaa
111781	attataaaaat	tattTggatt	aaaagctTTg	ctTaaacgat	taatctcctt	gtTaccCaag
111841	atacaaggTt	gataaaactc	ttTgaactaa	ctTTTTtatg	TgttattTTt	Taaaacacac
111901	gtcgtgTTta	aacaacggga	ggctTTcCcc	gaaggggaaa	actTTattat	aaaaaaaatt
111961	tataatTTtag	Taataaataa	Tcaaaaaaga	Tcgggcctaa	aaatagTatc	tagatattTta
112021	ctaactTTtcc	ctTTTTgtTT	ttTgtcgtcg	aacactgtTc	agactgtaga	caatgagTat
112081	aatctaaTTa	aatgtctaaa	cctctTgctt	TacaagcgtT	gctTgTaaaa	ctTTattaca
112141	aactTTgtTT	gtaattTtagg	TatgtgagTt	agTaatTTat	aaataagTtg	ctTattcatc
112201	aactagtGag	aaacgactaa	taggtgaaag	cctataaaatt	Taacccaagc	TTTTtctaag
112261	ctgtattTTct	aataaaTTTT	Tgaaaaaaat	atatcaaaac	TTTTtactag	acaaatcagT
112321	Tcaatgatta	TaaataaagT	Tcttagatga	tagTtctaca	atTTTTttac	gtTaatTatg
112381	atatctTTtcc	Tctctacaat	Taaatataaa	TattTgtgtt	aaatgtacaa	agTtaccaaa
112441	cacataaaaa	agtctacaca	gtTTgtTTtT	TgtTTggTaa	caaagTTTgt	agattagaaa
112501	ctatactTaa	aaattaacac	ctaataaaatt	TTTTTaaatg	ggatTTTcaa	caaattTatt
112561	aggtgtTaat	TTTTaagtag	aacaaatctt	actaaacatgt	actTaatTaa	aagatacaaaa
112621	TaattTgtat	ctTTTaaTTa	aaaatatagc	Tggaaaacag	Tacagctaat	acgaaaatta
112681	ataatagTcc	ccaatataga	ctagTaccggT	Ttaattcaac	agattgtTTa	TtcggattTg
112741	gtTTtagccat	atTTctTTtcc	Tatgaaaact	Ttctaacgct	gaataaaattg	cattgtcTgta
112801	atagaaccta	gaaaaaatac	agTtggTaca	gctaaagcat	gaacagcaag	ccagcgaaca
112861	gTaaaaaatag	gataagTata	agatgattTT	ttTgtagcca	TgtTctattg	TgacctctTt
112921	aattTactTt	gcaatcgctt	Tattacaaac	TaagTTtgta	ataaagTgtt	acctctTggg
112981	Tatgcgagcg	aagctcgcca	actTTctat	Tagaaaggtt	aagcctTact	ctagagcaag
113041	caaagTTtat	TcaaattTTt	TgtctaaGtc	Taatcaaaata	TtgcatTTtT	attactTTtT
113101	TtgtTctTTt	Ttagaagaaa	cgtaaaatag	TaggTaaCa	aaattTaaatt	Taggtctaaa
113161	TTTTaaaGct	aagctTTTaaa	TtagTttTtTg	actTTTcaaa	TtctaaagTt	Tatgtctaaa
113221	atacaatcag	aaattatatt	aggatctTga	aacaaagTtT	caagtctTTa	ctcagcctaa
113281	atTTTtaatt	TccattTtagc	aaagacacaa	attataaaat	Taaaataaaa	TTTTTTtctc
113341	gaaggcctaa	ggcctaaact	aacaactTTa	gtgatagTaa	cgaacaaagT	TctTaaGtcg
113401	TgtctTTtagc	cgaagcgtaa	aactTTatgt	TTTaccgTtg	ccctgcctgc	aacggcaagg
113461	gtctaaGata	aaacattTtag	gaactTacca	actTaaaaaa	aactTgaatt	aaccatgcta
113521	atTTTaaGgc	aaggtTaaCa	aatTTTaatc	gtTtaccCaag	TattacaagT	aaatccaaac
113581	TTTaaGctct	aaattTTTtaa	atTTTgattc	gtcctacgga	cgaagctTga	Tgtaaactaa
113641	aagTTTtacac	Ttagaaaattt	gtgcaccata	gcctcaccCC	Tcacctggga	gccctaaact
113701	accggaggTta	gTaaGgtTga	TaaGctTcgc	TtgtTTaccC	atgcctTcca	ctacgtggaa
113761	gtccaaaatg	TgtgctTgca	cacactaggc	TTTgcccgtT	gggcaaacct	TtattTaaCa
113821	acgaagTtgt	TaaattTtaga	cctaactctaa	acaaactatg	TTTgtTTtagg	atctTccctt
113881	tagggaaGtc	TtctTcctag	gtTcctaaat	attgtTgcca	acggcaacaa	TataaagTgt
113941	Taacaagcgt	agctTgtTaa	agcaggGcta	cgcccctTtg	agacaaggtt	atcaagctgc
114001	gctTgattac	ctaaaaaaa	ggTctattgG	ggTtaaagTg	caagatTTtT	TtattTTtTt
114061	TctattTcaa	ataaagatac	aaattgacagT	attattccgc	TaactgctTc	actTgtTcta
114121	atgcgtTgaa	acgatcagTta	ataagTggag	TTTctTgacg	atctTcagTta	aaatattcgt
114181	taggtcgagg	actaccaaatt	acgtcatatg	caagaccagT	actTacaaaa	agccatcctg
114241	caataaataa	agaaggaata	gTaaactgtT	gaattaccca	gtaccgaata	ctggTcagga
114301	Tatcagaaaa	cgggcgtTtct	cctgtagctc	cagacattat	aaaactcctt	aaatctTatt
114361	TTTattTTtT	aaataattca	actcatcaat	cagaaatatc	gctTtctaata	aagTttTtTgT
114421	tagTtataag	TTTaggTtat	aataaaggtT	Tccgtacaaa	TTTTaataTg	TgaatattTT
114481	TactagattT	TattTaaaaa	ctaaaaaaac	agagctTtaa	aaatgtggat	TTtagactTg
114541	agattaaaaG	Tctcgaccaa	atgtctTTtT	TtattaaTatc	TaacattTTc	TtactTTacc

114601	aaagttttatt	agaacttttgg	agaaatttcat	tttgttttat	aagctttttta	gaaactacta
114661	gactagaatg	aactttcatta	aactttttgaa	gcatcgaatt	atztatgaat	ttataactaat
114721	taattttttta	tttaaaaaat	taatagttta	aaaaaactct	ctttgtttta	tctattgaaa
114781	ttctctctac	taaaaaactg	agtttctaaa	ttacaaacaa	agtttataat	taagcttcac
114841	ccttaaaatg	aaacatctgt	aaagtcttat	ttattaatcc	ttatgaaaat	aaatgaattc
114901	tattttaaatt	tgaaaataat	tttaacattt	gaacataaaa	tgttcaaaca	agagaaagat
114961	gatgattttta	aaattttccta	tcaaaaagta	aataaaaattt	atttttattg	gtttcataaa
115021	gtacaatcat	tacaattgaa	acttattttt	aagtgaaaaa	cttgtagaat	ttagagaaaa
115081	cgaaattttct	taaattttaat	agcttactca	ccaattattt	caaaaaataa	attggaatac
115141	cattaaaacg	agtgagataa	tttttgttta	gacaaatgta	gtataacctt	gtctagaat
115201	agaatcaagc	tttacaaaga	gtgtccactg	ataatttcat	tttaaaacct	gtaattatac
115261	tattttgtta	agatttcatat	atttaactgt	tgaattttac	ctcaaattaa	gttaaagct
115321	taccttcttt	taaagacgag	ttctagaag	tgtgagatta	tttcttagtg	gctctgataa
115381	taaacgaagt	atttagtact	tattaaactt	cgtttaatga	gtggaagatc	ctaatagaac
115441	tgtcaaatct	attctaaact	tgttttacgt	attttttaat	tcttttataa	cgtttaattta
115501	ggctattttta	gttgaaatct	gatatacaag	ttttcctaag	catagctact	ataccgtagc
115561	ccgggaatct	agggctcccg	ggttttttta	aacgatagtt	tttaaatctc	gtagaattac
115621	cttttaata	aaatctaatt	agttaatttg	ggcgaacgac	gggactcgaa	cccgcgaatg
115681	atggagtcac	aatccattgc	cttaccactt	ggctacgccc	gccatatctt	ttaatataac
115741	atttttataa	caaactcgtc	tagtcttcgg	aacatctttg	aaatgtctaa	gatccatttt
115801	aaaaataagc	cattaaacgt	agtaataata	aactatttga	tgcttacatt	agaaataggt
115861	aggctaatta	atgagtattt	tgtatataat	ataaaaatcct	ttaaatttag	cggatgacgc
115921	gactcgaacg	cgcggcattg	gacttggaag	gaccacgctc	taccaactga	gctacatccg
115981	catatttata	attatatcat	atcgatttac	tactggcaat	atgtaaaaaa	tactttttatt
116041	taatataaat	ttagcttggt	attttcgagt	gaagcctaaa	gtgtacggag	tcacatttaa
116101	aaattgagta	aaaagtaata	tatagctctc	aaataaataa	cataggctta	tcacttttta
116161	tgtttttttc	gcttgagagt	ttcaatgaat	tttttagtaa	aatttcacaa	aagttcttta
116221	ccgtttaaga	acagttcatt	ctaattgtaa	gatttatttc	taaaaaataa	taataagaag
116281	tatgtacttc	tttctaaatt	aaaacaaaac	gcacggggta	tataagattc	agtcgaaaac
116341	ctaattataa	gctatgcttt	tagttcaggg	tctacttagt	ttataccaat	tcttttaaac
116401	aaaatttttaa	attttgtaaa	aataaaaaag	cggtgttcta	gaattatgtc	gtataaaaaa
116461	tttgttttta	cgtagatttt	tgtgccttga	gctattctcg	tgaaaaataa	taaatcattt
116521	ttttcaatag	aataaattaa	agtaaaaacaa	ctaattcttt	ttgtccttat	attatagcct
116581	ttggtttttg	agattttttt	ccattatacc	ttatcctgtc	ttttgtgttg	tcgtaagggt
116641	ttacgttttg	cgtaaaaactt	tatgttttat	cgttgccttg	ccgccgttgt	ctaaatttaa
116701	caacttcggt	gttaaatcaa	gttttaccca	aagggtaaag	atgaaaatgt	gaaaatgcaa
116761	tgttatccaa	aatacaaaaa	agtcttagaa	taatctctct	gaatctagta	agatctccca
116821	cttaattttg	tatttggtag	ccagggctct	aaaactttat	acgaagtctt	cttcttagat
116881	ctaaatttaa	aagcgtagct	tttaaatcaa	gttttacaac	ctttatttgt	aaaatctcgg
116941	gttatgactt	aaattttagta	aaaatcaggt	tacggttaat	tttagaaata	agattgaagt
117001	tataggcatt	atttaaatta	atagtaacaa	attctataaa	ttgtgtataa	tactataaag
117061	aattatatgt	tcttattttta	aaaattaatc	attattttac	ttttattttt	ggtatatatt
117121	aaaaataaaa	ataaaaagaca	gtaaaataaga	aaaataaata	atgcagattt	ttttactttt
117181	aataaaaaaac	taatacaatc	acgcttttaa	actctaaaaa	tttgttattt	tcaatttttag
117241	cttttttagt	tgattgggct	tagaaatctg	tagtgttttc	attctaaactt	ttattataga
117301	caaacatagc	tccagcgccta	actcagagtt	tacatgtaga	tatgtaagcg	aagctcaaca
117361	accttactac	cttaatttagt	ttagggatct	aggtctgtca	gatttgaatc	gagcgattat
117421	cttttttggt	tcttctgcac	gagaaccaat	tttaataatc	tataaataat	atttttttat
117481	ataaaaattc	ttttggagat	aaaatttatg	actctcattt	ttcaactaac	tctatttgca
117541	ttaattgggtg	tatcattttt	attagttggt	ggagttccag	ttgcttttgc	ttcacctgac
117601	ggttggacac	aaaataaaaag	tagcattttt	tccgggtcttg	gtttatgggt	tttcttaggt
117661	ttcgcagtag	gggtttttaa	ttcttttggt	atttaaaaaa	agaattttta	gagcttaggt
117721	taaactactaa	aaaactactt	ttgatttcat	tgcatgcccc	attgctaaat	gttttctgaa
117781	agataaaaca	tcaagcatca	cccaatcccc	ggctagacct	tcacccttgc	taccaaagat
117841	agcttagtct	atacctctta	gtaaagacca	agaggcttca	ctttaactaa	cgaagggaat
117901	aatgggtgaaa	cttaaatcct	taccgggtgtc	ctgtaagctt	tacctaaacg	ttttatctta
117961	gataaaaacat	aaagttttac	gcttggcgta	aagcctacgg	gtaaaaactc	attcaaaacc

118021 tctgcccctt gccattgcaa gggcaatgga agatggaaga tectggtttg tgctttgcac
118081 aaagcttgaa actttgtttt tgaatattt ctttaccctt tggtaagcaa gcgaagcttg
118141 ccaaccttac gcagaagact ttaacccttg ccgtgcccg gagggaacg ggcgaagctt
118201 aaagatccta atgtgtgcaa gcacacattt tggacttcca caaagtggaa gatcctaata
118261 gaactttggt ctattttgga cttttttttt gcgctacgct acccaaggta gtttaggcta
118321 ttgcttatta gcaatagcta atgagtaaag accaagggga caacggcaag ctacgcttgt
118381 tttacggaaa agttgtcaag cttcagttga ctaccgaagg gtgggccttc tgacaccgta
118441 acagataatg caaagaaatg tacaatttaa attaaatttt tatcaaatta aatataaaag
118501 gcaattttaa cactacattt tccctattga cttcaaaatc ttatttattt agctggtacg
118561 cctctaacat ggtaaacata aaatgcttgt atagttttaa gttacaattc actaaagtgt
118621 gaagtagttt tagacaaatt ttctcaaata aacaggactt gttttagcct tagtactata
118681 agtcatctaa gtttaacttt taaaattttt tttacatttg aaaaataaaa ttttaaatga
118741 atatttttcc ctagtttggt aataatggaa aaattggata tctagtata gagcctatat
118801 atttaaatat tctaaagttg acagaaaaac ttttaattaat gttatactta atatatatta
118861 tatattaatt cgatggggtg tcgccaagtg gtaaggcagc ggtttttggt accgccattc
118921 ggaggttcga atccttccac ccagttagt tcttgttttc aaatgaagtc cagcttttta
118981 actttatgag cttctaacia ttaaaataaa aaatttataa cctagattta gtttagtgtt
119041 tttttgtat gaccttagac taacttcggt tattctgata actgaaaaaa aacttgattc
119101 aacgacagct agttttttaa ctgattttat tattactgac ataaacactc ttttaataat
119161 acacaagtta atagaaatac tagaatatag gcattaaaac cgtgttccag aagctaggaa
119221 acccttcctt aagaaggtaa gctaagcttg tttccctcac cccgaagggg gatacgaaca
119281 aagtttttta agtacgactc cgtttgccct cccaaattag taggcacca gtcctctgta
119341 taaatagcat ttaaaaataa tgacttattt atcaggattt tatacttttt ttcgccatca
119401 aagcttacgc ccattatttt atcaataatg cagccttctc taatttgagt agggatagcc
119461 ggaggaatat ttgaagctcc acgtaaccaa aaaatcaaat acccacctat ttttagcatc
119521 ttccagtttt tcataattct ttttaccttt tataattgcc ctgctcctt accgtgtccg
119581 gtaggctttt ttaaatggtg ttcaacattt aaaaaaactt tgcggagcgc agggaaaaag
119641 gtcttctgcg tttaaaagca gagcttttaa acttagacaa cggcggcagg ctttccccgt
119701 aggcctttacc gtagctcggg aacctagggc tcccggggcg aagcgtaaaa ctttatgttt
119761 taccgttgcc ctgcccgtcaa cggcaagggg ctaagataaa acatttaggg aaaaactttat
119821 ttaacaacga agttgtttaa tttaggcagc gggcaaggtt gccaaagcga gcttgactac
119881 ccctttatat tgttttattg ttttttatcc tacttattca caaagcttaa tatcggcggg
119941 gacatgttaa ttagcttcgc ttgttttcga caccttatgt ttaaacttta gtgcaaaata
120001 atgagcaaca agtgaagctt gctaaccctc ctacctgctt tcttttagga ctattgtcta
120061 tttctaaatg cagacaatgg acttagtgat aatataaaat tattcaaaat tcattgagat
120121 ctgtaacttt gtccagaata tagtaataaa taaaatttta ttttttagaa aaaaatctag
120181 ctaaaaacgt tagtaaatgt taaaattaag aaaataaaaa tgcaaaccca agtcaacaag
120241 tttaaaaatt tttttgttg accataattt gaatttgggg aaactcgggt tcataaaaa
120301 atccttaata gtgctgtttt tgttgccggt tgtggaacaa ttaatacgat aacaagtatt
120361 gcaaaggtta gttgaatgat atttaacata aaaagtataa tattgattaa cgagtaaatt
120421 aataataaaa aagtacagat ttgatcttgt catgaagata aaaaatatgt tattctttta
120481 ataatatca agtgccggga tagctcagtt ggtagagcag aggactgaaa atcctcgtgt
120541 caccagttca agtctggttc ctggcaaaaa gtaatttctt attcagtatt aattaataat
120601 aactaatcta attagccaat ttaaaataaa ttttttctt aataatttta aattttattt
120661 taaaagtatt acaacttaat atttaaattt ctaattgtta acccgggagc ctagattcc
120721 cgggctacgc gaccttaatt tttttgtctt aaactagggg attcacttct tgataaagat
120781 cacaagttt cgtccatacg atgaaacttg aacaagctta ctttgtttta gtagagtttg
120841 tttaaaaata aaaatttgct cacgaacaat gctttactta acaaaaatat ttaaactctat
120901 atataaaaaa taattttttg gaataattca atttacattt gaaatgttta aggttgaact
120961 atttttcaag tttacattaa agcattatag aactaccaac taaaaatatt agtagacttt
121021 tttaacactg ataactttaa ataaattttt ttagttatat atctctaaat gagctgaact
121081 aaatgtgatg aactttctcg atagttaaatt acttaagtaa agactcctac tacgtagaaa
121141 atcctggttt gtgcaaagca caaaccttg agctggaatc ctctctata acaactcaa
121201 cactgttgaa ttaaatatatt tacgtttgat tgttttgaga agtaaaactga acttttattc
121261 agcaaatagg gtcatttttag taagattcag aacatacatt ttttaatttt gtttatttta
121321 gttaaaatta aaaaataata tgtgataaaa aacaaagtta tacgaaccaa agtgtccatt
121381 ttactagtca atctacgttt agtaagaggt gttcttttaa taaggcgaat ttaatacaaa

121441	tgatatat	tagtttgagt	gaaaaaccta	tgttttacaa	atTTTTtttg	taaaacttga
121501	tttaaaagct	acgcttttaa	atTTtagaac	atgggtgaaa	aaagttaaaa	agtgtatcct
121561	ttatctgaaa	ttaatactta	aattatgaaa	atTCagaaaa	atagtaatca	aataattgat
121621	ttagttaata	aaactcatgt	aacaatcact	aaacaaaatt	tagttatacc	tactcagttt
121681	atcttaatag	ctatttcagg	aggTcaagat	tcaatttgct	tatTTTTt	tttattacag
121741	ttaaaaaacc	aatggaagtg	ggattttggc	gttttatatt	gtaatcattt	ttggcaaact
121801	gattcatttt	atacaagcgc	acttattttc	agagtagggt	ttatattttc	aattccggct
121861	tattttaaatt	taccatcaga	aaacattttt	actgaacaaa	aatccagaaa	ttggagatat
121921	tatcaattta	aaagattaag	ctattttttat	aattatgata	ttattttttac	aggacataca
121981	gcaacgggatc	gaattgaaac	aattattctt	aatcttatcc	gaggaactag	tacgcgagga
122041	ctatctactc	taaactggat	taggcccatt	tctcgacaaa	gctttttatt	ttctcgttcg
122101	ttttgtatta	caaagcaaaa	acttgctaac	cttacagata	gatcaaataa	agtattttaa
122161	aatcaggact	tacctttaa	tgacgtatta	tttacagtaa	atTTTTgtaa	tcactttttt
122221	ttaaaattac	aaaataattt	gtttactttt	tataaacaaa	taaataatga	aagtTTaaag
122281	cggcggaatc	aatttgagca	aaatattgtc	cagtgcacaaa	ttcaagttct	ccaacacaaa
122341	accacaaaatt	ttaatcagag	ccctcccttc	ggctctttat	ttttaatagc	tgagactttt
122401	aagcgcgtcg	ttggcgtgca	aaataaaaaa	agaagtttgt	tttctatgaa	taatattttg
122461	gtttatgacg	acttcttttt	tttaacttta	gaaaaattcc	gaatacctaa	tgaagtatac
122521	gaattaaagt	ttttaaaata	ttattacttg	actgataaag	aatgtcaaaa	atgtttttatt
122581	cgacctttat	atTTtgataa	atctcaagtt	ttagttgaac	gtgaattaaa	tataagtaca
122641	attatttttaa	attatacaaa	tttaaaccac	cagtctaatt	tttatagatt	aaaatggaat
122701	ctctggcata	aaaaaaaaata	aaatacaaa	agtgggtttt	ttaatttagt	taatagctca
122761	agcaaagcta	aaaaatacaa	gttcggttaca	aaaaagataa	aaaaaattaa	gttagtaact
122821	agagcgggggt	tagtgaagggt	cccagtttaag	ctttgctttc	ctaaaataaa	atcccagatt
122881	ttattttaagt	cttataatgc	gcaaaaattct	tcttttatta	taaggcaaaa	agtatcagtt
122941	tttaattttta	aacctatttt	tttttcgaaa	aaaatttcta	aaatcataat	taagtttgta
123001	aaagtatcat	attggactca	atTTaaacaa	aatttttttag	aacttgaatt	tttgaaatta
123061	aaccctggag	ctttcaatac	cgcgctacgc	aatagcttaa	atTTaaagaa	cactgagagc
123121	tttaattttga	gcgctaaaac	tatgaatcaa	caaaatattt	ctaaatctat	tattacgctc
123181	aacctttttcg	ataaaccaca	aatTTcaaat	aaatttaagt	tgTTTTtacc	agaaaaagag
123241	aaatttgaagc	aaaaagattt	ttttaataaa	aaattttttta	tatttagaagt	gatcactttt
123301	ccgcctaact	tcagtaaaga	ttctgttagt	tacgtttatc	agatgaattc	ttttaatcaa
123361	atcagggatt	tgatcgtatt	tttatttttt	aggcaagctt	tagtttcaaa	ctttattcat
123421	atcataagat	acaataatca	aaatctgagc	tttattagtc	caaatacact	tttattttgga
123481	tcatgtcaac	taaccaataa	acaactttgt	gtgtttcggc	ctgttttatt	tttaaataga
123541	tttgacgtga	aacaactatg	tactttttgg	aaccttccag	tttatcctga	ttatacaaat
123601	aatcaattaa	catattatcg	aaatcggggt	cgaaaacaat	tattaccttt	attaagattt
123661	tttttttaatc	cgcaaattga	taaacttttt	ttacagttcg	cagaaatagc	aaatactgaa
123721	caaattttatt	taaattctgt	atcaactcgg	ttacaacaag	aatttcaaat	taaaaaaaca
123781	aaaagtttttg	aattaaatat	gtcattattt	cattttctcc	ctttagctat	tagaagacgc
123841	ttactgaagc	aatttctaga	tgactcttta	actaaaaaag	taaacttttt	tcaatttgat
123901	tttttactta	atTTtaattaa	tagaaaaaaa	aaataaattt	ggattctatt	acccttgccg
123961	cccctcctta	ccgtgtccgg	taaggcttta	ccctttgggt	aaaacttgat	ttacaacga
124021	agttgtttaa	tttagacaac	ggcggcaggg	caacggtaca	attaatcgct	gtgataatga
124081	cattttctcaa	cattaagtat	caaagcccgg	gaatctaagg	ctcccgggtt	ttagtttaa
124141	atTTtaattga	atctaaagttt	aagttttgat	gttgctctat	tatTTctctc	cttacgcggg
124201	gattttcata	acctaaacta	tcctcggcag	ttaaagaaaa	gcagtcaggc	tacgccccca
124261	ctcctccggg	gtaaggttgt	caagcaaaaga	ctgactaccg	gaggctctcac	ctaaagggtg
124321	aaactttaac	aagtttctgt	tgTTTTtagac	ctcgtgattt	ttactaaaac	gccaacgggtc
124381	tacatctaag	tttcaactat	tgTTaattta	gaaaatttaa	atTTTTttat	aaactaaaac
124441	ttagctataa	atagggcgta	gcctggcttt	aataagcgaa	gcttatttaa	acttgatatt
124501	gttcttcgaa	caatattttag	gaacctagga	agaagacttc	catecttacc	atgtccggta
124561	aggacaatgg	cgaagtggaa	gatcctagtg	tgtgcaagca	cacactttgg	acttgatatt
124621	ttcccatggg	gctcccaggt	ttataatttt	aaaacttaag	cttgaaaaaa	taacaatttt
124681	aaaaaacttt	tactagtttt	tagtttttatt	atatttttagt	atgatccaaa	tataaatgaa
124741	taaaataagc	agatattttt	taagctctaa	taaaaaagtc	aaaactcgaa	atattttttc
124801	taataaaaaat	aaattagaaa	ccccacaact	tctattttgt	ccaggagttg	ggactttttt

124861	tatcggatct	caacgtttaa	taattttatac	aaaatctatt	aatcaaacat	tttaggtaat
124921	ttgtgtcaaa	aaagcatttt	ttgagaacga	gatacattaa	tttctcaagc	tctttgttta
124981	atttaciaaag	ctagctgtgg	ggatatttcta	ccacccttta	actaccccag	cttcatttat
125041	tatcgtttta	gataaatagg	aaaattatca	aacgaaatft	ttctaacaaa	ccactaataa
125101	atatattaac	tatctaatac	tagatgtact	aaactgggtt	aacaaataaa	ttttatattc
125161	ttgggttacc	cgggactcga	accggaact	aatcgggtaa	aagccgaata	ctctaccatt
125221	gagttagcaa	cccttaattt	taactaatag	tatgttaatt	agctttgcgt	atfttttgcaa
125281	ctaaaaattc	cagaaaaaaa	agcttattaa	atacaacta	tttttatttt	ataaatcacg
125341	aatagctgaa	aaacgaaatc	ttagtttatg	caaagcataa	accaagtatt	accgtaattt
125401	tgccgttgta	gaagaacttc	atfttttaaac	aaaatataaa	ttgagagatt	acattgtatt
125461	aaaattgaat	agttttatatt	aaaacagggga	catagcaaaa	atgaggacat	aaaaagttac
125521	atcttactat	gacatcatta	atgatgtcat	agtaagatgt	aactftttttc	tttatattta
125581	ttaaaattaaa	aaataagttg	ttatftttatta	attacttgta	aagagcaatt	cctagacgaa
125641	gagcaagaat	accagttact	aaggcagtaa	aaagagcaat	aaaaatttga	ctatcggaaa
125701	gtatcataaa	tttggattta	tccttattttt	tgaaaatctt	aacattttatt	ggaaattttt
125761	aatftttttaa	aatattftaaa	aagaacaaat	taaagatttc	caccacggct	tgcaagtaaa
125821	ataataacaa	taggacctgc	tgctaaaaatg	aaaacaagtg	atgctaattg	gaaaataatt
125881	tctaaattca	ttatatacct	caactaaata	acaaaatatt	tagattttaga	aaagttttatt
125941	ttagttttgt	tttctaataa	aactacaatc	aaaataaacg	tttctaattt	aaattagaac
126001	ttcttaattt	aaaagccagg	ctfttttaaca	aagattftttt	ttgacatagt	cttgagaagt
126061	tttctaattg	tataaaactt	attcgaagat	taatgtctaa	atttaagaac	gaaattfttaa
126121	gcccagttta	gcctgggatt	ttagaatttc	gcttaaacft	tctacactac	caaaggtagc
126181	aaggttgacg	agttttactc	acgtacttaa	atfttaaaagc	tacgccatt	taaacataaa
126241	atgttcaaac	aagggttatt	tgtgttctgc	acaaactftt	aaatatagaa	aacttgaagc
126301	gaagcttcaa	gatcttaata	tactccttat	ttaaacaacga	agttgttaaa	taaaattttg
126361	ctctaccagc	aaattctatc	ggacaccggt	atcaagtagt	gtagcctgg	tttaacaagc
126421	tatgcttggt	aaaactfttat	attattggccg	ttagcaacaa	tatttacgaa	cctaggaaca
126481	aaacttccct	aaaggggaaga	tcttagtgtg	tgcaagcaca	cactfttggac	ttccactatg
126541	tggaagactt	taatacaata	atacaatctc	tgattgtagt	ttagacttgt	gttctfttccc
126601	cgttgccctg	ccgggcacgg	caagggcaaa	atacaatcgg	agattftttt	aggtcttccc
126661	acgtagtgga	aggcatgggt	aaacaagcaa	agcttgtcaa	ccttactatc	tccggtagtt
126721	taggtctttc	aggtaagcag	ctfttgttcc	atacctcttt	ggggtgaggg	tgtggftttt
126781	aactftttatg	taaggcgttg	gttatftaac	gtcataaggt	aagcttagag	acattfttaaa
126841	aaaaacctta	accacttctg	ctaattftaag	tataataata	agtcagctat	cattfttttag
126901	ctagtttagat	aagttttatc	taaagctaac	tgatgcttgc	catacaaatg	ctaatagtat
126961	aaaaagaaca	gggataattg	gtaacacatc	tacaattggg	tcaaaaggtg	cataagcttc
127021	aggtaatttt	gcaagtaaaa	aggctatagc	taatacaaat	ctcctcggaa	ataattfttta
127081	aattgttaatt	ggcgattagg	tgctfttggta	ctagttfttctg	aaaatattaa	ttataaactc
127141	taaaataattt	tgacacatat	ttagttftatt	taaagttftt	aattfttaact	aaacctcaaa
127201	atfttttgtat	tctacccaaa	gcttgtfttct	aagtttgcac	aaagagcaaa	actaattfttg
127261	ttftttactta	atftttfttta	ctfttcattac	gtgactfttgt	tactaaacat	gattftaataa
127321	ccttagatttt	tctcgatggg	aaaaacttgg	ttaaacaaact	tcgftftttta	acttagaaaa
127381	ttaaaatttt	gtftfttttca	agttftttaaa	tttgcftttag	tttagttfttc	caagtaaaacc
127441	ctaagttftaa	aatctacctt	tgagtaaacaa	ctaattftttt	tagtcctatt	ccgtcacata
127501	agcgaagctc	atcagcttag	gtatcaatat	agttaaacaat	ttagagtatt	aactftaaaag
127561	agaaattaac	aatatctaaa	tgagttgcag	gcattftttat	tcaactgcaa	ttaaaattta
127621	aaaaacataa	atftttctgt	aactfttcaact	tttaaaagct	ttfttttacta	cttctctgtc
127681	tgagctgtac	tagtttagacg	atfttttctta	cgggctfttat	aagcgtagct	tgtaaaactt
127741	gattacaaac	gcagaagact	ftaaccttctg	ccgtgcccgg	cagggcaacg	ggcgaagctt
127801	aaagatccta	atgtgtgcaa	gcacacattt	tggacttcca	caaagtgga	gatcctaata
127861	gaactctggt	ctatftttgga	ctftttfttct	acgctgcgca	aagtttgtaa	tttagaaaaa
127921	cttggttaac	aaactaggtt	tgfttaactta	gagatctagt	aaaaattcgc	aaaactagtt
127981	aataagtgat	ttfttattaaa	ttagtataaaa	aaatatgatc	aaccattfttt	tattctaacaa
128041	aaaaactftt	atftaaaatct	ataaattftta	ttfttaattft	tgctftttaat	aaagattctta
128101	ttftttaaaat	atgctftttta	ataataaaaa	agtacatcaa	tcatgftttta	ttfttaattt
128161	ttfttatttgg	cccaataaat	ttataagttt	ctftttftaaa	tttaaaaatg	acactgattc
128221	atcttattttc	taaattftaaa	agctatgctt	ttaaatcaag	ttfttacaac	aaggtttgta

128281	aaacctgtct	aaaaaaaaag	tttaacaaaa	aatcttctaa	aataaataaa	aaaaaaaccc
128341	cggtttacag	aaaattgaat	ttttatagga	agagcctaac	ttttcttatt	aataatctcg
128401	aaaagccggt	ttggtatata	tcactttgga	gaattatata	tgaaattagc	atattggatg
128461	tacgcaggac	ctgctcatat	tgggaccctt	cgtggtgcaa	gttcatttaa	aaatgttcat
128521	gcaataatgc	acgcgcctct	tggagacgat	tattttaatg	taatgcgttc	gatgtagaa
128581	cgagaaagag	atcttactcc	ggtaacagca	agtattgtag	atagacatgt	tttagctaga
128641	gggtcacaag	aaaaggtagt	tgaaaatatt	acaagaaaag	ataaagaaga	acggccagat
128701	ttaattggtt	taacacctac	gtgcacttct	agtattttac	aggaagattt	acaaaatttt
128761	gttgatcgcg	cggcaattga	atctgaatca	gatgtaattt	tagctgatgt	taatcactat
128821	cgtgttaatg	aattacaagc	tgcagatcga	actctcgaac	aagttgttag	atcttatatt
128881	gaaaaagcta	gaagaagtaa	tacttttagaa	ttagaaaaaa	cagaaaagcc	gtccgctaatt
128941	attattggta	tttttact	agggtttcat	aatcaacatg	attcaagaga	actcaaacga
129001	tttttaaatg	atctagggtt	tgaatattat	gaagtaattc	ctgaaggtgg	tttagttact
129061	aatttgcgta	atcttaacaaa	agcctgggtt	aacctaatc	catatcgtga	agtgggtta
129121	atgacagcaa	cttatcttga	aaaagaattt	aatattccat	atgtaactac	aacacctatg
129181	gggatagtag	atacagcaga	atgtattcag	gaaataaaaag	aaattctaaa	aaaaaataat
129241	tattataaaa	atcttgatga	atataattgag	cagcaaaaa	aatttgtttc	tcaagcagct
129301	tggttttcta	gatcaattga	ttgtcaaaa	ttaccggaa	aaaaagctgt	tgtatttggg
129361	gacgcaactc	acgcttcctc	tatgactaaa	attctagcta	gagaaatggg	aattcgtggt
129421	tcttggtccg	gaacttattg	taaacatgat	gcagattggg	ttcgagaaca	agttcaaggt
129481	ttttgcatg	aagtgttaat	tactgacgat	catagcaag	ttggtgat	gattgcaaga
129541	attgaacctt	cagctatctt	tggcacacaa	atggaacgct	acattggaaa	acgtctagat
129601	ataccttgtg	gtgtaatttc	tgctcctggt	catattcaaa	atcttccatt	gggttatcgc
129661	ccatttttag	gttatgaggg	gacgaatcaa	atagcggatt	tagtttataa	ctcttctact
129721	ttaggtatgg	aagatcatct	tttagagatt	ttcgggtggc	atgatacaaa	agaagttata
129781	acaaaatctc	tttcaacaga	ttcaatttta	aattgggctc	ctgatggctt	agcagaatta
129841	aataaaaattc	caggttttgt	tagagggaaa	attaagcgt	atagggaaa	atctgagaga
129901	gaaaaaaata	ttaatgagat	tactattgaa	gttatgtatg	cagcaaaaaga	gtcagtaggg
129961	gcttaattaa	atgtggctca	tttaaaaagt	ttatgtggca	aattatacag	aaaaaggggtg
130021	cacaacaac	agtttttatg	tttggctcct	tctacttttt	tagataagct	tagtgaattt
130081	cgatgaccga	tttataaaaa	agcatataga	cttattttta	atctactagc	ttttacgcct
130141	cattattcaa	atcttagtaa	gtttgctaaa	aaatggtggt	atctaaaaaa	gtttaatagc
130201	ttttaacccc	taaagcgctt	tttgtttatg	taaaaaattt	aaaataagac	gatactttta
130261	aagtattaat	ttttttgtaa	ataacaattc	aaattaaaga	aaaattgata	tactctaatt
130321	aggtttaatta	agatagattg	atctttgttt	taaaaaatct	tccttgtttt	tagtaaaagc
130381	ttaaaaactt	aatagtacat	aattttttac	tcatttgatt	tttttctata	tactttatta
130441	aacttagtac	tttacctaaa	ataaccaa	ttggatgggt	ttttaagttt	aaaacaaaa
130501	agttaaataa	aactcgtctt	tcttttagtta	taacttttgt	ttgtaacgaa	gatttggcta
130561	aaagcgaagc	tcgttgctct	aagtttacca	aataactttt	ggtaaccaag	gttcaaat
130621	aaaagctcta	catttgaaca	ttttgtatcc	aatcaagggg	aagatttttt	ttccttgcat
130681	gagctcttgc	tttagtcttt	actcttctac	caaagggtaa	gggctcatct	acactatctt
130741	atctccactc	tgttactgca	ataactaatt	aagatttaac	aagctacatt	tgtaaaaact
130801	taatattggt	tgtattttac	ttggtcgaaa	ggcggtgaga	acaacattta	gaaactattg
130861	gtttcaatat	ctcgtatac	ttatgctcag	tggaagcaga	aaacacaaaag	gtttgaccag
130921	agggtaaagc	ctagtttaata	aacctgggaa	gctttgttta	aattttcag	gcttacctaa
130981	gctaccaaaag	ctagaaaggc	tgatgtattg	cttatagcaa	aatgaataac	cctagattct
131041	ttataaagta	aagtttaatt	tatcctatat	ccgataggct	ttctttttaa	aggaacatta
131101	gcacaacaat	aagtctaagg	tttatattta	aataaaaacta	tggaaaaaaa	ggtaagtcca
131161	aaatacactt	ttgattgta	ccgttgccct	gccgggcacg	gcaaggggta	aattttataa
131221	ttataagcta	ttgctactga	ataatcgcaa	agtgggggtc	ttatttctta	tttcttattt
131281	cttaattcca	gagccacagt	ttgtttgtta	acttttgatg	ctgcttagat	ttttatccat
131341	aaacataaaa	atcttaattg	aatatatatc	agtatgtatt	caaacaatag	gtaatcaagt
131401	ttcgcttgat	aaccttccct	taaagggggg	aggcagttgg	tagaaaaaac	ataagtgaaa
131461	tttatataat	atctgtttta	ggagaaattc	gatgacaatt	agtccaccag	aacgagaagc
131521	aaaaaagggtg	aaaattgtag	tcgatcgaaa	tccagttgaa	actagttttg	agaaatgggc
131581	taaacctgga	catttttcaa	gaactttatc	taaagggcca	actactacaa	cctggatttg
131641	gaatctacat	gccgatgcac	atgactttga	tagccatact	agtgatcttg	aagagatttc

131701 tagaaaagtt tttagtgctc attttggcca attaggtata attttaattt ggcttagtgg
131761 aatgtatttc catggagcgc gtttttctaa ttacgaagca tgggttaactg atccaactca
131821 tattaaacca agtgctcaag tagtttggcc tattgtagga caagaaattt taaatggtga
131881 tgttggaggg ggctttcaag gaattcaaat tacttctggt tttttccaac tttggcgagc
131941 tagtgggtatt actagtgagt tacaacttta tagtactgct attggtgggt tagtattagc
132001 ttcagcgatg ttttttgctg gttggtttca ctatcataaa gctgcaccta aattagaatg
132061 gtttcaaaat gttgagtcta tgttgaatca tcatcttgca ggcttattag gtttaggaag
132121 tttagcttgg gcaggccacc aaattcatgt ttcattacct attaataaat tacttgatgc
132181 gggagtagat ccaaaaagaa ttccattacc tcatgaattt ttattaaatc gagatttaat
132241 agcacaatta tatccaagtt ttcaaaaagg tctggctccg ttttttacta taaattgggg
132301 tgaatatagt gattttttta cttttaaagg tggcttaaat ccagttaccg gaggcttatg
132361 gttaacagat acagctcatc accattttagc cattgctggt cttttcttag ttgcaggcca
132421 actgtataga actaattggg gtattggaca tagtatgaaa gaaattttag aagctcataa
132481 aggcccattht actggggaag gtcataaagg actttatgaa attttaacaa cttcttggca
132541 tgctcaatta gctattaact tagctttatt tggttcttta tcaattcttg tggcacatca
132601 tatgtattca atgccacctt atccttattt agcaactgat tacgcaactc aactttctat
132661 ctttacacat catatgtgga ttggtggttt ttgtattggt ggagcagggg cgcagccgc
132721 tttttttatg gtacgtgatt acgatccaac taataattat aataatttat tagatcgtgt
132781 aattcgtcat cgtgatgcaa ttatttctca tttaaattgg gtttgtattt tcttaggatt
132841 ccatagtttt ggtttataca ttcataacga tactatgagt gcattagggc gaccacaaga
132901 tatgttttct gatacagcaa ttcaattgca gcccgatttt gcacaatgga ttcaaaatc
132961 tcattttttt gcaccagaat taactgcacc aaatgcttta gctgcgacta gtttaacttg
133021 ggggtggagat gtagttgctg ttggtggaaa agtagcgatg atgccaattt ctttaggaac
133081 agccgattht ttagttcacc atattcatgc tttcactatt catgtaacag ttttaattct
133141 actaaaaggt gtactttatg ctcgtagttc acgtttaatt cccgataaag ctaacttagg
133201 tttccgtttt ccatgtgatg gaccaggctg aggcggcaca tgccaagtat cagcatggga
133261 tcacgtattht ttaggccttt tctggatgta caattcatta tctattgtha tttttcattt
133321 tagttggaag atgcaatcag atgtttgggg tactgtaaat gcatctagtg tttctcatat
133381 cacaggtggg aattttgcac aaagtgcmaa tactattaat ggggtggctc gtgacttttt
133441 atgggcacaa tcttctcaag ttattcaatc ttatgggtct gctttatctg catacggctt
133501 aatgttctct ggagctcatt ttgtttgggc attcagtctt atgttcttat ttagtggacg
133561 tggttatttg caagaactta ttgaaactat aatttgggct cataataagc taaaagtagc
133621 cccagcgatt caaccaagag cattaagtat tactcaaggc cgtgctggtg gtgtagctca
133681 ttatttatta ggtggtatcg caactacttg gtctttcttc ttggcaagaa ttatcgctgt
133741 tggttaattht agcaaattht ttgctgctta gtaatcgagt aaatatttga acactatctt
133801 tacagttacc aaataaagtt tattgaccaa gggccaaatt taaaagtcta atctttaa
133861 gaggttttca ttagttgata attttgtggt caaaaaaaga gttaaatatht tcatccttaa
133921 catgtctggt aagacaatgg ccaagtgaaa gtccaaaata gactcaaaga gtctattagg
133981 attttgaagc ttcgcttcaa gtccaaagtg tgtgcttgta tacattaggc tcttccctaa
134041 aggcctttacg cgaagcgtaa aactttatgt tttatctttg ataaaacatt taggaagtct
134101 tcttttaagg tgtctaataca ttggtcgaag aacaatatgg ggttttaaca aacttcgtht
134161 gttaaagcca ggcttcaatt ttcgactgta ttcggatttht ttactaagta gaaatttgha
134221 tctgtcttht aatactgggc aagaagaaaa agttactaaa aatcttagaa taaagttaaa
134281 tttgtttttt taagaaacgg aggatttcta aaagactatg gcaacaaagt ttccaaaatt
134341 tagccaaggt ctggcacaag accctactac tcggcgaatt tggatggaa tcgcaactgc
134401 ccatgattht gaaagtcatg acggtatgac cgaagaaaat ctttaccaaa aaattttcgc
134461 atctcatttht ggccaatttg cgattattht tttatggact gctggtaatt tatttcatgt
134521 tcatggcaa ggtaaccttht aacaattggg cagatcca ttacatgtht gcccaattgc
134581 agcatgcaatt tgggactctc atttcggtht accagctgtht gaagcattht cctcgtggcg
134641 agcatctggt ccagtaacaa tttcaacttc tgggtgtht cagtgggtggt atacaattgg
134701 attacgtagc aatcaggatt tatataatgg atctattht ctgattctag gtgcgggctt
134761 ttttttattt gcaggttggc ttcatttht gcaaaaatt caaccttctc tttcatgtht
134821 taaaaacgca gaagcgagat taaatcacc cctatctgga ctatttggtht taagctcact
134881 agcatggact ggtcatttht tacatgtht tatcccagaa tcgctggtht agcatgtgtht
134941 ttggaataat tttttaacaa ccttaccaca tccacaaggt ttaggtht ttttctgtht
135001 aaattgggct gcttatgctc aaaatcctga ttttcaaat catattht atacagctga
135061 aggttcagga aatgctattht taacgtht aggtgggtht catcctcaaa ctacagctt

135121 atgggtaaca gatatcgctc atcatcattt agctattgca gttgttttta tcttagctgg
135181 tcatatgtac cgaactaatt tcgggattgg gcatagtatg agagaaatct tagacgcaca
135241 cacgcctcct gctggagggc ttggagctgg acataaaggc ttatttgata cagtaaataa
135301 ttctttacat ttccagctag gattagctct tgcggcagtt ggcactgttt gttcattagt
135361 agctcaacat atttattcgt taccgccata tgctttttta gctcaagatt ttacagctca
135421 agcatcactt tatacacatc atcaatatat tgctgggttt attttatgtg gggcatttgc
135481 tcatggagct attttcttta ttcgagatta tgatcctgaa caaaataaag gaaacgtttt
135541 agcaagaatt ttagatcata aagaagctat aatttcacat ttaagttggg taagtctttt
135601 cttagggttt catactcttg gactgtatgt acataacgac gtagtacaag cttttggaac
135661 tcctgaaaaa caaatcttaa ttgaaccagt atttgctcaa tggattcaag ctgctcacgg
135721 gaaaaccctt tatgggtttg atctattatt atcttcatca gcaagtcctg ctttcacggc
135781 aagtcaaagt ttatggttac ctgggtggct agatgctatt aataataata ctaactcttt
135841 atttttaact attggacctg gtgattttct tgtacacct gctattgctc ttggacttca
135901 tactacaaca ctaattctag taaaaggggc tttagatgca cgtgggtcaa aactaatgcc
135961 agataaaaaa gatttcggat atagtttccc atgtgatgga ccaggctctg gtgggacgtg
136021 tgatatttca gcttgggatg ctttttatct tgctgtattt tggatgttaa atactattgg
136081 ttgggttaca ttttattggc attggaaaca tttaggaatt tggcaaggaa atgtaaatac
136141 atttaatgag tcatcaacct acttgatggg gtggctaaga gattatttat ggttaaattc
136201 ttcacaatta atcaacgggt ataatccatt tggatgaat agtttatcag tgtgggcttg
136261 gatgttttta tttggccatc tgatttatgc tactggattt atgttcttaa tttctggag
136321 aggttactgg caagagctaa ttgaaactct tgcttgggct catgaaagaa caccacttgc
136381 aaacttagtt cgttggcgtg ataaaccggg tgctctttca atcgtacaag caagattagt
136441 tggactagct ctttttcag ctggttacgt ttttaacttat gctgcattct taattgcttc
136501 tacatctggg aaattcgggt aaacacttta ttttaaagtt gtcaaattca aattattaag
136561 tttttttagg ttgtctctaa atgaacaaac ttgtttaaca aacgaagttt ttcctaaaca
136621 aacatagttt gttcagatta ggtcctaaat ggtttaccgt tgcctaagtt taaaagtttg
136681 tgcgaaatgca caaaccaatg ggtatgcgaa cgaagtctgc caacctttct acctccggta
136741 gtttaggcag agctttttaa ctttctaaat gttgttcaac atttagaaaa gcctgcccgt
136801 ttaagatgct tacaagtaaa catctaagca agggctcggag ataaaaaata aagttttacg
136861 tcttgccccg ggagccttag gttcccgggc tacagtaaaag cctacggaaa aatcgttagt
136921 tttcaccaaa aaataaagat taagttttat tttgtgctta atgtcctggc taagcacaaa
136981 ataaaactta gttatgggta agaaggaatt tccatgggtc tttgcatctt tttacgttgt
137041 tttctttatt gcttataaag ctaagcaaaa gtttatttta tcaataactt ttgattagct
137101 tctaaattta aagtcagggt ttgctcatag aagataaatt aaataattaa atttgggaac
137161 tttagggttc caagctaggg attactaatt tttaaatata atatttgcta aaataaataa
137221 actaaaatat attaaactta gcttttttaa aattaattct aaaaaagaag aagtattaag
137281 ccttttagacc agtcttctaa atttatgtta agacttttac tttggaaaag ctctagtca
137341 ataaatagag tctattaatt ttaaaaagta gttaaccaat ttgttaatgt aaagggattt
137401 aattgtataa aacaaatttt attattaaaa tatctcagtt ttaataataa aatttgtttt
137461 acttctttat ttttaagcta atcattattt ttatgaaaaa agtagtaagt tgaattttta
137521 tacaattcac aattataata aaaattgtct taacaaaaca aaaaatattt gatttataga
137581 agtttcataa tttaaattaa aaattttggg tattagtaaa taataaaaaa tttttattat
137641 taactcaaaa acatttgtat ttttaatttt tataatattt tggagattta cgctatgact
137701 atagcgattg gtaaaaactga agaaaaactg ggtctatttg atctagtaga tgactgggta
137761 cgtcgtgatc gtttcgtatt tgttgggttg tctgggtctt tattatttcc aactgcttac
137821 tttgctctag gtggttggct tactggctact acatttgtaa catcttggtc cactcatgga
137881 cttgcaagtt cttatcttga aggttgtaac ttcttaacag ctgcagtttc tactccagca
137941 aatagtatgg gtcattcatt actattttta tggggaccag aagctcaagg tgattttact
138001 cgttgggtgcc aattaggcgg cttatggact tttgttgctt tacatgggtc ttttgctta
138061 attggattta tgttacgaca gtttgaaatt gctcgttcag taaaattacg tccatataac
138121 gcaattgcat tttctgcacc tattgctgggt tttgtgtctg ttttcttaat ttatccttta
138181 ggacaactctg gttgggtttt tgcacctagc tttggagtag ctgctatttt ccgatttatt
138241 ttatttttcc aaggatttca taactggact cttaatccat tccatatgat ggggtgtagct
138301 ggtgttttag gggcagcact actttgtgct atccatgggt caactgtaga aaatacatta
138361 tttgaagatg gtgatgggtgc aaatactttc cgtgcattta acccaacaca agcagaagaa
138421 acttattcaa tggtaactgc aaaccgtttt tgggtctcaa tttttgggtg agctttttca
138481 aataaacggt ggttacattt ctctcatgct ttcgttccag taacaggtct atggatgagt

138541 gctattggag tagtaggttt agcgttaaac ttaagagctt acgattttgt ttcacaagaa
138601 attcgtgcag cagaagatcc tgaatttgaa actttttata caaaaaatat tcttctaata
138661 gaaggtattc gtgcatggat ggctgctcaa gatcagcctc atgaaaaact tgtattccct
138721 gaggaggttt taccacgtgg aaacgcctct taatggtaca ttaactgtag gtggctcgaga
138781 tcaagaatct actgggttttg cttgggtgggc tggtaatgct cgactaatta atctttctgg
138841 aaaattatta ggtgcacatg tagcacatgc tggacttatc gttttttggg ctgggtgctat
138901 gaacttattt gaagttgcgc actttgttcc agaaaaacca atgtatgaac aaggtctaata
138961 tttattaccg catttagcaa ctttaggata tgggtgtaggc ccaggtggag aagttattga
139021 tacttttcca tttttgttt ctgggtgttt acatcttatt tcatctgctg ttttaggggt
139081 tgggtggtgtt tatcattctt taattggacc tgaacttta gaagagtctt ttccattctt
139141 tggatatgta tggaaagata aaaacaaaa gactactatt ctaggaattc atcttggtat
139201 tcttggaaac ggtgcttggc ttctagtatg gaaagcattg tactttgggt gtggttatga
139261 tacttgggct ccaggtggag gtgacgttct tctaattact aaccctacta ttagtccctgc
139321 tgttattttc ggttatatct taaaatctcc atttgggtgga gatgggtgga ttggttagtgt
139381 tgataacatg gaagatatta ttgggtggtca ttttggatt ggtacactaa atattttagg
139441 tggcgttttg cacattttaa caaaaccgtg ggcttgggct cgtcgagctt ttgtatggtc
139501 tgggtgaagcg tttttatcat acagtcttgg agctatctca ttaatgggct ttattgcttg
139561 ttgcatgtct tggtttaata atactgcata tccaagttag ttttatggtc ctacaggtcc
139621 agaagcttct caatctcaag ctttactttt ccttgtacgt gaccaacgtc ttggagctaa
139681 tgttgcattc gctcaaggac caactgggtt aggtaaatat ttaatgagat caccaactgg
139741 tgaatatttt tttgggtggtg aaacaatgcg tttttgggat ttccgaggac cttgggttaga
139801 accattaaga ggacctaatg gtcttgatct taataaacta aaaaatgata ttcaaccttg
139861 gcaagaacga cgtgctgcag aatatatgac acatgctcct ttaggttcac taaattctgt
139921 aggtggtgta gcaactgaaa ttaacgcagt taactttggt tgcgctcgta gttgggttagc
139981 ttgttctcat ttctgtcttg gattcttctt tttcgtagga catttatggc atgctggtag
140041 agcacgtgct gctgctgctg ggtttgaaaa aggtatcgat cgtgataacg agcctgttct
140101 ttcaatgctt ccattagatt aatctaatag agttttttac ctaagacttg gtcccttgcg
140161 ttgttttaaa tagcaaagta ataaggtgac caactaagtt cgtatacccc ttgctttacc
140221 cccttgggatt gtcccttggg aagcaagggg taacgagcaa ggctttgctc cttcgcattc
140281 tttttttttt agctaataaa taactaataa aaatgtgtaa taaacagaca attccaattga
140341 taaaatcgcg atatccttta tcgctctttt tagaaaaagg cgatttaccg tggccttctg
140401 aatttgaatt tatagatgca ggagtttttc tttttacaac tgagctttat atacatttg
140461 attttctctt agaaagggtt tttttgagag ataaagatgg gttagttaaa tcataataat
140521 ttaactgtaa gagagtgcag cttttgtgaa ataatagttt cggcaactgt tttttatctt
140581 gaaacggcta gtgatcacia gggagaataa gcctttttta actggccacc agtcataatt
140641 cttctttttg aatcaccttg cagtcttttt ctttttataa tactaaaaca acttcaatat
140701 cctaagcact ctgttggaca ctttttatca agcacaacaa cgtatttatt agatttggag
140761 aaaaccttag taaatcaaga aaaacaaaat aaagttagatt tcaattctct acctacaaaa
140821 gatcatttca tacgagtctt agatttagta ttagaagaag atttaataata aaatctagtg
140881 gatacttttt ttagagtagg taattataaa ttaacgagta tagatttaaa atcaattaga
140941 atttatctag aaactttaat taaaagtaac cctaaacatt atttttaaaa atagtattta
141001 ataacaaaat aattcaaaaag ttaattgggt ctgtcagttc aattctattg tcagcaacac
141061 tactaacact tactttgggt aatgtaatct tttttccaaa aagagcttat tctactaaat
141121 gaggaacaaa gcaagaatta ccttgaagaa aaatggtgga cagctctcaa acaatgtgta
141181 gtggctgaaa tacaagaagc agataataat ctacgcaaat ataaaatata ttatacaatc
141241 gaaaagaaat acgggaaact agacggtgct attctctata tttttcaaac aaaaacaaat
141301 gatgctagtg ataaggaaaa tcaagcgaat aaagctaaag ggaccaaaaga aatcatccct
141361 tatgaaaaac tttatttaac caactctaca ccagtatcgc agaaaaattc tgataagttt
141421 agacaacgta cacaaaacgg caaattattt agatacaatc aaaaacaagcc tttcaaaatt
141481 ttaaaacttc ttttcgctat gcttgttttag gaaagatctt aatagagcgt agctctatta
141541 tggacttttt tttctgtgct acgctacgaa taaatcttgc cggaccattg ttcggttagtt
141601 tttaccttgt cctaaaacaa gctgcgcttg ttgaagtttc acccataggg tgaacgcaga
141661 agactttttt tctgcgctac ggtcttcggg gtgaggctaa ggagcacgtc aagggccccg
141721 tgttcaaaat aaaactttct tcaactttat cttttgactt tccaaattgc tctatatatt
141781 tttttaatta actttgttat aatcacttga taactggtgt ttaatatgta atactactat
141841 ttaccaagtc ggaataaaaa aacttataaa tttcgcagaa gatttttttt tctgcgcttc
141901 gctacgagtg gcaacaatat aaagttttaa caagctatgc ttgttaaagc caggctacgg

141961	aagcttagat	tatgtaagag	tatccggttt	tgagaaataa	aatTTTTTTT	agaattaaaa
142021	attattatga	ctgctatTTT	agaaagacgt	gaaagcgcta	gcctatgggc	tcgcttctgt
142081	gaatgggtta	ctagcactga	aaaccgtTTA	tacatcgggt	ggTTTTGGTGT	tctaattgatc
142141	ccaactttat	taactgcaac	ttctgtatTT	atcatcgcat	tcatcgctgc	acctccagta
142201	gatatcgatg	gtattcgtga	gcctgTTTTct	ggttctctac	tttacggaaa	caacatcatt
142261	tctgggtgctg	ttgttccaac	ttcaaacgca	attgggtcttc	acttctaccc	aatTTGGGAA
142321	gctgcttctt	tagacgaatg	gTTatacaac	ggTggTcctt	accaacttat	cgTTTgccat
142381	ttcttcttag	gtatctgctg	ctacatgggt	cgTgagTggg	aactTTcttt	ccgTTtaggt
142441	atgcgTcctt	ggattgctgt	agcttactct	gctccagTtg	ctgctgctac	TgctgtatTT
142501	atcattttacc	ctatcggTca	aggTtctTTc	tctgacggTa	TgcctTTtagg	tatTTctggT
142561	actTTTcaact	tcatgattgt	attccaagct	gagcacaaca	tcttaatgca	cccattccac
142621	atgctTggTg	ttgctggTgt	tttcggTggT	tctttattct	ctgctatgca	cggttctcta
142681	gtaacttctt	ctttaatccg	Tgaaactact	gagaatgaat	ctgctaaccg	TggTtacaAA
142741	ttcggTcaag	aagaagaaac	ttacaacatc	gtagctgctc	acggTtactt	TggTcgTTTA
142801	atcttccaat	atgcttcttt	caacaactct	cgTtctctac	acttcttctc	agctgctTgg
142861	cctgtagattg	gtatctggTt	cactgctTTA	ggTatTTcaa	ctatggcatt	caacctaaat
142921	ggTTTcaact	tcaaccaatc	TgtTgtagac	tctcaaggTc	gtgtaatcaa	cactTgggct
142981	gacattatta	accgtgctaa	cttaggtatg	gaagtaatgc	acgaacgTaa	cgcgcaaac
143041	ttccctctag	acttagcttc	TgtTgaagct	ccttcaatca	acgcataatt	ctaattatgc
143101	ttaatctacc	ggTtccgggA	gcctaaaaca	aacgaagTtt	ggTcaagTtt	tataacgcag
143161	aagaccatgg	taagaaaaca	agtccaaagt	gtgtgctTgc	acacactagg	atcttccact
143221	tcgccattgt	ccttaccgga	catggtaagg	atggaaagtct	tcttactTTT	tttctgcgct
143281	acggccttcg	gcttataaag	cctaggTtcc	cgggctacgg	tctTTTggT	cttTactcat
143341	tagctattgc	taatgagcaa	tagccaaaag	gtaaagacta	aagaatatct	ggataataaa
143401	aataaatgtg	ataatccgcc	attgccaaagg	gcaatggcga	agcagaagac	catggtaaga
143461	acacaagtct	aaaatacaat	ccttccTTaa	attaccttag	cgtagcccgg	gaatctaggg
143521	ctcccgggta	atTTaagaaa	gtatattTTca	tatcagagat	Tgtattaaac	ccgggagccc
143581	taaattactc	cttagctatT	gctaattgagc	aatggccggA	ggtagcgcag	cgcagaaaaa
143641	aagtaagaag	acttccatcc	ttaccatgtc	cggtaaaggac	aatggcgaag	Tggaagatcc
143701	tagTgtgtgc	aagcacacac	ttTggactTg	Tgttctgtgc	tacgctgcgc	ttcgctatgt
143761	tTgtttagga	tcttccctaa	aggcttTtacg	tcttaccctc	ttggccgTtg	ctcattagca
143821	acggctaatag	agtaaggcct	tccaccattc	ggTggaagat	cctggTtacc	aaactTtgTt
143881	Tggtaacctt	ggaccaaggg	gcgaagcgta	aaactTTatg	TTTTatcttt	gaccctTgct
143941	tagatgtTTA	ctTgTaaaca	tctTaaaccg	gcaggctTTT	ctaaatgtTg	aacaacattt
144001	agaaagtTTA	aaagctctgc	ctaaactacc	ggaggtagta	aggTtggcga	acttccgTtcg
144061	cataccatt	ggTttgtgca	ttcgcacaaa	ctTTTaaact	taggcaaccg	Taaaacattt
144121	aggaagtcca	aagtatacaa	actTtgtTtg	tatactagga	tctTaaagct	acgccctagt
144181	aatTTaagga	aggagtctat	ttTggagcaa	TggccgggTt	TggactTTat	cctTgtTTTg
144241	cttaatcctc	gattTTTTTTA	atataaaact	cgcaagaatc	cactTTaaat	attacttctt
144301	TgtTTTcaaa	attacaaaca	gcataatcaa	aaggaggctt	ttccaaatct	ggcatcgact
144361	cattttcaga	aagagtcacg	aatgagaaa	attgaaaata	attctgttcc	ggatTTTctc
144421	ctTgcgtTaa	ggtaaggcaa	cgggagctag	atcgtagccc	aggatctagg	gctcccggct
144481	tatTTaaaaa	gattTTTTTTT	Ttataaataa	actcgaagaa	aaattcatta	TTTTtcggTa
144541	aacaaaaaat	ttcgtgagTt	gattTTTTTTA	catccggaaa	aatgctcgt	gggcttctaa
144601	aataaatTTc	cgcaccactt	aaccaatatt	gctgaatata	aacggtagTt	TgcagTttTg
144661	tagctaaaag	gggctcgagt	agaagctccc	agattgcttc	aatacagTtt	tctTTacaat
144721	agtaaggccg	atcgtcatct	cctagctTTT	Ttaaagacgg	ccaagatagg	agaataaaat
144781	cacgaaattc	Taaaacaaaa	tCagaaatTT	gaatcattTT	TTTTatctct	tttagTaatc
144841	taatatatac	atataacata	aagaagcaa	accgtgcgta	gcccggccat	Tgctcattag
144901	caatgcctaa	Tgaggaatct	agggctcctc	attaggcatt	gctaattgact	tctactTtTg
144961	agaagatcct	aatagactct	acgatatgaa	atatactTtc	Ttaaattacc	cgggagccct
145021	aaactacctc	cggtagtaag	gTtgacaagc	TtcgctTgtt	Taccgcagcg	cagaaaaaaa
145081	gtaagaagac	ttccatcctt	accatgtccg	gtaaggacaa	Tggcgaagtg	gaagagccta
145141	gtgtgtgcaa	gcacacactt	TggactTgtt	Ttcttaccat	ggTcttctgc	gtaagtataa
145201	atagctggaa	gtacggTgtg	TgctTgcaca	cacctatTTA	TacaaagTtt	Tataagcaga
145261	gcttataaag	cctagattcc	tCattaggca	Ttgctaatga	gcaatggccg	ggctacgccc
145321	tagtaatTTA	aggaaggagt	ctatTTTgga	gcaatggccg	ggTtaccTTT	TTTataaatt

145381 gatgtctctg actttaagtt ggtacctgaa ttttcgtaca aagtaacagt cttagggcgaa
145441 aatagttgtc ttaaagtttg caccaattgt gtttccccag gagtaacctt acctttgtca
145501 gtgctatttt cataccctac cgctctctcg tgaagctcga attgcttggg ttcaaggcgc
145561 tttcgagttt cattaaaacg aaatattttt tttcttactt ccgtaataatc ccaacattct
145621 tcagaatgtg tgttaaaata ctcacgttg tagttctttc tttctttttg atctttcgcc
145681 aacaatttgt ctgctgtttc aaaattagca tcttttctta gcttacaat agttacacca
145741 agtttcccat caagcttacc attagtagtc gtgtttcgga atgcaagcaa atctgcatcg
145801 cggatgaaag ctactaatc acgatcatcc gatgacgta atatttgtat tgcaaagtta
145861 accaaagtat cttgactaga gacagctaga gttggaatct caccgtgata ttgaagttca
145921 ttaagtttta ctttcggggt taatgtgctt ttttttagct tgaattgaac agcttgagta
145981 cttaatgata gcatcctttg tgctcgtgga gtttctactc gacttgagtg gcttaatacg
146041 ttactctctc ctctacgtgt ttgaacgcta gtacggcag ctctacaga ccttgctctc
146101 ggagatgaag aagacagggg attttctgct aatgaactag caaaagaaga cctaggggga
146161 aatgggccag acatcgaaaa gggagatgcg tgtgtacgtg gcggctggtc atccgggtgc
146221 ggagctatag gcgaaggaaa gcagcaagtc atccagttaa aacacgacaa cggagtgtgc
146281 tcagttgttg gttccatcgg cgggatccct gaagtgggtg cacccgagaa ctctgataga
146341 tgagagtgat aatcagcttg gtttggggag gacgcttcag cgctgttgc ccaaactcga
146401 tctggttgtt gcatagcata tgccgattta gattgagtta agagaactaa accaagccca
146461 atgaaaatac taaaattgaa tacaagtct gatagaagtt tttttttttt tttatgacaa
146521 aaatttttgt aaaaatgatt tgaatcttca aagtaatatc gaaatataaa aatttghtaat
146581 tcgtcttcac taaaattttc catactttta attttttgaa cttactaaa atttttcata
146641 cttctaattt tttgaacatt ttacacgaat ttcactttag acttgtgttc tcccctaaa
146701 taaacgcaga agacttaaag catagcttta agatcctagt atacaaacaa agtttgtata
146761 ctttggactt tttttcacag aagaccatgg taagaacaca agtccaaagt gtgtgcttgc
146821 acacactagg atgttccact tcgccattgt ccttaccgga catggtaagg atggaagtct
146881 tcttacttgt gttctgtgct acgctgcgct tcgctatgtt tgtttaggat cccgttggcc
146941 tttgaccaac gggcaaaggc aaaatacaat cttcgattgt acccttgcct agatgtttac
147001 ttgtaaacat cttagcccgg gaggttttcc ccgaaggctt tacgccaagc gtaaaaactt
147061 atgttttacc gttgcctaaa ttttaacaact tcgttgttaa ataaagtttt ccctaaatgt
147121 tttatcttag acccttgccg ttgacggcag ggcaacggta aaacataaag ttttcccta
147181 cggggaaagc cttcggggaa agcctgcccg cgttgtctaa atttaacaac ttcgttgtta
147241 actcaagttt tacccaaagg gtaaagcctt actcattagg cgttgctaag gagcaacggc
147301 cggacacggt aaggatcaac ggcaagggtc taagataaaa catttaggta aagccttact
147361 cattagcgt tgctaattgag caacggccgg acacggtaag gatcaacggc aagggcttaa
147421 gataaaacat ttaggtaaag ccttactcat taggcgttgc taatgagcaa cggccggaca
147481 cggtaaggat caacggcaag ggtctacgat aaaacattta ggtattccat gtaatggaag
147541 gcatgggtaa acaagcgaag cttgtcaacc ttacgcagaa gactttctta aaagaaagat
147601 cttaatagag ctacgctcta ttatagacta tagacttttt ttctgcgcta cgtaacctcc
147661 ggccattgtt cattagcaat agctaaggag tagtttaggg ctcccagggt aggggtgggg
147721 ctacgcccaa ggggtcaaat aaaaatatcc ttgctttatg cgaagtatac ttcgcataaa
147781 atctcgtggg gtcttgatta ttgagactaa aaaatgtatc ctactattta ttcagtgtga
147841 aatttaatat cactatgtaa aacgtaaac taaaattgga gtgggtgtct agtggccgaa
147901 agagctcgat tgctaatega gtgtacagct acctgtaccg agggttcgaa tccctctcac
147961 tccgtacttg taactaaatt gagctttggc ccagaagact tctatataat agaagatcct
148021 accttatgct acgcataagg tttggacttg aagcgaagct tcgagatcta aatgttttac
148081 cgttgcctaa gtttaaaagc tctgcctaaa ctaccggagg tagtaagggt ggcgaacttc
148141 gttcgcatac ccattgggtt gtgcattcgc acaaactttt aaacaaagt tttctaagt
148201 ttgttcaaca tttagaaaag cctgcggggt acggcaaggg tcggagagaa aacataaagc
148261 cttgcctaaa ctacttaatc aagcaaggtg tgattaagta gcgtagcccc cgttgccctg
148321 cggggcacgg caagggggcc attgtcatt agcaatgcct aatgaggaat ctagggtcc
148381 cgggtaagggt tggcgaactt tgttcgcata cccgtagggt tgggtgagct tgcctgtcta
148441 cctaaattta aaagcgcagc ttttaaactt agacaacggc ggtaagacaa tggcaagcta
148501 cgcttgtaa aactttatat tgttgccgtt ggcaacaata tttaggaacc taggaagaag
148561 acttccctaa aggccttaag cgaagcgtaa aactttatgt tttatctttg acccttgccg
148621 taaccggcag ggcaacggta aaacatttag gaagatccta aacaaacata gcgaagcgca
148681 gcgtagcaca gaacacaagt aagaagactt ccatccttac catgtccggg aaggacaatg
148741 gcgaagtgga acatcctagt gtgtgcaagc acacactttg gactttcaca tagtgattta

148801 accggttgect tgccgggcta agatgtttac aagtaaacad ctaagcaagg gtacaatctc
148861 tgattgtatt ttagacttgt gttcttcccc cgttgcctaa gtttaaaagt ttgtgcgaat
148921 gcacaaacca atgggtatgc gaacgaagtt cgccaacctt actacctccg gtagtttagg
148981 cagagctttt taacaaagtt tttctaaatg ttgttcaaca tttagaaaag cctgccgggc
149041 taagatgttt acaagtaaac atctaagcaa gggcaaaata caatcggaga ttgtattaag
149101 gtcttccacg tagtggaagg catgggtaaa caagcgaagc ttgtcaacct tacgcagaag
149161 accatggtaa gaaaacaagt ccaaagtggt tgcttgcaca cactaggatc ttccacttcg
149221 ccattgtcct taccggacat ggtaaggatg gaagtcttct tacttttttt ctgcgctgcg
149281 ctacctccgg ccattgctca ttagcaatag ctaaggagta gtttagggct cccagggtgag
149341 ggggtgagact acgccc aaag ggtaaagact acgcaaagtt tgttttagag acaccgtata
149401 aagcctcttt aaataaagta ggattagaga gagggcgact ggattgacgt aaataagctc
149461 tgtaggatata tcttgaagaa gagaaaagtt aaaaaataat ttaaacacag aagacttcta
149521 tataatagaa gatcctacct tatgtcagc ataagggttg gacttgaagc gaagcttcaa
149581 gatctaaatg ttttaccgta gtttaggtga agcctagttt gacgtcaagt gtaaagcctt
149641 tggacaaatc ctaatcaggt aagggagtat agctcagttg gtagagcgtc gcctttgcaa
149701 ggcagatgtc agcgggtcga gtcgcgttac ttccaccagc ttattaatta atagctggtc
149761 cattgggtttg tgccctgcac aaactttcta acgcagaaga ctttaaccct tgccgtgcc
149821 ggcagggcaa cgggcaagc ttaaagatcc taatgtgtgc aagcacacat tttggacttc
149881 cacaaagtgg aagatcctaa tagaactctg ttctattttg gacttttttt ctgcgctacg
149941 atctttgttt aaaaacttcg tttttaaact tagaaaagtt gaacaagctt tgcttgtttt
150001 agtaaagatc aatgaagaa aagcttacgg tggataccta ggcacctaga gacgatgaag
150061 ggcgtgga aa ccaacgaaat gcttcgagga gctggaaaca agctatgatt cggagattcc
150121 cgaatagggc aaccttttat actacctatt gaattcatag ataggaaaga gacaaccag
150181 tgaactgaaa catcttagta gctggaggaa aagaaagcaa acgcgattcc cttagtagcg
150241 ggcagcgaaa tgggaacagc ctaaaccaac tcgatctgtc gagtcggggg ttgtgggaaga
150301 aagtctatgt ttacctaga aaacaacctg ggagccctag gttcccaggc tacgggttat
150361 caagcaaagc ttgattaccg gttgcctaaa ttaacaaca aagttgttaa ataaagtttt
150421 ccctaaatgt tttatcttag acccttgccg ttgacggcag ggcaacggta aaacataaag
150481 ttttacgctt tgcgtaaagc ctacggggaa agcctgcccgg ccacgggtctt taccttttgg
150541 taaaaaccaa aaggcaaggg ccaggtaaaa ttttatgaag aattcactta actaggtgaa
150601 cgcagctgaat cctgcacat agatgggtgaa agtccagtag ccgaaagttc agtगतatata
150661 ccaactaaaa accttattaa taaggtcagc cttgattggt tgggcccctg ctcttagtct
150721 tttacccttt ggccgttgct cattagcaac ggctaagtag taaagactaa tgagcaaggg
150781 ccgggctacg gatcaactaa atttcttate ccgagtagca tggggcacgt gaaatcccgt
150841 gtgaatcagc gaggaccacc tcgtaaggct aaatactcct aggtgaccga tagtgaaata
150901 gtaccgagag ggaaagggtga aaaagaacc ctgtagggga gtgaaataga acatgaaacc
150961 gtaagctgtc aagcagtggt agggcgaaaa gcctgaccgc gtgctgtttg aagaatgagc
151021 cggcgactta tagggagtgg cgtgggttaa gaaaaccatc tgaagccaaa gcgaaagcaa
151081 gtcttaatag ggcgctaag tcaacttctta tggaccgaa cccgggtgat ctaaccatgg
151141 ccaggatgaa gcttgggtaa aaccaagtgg aggtccgaac cgaccgatgt tgaaaaatcg
151201 gcgatgagc tgtggttagg ggtgaaatgc caatcgaact cggagctagc tggttctccc
151261 cgaatgctg tgaggcgag cggttgatga cttgggctat ctaggggtaa agcactgttt
151321 cgggtgcccg tgcgagagcg gtaccaaate gtagcaaact aagaatacta gatatgcttt
151381 ccatcaacca gtgagacggt gggggataag ctatcatcgtc gagagggaaa cagcccagat
151441 caccagctaa ggcccctaaa tgaccgctaa gtggcaaagg aggtagaat gctgaaacaa
151501 ccaggagggt tgcttagaag cagccacct taaaagagtg cgtaatagct cactggtgga
151561 gcgttcttgc gccgaaaatg aacgggacta agcgggtctgc cgaagctgtg ggatttttaa
151621 actaaacaac cataatttga ttccatcaaa ttatatacgc agaagacctt gggatctcga
151681 cgtagctcgc ccaaccttac tcttgaact tgttcaagta gtttaggaag aacacaagtc
151741 caagtttgt gtgaaacaca aactaggatc ttccacttcg tggaaagcca aaatgtgtgc
151801 ttgcacacat taggatcttt aagcgaagct taaagtcttc ttactttttt tctgcgctac
151861 gacctatggt aaaaatcggg aggggagcgt tccgctctag ggtgaagcat aaacgcaagt
151921 taatgtggac aaagcggag tgagaatgtc ggcttgagta acgcaaacat tggtgagaat
151981 ccaatgcccc gaaaacctaa gggttcctcc actaggctcg tccatggagg gtgagtcagg
152041 acctaaggca aggccgaaag gcgtagtcga tggaaaacag gttaatattc ctgtactaat
152101 tcttgattag taccgagggc cgaagaaagc aaagggttaat caaaattgga tttttgggtg
152161 aaacgttcaa ggtgtcgaga ggtagaataa aaactaacct tgagctgaga cgtgatacgg

152221 attcattccc tggattagtg gtaagctagt ttcattagtg aacttaacta gttgttaaac
152281 ttcctagaaa agctcgaact actaataatt aagaattacc tgtaccccaa accgacacag
152341 gtaggtgggt agagtatacc taggggcgcg agataactct ctctaaggaa ctccgcaaaa
152401 tggccccgta acttcgggag aaggggtgcc tctagaaata gaggccgcag tgaccaggcc
152461 caggcgactg tttacaaaa acacaggtct ccgctaagtc gtaagacaat atatgggggc
152521 tgacgcctgc ccagtgccgg aaggtgaagg aagttgggta ttttattctt cggaataaga
152581 aaagctgaca accgaagccc cgggtgaacgg cggccgtaac tataacggtc ctaaggtagc
152641 gaaattcctt gtcgggtaag ttcgcgaccg cacgaaaggc gtaacgatct gggtagctgc
152701 tcggagagag gctcgggtgaa atagacatgt ccgtgaagat ggggacttcc tgcacctgga
152761 cagaaagacc ctatgaagct tgactgtagc ctggaattga attcgggctt ttcttgcgca
152821 gcctaggtgg gaggcgttga taattctctt ccgggagggt tggagccatc agtgagagac
152881 cactctggag aagctagaat tctaattggcg atccttgaat caggacggtt gacagtttca
152941 ggtgggcagt ttgactgggg cggtcgcctc ctaaacagta acggaggcgt gcaaaggttc
153001 cctcaggctg gacggaatc agccgaagag tgtaaaggca gaaggagct tgactgcaag
153061 acctacaagt cgagcagggg cgaaagccgg ccttagtgat ccgacggtac cgagtggaag
153121 ggccgctcgt caacggataa aagttactct agggataaca ggctgatctt cccaagagt
153181 tcacatcgac gggaaggttt ggcacctcga tgcggctca tcgcaacctg gagcggtagt
153241 acgttccaag ggttgggctg ttcgcccatt aaagcggtac gtgagctggg ttcagaacgt
153301 cgtgagacag ttcgggtccat atccgggtgta ggcgtaagag cattgaaagg agtcttccat
153361 tgtacgagag gacctggaag aacataccac tagtatacca gttctcctgc caaggggata
153421 cgctgggtag ccaagtatgg agaggataac tgctgaaagc atctaagtag gaagcccacc
153481 ttaagatgag tgctctcgtc tagttttcta gacaaaggtc acggtaagac taaccgttcg
153541 atagctatca agtgaaggc tagtaatagt tgcagctgag atagactaac agaccgtatg
153601 attttgatct aaatTTTTCC aaaatagact taaaaagtct attaggatct tgaagccaag
153661 cttcaagtct tagattatct ggcccttgc ctaggtcca aagtttgtgt aaaacacaaa
153721 ctaggatctt ccacaaagtg gaagtcttta ccctttggcc gttgctcatt agcaacggct
153781 aatgagtaaa gaccaaaggg caagggtctaa tgagtagtca agctttgctt gactaccagg
153841 tagtttaaga taaaccttat ccaacttcgc cattgcccctt ggcaatggca aggggttggat
153901 aatcctgggtg tccttgTTTT ctaagcctgta caccaaccctc tctcgaactt ggttgtgaaa
153961 cgtagaaaga gtgacaatac ttaagccgtg gggctttgga aagataactt gatgtcagga
154021 taagatgata agagctgtca tggataaatt ataagtaaaa cccggtagcc ccataaagggt
154081 ttgcccttgc ttgatgTTTT acttgtaaac atcttagtct ttaccctaaa caaacgtagt
154141 ttgttttagat taggtctaag tttaaacaaca tagcgaagca tagaaccata gaaaacttgg
154201 gttctatgct tcgcccagga gccctaagct accgaaggta gcaaggttgg caagctttgc
154261 ttgcttacct atgccttcca tgtaatggca gatctaaata caatctctga ttgtattttg
154321 cccttgctta gatgtttact tgtaaacatc taagcaaggg ctcagggtctg tcaacttttg
154381 acagtctcct tagccattgc taaggagcaa tggccaaagt gtgtgcttgc acacactagg
154441 atctttacct gggagcccta aactacttga acaagttcaa gtagtaagggt tgacaagctt
154501 cgcttgTTTA cccatgggaa gaacacaagt ctaaaataca atcgaagatt gtattaaagt
154561 cttccctaaa gggaggtcca aagtgtgtgc ttgcacacac taggatttgc ccgtaggctt
154621 tacgtcctta ccccttggtg aggaccaagg ggcgaagcgt aaaactttat gttttaccgt
154681 tgccctgccc ggcacggcaa ggggtctaaga taaaacattt aggcaaaact ttacttaaca
154741 acgaagtTGT taagtttaga cctaactctaa acaaacgcag aagacttaaa gcatagcttt
154801 aagatcctag tatacaaaaca aagtttTGT actttggact ttttttctgc gcttcgctat
154861 gtttTGTtag gatcttccac ttcgccattg tccttaccgg acatggtaag gatggaagtc
154921 ttcttcttag gttcctaaat attgttcgaa gaacaatctc aagtttTAA aagcttctgc
154981 tattaaagtt taaacataaa gtgtttaaac caatgggcca ggctacgagc ttcgcttaaa
155041 gtcttcttcc tatacccaaaa attttaatt taaatgatta aacttctcta aatcaaaaag
155101 aaaaaggcca tggtaagaac acaagcttct ttaactttaa gttgtatttg taacattagc
155161 tatttactaa tttaaattac agcttaaagt tgtttggaac tttaccgcca aactaaaatt
155221 gaaaattTGT caaattgtag agacaatcat tttgttTGT attagaagaa ctacactgaa
155281 ctcatatcaa taatttatat ctttctaaaa gaaagataag cttaaataag tataaaatat
155341 taatttTGT atattaaaa attaaataga aattggtaat tatttaatta attttttaga
155401 atcatcgaat aatttttatt taatatggcg tattactaag tgcctgata taaaattagt
155461 gtaattaagg ataatttaaa tgacaaagag aaatttattc gctaaaatta attcggaaaa
155521 ttctaacaaa tcccgttcat ctgaaactgc cgaagacaca atagctcagc ataaagtccc
155581 aaaacccaat attagtTGT atcccactcc tcaagaggca ccgggaaact ctaacgaatt

155641 tcgtccctct gagaaaacag aacaggcaat agctgagact gaagtccgaa taaccctgg
155701 tttgggtgag cccaagattc aagggggatc gagaaattct gccaaatccc gttccgctca
155761 agaggcagaa cacacaatag ttgaggctca atcccccaata agctctgctt ttgttgagtt
155821 caacactcaa gggggaattg ttaatcccga actcggacag cagcgttacc ctccctagtc
155881 ggaagtaaag cctaatagata ctcgatacgg caatgttttg acttttgttt ctttactaag
155941 ttttaataagt ttagtcggag ttgctattta ttttcatcta actcgtcata acgattcaaa
156001 aaaaacaaaa aaacgaaaaa ctgaagtact agaattgact gtcctttttc cgtttgaaaa
156061 aaatacacc atacttatgt gtaagctaca tccaaaaata aaattaggac aaattggtaa
156121 agttaagaaa aaattatcag caaacttagc ttcagttgaa tttcagaaat tagaatttga
156181 gaatagccca atttctaaag atttaattgg gttagctctc agcagtaagg gggagggtt
156241 ggtgttttta taaataaaac aaaagaagct ttgccaaaaa ctttaagattt tagttaaatt
156301 aaaaccttta gctataccta agtttaaaag ctttgcctaa actacttaat caagcaaagc
156361 ttaattaaag agtaaggttg gcgaacgcag aagacttcta tataatagaa gatcctacct
156421 tatgctacgc ataaggtttg gacttaaagc gcagctttaa gatctaaatg ttttatctct
156481 gacctaaaac aaacaaagtt tgttcaagtt ttataagcct tcggcttata aagccttacc
156541 gtgtccggcc gttgctcatt agcaacgcct aatgagtaag gctttacctt aatgttttat
156601 cgtagaccct tgccgttgcc ggcaaggcaa cggtaaaaaca taaagtttta cgctttgcc
156661 cttggtccaa ggttaccaaa ctttgtttga taaccaggat cttccaccga atggtggaag
156721 gccttactca ttagccgttg ctaatgagca acggccaagg ggtaaggacg taaagccttt
156781 gggtaaaact tgatttaaca acttcggttg taaatttagc ctttgcctgt tgggcgtagc
156841 ctgggaatct agggctccca ggttagatat ttattcataa atattttacc caacgggata
156901 acggcgagcag gcttttctaa atgttgaaca acatttagaa agtttaaaag ctctgcctaa
156961 actacttaat caagctttac ttgattacgt agtaaggttg gcgaacttcg ttcgcatacc
157021 cattgggtttg tgcattcgca caaactttta aacttaggca acggtaaaaac ataaagtttt
157081 acgctttgcc ccttggctct taccaagggg taaggacgta aagcctttgg gtaaaaacttt
157141 attaaaaaac ttcgtttttt aatttagaca acggcggcag gcttttctaa atgttgaata
157201 atatttagaa agtttaaaag ctctgcctaa actacttaat caagctttac ttgattacgt
157261 agtaaggttg gcgaacttcg ttcgcatacc cattgggtttg tgcattcgca caaactttta
157321 aacttaggca acggtaaaata tcttagccaa cggggcaacg gttttaggta aaaattatct
157381 tgcattagag ctactaaata aactagagaa aagcccgcga attaaacaat tgtaaattct
157441 aaccgagtaa aacttttttg gtttttctt ataaatggtt attagccttt aaatctaagc
157501 ttagataaaa taaataaaaa ttttttatta tttgctaata cccaaaatft tgatttaaat
157561 caaacaact ttgtttgttt taaacttcta taaattaaaa aaaaaagac catggtaaga
157621 acacaagtct tcttacttta aagttgtatt tctaaccattg gctactgact aatataaatt
157681 gcagcttaaa gttgtttgga actttaccgc caaactaaaa tcgagaattt gtcaaatftt
157741 ttgttgata attttatata taaatttatt tttcaatttt aactttgftt attttaacca
157801 aaaactatgt atactttatc atcatcttta aacgcagcaa tccctttaca tagacctgac
157861 atgcctttaa ggagacccta tgagtttgga aatatactta aaactggaca aataaaaaag
157921 atccctaaac cgatttgccg ggctatacca actttggcta aagatcaaca atccaatcct
157981 ttcaaccaat tagcagctct ctcttcaagg ccaacatggc ctccagcatc tccatatcca
158041 aaatttcagg cagagcaatg gtatacttgg gttgaaaacg taccaggtcg ttctataaat
158101 tacctagata gattaatatt gactgtaggt aactttttac cttccctgga ctttcaatat
158161 gattttccag cctgggtttaa tagaattttt gaacattgtc tagaaacaaa gcatattfta
158221 gatttgattg aaatttgcaa cctaccaa atttataaat ccacgacctt ttataatagc
158281 agaacatctc aattgacgga aaaaggagaa aaattaaaga gtttatttac gttgtttttt
158341 atggtattca ttttcaacga aatttatgaa aaccctaata cagaaactaa ctttaagtaca
158401 aattacgaca attttcaaaa ttttaactgcg aatgatctcg actttattgc ttactccaaa
158461 aataataatt taaaaactcc tcagattctt actgcgagat ttcgccatga cgttcatcgt
158521 gaaataagtg tggataaaaa actaatgtgga ggttcattta cattgattta taccatttgt
158581 ctttttaacg agttgtggtg gatcgagccc gtaaccgatt ttgttaaggg tggaaaccaat
158641 cagtctaata aaaaccgata taaacaagta ggctcgtgga accaaaatft ttaccaagaa
158701 ttaagaaaaa ttcttgaaat tttaaaaaag ttttataatt ttaatgatgt attatcaag
158761 tgtacttatt taagtgaaca acaaaaaaag gaggttaca atactttata taaagtattt
158821 aacgggaaag atccgtttta tgatactgat ttttaaggcta ttttaacttc tttagacatg
158881 taaaccatta actaaatcaa tttagaatga aaaaacaaac agaaaattta ggttacgaaa
158941 gtaagaaatc aaaaaataat aaagcaaagg agaatacgtc tattgccttt gaaatatata
159001 gcagagaacg tcttatctca cctggggatt atgttgcatt gcgtacagat ttaattgata

159061 aaaaaaccct ttcttgtttc tttatagggc agataggtcg agttggacca aaagtaaagt
159121 tagataagca ggaatcagaa acatatttta taaattttta tgatgaatta aataaactga
159181 agaatgttaa ttcttatgac gcttttgcaa ctccaattct tcacttagat gagatactta
159241 aacttcattt taaacagcag atagcaaagt attatcccgt taattcggta aaaaaagctt
159301 tagaagaagg tacaacgtta ttagaagaat atcaagaaga tttagactta attaataaag
159361 aagaaaacat tattattaac cataatttta tgccttctag atttgaccct ggaaatthag
159421 taaccattgt ttatgatttt gctataaaaac aagagaacga gaacatagta attttttctg
159481 gtcaagtagg tatgataaaa aaagtcttat caaataataa cgtatctgta gtatthagac
159541 aattaggatt aagccagtta ctcgataaag aaacgaattt agaagagata caagatatat
159601 cagcactaaa aaaagaagaa aactgggtaa ttgaaacgac aattcataaa agctttcttc
159661 ttctctttta ttatactaaa ttacaactaa ctgataaggg gaaaaaccaa cttttgaaatg
159721 cctaatgagg aacttaggct ttataagctc tgcccataaa tttgtataaa tagctgaaat
159781 tatacataga aaacttgggt tctatgcttc gccaaagtttt ctatgtggct ttattatggg
159841 ctatgctgaa ttcgaatcag cgacctcacc cttatcaggg gtgctgctca accaactgag
159901 ctaatagccc ttataacact atgggtatact aaaactaatc agtctgtcaa ccctcattag
159961 cccttgccct ttggtcttta ctctaaggac ctacactcct acgtagtagg aggagtaaag
160021 accaaagggc aaggggatcc taaacaaacg cagaagacca tggtaagaat acaagtctaa
160081 aatacaatcc ttctttaa ataccttagcg tagcccgga atctagggct cccgggtaat
160141 ttaagaaagt atatttcata tcgaagattg tacccttgcc gtgctctaa aacaaacata
160201 gaaaacttgg cgaagcatag aacccaagtt ttctatgggt ctatgcttcg ccaagtttgt
160261 tcaagtttta taacgcagaa gactttaacc cttgcccgtgc ccggcaggct tttctaaatg
160321 ttgaataaca tttagaaagt ttaaaagctc tgcctaaact acttaataca gctttacttg
160381 attacgtagt aaggttggcg aacttcgctc gcatacccat tggtttgtgc attcgcacaa
160441 acttttaaac ttaggcaacg ggcgaagctt aaagatccta atagaactct gttctathtt
160501 ggactttttt tctgctctac ggccttcggc ttataaagcc ttaccgtgtc cggccgttgc
160561 tcattagcaa cgcctaata gaagagcttt acctaaatgt tttacaaagg ctaataaaat
160621 agaaagttaa taaacctgct tttttcttac aaaaaatgtg caaaagccca cctatttcgt
160681 gctaaaaata ccacaagcat gaatactgta gagcctgaag gatttttgac agcattttgt
160741 aaaaaagtc gcgagttttg aatttgagct ttcatgcctt gttcatcaac taattcttca
160801 ttcgaaagtt catccataaa tgcggtaacca tcttctggta aaaacagttg atgaaagtca
160861 aataatagtt tgaaaagtat tgcctgcgaa tagattccaa acatataatc ttgtcttctt
160921 ttgtcgtata taccgttagt atgttttgaa agtttttcaa ccacgtcctc aattttcatt
160981 gcgttggctt ttaggcaatg ggttaaacia tgttctttcc atgactttgg gtggccaacg
161041 catgggggta ggttacacgt gcagttctatt gtggatacga gttctcctcg ccgtccctga
161101 tctataaaat gcgattgtgc cttagcataa tacgtttgaa gtctaacatc tcgagctcaa
161161 ttttcattat cagaatcaga aattttaaaa tctgtatacg gggactcca aggtgttggg
161221 atccttctcg ggggagcaac tgcctgtgcca ccccgcttc gtctgtttag attaccactt
161281 gaattaaaat cgtcatcagg tagagataaa ccgttgtttt ttaaaacgct gagaagttct
161341 cgccatgttt cattagttag tagttgtaaa cgcctagaat gtaagctctg tattactgca
161401 cgattcttaa tggtagccca atgtcggaaa ttataagtga ataatggtag atccagtata
161461 ttgactgcca taaaatcatg gaaattggta taataattag gaaaatattt ttttttaagc
161521 tccaagaacc taatgaaata agcttcaatc attttttcta gccaagactc agcggcaggc
161581 ggaattttaa ctttttgagc ttcgttcaaa tatgctgtta gacgagcaga acgagtagca
161641 tegtctgtat taattggtaa agattcagac gacgaagatg atcctgtggg attatttaaa
161701 ttttgcataa tatttttaaa ttgtaaaatg tcaaaaaata tattaaaaaa aagctttcga
161761 gtttaaaaca aaagcgcaac cgaaccaata ggcgatcata caccataata agtaagcttc
161821 caggttttcc agacgttggga ctgtatactc aacttgagta atatagtgta tatcttgtgc
161881 aatttgttta ttttgtaact cagggtttttg ttttaaaatt tctatattgc gtttcggcaa
161941 tttagaaaaa agatgaatag gttttcttcc aaaataattc ttggaattag aaattatgat
162001 accatcaacg tattcatcct gcataaaact aatgataagt tctcgataat aagaaatatt
162061 ttttaaaaaa tttcgacaat aaaaaagggt tagtttctact ccgaggataa aataaaaaaa
162121 aggtttttta aagtagcttg ggcgaattat attaatataa aaattgctat agcttttaga
162181 cgctggaaaa ggcaattcta ataaattaac ttttatgctg aaataaattt gataactgac
162241 agttagcaag acattgctaa ttatttgatt catattataa aattgatagt aaaagtttag
162301 tctttaccct ttggccgttg ctcatagca acggctaatt agtaaagacc aaagggcaac
162361 ggctaattgag taaagaccaa agggaagtga tgttgatttt tagtttaaaag taagaattaa
162421 aagaaagttt ataaataaaa ttgctaactt tattttacca taacattaa taataactac

162481	ttttatttat	aaacttttatt	gtttataaag	aaggtctaata	agagaaggggt	agtaataaac
162541	caattttatta	gaacttttgaa	acattgtttat	aatataaccag	aaataaaaaat	atagaattaa
162601	aagaatttttt	ataagataaaa	cttactagac	ttatttttaac	taaaaaaata	ataaatacta
162661	acaaaggttag	tttaggcata	gcttttaaatc	aaaatcctttt	gcgtagcata	agataggatt
162721	agaataatac	ctattatata	taagtccaaa	ataaaatcac	agctttttatt	aggatcttta
162781	agcttttgctt	aaagtccaaa	tcttatgcat	agcataagat	aggatcttca	actaattaa
162841	cttctttttta	cgcagcttac	cgtaaacttc	tacaatttta	tctagttctg	ctattttatc
162901	agtaccaa	gctgctacaa	atttacccat	atgtaaactc	agatactggt	tatcgttcag
162961	gggtgcgcta	gctgattttt	attcggctta	tgtctctggc	aaagaccaa	aggtagaggg
163021	ttggttaggg	gtagaggggt	ggttagaggg	cgcagtagat	tgagtagcct	tatttccggg
163081	taggtagtgc	cctagtctct	caaaatggtt	atataaaagc	tgcgctgagc	tgaaatcatc
163141	gggtgccaag	agttggaaag	ctcctgacat	gaaaaaatca	ataagtattt	tttcgttggt
163201	ttttccttgt	agtacatatg	caggatttaa	taatgtgtaa	gccaatttta	aattcccagc
163261	tggatcta	ctgctatctt	ctttgtattt	atcctggtca	caacgatgaa	atctcgcaat
163321	tttagcggtta	tcgcttttct	cgtattcttt	tagtctgtca	acgcagaact	gatcaaaatc
163381	tgatgtttgt	tcctgtgaca	atttatagtt	ttctgaaaat	ccatgctcct	tataatattg
163441	tttttttatt	tcttttgaaa	ctgtagaata	atcttctgct	gttacgagct	ttaattgtaa
163501	tgtttgccgg	gctttgtgaa	gtgtttcttt	ctggagttcc	cacattatgt	ttaaaattgg
163561	acgaaaaacg	ctctgggtcaa	gaccttctgt	aaattttgtc	ctacagtaat	tacagctcag
163621	cttattacac	gaattacaat	cgttattttc	ttgcttatcc	ttaagatcct	gaagaatcgc
163681	cttttgcttt	tttttttctt	gtttattttt	atttttataa	aaaatgcttt	tacctaaatt
163741	tttagtaaaa	tttgctttta	gatatttcag	gaaatctggg	agaaaatctt	cgtaaagat
163801	agacataa	agtaaattaa	attaattgat	tatagcccc	gttgtctaaa	ttaaaaaacg
163861	aagtttttta	ataaagtttt	acccaaaggc	tttacgtcct	tacccttgg	taaggaccaa
163921	ggggcaaacg	gtaaaacttt	atgttttatc	ctagataaaa	catttaggga	aatcccgttg
163981	ggtaaaat	ttatgaataa	atatctaacc	tgggagtaaa	tgcttttgta	agcattacaa
164041	aggcattaac	ctgactacct	ttggtagtct	tggccctaga	ttcctaaaca	aacgcagaag
164101	accatggtaa	gaacacaagt	ctaaaataca	atcagagatt	gtaccctaaa	caaactacgt
164161	ttgttttagat	taggtccttg	cttagatggt	tacttgtaaa	catcttagcc	cggcagggca
164221	acgggttaaag	tcttccactt	cgtggaagtc	caaagtttgt	gtgaaacaca	aactaggaaa
164281	gcaagctttg	cttgcttact	ttgtttaaca	acatagtgtg	taaacttaga	cctaagctaa
164341	acaaacgtag	cgaagcgcag	aaaaaaagtc	caaaatagaa	cccgggagcc	taagtataaa
164401	tagctggaag	tacgggtgtg	gcttgccacac	acctatttat	acaaagtttt	ataagcagag
164461	cttataaacg	ctagatcccg	ggctacgcaa	agttctatta	ggatctttta	gctttgcccg
164521	ttgcctaaat	ttaacaactt	cgttgttaaa	taaagttttc	cctaaacaaa	catagtttgt
164581	tcagattagg	tcctaaatgt	tttatcttag	acccttgccg	ttgcccggcag	gcttttctaa
164641	atgttgaaca	acatttagaa	agtttaaaag	ctctgcctaa	actaccggag	gtagtaagggt
164701	tggcgaactt	cgttcgcata	cccattgggt	tgtgcattcg	cacaaacttt	taaacttagg
164761	caacggtaaa	acataaagtt	ttccctaaat	gttttatcta	agataaaaca	taaaggttta
164821	cgcttcgcgt	aaagcctacg	gggaaagcct	tcggggaaag	cctgccgggc	taagatgttt
164881	acaagtaaac	atctaagcaa	ggacctaatc	taacaaacg	tagtttggtt	agggtacaat
164941	ctctgatatg	aaatataact	tcttaaatta	cccgggagcc	taagtataaa	tagctggaag
165001	tacgggtgtg	gcttgccacac	acctatttat	acaaagtttt	ataagcagag	cttataaacg
165061	ctagattcct	cattaggcat	tgctaatagaa	caatggccgg	gctacgctaa	ggtaatttaa
165121	ggaaggattg	tatttttagac	ttgtgttctt	acctgggtct	tctgtgtata	ccgggagaaa
165181	acaaattaaa	aagcatcgta	gcctgggaac	ctagggctcc	caggttttta	atttttacca
165241	gaattattta	gattctgaga	aatattagtt	aaaagatatt	ttcgtttttc	ttcaggtagt
165301	ccaaatttag	ttaataatac	ttctaactcg	gttatattga	ataaatttagg	tgtgggtaag
165361	ttataacttt	ctaagaaatt	aattattttt	gagaaacttc	ttgttttaaa	agtttcatta
165421	actagcagct	ggatttgagc	ttctttaagt	tgtaaaagtt	tgagaatagt	atataaagtt
165481	tgtcgaggag	cttgatatata	ctgaaagaat	aaatctggta	atctttcttg	tctaagaaaa
165541	aagttttttg	ctccgttaat	agcaaatca	gggtaaagct	caaaaaacca	attacgtctt
165601	ataacattaa	gtaattttat	ttgctttcca	gtatcgatga	acattaaact	atatagcaac
165661	ttggaccaa	tatttgata	gagttttcct	gtcgtttcaa	gattatcata	aagaataaaa
165721	aatagagcag	aaccaaaaat	ataatcagga	atccattccc	aagccggaga	ttgttttaaa
165781	acgtctaaaa	tttcatcttc	atcctgaaaa	gattcatcag	aaggtaattg	attccactga
165841	ttaaacagtt	gttctaaatt	taaattagaa	gaagcttcat	caaaagcata	ataagcga

165901 tccagcgaga atgtccaatg gtatggacaa tttgggtccaa attccttttc aaataaccct
165961 aaaatttcag atataacccc ccataaggca ttagtttcgt ttgtcatata aatttctaaa
166021 ttgaaaataa ttaattataa gctttatfff tttacatata ttttggatta tcaagtaatc
166081 caataccttg aattaatata cctttagaac tttttacttc aacaacaccc tcaatttcaa
166141 acttacttac gataaatatgt attaatatff ctttaaaagc acgtcgccca atctcaggat
166201 tattgccact gtatfffgaaa aagctcatgt aaagatcact taatcgctca gtacaagcag
166261 gttccagcac aattctatcg gtaacaaagg catctactag ttttttatct tttaatcgga
166321 gtgctfffag tttaatcata ttatfffct caagacgagt ctggctctgtg cggcttatac
166381 tctccattgg aaatftacac tgtagtatat ttatatattc aatgcagggt gaaaccgctg
166441 taattaaggt taaattctga ttatatfta ccgaaatagt tatatftaaa caagaataag
166501 tftftaaacct aagctctaac acattgaaatg tgacccttgc ttagatgftt actftgtaaac
166561 atcttaatcc gacaggctft gattagaag gaattggtaa atftattata ttatacatag
166621 aaggactfta agctftgccc cttgattfta aagfttgtgc gaatgcacaa accaatgggt atgcgaacga
166681 atftgftaaa cttgactgag aagacctagg taagaaaaca agtccaaagt ttgtgtgaaa
166741 agttcgccaa cttacgcag aagacctagg taagaaaaca agtccaaagt ttgtgtgaaa
166801 cacaaactag gatcttgaag cttcgctfta agtcttctta cttftftct gcgctgcgct
166861 acctccggta gtttaggcgt agctftfta tftagactta aacatagaaa acttgggttc
166921 tatgcttgc aaagftfta atctftfta acaactftg tftgtftaatt aaggaggftga
166981 tccagccgca cttccagta cggctacct gttacgact cactctcatt accagccctg
167041 ctttaggcgc tctcgctcct tacgggttag agttacgact ttgggccaag ccgactctgg
167101 tagtgtgacg ggcggtgtgt acaaggccc ggaacgtatt caccgctgta tggctgacca
167161 gcgattacta gcgattccga cttcatgcag gcgagttgca gcctgcaatc cgaactgaga
167221 ttaggtftft gaggttagct tgctctcgcg aggtagcatc tcattgtcct aaccattgta
167281 gcacgtgtgt cgcccagggc ataaggggca tgctgacttg acgtcatcct caccttctc
167341 cggttatca ccggcagctt ttctagfttc cttgacatat cgtagctftt accaaagtgt
167401 aggtaaagac caaagggcca acggtaatca agcgaagctt gataaacctg aggcgtagc
167461 ctggctftta taagcgaagc ttatftaac ttgatattgt tcttcgaaca atatfttagga
167521 acctaggaag aagacttcca tctttacct gtccggtaag gacaatggcg aagtggaga
167581 tctagtgtg tgcaagcaca cactftggac ttgtattctt cccatggggc tcccaggftat
167641 ttatacttag aaaagcaact aaaaacaagg gttgcgctcg ttgcgggact ttgaccaaca
167701 tctcagcaca cgagctgacg acagccatg accacctgtg tccacctcta aaaaacaaa
167761 aggttctag taaaaacctc tftttcaaga gattaaatgg catgtcaagc cctggtaagg
167821 ttcttgcgct tgcatcgaat taaaccacat gctccaccgc ttgtgcgggc ccccgtaaat
167881 tctftttagt ttcactcttg cgagcatact ccccaggcgg gatacttcac gcgttagcta
167941 cagtactgga cgftfttagt cgcacagcac ttagtatcca tcgfttacgg ctaggactac
168001 tggggtatct aatcccattc gctcccctag ctttctctc tgagtgtcag tttcgccca
168061 gcagagtgtt ttcgccattg gtgttcttcc cgatctctac gcatttcacc gctacaccgg
168121 gaattccctc tgcccctacc gtactcaagc tcattagftt ctactgccga ctcagggtta
168181 agccctgatc tftaacagta gactftaata accacctaca gacgctftac gccaatcat
168241 tccgataaac gcttgcattc tccgtcttac cgcggctgct ggcacggagt tagccgatgc
168301 ttattcctca gataccgtca gaacttcttc tctgagaaaa ggagfttacg acccacaggc
168361 tgatcatctc cacgcggcat tgctccgtca ggctftcgcc cattgcgga aattcctcac
168421 tgctgccttc cgtagaagtc tgggcgctgt ctcagtecca gtgtggtga tcatctctc
168481 agaccagcta ctgatcgttg cttggtagg caattacccc accaaccagc taatcagacg
168541 caagcccatc tftaggcaga atactattcc tfttactcct cagcttatgg ggtattagca
168601 accgtftcca gttgttgtcc cctcctaaa ggtaggftct tacgcgttac tccccgtcc
168661 gccactaaga atftaattft tftaaacct aagcctaaaa aatftaacct ccgtctgact
168721 tgcatgtgtt aagcatgccg ccagcgtfta tcttgagcca ggatcaact ctccatgaaa
168781 ctttttagt tagaccgtag cacagaaaaa gtgtcttctg tgaaaacaaa ccaaaatc
168841 actcagatag tggattagga tcttctacta tctagaagft tctgtgttg aagtcagaga
168901 aatcttagag gftttactft aagftftcta tgtgttatac gataacctac aacttatatt
168961 tacgtcaatc cagtcgcct ctctftaagc tfttcatagc cgggfttga atcatcatct
169021 agtctftggt cttagaatft tftaatctta cgtagtatag atcatcacta attcaagtct
169081 agtagctfta tgaatcccta aaacctta tfttggccag tfttgaatat tgacctgatg
169141 tattccataa attgaggctc ttacatatta gttgtagctg atctaaatcc gcctfttctg
169201 cgattataga aatcttgtft ttcaatgcgt ctactcattc tftacattgt ataaacattt
169261 taagctattc cacatftfta taactaaatt ccatggcgtc aaagtatcta tcttagftat

169321 ttaaaaaaag tgtatgcgct tcttgtctac atctttttaa ctcatttgat agctttacat
169381 tagttttata aatatttgta gctctgtcga gaaattgtat ttggtgctca aaattttaa
169441 aaaactcatc tacatattga tatgaaagca gttgggtgtt tattcttact ttttctaaag
169501 ctaacacatc agcaagaata cccttatgta ctttgaaagc gtgatacgct ggcaaactcc
169561 aacctttttt aatttggcgt ttctttataa acatctcatg tatgtcattt tccatcatca
169621 tagaagctat tacttttagca tagtgaattc gctgtctata attctcagca aaaccttctt
169681 tttgatgatc atcttcaccg tctttcgaat taaaagatgt aataaatcct tctcttagaa
169741 gtacatcaat atcaacacct tcttcttttag catgagcata agttatttca cttaaaaact
169801 tattctgacc atcagttcta tcattccttt tctactaac ccgaatagtt atccccaaat
169861 cataagggcg ctctaaagct tcgcaatctg gtttttctaa caaatgggtct gcttctgtat
169921 atataatgct tctacgggat ttatcatagc cacaactctc aaaaaatatt ataagtttct
169981 taataaattt ttctgcctcc ttataagatt ctactcgagc tgcagcatct attactacct
170041 gtccatcatg aaactttatc caaataactt gacctctaaa tttctaagga atatatctct
170101 cgctttgggtg caatgtaaaa ttctgaatca aagataaatc ctgaagcgaa tcatacacag
170161 ctttactcctt acgaatagac tgttttgata tgataatatt ctggtagttt actaagtctt
170221 tcatatttgg ttgtatttta ctttctgaaa atgtcaataa acctgattct accccacttt
170281 tttgagatgt caataaactt ggttgttcta aagagctgga tattgtttcg cttgctgtat
170341 ttcttgccca aaattttcta aaaaattgaa aataattaaa aaacttcac cagttatttta
170401 caagagggtt aaaaaaccta atggatatgc gaacgaagt cgccaacctt acttattgggt
170461 ttaggcgtag cctggcttta caagcaaagc ttgcccttgt tttactttc gcgttcaaat
170521 aaaaatttat agtatatcag cttagctgat atactattta ggaatctagg ctttacaagc
170581 tctgcttgta aaactttgta taaataactt gttatttata cttaggccaa gactaccaac
170641 ggtagtcagg ttgatgcctt tgtaatgttt acaaaggcat ttactcccag gttctatcct
170701 ttgcttcctt gtttgaacat tttatgttca aatgagttct atactactgt tgccttaata
170761 aaaattgtat aataatgaat ttaaacccaa aacaaattta ttcattttca aaaagggaaa
170821 tttacagtaa ttatcctttt actgacgaat ttctttttac attagttaat gacccttatg
170881 tttctgtatc gaaaagctta gaagaatttc tttcgtttca tttagatatt gattctgaaa
170941 tcaacattaa ttgtggagtg tttaaaaatc ttgatcgctt agataaagag gatctcgaaa
171001 tagattttaga acaagtgtta aaaagtccaa atttgattta tgatttagaa aacgtgtata
171061 agattaacgg gctcatttat tattataatg caaagtcaca agtttactat aatttaacaa
171121 atattaattt tagtgatec taagttttat ctagttttaa ttatttttt atacaaacta
171181 acaaggatgt aaataatcag aatcgttggg aaataattac tgaatacttg tttacaaaa
171241 aatttcgaaa aaaaaatttt tggatattga aaaaatttca aatttttagaa gaccctacta
171301 ttaaaaaaca atttgttcga aatgttagta gtcgatgccc gcaatacttg tccccactcg
171361 catcttcgtc aacagcgcgg tatgcttcta caatgctttt tgatgtatta aattttattg
171421 gaaaaaataa aatgaaatcc ttagaactag tttcaaattt taaaactaca gcaaatccaa
171481 agcctatttt aagatatatt taagaattgg gtagtgaat ttctggaatt taatccccct
171541 ctagaatcgg atattcttgg agtaaaatcg gacaatactg aattggataa tttggtagct
171601 cccgacattt caacaaattt gaatctatct gctctagacg atataagttg gtcagattct
171661 agcaaaacta tattgttaaa gtgggcactt attcacacca atattggatt ataaaagtaa
171721 ctaattttatt acttcagaga aactagaaat ccaaaactac taaaatgatt aattaatgat
171781 ctacaaactt taaattacac aataccaatt atttaaattt aaatttgggg tttaaagctt
171841 taatttgtat ttatatatac gtatataatg ttatatattt aacttgcgta aaactataaa
171901 aaaaatataa aataataatt aaagcctact tagctcagtt ggtagagcat ccgtctcata
171961 cgcggtatcgt cactagttca agtctagtag taggcatttt tattaaataa taattatttt
172021 attgactttg ttttttagtt tttattaatt aattggaaaa gaaagttcag atactagaag
172081 ataaaaatta atgaataaga cgatccacta atagacaact agatgctgta tactaaatat
172141 tactaaatat tactaaatat cattagatgt ttattatcat aaaaatata agaataattt
172201 tttgattttt ctatatattc tcgattattt tagtaatatt tagttatttt tctttataaa
172261 aaacattaaa gatcactgat tgcgcgcgga attgcctccg gtcggactcg aaccgacacg
172321 caattacgca ctagttccta aaactagcat gtctaccaat tccatcacga aggctatatt
172381 ttattaaggt atatacttat ttaataatac gaaataacat aattgatgtc aatactcttt
172441 aggtttttta aaagactagg tttctccgta ggttcaaact tgaaaaagtt ttgctgattt
172501 taaacaataa cataatttat tcaaatttcg ttgttatcta aattgtattt tatttgggcg
172561 ttattcaagc tttagcacia aacgtagccc cagattttaga taaaaatttt agaaagttaa
172621 aatctttgggt ttatgtttta ttttctaatt cctaaaataa acaagtattt tttaaaactg
172681 aacaatttta gtagtaaaat aaacgcccgt tattaaaaga ataaaagaat taggaattct

172741 gataagctta tataaattaa accttcgggt atataagatt taccggggtc ataataattha
172801 ttatcttgcg aacttttaaat tgttaatttt ttaaagtca tagcccccgt tgcctgccc
172861 ggcacggcaa gggaccattg ctccattagca acgcctaag aggaacctag ggctcctgag
172921 tttattttgt atctaaatth cgcgctaata ataagataaa acaaaaagtt ttaaataat
172981 cataaggctt agaccgtgtg acgataacaa agtcttttta actttgtcaa aaaataatth
173041 taaatgctta aatttcttaa acttaaact gaaagttatt tccactaaac tacgcctaaa
173101 attttctctt agcgaataaa atttagttatt ctctgaatat taacctgaa ttgagtagct
173161 tttttatgca aattagaaaa cactagaaaa cgtataatat gttcaataaa aacataaatt
173221 taatagtaga accaacaag aaactttgtt acaattaaaa tataaaagtt tttataaaat
173281 aattaatthc atatthttta agaaaaatth aaagtttagt tatatcaatt aataagatct
173341 agtacttaat ttagattcta taaaaaaat aatthttattg tgtcataaag aataacaaaa
173401 aagatttata aaatthtcga aaatthtaag acaactgaaa ttaggctgtc gcagtttaat
173461 caagaaaaat aatatacaa ttaaacttta ttcctcgaac aaaatthtaa cttagctata
173521 aaacatattt tcgatataaa ctaaattcta tttaaaaatt aaagaatcct aatctactct
173581 ctataaagaa aacttgctct ccaaaataca accgtaagtt ttattagaat cttgaagtat
173641 tgcttcaagt tttacataa agctthtaag ttatagcttt ccaaggtttg tgcgaagcac
173701 aaacaaaaat cttagactaa ctagaagtct taaaaaatt taattcaatt taaatattaa
173761 ttgccttcag gcaaaggaat aaaaaacatg gctgtacca aaaaacgtgc ttcaaaaaaca
173821 aaaacaaaga cacgaagaac tttttggaaa aaaaaagctg aaaaacaagc taataaagca
173881 ttttcattag ctaaactcagt attaaaaaac ctgctcaatg aagcaggaaa taaataatta
173941 tcaaaatcat aatthtttaa atttagaat taagaccctt ttctaagtgt gaactthtag
174001 ttcacactta gaaaagatcc aagtaattht tghtaatatt agtagaattt taaaacaagt
174061 cagtatatac gctthttaat aaatthttaa gtgatacatt aactthgaaa tagatthtag
174121 agtaaatagt aataacgtgc attaatthtatt attataatct gaaatthttgt atatthgaga
174181 aaaattagaa tttctcgtaa ctttatactc gtctattaca ttcatthttt accgtagggg
174241 atgttcaaaa agtaaaagtg aatthttgaat agcgaactag gagtaggcac tttgtgttta
174301 aatthttaagg gagagcaacg aagctthttta aaaaagatta ttagcttcaa aattataacg
174361 taataattac aggtcaatga atcagtaaac tggagaaaaca ttgtthttta aaaaactatt
174421 ttaaagaact aatatcgaat tctcgaactt ttgtctcgac ttttgcacaa ttaaatagca
174481 agtthttattg ctttaactg actthtaagat aatattgact gctctgttht attattctth
174541 ttaacaaaa acaatthaaa gtctactatt ttaatthttta gaaaaaaaat gctthaaaact
174601 aattataccc ctctthttat tttttgaata catttcagtt ctagtttggg ttttactac
174661 gtgaaccagc gtatthttaaa ttttagagaa acttcgggac atttccacat acttatatth
174721 tcaaatataa cccgaaacgc tatatagagt tagttattag ctaatthttta tattgtgcac
174781 tatagaagtg aataatthtag atgagtcgga gtaatthttc taaagaaatt caatagcgta
174841 ggcagaaaa aaagtccaaa atagaacaga gttctattag gatcttccac tttgtggaag
174901 tccaaaatgt gtgcttgcac acattaggat ctthtaagctc ggctthaaagt cttctgcgth
174961 atthtaacat ttctaaatga aaaatthtaat tataaattth gttthaaattt taaaaaaatt
175021 atataacatt caaggagaag tttatgacgc ttttcaaaaa ctcatggcct tggctthttta
175081 ctthattgta tctthtattt atgcttattt tgcctataac atcgctatta actacggcaa
175141 gtcaaacatt agtgatacga ttttgggaaa tagcaactga accggtagct gtttcagctt
175201 atthttgtaac aatthcaatg gctthtattg ctthgtttct taatgctatt tttgggttta
175261 ttcttgctgt ggthtttagtc cgtataaatt ttcccggtaa aaagttatta gacgcagctg
175321 tagatthacc ttttgctthta cccacatctg tagcgggtth aacattagcg actgthttata
175381 gcgatcaagg ttggataggt tcaattctaa atgattthtg atatacaaat gthtttacac
175441 gthtaggcat tacaatagca atgattthttg tttcatthtc tttgtttgt cgaacattac
175501 aaccagthttt acaagaaatg gaaattgagt tagaagaagc tgcttggct ctthggcgcac
175561 cacctthgca aactthttata aaaatthttt ttcctccatt agtaccagct ttataaccg
175621 gttaacatt agctthttca cgtgcaattg gtgaatatgg gtcaattgtt attgtctctt
175681 ctaatthttc tttaaaagat ttaatcgct ctgttctaatt ttttcaaac ttggaacaat
175741 atgattatat aggcgcaaca gtaattggaa cagtgggtgt agttatthct tttttttac
175801 ttattagtat taatthattt caatcttggg atcaaaaatta caaaatgaa taacagcgta
175861 ttcttgatac gtcaagcaaa gcttgactac cgaaagggtc aaattattaa gtgtgtttta
175921 ctaacgttaa ggtatthacc tttgggtgtt tattgtthtt atagcagcta tttaaattht
175981 agggctthtt ctaaactgtaa actthtcagtt tacacttagt aaacatataa tttthttaaa
176041 taaggttgga atthgtcaag gthtttctaa aaaaatthta ggaatcttac tagataggt
176101 aaaaataaaa gaccttaagt tttatacaac ctaaattgta tagttacaaa aatattaact

176161 cgaagtataa atcttttggtt tacagaaaat tttcttcata gaatacagtt tcatcaaata
176221 ccattaaaaac ctgagagcaa acctacagct ttagagttat atacctttga tttaaataaa
176281 ttggcaatat tggggacctt tactgatata taaacgtgga aaaaatcttc tttttagatc
176341 cctgaattcc gatttacaaa agattatggt tctaaatggt ttattaaagt tttattaagc
176401 atagcttgct tgacttgaaa ggtgaaacct tacgaaacac aacgtagcac agaaaaaagt
176461 ccataataga gctacgctct attaagatct tcccgaatca aacatagcga agcgcagaaa
176521 aaagtccaaa gtatacaaac tttgtttgta tactaggatc ttaaagctat gctttaagtc
176581 ttctgcgttt gtttagatta cccgggagcc ctagattcct cattagcat tgctaatacag
176641 gaatctaggg ctcccgggtc taaattgaac aacaaagttg ttcaataaag ttttcctgt
176701 cgggctttac gcgaagcgta aaactttatg tttaccggt gccctgccgt caacggcaag
176761 ggtctaagat aaaacattta ggaaagcctt tagggaagtc taaagtgtgt gcttgcacac
176821 actagatct ttaagcttcg cccgttgccc tgccgggcac ggcaagggtt aaagtcctct
176881 acgtaaaaaac tttatgtttt tacttgggga aagctttgat aaaacatcac ctttactctg
176941 aagggttaag ccttgcaatg tttacttttt ggggaagcaag ctaagttcat agagccacag
177001 cacacgatat tatattttat gtttagcttg gttagcagtt tataaataga tttgttagat
177061 aactagagtt cttagtttct gtaaagtaca aacttagggc acatactttt atgcaggaaa
177121 aaataaaaca tttaaaaaaa ataacgcaa aatattttgc caactaaaaa ctaaagggtg
177181 ttatcaaatt tgatctttta ggttaaataa gttatgatta caaaactgaa acaagcttca
177241 cttatttttag ataccaatth tcaacttttt tttcgcaggg cttttacaag tgtactttat
177301 cattgtttga agttgcttgt tttcaaatga aagctttatc atttatttaa ttgcaaattc
177361 cacacttgtc attgtttaag ttttaagaaga aaattattaa acttagtttt ttcagtaagg
177421 aaagacctac taagtatgat aacaagtaaa gcttgggaaa gagggcgggg tcaaggctac
177481 aggttttcat aaaaggctga ccgtagagtg acacgtaatc ttattctacg attaataatt
177541 aaagtataga taccagataa accattcaga tatatgatta aagctcacta acctaattat
177601 ctccgaccct tagtcattag ttttttatct tttgagcttt accaaagggg attgggctaat
177661 gattatttta aagtttccac ctttgtcccg cccaaaggct tcaactccga ggctttagcc
177721 ctcacctggg agccctcaat taccggaggt agtaagattg acaagcttcg cttgtttacc
177781 catgccttcc actacgtgga agaccttaac cgttgccctc ttgggtaaga tatttatgaa
177841 taaatatcta acccaagagg caaaggctaa atttaacaac gaagttgta aataaagttt
177901 tccccgaagg ctttacgcca agcgtaaaac tttatgtttt atctacgata aaacattttag
177961 ggaagcctg ccgggctaag atgtttacaa gtaaacatct aagcaagggt acaactcgaag
178021 attgtatttt gcccttgccg tccccggaag ggcgacgggg gaagaacaga agtctaaaat
178081 acaatcagaa attgtattaa atcactatgt gaaagtccaa agtgtgtgct tgcacacact
178141 aggctttgcc cgtagggcaa acctttatth aacaacgaag ttgttaaat tagacctaat
178201 ctaaacaac tatgtttggt taggatattc ctttaggga agtcttcttg acaggttcct
178261 aatatattgt gccgttgga gcaatattta gttttaacaa ggggagcttg ccgttgccct
178321 acccccgttg tctaagttta aaagctctgc ttttaacgc agaactttt aacccttgcc
178381 gtgcccggca gggcaacggg cgaagcttaa agatcctaat gtgtgcaagc acacattttg
178441 gacttcaca aagtggaaga tcctaataga actctgttct attttgact tttttttg
178501 gctacgttac cgaaggtggt ttaggctatt gctcattaac aatagctaat gagtaaagac
178561 cagggggaca acggcaagtt caattagccg ttgcggtgcc tccttaccgt gtccggtagg
178621 atttgctcta ggagcaaac tttatttaat aacttcggtt ttaaatttag ctttgccctg
178681 ttgggcgtag cctgggaatc tagggctccc aggttagata tttattcata aatatcttac
178741 ccaagagcgc aacggccaaa attatcaagc aaagcttgac taccaaaggg taaacttgaa
178801 caagcgtagc ttgttttaaa gcaagattga caaacatagt ttgtctacca aggagggtca
178861 ggtatgtttt gaaatatggt attttaaaag atataagttta ctattaaaaa ttaattatgt
178921 ctttggttac agcaataaaa gattatattg acgtattaaa ccatgtatat gattccgctg
178981 gtaatgtaac tatccaacaa atttttcagc aaacttttgt atatattttt tcaagcatta
179041 agtttctagc tggttatttg ttaagttttc aatggctacg tgatttagca tatttacctta
179101 ttctagttcc agaaatttct acaaccattt taaaaggaaa ctatttttta gatgcccgt
179161 tatctaattt ttttacatta ttagaaacac caacatatga aaataataaa tttttaattg
179221 gctttttaaa tagctttttt ttgtgtttac cattgtctac aaccattttt atttctataa
179281 gacgtttatt agttcaaggt ataccagctg gtattgtttc taatatagga acaatttttag
179341 gatatacttt ttttattttt tctgttctat ttggtttacg attggttata attccctggg
179401 tagcatatga accttttagc tatattattg gagtcttatt tactttaaca attgtttatg
179461 atatgactca tgaacgtact attcgtaaat ttgattgggt ccaaaaatct accttaatac
179521 aaattttttt tattaatttt ttattaactt ggactgaaca aagctgtatt tttagttatt

179581 ttggaaattht aacttttaaac ccacaaccta ctatthtttga taatthtttct tcaactaacia
179641 aatcaggatc tathtttagct catagtacat atthtaattgg tatattatta ggaagtgtct
179701 tthtttactag tthttttact cttatatgcc tacaattaag caatctatca ttaaaattat
179761 ctgcagttcc ctactctcgt tggttgagaa gaactaattht tgtcttatta acaacaataa
179821 tagctthttac tthttgcctcg attccatthtt atagthttaga ttatctthttt ggtgggtccac
179881 ttggthttcat ttctcaggat aaagctthttg acaaaactat tthgtcatcc aagaatgtag
179941 aagattctac tcggatgtta ggagaaaatt ctaatthttaa atcactcggg acagatgtta
180001 ctctthttcga tagaggcagc tathttattgc cagattcaaa tgaaaatthtt gaagatthta
180061 attatcaagg ggaatacgtc tggacagctc gtaaagatca tcgagctgta tatggtagctg
180121 agaaatctcg taaattattht tcgaatthttt ttaaaaaaac tgmtcaaaac ttgagtcaaa
180181 aacaagatga aatacaaaat ccgattthctt ctcaaaaccg atattcagca cctaagaaa
180241 aaaaactctcc tctcgcggat cgcagttcag aacagaaaatc tthttgacttg aaaaataaag
180301 taaaactcaa tcagttagat tctaactcag acctaataga aaaaagcgat tcagaaaaaa
180361 tagtaggaca aaacgaagtc aataactaatg aatthtcaga taatthtatct gatttcgaaa
180421 ccgacgattt atatggtagt tctcaaaagg gtattccaag ttcttathttt tataagctta
180481 atcatagaat taatactgac tthtaaaataa atcaatcttht acaacctthta ttagatacaa
180541 gthttttctcc acaatthtttt ggagaaacct taaataccct agatcctgaa ttagaaaaaa
180601 ctcttaaaaa gagatactat tcaaatcctg tthataaaac tthtactaagt gctgatattg
180661 atthttthttt acgtcgacaa ccttcttcat atthtattaag cccgaaagaa gaaaaaaatt
180721 tathttcaaaa acgattaatg ttatcaaaact atthtgatag cctccgggat tatcaaaaaa
180781 tgccgtatac aaaagaattht caaaatthttt ttagtggttc aaaaagttac gcagatcggg
180841 tthataatca acaattcaaa ggaacattaa aagtcgthtcg tcgthttattht tctattactt
180901 ttaatgaaga agaaaatthta cctgaaaaaa ttacattaaa atthtgatcaa cthttatata
180961 aaaaattgaa acaaaataat aatthtaattg gacatgaaga atthagtcgta acaaaaaaaa
181021 ataaaccttht tattgaatta gtaaatacaa tthctthttta tactggthttg gatgagcaat
181081 taagaaagct tgttataaca aatagattat taccacgthc aacagctgga tataacatgc
181141 aatthtgacag taaaaaagag ttgacagatt acctcaaacct tagtgagctt ttaaaatcaa
181201 cgaaaaaagt agatthttact acatggccat taaccaaaac taacttagaa aagccaaaag
181261 aggattthtaa aatacctthac aatgtacttht atgagcaaat aagtgaccca caaaataaac
181321 thctagcaga gtcattaaaa gacctcgaag gatcaccttg gacatttgaa actgtgccac
181381 ctaatthtaaa aaaaattgac cctgataaaa ttaatgacgt thtaccctct actcgaggag
181441 gattcgttht gtcaggtcat tcttctthtaa aaatthattt taaaaaaatt tthaaagaag
181501 gaataaataa atthttatacg tthaaaatgc cthttttctca gctttaatata agtccgggaa
181561 tthaaaaaag agatthtttcc cthttttagcc tthtaatttg tacataagat atthtaatta
181621 aaacttgata tthttctctaa gthtaaacctc tgcattthgct gthgtctaaa ctaaaaaacg
181681 aagthttthta ataaagthttt aaccaaagg gtaaatccctg tgatctthtacc cattagcta
181741 ttgttaatgg ccaataactt aaacagcgtg gcgcagaaaa aagtccaaaa tagaacggag
181801 thctattagg atcttccact ttgtggaagt ccaaaatgtg tgcttgcaaa cattaggatc
181861 tthtaagcttc gcccgthtgc cgtccgggca cagcaagggt taaagtcttc tgtgtaaagt
181921 tggcaagctt cgtctgcttg ccaagggcta aactacttga acttaccggg gtccgttaag
181981 gthtaagtagt thtagcccttg ccgtgcctcc thaccgtgtc cggtaggatt tccccgtagg
182041 cthttacgthc thtacccttg gtaaggacca aggggcgaag cgtaaaactt tatgthtttac
182101 cgtthgcccctg ccgtcaacgg caaggttcta agataaaaaca tthtagggaaa actthtattta
182161 ataactthcgt tgtttaaattt agcctthtgc tgtthggcgt agcctgggaa tctagggctc
182221 ccaggtthaga tathttattca taaatatctt acctaatagg gcaacgggca aggttataaa
182281 gcgaagcttg attaccgaag gggcgaagtt gacaaaggca atthtatctac caaagaggtt
182341 tagagtctat tthttthttaa acaaacatat tthtaagcaaa gtaatthttt ccttgcaatt
182401 thctthtaaaa caaagtaaaa thtagatagt ttaaaaaatta ttgcccgtat agctcagthg
182461 tagatcatct ccttgccaag gagaatgtcg cgcgthtcgaa tcgctgtacc cgttgatca
182521 atgagaacct gcataaacta agatctthgtt gthgaacaag tcttgaaaag aaaaatgaaa
182581 tatgcataat ataactatat aaaaaataaaa aataaaagg gactgtagthc aattggthtag
182641 agcactgccc tgtcacggca gaagthtgcgg gthtcgagthc cgtcagthccc gtataagatt
182701 atthtgaacta tathgtthtaa cccaatthta cthtacttht tathtttgtht aattaattct
182761 attacgatta tactgtthttt atthgtactta ctagcaatct gatctcaaat gthttttataa
182821 atagthgttht tthttatcttht thattaacat gtagctthtag tthttaaaatt taaagccagt
182881 tgaaaaaaag tthctthttca atgctcctgt tthtaaacag actthtagthc ccaactthta
182941 tthtaagggt tgactthtaaa agatacatgt ttaaccatac tatctthtcat taatagaaga

183001	acttttccatt	gggttgccac	ttgctttacc	ttaaaaat	gctcatagag	caaaatttga
183061	atgtgaacaa	agcactagaa	atcacgagacg	aggatatttt	atgttcaa	tttagggcaa
183121	agcaacggct	gaaaaagggt	ttgttcaata	aataaaactg	aaatttagat	ataaactcgg
183181	tataaacgta	aatgtaaaaa	cttagatcgt	agaccggggc	tttagtttt	ttaaatataa
183241	aatctagagg	attaatataa	aaaagctggg	tctaaactat	aagggtta	tgaaaactga
183301	cggtaattat	tcctcttgac	tagcagtttt	tacatataaa	attaataaaa	atgaagttgg
183361	aattaaaata	aataatgcgg	ttgcaattaa	tccaagaata	tttacttcca	tgataattct
183421	ccttagtgat	atrttagta	tattttttt	aaaaaaatag	ttcttgattg	aaatttcacc
183481	ccttatataa	actattagct	cacaatagtt	tggtctaagc	ataaacaaaa	gctttacgct
183541	aaattttaca	agcaaagggt	gtaaattctt	taccattagg	tttttacact	taagaccttt
183601	aggctcgacct	tgaaattcta	atrttagccta	ggttccattg	caagccaagc	tagcgtgtta
183661	cttgctaaat	aaagtataaa	ctcattatta	ggttctagcg	ttgtaaaaaa	caaaaatttg
183721	ctgtaagttt	tgatttaagt	tttctaagta	accaactaaa	ttttttctt	ctttaaatca
183781	caaatcgaat	ttataaaaa	gttagaaca	taatttagat	ctgtaccggg	gatttgggtca
183841	tagtaaaaga	tcctcccatg	ttcacctgtt	ctgggaagac	aacggctccg	tgaggagcata
183901	atgctaattt	aatacaaaa	aggaaaaaga	ctctctaaaa	caagctacgt	ttgttcaagt
183961	tttacccttc	ggtagtcaag	cttcgtttga	taaccttgcc	tattaccgtg	cctccttacc
184021	ggtgtccggg	aggatttgct	ctacccttg	ccgtgcccgg	cagggcaacg	ggagcaaaac
184081	tttgttta	aacttcggtg	ttaaatttag	tctttacccc	ttggtaagca	agcaaagctt
184141	accaacggtta	cgcagaagac	tttaacctt	gccgtgcccg	gcagggcaac	gggcgaagct
184201	taaagatcct	aatgtgtgca	agcacacatt	ttggacttcc	acaaagtgga	agatccta
184261	agaactctgt	tctagtttgg	actttttttc	tacgctacgc	taccaaggt	agtttaggct
184321	attcctcatt	agcaatagct	aatgagtaaa	ggaggacaac	ggcaagttca	agtagtaaa
184381	aactttgttc	gtatacccct	tcgggggtgag	gatgaaacct	tttaccattg	ctttttcatc
184441	tttactctt	gtggtataag	agccaagttta	aagtgagggtc	aatagaggac	tgaaatgggta
184501	taaaaactaa	attgattaag	cttactataa	atrcagttta	aagcttaaat	aaatttttaa
184561	atctagatac	gaattttctt	atataaaaa	atataatcaa	aaatttaaaa	ttttattttt
184621	tatctttatc	aaaaataatt	cttataagat	acaaacgtat	aaattgagca	ctggcaaaaa
184681	acaactaaa	gtttttttgt	aagaatccct	ttattatttt	acaatatatt	tactataaat
184741	acttttaaat	tttttattaa	atrttagaat	tgaacttaaa	atrttaactca	tcataatttt
184801	ctagtaaat	gtatgacaac	gaggcttatt	ctaaaaatgtt	ttttaaggca	ttttaatact
184861	ctaacctaaa	cattataaaa	gctacctaga	gcaaaaat	aaaagctaaa	ataaagtttg
184921	aaattttaa	accaaacc	ctttttttga	cgataaagta	ggggtagcaa	aacttactac
184981	ttttacc	aagactttta	cttgccattg	tccttgga	tggcaggggt	gtaagaatct
185041	atrttga	tatgcctgg	actaagctac	taaaaaggac	gaaaaactag	tcaaaatagg
185101	ctttgttcta	ataatagaac	ttaaagttag	atctaaactt	agtcaattca	atctatttat
185161	taacgttttg	tttcaagctt	cggatattta	tttaaaat	taattttttt	tacagcttat
185221	aacagttgaa	tttttagagaa	ttttgtcta	aaaacacgat	tcttacaata	cttgttctag
185281	tttaaat	tataacc	gcatgactcc	tcacctggga	gccctaaact	accagaggta
185341	gtaagggtga	caagcttcgc	ttgtttacc	atgctttcca	ttacatggaa	gaccttcatc
185401	cttgccctct	tggttaagat	atrttagaat	aaatatctaa	cccaataggc	aaaggctaaa
185461	tttaacaact	tcgttggtta	ataaagttt	ccccgaaggc	tttacgcaa	gcgtaaaact
185521	ttatgtttta	tctacgataa	aacatttagg	gaaagcctgc	cgggctaaga	tgtttacaag
185581	taaaca	agcaagggtta	caatcgaaga	ttgtattttg	cccttgccgt	gcccgggaag
185641	gcaacggggg	agaacacaa	gtctaaaata	caatcagaga	ttgtattaaa	tcacatagtg
185701	gaagtccaaa	gtgtgtgctt	gcacacacta	ggctttgccc	gtagggcaaa	cctttattta
185761	acaacgaagt	tgttaaattt	agacctaac	taaacaaact	atgtttggtt	aggatcgtcc
185821	ctttagggaa	gtcttcttcc	taggttctta	aatattgttg	cggttgggta	taataataaa
185881	tttaacaag	cgtagcttgt	taaagccagg	ctacgcctct	tcgagacaag	gttatcaagc
185941	gaagcttgat	taccaaaaag	gtagacgaac	aaagttcctt	actactgaa	cttgccgtgg
186001	tcctccttta	ctcattagct	attgcta	aggaatagcc	taaactacct	tggttagcgt
186061	agcgtagaaa	aaaagtccaa	actagaacag	agttctatta	ggatcttcca	ctttgtggaa
186121	gtccaaaatg	tgtgcttgca	cacattagga	tctttaagct	tcgcccgttg	ccctgcccgg
186181	cacggcaagg	gttaaagctt	tctgcgtaac	gttggttaagc	tttgcttgct	taccaagggg
186241	taaagactaa	atrttaacaac	gaagttatta	aacaaagttt	tgctcccgtt	gccctgcccg
186301	gcacggcaag	gggttagagca	aatcctaccg	gacaccggta	aggaggcacg	gtaataagca
186361	aggttatcaa	gcttcgctta	ataatctata	aaaagtattt	ttttgtgagt	atgtgggata

186421 aattcattaa acttactacc tccggtagtt caaggcttac gggtttcgaa gttttgtatt
186481 tatccattac cacctaaaaa aaattcttga actttttgaa aaacacctt atacggagtt
186541 gttaaatacga taaaataatc ttgttttccg ataagtcttc gagctgcatt ttcaaaagca
186601 atacctgata atgttagctt cttttttaat acaagagggt caccacgatt ggtagaaata
186661 atcacatgat tatcttctgg tattgcacca agcaagggga ttcctaacat ttcttgaaca
186721 tctctgaccg acatcatatc atttcgttga atcatatccg gacggactct attcactaaa
186781 agttttacat tataaatacc atttgcctct aaaagcccag ctacacgatc tccatcacga
186841 attgctgtaa tttcaggagt cgttacaatt aaagcttcat gagcagggtga aatagcatta
186901 ataaaaccta catcaattcc agctggacag tcaattaaaa taaaatgata acctagcgaa
186961 gaaatggaat tgaccagatt ttccatattt ttctgagtaa cattatatcg ctgtctattc
187021 ttagaaattg aaagtaagga caaatttttc catcgtttat cacgaattaa agcttgatca
187081 agtctacact gaccttctaa aatatccatt gcggtatata ggatgcgatt ttctaaacct
187141 aaaagtaaat ctaaatttct taagccaata tccgcatcaa ttaaagccac tttataacct
187201 aatcttgcta tcgacattcc taaattagct gtggcagttg tttttccaac tccacctttt
187261 ccagatgtaa ttactataat tcttgactct acagtattac tattctccga aatagatatt
187321 tcatttcctg atgatatttt tgttagaaat tttttttta attgaggata tttctcacct
187381 tgttgactct ccctataaga ttttttaaaa atgcttcttg aaaaagaaaa attggtgata
187441 ctacttgatt tatattttaga taaagataca ccaatacctg attgtacatg gaagggttta
187501 caaaccttgt ttgtaaaact tgatttaaaa gcatagcttt taaattttaga aaacatctta
187561 tttcgtggag ttattatctg tgacaacgct gaataggaag gataattttt ttttaatcca
187621 ccatagatgc attgctttac acgttgtctt gccggttggg ttgcagcaaa gcatgggata
187681 tacgagtaaa gctcgcta at cttactactt tccgtaattg aggcaataga acaggggaag
187741 agaagccgag aactccattt aaattttgaa tttatcataa taaatacaaa ttaactttaa
187801 taacctaaaa aataaaattc ggtattctta ctaaattagt gtttaattttt atttttaatc
187861 ataaattatg atggtaattt tactaataga taaatattca aaatttttaa aaaaatttga
187921 tttaaatcga agaccagatt tttaaacgca accgttttga tttattctat atatttttat
187981 actgtctaaa tttaaaagct atgcttttaa atcaagtttt acaaaacaagg tttgtaaaac
188041 ctcgatctct aagtattgtg tgttaaacata aaaatgtttt tggattctct atcaaaaggt
188101 aaaaactgaa gcttttgatt attgtttgac tcaaattaga attaaagttt agatttataa
188161 attactaaaa aatctaaatt atgacactta acaaaaaagc aaattttcac atacaaattt
188221 aacttatttt accttcacat actaaagcag aagaatgct aagcaatatt tttagaatat
188281 taaaacccta caaattgagt atttttaaaa agcgggag gagggattcg aacccccgac
188341 accgtgggtc gtagccacgt gctctagtcc actgagctac acaccctaa acttatttta
188401 ctctgatatt gatgatcttg caatttttat aaatactatt tctatttgat aaattgtcac
188461 cagacaaagt cttctctagg aagccttacc gttgccccgc cgccattgtc taaattaaaa
188521 aacgaagttt ttttaataaag ttttacccea agggtaaaat atttatgaat aaatatctaa
188581 cctgggagcc ctagattccc aggctacgcc caacaggcaa aggctaaatt taacaacttc
188641 gttgttaaat aaagttttgc tcccgttgcc ctgcccggca cggcaagggg tagagcaaat
188701 cctaccggac acggtaagga ggcacggtaa gggcttaact ttaagaacaa aaaatattcc
188761 gaagtttaag aaaaagactc gcagcgaagt tccaaaacct taattaccaa gcacagaaaa
188821 gttcaagcaa attgaagtat tgggtttgag aacagaagag ggattttata ctcactagcc
188881 agaaagatac aaagagactc gaaaagtcta ttttaaaact tttgtttact ccgtttctaa
188941 atttaaaagc atagctttta aatcaagttt taaaaattt atttgtaaaa cattcgataa
189001 tcttcaattt tcacgaaatg gtcattgtgt attttgataa ctgtaaaatta aaactagata
189061 ggcctagggt ttgagctacg ctttatattt gatataaagc gttagctcaa acctagttta
189121 aaaattgagg tgaaccatc catagctatg gagccctttt cctaataaat taactcctaa
189181 ataacagatc caaacaatga caaagccagc acttgcaata attgctgggt ttttaccttg
189241 ccaaccctta gtaattctag tatgtaata aattgcaaat acaagccagc taattaaagc
189301 ccaagtttct tttggatccc aactccaata tgaaccccag gcttcatttg cccaaactgc
189361 tcctgacaat attcctattg ttaaagtgg aaaacctata cctaaagttc gataacttaa
189421 attgtctaaa attttagcta atttatattc actacctcg tggtttatga aatatttttt
189481 aatatttttc tctaaactag aactagagtt ctcaataaca tactcatctt tagaagaagt
189541 atctgctgca ttactagaaa ctgcccgtaa aaattgtttt gtatttccag tattgtattg
189601 taactcttta aaagaaaagt ctaaattgga caagtttgct tttgaattta tatcattgat
189661 tgaatttgta gaaactaacc tagtatcaag ctctgtttga taaaaacctt gaaattttac
189721 ccggttgctt ccggcaacac cagggctcag tccaaggcta tttcctgtta ggtcaatggt
189781 ttttccaaaa gtaataatta aaaaagcaat tgctaacaac gaaccgccta atagtgtctg

189841 ataacttaac atcattactg taacatgcat cattaaccaa ttagattgca gtgcaggaac
189901 taaaacacta gctttttgca tttcttttgg taaagaaaaa gaggcaaac tattagtaaa
189961 taacgcaatt ggtgaaataa tggctcctaa aaaaatattt ttacttagtt tttccagaac
190021 cagttgaagt agaagtagac tccaagaaag aaagattaaa gattcgtaca tattacttag
190081 aggaaaatga ccataattaa accatcgcgt gattaataag agcaatagac taaaattagc
190141 gagaattggt gctgaaaaag caacaaaatg aaatntagag aatgtcgaaa aactaacttc
190201 aagccaataa aaaatcatga ttccaaatac taaaataaac gaaaaatttc ctaaaatgat
190261 ttcaaaattg agatgtgtca taaaattgat aaattaattg gatatacaga actaaaatct
190321 gtttttagtag taaatatact tttatatttt taaaaattct aaattagaga gtaaaaaata
190381 tgctaacaat ttcttaaaaa aaaattttatt taacaaaaaa tatttttaaca ataattcgaa
190441 attaaataaa tttttcaaat gccatttaaa tgtcttagac ggttaaagtt ccttactact
190501 tgaatcctta cgggtgtccg gtaggatttg ctctaccctt tgccgtgcc ggcagggcaa
190561 cgggagcaaa actttgttta acaacttcgt tgttaaattt agtctttacc ccttggtaag
190621 caagcgaagc ttgccaactt tacgcagaag actttaacc tggccgtgcc cggcagggca
190681 acgggcgaag cttaaagatc ctaatgtgtg caagcacaca ttttggactt ccacaaagta
190741 aaagatccta atagaactct gttctagttt ggactttttt tctgcgctaa gttaccgaag
190801 atagtttagg ctattgctta ttagcaatag ctaatgagaa aagaccaagg gggacaacag
190861 caagttcaat tagcccgttg cgggtgctcc ttaccgtgtc cgataggatt tgctctacc
190921 cttgccgtgc ccggcagggc aacgggagca aaactttatt taataacgaa gttgttaaat
190981 ttagcctttg cctgttgggc gttagcctggg aatctagggc tcccagggta gatattttatt
191041 cataaatatt ttaccacacg ggataacggc ggcaggctct ctccgtaggc tttatcgtag
191101 ctccgggaacc tagggctccc ggggcgaagc gtaaaacttt atgttttagc taagataaaa
191161 ctttaggac ctaatctgaa caaactatgt ttgtttaggg aaaactttat ttaacaacga
191221 agttgttaaa tttagccttt gcctattggg ttagatattt attcataaat atcttaccca
191281 agagggcaac gggcaaagtt gtcaagcttc gcttgactac cgaagggca tatttagaga
191341 agcttgatt ctattttgga tcttctattc cttttacgtc gacgcaaggc ttcacccttc
191401 gagttaaact tgatgcttta tctgtctagt gcgtagtctt tacccttga taaggaccaa
191461 ggggcagtat aggtcaaaga taaaacattt agtaagaggg gctacgtgac tagaagaagt
191521 tctacttatt caattagttg aagctttgac tcaagggttc aacttaaagc tttgtcttat
191581 tgtcgtaat gatctttaat aagaagccta gagttaaaat gctaactcgt catacatgta
191641 ttcattttat aaaacaaaat aacttttagta ctaattttat tattaatttc gttagatcta
191701 cttttatttg cttcgacatt aaatcgtatt tatcgtacct agtctatctg actttttact
191761 caataaaaaa ctggaaaaaa gtttgcatta tgcaaaaaa gtcttacctt gtttagaact
191821 aatgagaaat attaatcagc aaagtgtgat aactatatct taacttattt taaaatttaa
191881 agttatatta ctataaacta aataaaattt ttaactatag tttacttaca aataaaaaact
191941 agatcataat ttatatgatg taaaattttt aagtataata agtaaacaaat tataagcttt
192001 aattctttta aaaaaattta ataaaaaagc gaatatttag agaatttatt gtaataagat
192061 aaaaagattt tcttaaaagt aaatactaaa tattaaaaaa gctttttaag gcaagcgtag
192121 attgttcaag ttttactttt taggtgaaac ttttattctc aattttaact ttttatagta
192181 atacagagta agaaaattcc atgtaaaatc taagtcttag tatatgtgat tatatgaagg
192241 attatattgg cttttaattg aaattaaaac atttgtaatg tataaaatct tgagtggtca
192301 caaatttcta actttcattg gtaataaaat aaacttgaat aagctcccat aagacctttt
192361 ttataggata cgctgcttgt ttcaactact taagcttaaa atttatacag tatttttcta
192421 atcttttaaa agtttctcta aatttaaaaag cgtagctttt aatcaagtt ttacaaattt
192481 tttttgtaaa acattaacat ttttcaaatt taaacaggag tatatataat gaagtttagca
192541 gtttatggta aaggtggaat tggaaaaatca acaacaagtt gtaatatttc gatagcttta
192601 gctcgacgtg gtaaaaaagt tttacaaatt ggttgtgatc caaaacatga tagcacattt
192661 actttaaccg gttttttaat cccaacaatt attgatactt tacaactcta agattatcat
192721 tatgaagatg tttggccaga ggtgtttatt tatcaagggt atggagggtg tgaatgtgtc
192781 gaagctggag ggcctccagc aggtgctggg tgtggaggct atgtttagg tgaacagta
192841 aaactattaa aagaattaaa tgcgttttat gaatatgata taattctttt cgacgtttta
192901 ggggatggtt tttgtggtgg atttgctgct ccgttaaatt atgctgatta ttgtattatt
192961 attacagata atggttttga tgctttggtt gctgcaata gaatcgttgc ctctgttcgt
193021 gaaaaagcgc gtactcatcc gcttagatta gctgggttaa taggaaatcg aaccgcaaaa
193081 cgagatttaa ttgataaata cgttgagagt tgtcctatgc cagtgttaga agtattacca
193141 ttaattgaag atattcgtgt atctagagtt aaaggtaaaa ccttgtttga gatggcagaa
193201 tctgaaacag ctttaacata tgtttgtgat ttttatttaa atattgctga tcaacttttg

193261	tctcatccag	aaggtgtggt	tccaaacgaa	ttacctgac	gagaattatt	tagtcttctt
193321	tcagatTTTT	atttaaattcc	acaagaaaac	attactgacg	ctcataatga	cagcttagat
193381	tttatggtag	tttaattaaa	caaagggttag	agcctcactc	ttactactcc	tcctactacg
193441	taggagtgta	ggtccttagc	cattgcctca	ccctaaaaca	agcgtagctt	gttaaagctt
193501	caccctacgg	gtgaagcctt	cggggcaagg	ttgtcaagct	tcgcttgact	acctttggtc
193561	tttgcttatt	agttattgct	aatgagcaat	agccaaaagg	tagagattaa	tgagcaatgg
193621	ccgaaggtag	cttagtcttt	accttttggg	aaaagccaag	aggggcaagg	tttatcaagc
193681	ttcgcttgat	taccctgttg	tagcgggttt	tatcgtgctc	acgaccttgc	gggaacgcaa
193741	ggtttaaagc	atggcttttt	ctaaaattta	aattctaaat	cttgtgttta	gtttatatgt
193801	tcaaaacaag	cggagcttat	tcaagtttca	gccgaagggt	attcgttttg	cttctagcaa
193861	aacgtcaacc	ttactcagaa	gactttaacc	cttgccgtgc	ccggcagggc	aacggggcgaa
193921	gcttaaagat	cctgccctaa	ttagggcagt	gtaggctcta	atgtgtgcaa	gcacacattt
193981	tggactttcca	cgaagtggaa	gatcctaata	cactttctaa	gtgtattttg	gacttccatc
194041	cttaccatgt	cggccggttg	ctcattagca	acgcctaata	agtaaggctt	tacccttcgg
194101	gtaaaacttt	attaaaaaac	ttcgtttttt	aatttagcct	ttgcctgttg	ggcaaagccc
194161	aacgggacaa	tggcgaagtg	gaaaatccta	atagaactct	gttctatatt	ggactttttt
194221	tcacagaaga	cttgtgttct	gtgctacgct	gtgctacgct	accttcggcc	atggctaatt
194281	agcaatggct	aaggaatagt	ttaggtaaag	cattctccaa	ggttttaatt	tttagtaaaa
194341	ataaataaaa	gaaaaaaatt	agaggaatta	aattatgtct	gaaacattat	ctgaaacctt
194401	tacatttgaa	tgtgaaactg	gtaattatca	tactttttgt	ccaatcagct	gtgtttcttg
194461	gttatatcag	aaaattgaag	atagtttttt	tttagtagtt	ggtacaaaaa	cttgtgggta
194521	ctttcttcaa	aatgctttag	gcgttatgat	ttttgcagaa	cctcgttatg	caatggctga
194581	actcgaagaa	ggcgatatat	cagcccaatt	aaatgattat	aaagaattaa	aaagactttg
194641	tattcaaatt	aaacaagatc	gtaatccaag	tgttattggt	tggattggta	cttgtacaac
194701	agaaataaatt	aaaacagatc	ttgaaggaat	ggctcctcgg	ttagaagccg	aaattggaat
194761	tcctatagta	gttgctcgag	caaatggttt	agatcacgcc	ttcactcaag	gggaagatac
194821	tgtattagca	gctatggcac	atcgggtgtcc	gactcaagac	tctaagattt	tggtaggtac
194881	aatctctgac	tcggaaaaaca	atgatcagtg	gccttttaaat	aaagtgttga	atagcaaaat
194941	tgtaaataac	ttattttcaat	ttaataaaaa	aacacaaccc	aatgattcaa	ctaattatta
195001	caaccatcct	ccttttagttt	tattttggatc	attaccaagt	actgttacia	cacaattaaa
195061	tctagagctt	aaacgtcaag	gaattaaagt	gtctgggttg	ttacctctc	agcgatatat
195121	tgatttacca	acttttaggtg	atgggtttta	tgtttggtga	gttaatcctt	ttttaagtcg
195181	tacagcgact	actttaatgc	gacgaagaaa	atgtaaatta	attggcgctc	catttccaat
195241	cggcccagat	ggtacacgtg	cctggattga	aaaaatttgt	tctgtttttg	gtataattcc
195301	taaaggttta	aaaactcgag	aaaatcaaat	ttgggaagg	ttacaggatt	atatagattt
195361	agtaaaagga	aaatctgttt	tttttatggg	agataatctt	ttagaagttt	cttttagctag
195421	attttttaatt	cgttgtggca	tgattgttta	tgaagttaga	attccctata	tggtataaacg
195481	ttttcaagct	gctgaattag	ctttattaga	agacacttgt	aatgaaatga	atgttccact
195541	tccacgaatt	ggtgaaaaac	ctgataatta	taatcaaata	caacgaatta	agaattaca
195601	acctgattta	gccattacag	gtatggcgca	tgcgaaatccc	ttagaagctc	gtggaattag
195661	tacaaaatgg	tctgttgagt	ttacttttgc	acaaatacat	ggatttacta	atgcacgtga
195721	tatttttagaa	ttaataactc	gaccattacg	acgaaatcaa	agtctagaag	gtttaggctg
195781	gaccaaaatta	gtcaaatcag	cgacttctta	aataattttt	tagaactatc	taaatttaaa
195841	agctacgcct	aaactacttt	ttagctattg	ccctttggtc	tttactcctc	ctactatgta
195901	ggagtgtagg	taattagccg	ttgctcatta	gtccaaaata	aaaccagagg	ttttattagg
195961	atcttccatt	aattaaacat	agcgaagcgc	agaaaaaaag	tccaaagtat	acaaactttg
196021	tttgtatact	aggatcttaa	agctatgctt	taagtcttct	gcgtttaatt	aatagaaaat
196081	acttctttta	ttggcaaagc	caataaagga	aggatcttcc	ctaaagggaa	gtctttacc
196141	tttggtaagc	aagcgaagct	tgccaacctt	acgcagaaaa	aagcttctcg	gaaatttgcc
196201	aatcttgttt	ttctactcct	tgccattg	ttcaccctta	ttgccaacg	aaattgagcg
196261	taggggtgag	gttgtcaagc	tacaaggtga	tagattatta	tttggagaaa	aacattgtct
196321	gaaaaaaata	tacttaaaaa	gattaacaca	attatagatg	acaaaaatat	ttcatggaag
196381	tggggaattg	aaaatgcaca	tgttgtgtta	gaaaatgcat	ctatttttac	ggctatttct
196441	gatatgctct	tcgattcaga	aatagaattt	ataggcccat	taaagttaaa	aaatacctct
196501	aacaaatcta	actgtctaga	ttttctccat	atagaaaaat	ggtcttacet	tgattttgtg
196561	ttgtctgcgc	tgcttttttc	tattaactat	gaattgcaac	acaatactcc	tcattgggat
196621	tatttgtcta	aaacacttag	ttttttacta	atagagggaa	acttattagg	ggaaaagatc

196681 ttttaatatct tacaaaaaat agatattttt tttattttta aaagagattg ttgaagtgca
196741 ctacttttatt tttttaactc tttttatctt tttcgtggat tcaaagtact actaagtttt
196801 cgtgttatag agcgattttt aatattatct cattttccaa aattcttcat acgcgaatgt
196861 attcatgatt gtatgaaaga aaaaactgct ggaccggtgc ttttagcctt atctgcatat
196921 gaagctagat gttcgaatag ttctactttt gatcctgcct tcacttttaa gaaaaaaggt
196981 tatcaaaaga taaatgaaaa tttacaactt gggatccagg ctacggttca acaaccggaa
197041 gatctttttag ctttcaacag ctttaaaaaag tatctaattc aaaagaaaaat gttatctcga
197101 attttttttt ttaagctttt gtttaaaaaa aaacaaataa gatattttta cgacttgagc
197161 aatgaaattc cagtaactca actttctgca gtcggagtct ggacaattcc tgttgattg
197221 ttggactagt ttcattttcc gatttagcaa gttcaagttc cggtttagca aaattaccgt
197281 cttcagttgt ccaacaaaaa caagtattta aattggattt taagcgtcct tcacaaatca
197341 gctcactgat acgaccaata agtaacgagc taaaaacgag tattcgagac acaagaagaag
197401 cagtaacatt ggttactttg ttgatacaaa actttccgga tatcccaat agtacaacaa
197461 atccatgggt gttgtgtaga aaaagtttga atattaggca agggcctcaa ggacttttag
197521 agctatctac taacaaagaa gttacagctg taatgtctga tttttgatgt ataatcgcca
197581 acgacaccga caaagaaaca atttccttga ttttaaccg catttcccct aatttaaattg
197641 aagatggttt tgaacaactt aaaaaaattt ggttggatt ccttccaaaa acggtagata
197701 aactcaactt tatagaaaat gcttgcgaaa agtcaaatac atatgaaaca aacgatatta
197761 cggattatac atctcgaggg ttgggcataa taaattatta tcaggtattc aaattaattg
197821 ttaaggcaaa tagtaacca aacaacgaaa atcaagaacg aagcgaaaaa aaagaactag
197881 accatccagc aagccgtaaa gaaatgttga ttttccgcca tcgaagtagc cagactactt
197941 tagatgattt acgagaatct acttacgaaa aagctaaagc aatgagaatt ttcagtacgc
198001 taaatcaaac tcgcgccata tacgggtgctc aggaacggaa agaggactta aagaatacag
198061 atttattaac aaataggtgg gaggcttttt ggcagacact tgatcatgat aaatttattt
198121 atcttctggt atcaaactgt gaagggaaatg aaaatgttgt taaagaacta ttgaaaaggt
198181 atagttagtt atatatctca tgctatttgc ctaacgtaat actttggcta ttaaaatctc
198241 aagtacccaa agaactctta tttagctgctc acgaaatgtt tgctgggac actcgtttat
198301 taacaccaat ccccaattta gaaagtttgt gggaaacttt taccaaagat aaggatagaa
198361 cagaacaagc agactttatg attaactcta tacataattg tatagtgtat aaaaatatac
198421 ttgtacaagc ttgtaaaaaa ttttttctta cacagtacc cgttttttaa acactttttg
198481 aggatttaat cactagttca ataaactctc ttgaagaatc taaaaaaaaa atgtactcca
198541 agtttttaaa gtcatatcga aagcttttaa gaagcttatt cttttgacgg gctgatattt
198601 aatgatgata gttttaaaat tacaaaatag gatattagt atcctgatgt ttttaaagac
198661 acgctcaata acgcttttaga cttttagat gatatgtcta aaaaaatggc attgaatcat
198721 tatactatta gtaataaaaag gttattacta ggcaaatggg ggtttccaga agaaacgatt
198781 aaatctgtga aaaagaattt taatttcacc gaggtagcga aaacttactg tgagaataac
198841 ttaaaccgga aatgggattg agaatttgat agtgtttaag tttttattta aagataagtt
198901 caaaat AACAAaatt gcttcacgca gaagatttct actacgtaga agatcctacc
198961 ttatgctgcg cataagggtt ggactttaac cgaagcttgc ctagtgtatg caagcacaca
199021 cttcggacta tcactttttg acagacatga gcccttgctt agatgtttac ttgtaaacat
199081 cttagcccgg caggcttttc taaatgttga acaacattta gaaaaacttt gtttaaaagc
199141 tctgcctaaa ctaccggagg tagtaagggt ggcgaaactt tttcgcatac ccattgggtt
199201 gtgcattcgc acaaaacttt aaacttaggc aacggttaat aaacgaagt ttgtaacttt
199261 ggtctttttt aaaggaaagc aaggaaaggt tttaatcgta gattaaaact tgaacaaact
199321 ttgtttgttt tagaaaagca gcaaagcttg cttctttttt tcaacataac aagcagtatt
199381 aaaaatagcc tcgggtatca tttcagagaa taaaacagta aacgagtaaa ttaggtaatt
199441 gcctacattt gctgcattac attaaaaatag gtttatagac ttttatgtaa tgctgcaaaa
199501 aagaactttt tgcagcattc cggtttaggac tcattagact tgcaactaa aaaacatact
199561 atttaaggaa aatatatatg aacttagact cattatggta ttcattagag cactggatag
199621 aatctaatat agaaatattt aatgacagct gattaatggt atcaggtgaa aattccttat
199681 attctaacta tatagatttt attcgtgac gaaatgaaaa tccaatttca ttaaaatctt
199741 tttgtactaa ttttttagta ttaatagcat atcatcataa gaaaatatta agaaaagtaa
199801 gaatgactaa aggtattcgt atatatgggt gtaaaattgaa acatcaacct gctagacatt
199861 cagctaagaa gaatcctggt aaaaaattgg atgagataca tacataacaa ccttgaaaca
199921 aacttatatt aactttttaa acaaaataaaa caaaagagtt gtttagacttt gcaaaaatat
199981 gtatcataga tcatgatggg gctttatttt tttatgatta tttcattata ttgtcagttt
200041 tttgtaaact aacaaataaa ttctaggtat ttttacttgc tagtttactt tgaacttctt

200101	ttattagttt	aagattat	tctaagaca	atthagatta	tctgtacaaa	acaaagcaaa
200161	agatagtgta	gcacggctg	tcagtatgga	atgggaaatt	agtaataaga	ttaaagaaaa
200221	aaacagttat	ttaaaataga	atattgcagc	tgggtttttt	ataaattcta	acagaaaatt
200281	aggatattat	caatttataa	aggatgaaga	tctttttaga	ttaatgattg	actttatcaa
200341	agataatcct	cattaccagc	gtgttgaaac	tacaaaagta	atTTTTgcag	aaaacgtact
200401	taagcattta	cgtacttcta	taggtaaaaa	aatagaagta	tgTtcagatt	tatgTttctt
200461	aatggtaga	ttaaacttta	aaactcaaga	attaagtgat	tatactccaa	atgTttattg
200521	ttttaatttc	aaaaacatgt	attacttgaa	aacagaagtg	ctTTTTgact	caacaataaa
200581	gtttattaca	aatttatgta	acaaaaatcg	attatatata	aattgTttgc	gtattTggct
200641	tagaaattgt	ttaattcggt	acatttcaga	tgaatatggc	ttatattTgt	tcgtatactc
200701	ttctcataat	taagtgaaac	ttgataactt	taccgctttt	tctctttacc	aagaagtaaa
200761	gactaatggg	tagcggagggt	gaggggggtga	gtagtgaagt	ttttaacaaa	gggtatagac
200821	tatgctgcac	taaatccctt	gattattcgt	agTtgTttag	gcttagctc	ccggggctgt
200881	tttttaggta	ggggTggcgc	tgcatTTTTat	cagTTTTtga	atatctTTTT	taaatctatg
200941	tttatattct	ggagTtaaaa	aatagTtgTc	cggattgagg	ttttttattc	tttctataca
201001	ctctatacac	ttatttgaat	ttgacaaaaat	tactattTat	gattTttTaa	aactgTtttt
201061	acaaataagt	gaagTtgata	tttaaatttc	ttttttTaat	agcaagccgt	aaattTaatc
201121	gtttttataa	caatatcaca	acaaattact	aatattTtag	tattctctat	acataaatgt
201181	tcttgcaaat	ttatgggaagt	ttaattataa	tattttataa	ttatactatt	tttaataaga
201241	ccattgaata	aaaatagaat	ttttacattc	agTtaactct	ctTgtcaaac	ttaaagatta
201301	tttagacgat	gagcgtat	taggggagag	gtggcagagc	ggTtcaatgc	accggTtttg
201361	aaaaccggcg	tagTgtcaaa	actaccgagg	gttcgaatcc	ctccctctcc	gTtttatttc
201421	tttattctaa	atctatcttt	taatctcatt	aaatagaatt	tttttatatt	attattTtga
201481	ttatttttta	aaaaataatt	tcttaaggTt	ttacaaataa	agTttgtgca	gggcacaaac
201541	caacgcaacc	ctTgTttgaa	cattTttatgt	tcaaattggTc	gtagctTttta	aatTttgaaa
201601	gctgtaaaaat	ctTtgcgacc	aaattTttatt	ttttatctta	ttttttacct	tgcgtaaacg
201661	caaagctTta	tccgggaggT	aatattTttaa	aaagctTttgc	ttattTttagg	tagctTtttat
201721	ttccattTta	gggtaaaaca	ctagaactac	gaggTttTta	aataaaaggTta	agactTtaac
201781	ccttgcttag	atgTttactt	gtaaactatc	tagcccggca	ggctTttctta	aatgTttgaac
201841	aacttttaga	aaaactTttgt	ttaaaagctc	tgactaaact	accggaggTta	gtaaggTttgg
201901	cgaactTcgt	tcgcataccc	attggTttgt	gcatttcgcac	aaactTtttaa	acttaggcaa
201961	cgggcgaagc	ttaaagatcc	tgccctaatt	agggcagTgt	agggcctaatt	gtatgcaagc
202021	acacattTtg	gactTttctat	tttttttata	caaagTattt	acaaacagta	tttacaagca
202081	gtattaacaa	tttaaaagaa	atTtagTtat	ggattctaat	acaaggatat	ttaaaaaagt
202141	tattcaattt	actaatattg	gtaaaaataa	actattTttct	gcaaaaaaca	actTgcctat
202201	aaaaattTtt	ttattattTtt	ctggattTttt	aataggaagt	ttattTtggaa	ctTttttacc
202261	aagTttgcct	gaacaaatta	attcttatag	tgTtattata	ttgattacta	ttaccgctat
202321	tgaaatagtc	aactattTtag	tttacagctc	aaaaaaaaaga	cagTttttct	tcggctcatt
202381	ttttggaatt	ctacttcaat	tttttctatc	gataaaaaacg	aaattagTtt	ctTttttttt
202441	taatagagac	ccaacacttc	aatcatatga	agctcaaaaa	aataacaaaa	aagTttcttc
202501	tgagTtacia	attcaacta	tggTacccca	atTtgaacac	ggtagcatta	ccccgaagg
202561	ggaaggTtat	caagcgcaac	tggactacc	ccttattTttta	cattctaaag	aggattctag
202621	tttttagagTt	aaacaatggc	agaatacaaaa	taaaaattTtt	tataaaaaatt	taaattcgTt
202681	caaaattggT	attatgTtag	gtTttTtttat	tgatgctTtt	aaagtgggca	gttagaacia
202741	attcttaatt	taccacattt	ccaaaaacaaa	aatatagctt	ttaaataagcc	aggcttatta
202801	aaactTttctt	taaactgagg	ctTtgaactt	agacctaatc	tactaaaaaa	atTtgtTttat
202861	ttacaaaaaa	cagatattTtt	ctTttattTtt	aaaaaagatt	gtTtagccgg	agtagTttct
202921	tgTtttagat	ttacctacct	ttTgcgaaga	tttaaagTtaa	tattaagTtt	gcatattata
202981	gagcgaTttt	taattattatt	ccattttccca	aatTttcttca	tacgcgaacg	tattcatgat
203041	tgtatgaaag	aaacaacggc	tggcccctta	ctTtttagcct	tacatgtaga	tcaaaTtaaa
203101	tgccacaaga	attctctaa	taaattTtaac	agtaccgcat	attccaacag	gtTtcaattg
203161	ttcaacaact	ttaaaagaga	ttTggTtTcga	acagaaaaatt	tatctcgaat	ttttttttca
203221	aactTaaatt	taaaaaaaaac	aaataaaaaa	tttttataat	ttaaacgata	caattcattt
203281	gaaacaagcg	tctgcgTttc	cactTtttatt	tttaatatTtt	tttactTtga	ctacacgagT
203341	tgatactagg	cctaagTatc	ctgcccagcgc	aaatagTgaa	caaccaagat	tttctTtttt
203401	cagaagctcg	aattgaaatg	aatTttTgtaa	agctcaataa	ctTaaacaga	accaacgatt
203461	ccagagaatc	atTttattgcc	ttttTaaatg	aactTtttatt	ccaagccaga	aaaaatgatt

203521	tggataagtt	gaaaactcac	atacagaaat	atcaagttcc	tcaaaaacat	cagctaataa
203581	gcaaaaatth	tactattaga	cagtctatag	taggtgttac	aatgcatctc	caagttgctc
203641	acaacctaga	atthttttatt	tattctccta	catataaaca	aaaattaaat	gcgcagcatc
203701	tcgccttaac	taaagaatth	aaggggctth	tggaacaatt	tattgaagta	aataaaaaata
203761	ctttatatth	taatgcagag	attgctaata	aagattthaa	agaattthaa	aaacaaaaac
203821	cagatcaaaa	cttcgaactt	gaaataatat	gtaataactga	cttggattta	ggaaattcat
203881	atggtagaaa	tgagagagatt	ttggattthaa	ggatacaaac	ggctattaat	ctagcaaaaa
203941	tataaaacag	acctaaagac	ttctatatag	aaaaaccgac	taaccagca	cttttagaaa
204001	aaaggtttgt	tccggattta	gctgatctth	gtcttcaaaa	tactgaaaca	gcatctgaat
204061	ctctatcatc	ttacgccaga	gaaagactta	tgaagccttc	tgctcaaacc	gacttaggtt
204121	taataactac	gctgtctgat	gatgaatgga	tgactgaaat	agttgatcgt	ttgaataagc
204181	aaagtcaaaa	tgataaaaagt	ggagacaaga	atagtataa	gtttatthaa	aatthttcag
204241	aaagtgaaga	taaaggctga	tattttgaaa	atgaaataga	cttaccaaca	gtgtttatac
204301	aagacctacc	tgactacaaa	ccacgtcgac	gttcacaaaa	agaagaattg	ccaaggttac
204361	ctcgtagaga	aacacctgat	gaagaacgac	ttcaacattt	acaagaagac	aaattgccag
204421	gggtgccaaa	cacacctgat	gaagaacaac	tttaaagttt	acaagcatca	cataatacag
204481	aagtagaagt	tgattttctag	ttttttacct	tttttgaaat	gaggtgttat	ttttgaatct
204541	gccccattht	ttatccctca	aactaataaa	aatgtttgat	ttaaacgata	tgacaaagcc
204601	caaagaatth	caagacagtt	tttaatttht	ctgagatctc	aataccgttg	aaaactgtcc
204661	tttgcaaggc	tgattaaatt	aattaatcaa	gcttactcaa	aaaaattht	atthtaagaa
204721	caactthaaag	atagcaaaat	atccactatc	attctthaac	gagatcaaag	ccaaaaacaa
204781	atthtagaacg	gtaagcgaaa	aataaattht	aaaggtcgaa	aaagtaaagc	tgaattttcta
204841	agatttagatt	ataaaaaaaa	atthttgattt	tcacttagta	taagtcttca	tttaaaaaaa
204901	ttaaataaatt	gttccttttag	tgctcgtcta	aaagcttht	tagatataact	taaaaaggaa
204961	actatthttht	cggatgaaac	aattgatata	ttttgttht	tgctggcata	agcataaaat
205021	ggcataaaaa	aagcttggtta	tgthcaaat	ctcaagcaat	ccactthaaat	aattagaaaa
205081	gtctthtcca	cttgataaaat	ggctthtattc	tctctacgat	aattgacata	ataaatgtca
205141	aaactgcagct	aaatattthaa	aatcaatata	atthtaattth	ggtthttht	gactthttaga
205201	ctthaagtcta	tagthtaaatc	tttaaaatga	atthgtctga	ttttacaaat	ttattctgtt
205261	gaagtgaaaa	ataaaactth	tatgaaatga	aaaaactthg	tttatattth	tttgcatttg
205321	tataattaat	taaaagctaa	tcttacttca	taattatatt	ttggagttht	atthcattat
205381	gtctcatgct	gtaaaaatth	atgacacctg	tattggttgt	actcaatgtg	taagagcttg
205441	tcctactgat	gtcttggaat	tggttccttg	ggatggttgt	aaagcgagcc	aaattgcgct
205501	tgcacctcga	actgaagatt	gtgttggttg	taaacgatgt	gaatctgctt	gtcctactga
205561	ttthtthaaag	gttcgagtat	atthtaggac	tgaacaaca	agaagtatgg	gcttagctta
205621	ttaaaggagct	ttactthtth	gaaagctthg	gggtthttag	caatatctca	acatthttht
205681	ttthcttacc	cattgacaag	caaggttat	tcaagtgtca	cccgtaggcg	gaagccttgt
205741	gttaatgtaa	agaccaagct	aaaaaaacta	taggtcaaag	gtaaaaatag	aatthcactc
205801	tatgggtaaa	gctthtatagt	gtthaccctt	taacttatt	tattaaaggg	taaacactaa
205861	tgaggaacaa	aagcctthaa	atctattgtt	ttaaacaatt	tatgaatata	taactthtth
205921	agaaaaatga	ataataaaaa	actatcttht	cgthtaaaaa	ttataaatth	ttaggtctcc
205981	ccaggttagga	cttgaacctta	cgaccaatcg	gttaacagcc	gattgctcta	ccactgagct
206041	actgaggaaa	aatthctata	ttthtatata	atathttaca	agthtttagc	tttaaaatca
206101	agthttctagt	gttctagctc	acaacctaca	tgtaaactaa	ataaaagctt	attattctca
206161	tatgtcagta	tgtctgtatg	tcagtaaagc	tattgaaatt	tattacttht	ggtattagat
206221	actcttggtt	tttattctat	ttthtaaaaa	tttgaaaaat	ctaaaaactt	cataaataaa
206281	ctatagaaga	atattaagat	gggtthtactt	tattaaaaaa	aatthataat	taaatatgaa
206341	agacaaaaata	caaaatagtt	ttagaataat	tactcaattg	tttgaaactg	agthtaactc
206401	tagttgcaaaa	ttthgggtggc	ccttctcatt	agaaatggct	gaatacgttt	ttacagctgg
206461	ttcttcttgt	acaagthtthg	agcaaaact	ttctataggg	aaatthtaca	cgagacaaga
206521	actthttagaa	aaacaaaaatt	ggttgaaatga	aggtaatgaa	aaaggtaatg	agaagatag
206581	taaaacttgaa	tatatagaca	atthttgtth	ccaaattcga	aactctgaag	tttataattg
206641	gccatacaat	aaacttctat	gttctgtgat	ctthtttgtt	tttcataaaa	agthgtgctc
206701	gtctgggtggc	atagattctg	tttcatgggtg	tcaactaata	actattctaa	cctcggtth
206761	ttcggcgatt	agaaatgaaa	tttatgatct	tatacaaaaa	aacggctatt	tccatthtaa
206821	tcctgaactt	gctattaata	gtgctagaga	ctatthtth	ttthataaag	aaactacgct
206881	ttthttcga	tttaaaagata	gcaatcagca	aaaagaaaa	ttgacttht	ttthaaagact

206941 tttacaaatt caagaatctc aaattcagtt gttagtttca aatgctttta aaaacaaaag
207001 ctttaattca ataattgaat ttttaaaaaa ttatagtttt gaaaaacctt aattattaac
207061 tttcaagcag ttacaagttt ttttaaatac aatagattta agctcttcag aaaaaaaga
207121 aattatacaa gatgtttcga ctgaactaga atctatgaaa tctatagatc tgctttccac
207181 gtgggacgat ctaactttta gttcaaagt tgcaattttt gaattctata atgagctgcc
207241 tcaaaggcaa aaattaaaac ttttaaataa agaaattatt gctattgtga cgacattaat
207301 tgctctagct aagtttgttg aaataaatac tcctaagatc attctttatc cgcccttaca
207361 aattatacaa aatcgtttat cttcaatttc atctgaaaat ttagccaatc tttatgaatt
207421 atttgttcta gacgaaaatg tcttagatac atctagcaaa agtcagctta aagagggact
207481 ttttaattcgg caggtaatac tatgctctgt atttggtaca cttattacta cttctctcgt
207541 aggtacaagg tcagccttcg cagctcaatt gccacaaggt caagccggga ccgtagccgt
207601 agcacgtcca gagaaaaact ctcaaatgca agccaaaata gtaacccaaa ataaaaatga
207661 ggctttaccg gctagtttac gaagcggaga aaaaaaagtg aagcggctta cagtcacaag
207721 agggcgttat ggttttagatg ttttcaaaga atatgattta ccagaagaac gtaaacatga
207781 gacttatgac attaaatacg aagactctta tgaaataatt atggatggta ttgctgttca
207841 agtttcattt caaacagaaa tagacaaaaa gacaaaacaa cccagaaagc tgcttagtat
207901 ttatgtaacc gaagcccctt atgggcatct gcaaacgta attaagcaat tccactttt
207961 gtttggtcct aatgtctctt ttaatagaga gcattgtcaa ttggatcatt tgtgcaccgt
208021 ctggctccaa aaagcttttg gcattcctgc aagtagggcc ggagttatca ctgtgttaac
208081 accccataat ccgtctagaa cggggggtga cgatgggtta aatgacctag gtcgttcttc
208141 tagcattact tatagggttg ctattcaaga cggattacat cctgttgaag gcttagcgcc
208201 tggaaacta aagtcttttc atagagtggg aaaaaaccca accgaaattc cgtatgaaca
208261 agaaaatata gagaaagttt tatcacaagt caaacagca attagcgata ttacggggac
208321 gttgactcat tctttatatt tccttatgct gtcccacaag gctatggttt tatgggtgatg
208381 gaagcttgtg ggtcagcggg cactctgggg cgtaccgcta cctttatgct agatatgatg
208441 aaacaagcac ctgaaaaatt tccgaccagt tttttataa atcaaggtgt aattgagcaa
208501 attcagagaa attaccaatt tgtaagaggc gttcaagagt ccgttaaatt agcaggttta
208561 tctactgaaa ctagcgacta ccttgagagg agattaaata gagcaactca agttttatct
208621 actataccta gtgaggttag agcgcacaga tttaggaaga gtagaatttt tgagaatgag
208681 aagcttgaaa tttctgagaa tgagaaggcg ttcgaggagc tttttaacaa atttggaccag
208741 gcgataaata gtataggtaa cccatatatt gagcaattag aagaattttc aaggaaactg
208801 caaattgata aacatgcagg cccacagagt tgtccgtag ctagcggttt tatagaacc
208861 caagattggg gagtaactca aagtagacta ctagatcaat atggtaatcc atttggagat
208921 ccattaccag ttagaagtcg attagtagat taccaaggta atccattaag ttcttatgaa
208981 actactgtca ttcgaaatag agtggaaact aatgataaat caatttattt gtagtaaatt
209041 tattaaattg atggaaaaaa ggaattcca atgagtaact tgaatatatt ctttaaacgc
209101 tttatgttta aaaatagcat tttttattta aagagaggcg gactggattg acgtaaatat
209161 aagttgtagg ttatcgtata acacatagaa aacttaaagt aaaacctcta agatttctct
209221 gacttcaaca cagaaaactt ctagatagta gaagatccta atccactatc tgagtgtatt
209281 ttgggtttgt tttcacagaa gacacttttt ctgtgctacg gtctaacaat aaatgtttca
209341 tggagagttt gatcctggct caggatgaac gctggcggca tgcttaacac atgcaagtcg
209401 aacggagggt taattttttt aggcttaggt taaaaaaat taaattctta gtggcggacg
209461 ggtgagtaac gcgtaagaac ctacctttag gagggggaca acaactggaa acggttgcta
209521 ataccata agctgaggag taaaaggaat agtattctgc ctaaagatgg gcttgctct
209581 gattagctgg ttggtggggt aattgcctac caaggcaacg atcagtagct ggtctgagag
209641 gatgatcagc cacactggga ctgagacacg gccagactt ctacggaagg cagcagtgag
209701 gaattttccg caatgggcga aagcctgacg gagcaatgcc gcgtggagga tgacagcctg
209761 tgggtcgtaa actccttttc ccagagaaga agttctgacg gtatctgagg aataagcatc
209821 ggctaaactc gtgccagcag cgcggtaag acggaggatg caagcgttat ccggaatgat
209881 tgggcgtaaa gcgtctgtag gtggtttatt aagtctactg ttaaagatca gggcttaacc
209941 ctgagtcggc agtagaaact aatgagcttg agtacggtag gggcagaggg aattcccggg
210001 gtagecgtga aatgcgtaga gatcgggaag aacaccaatg gcgaaagcac tctgctgggc
210061 cgaaactgac actcagagac gaaagctagg ggagcgaatg ggattagata cccagtagt
210121 cctagccgta aacgatggat actaagtgct gtgcgactca aaacgtccag tactgtagct
210181 aacgcgtgaa gtatcccgcc tggggagtat gctcgcaaga gtgaaactca aaggaattga
210241 cgggggcccg cacaagcggg ggagcatgtg gtttaattcg atgcaacgag aagaacctta
210301 ccagggcttg acatgccatt taatctcttg aaaaagaggt ttttactaga acccttttgg

210361 gtttttagag gtggacacag gtgggtgcatg gctgtcgtca gctcgtgctg tgagatggtg
210421 ggttaagtcc cgcaacgagc gcaacccttg ttttagttg cttttctaag tataaatacc
210481 tgggagcccc atgggaagaa tacaagtcca aagtgtgtgc ttgcacacac taggatcttc
210541 cacttcgcca ttgtccttac cggacatggt aaggatggaa gtcttcttcc tagggtccta
210601 aatattgttc gaagaacaat atcaagtttt aataagcttc gcttattaaa gccaggctac
210661 ggcctcaggt ttatcaagct tcgcttgatt accgttggcc ctttggctt tacctacact
210721 ttggtaaaga ctacgatatg tcaaggaaac tagaaagact gccggtgata agccggagga
210781 aggtgaggat gacgtcaagt cagcatgccc cttatgcctt gggcgacaca cgtgctacaa
210841 tggttaggac aatgagatgc tacctcgcga gagcaagcta acctcaaaaa cctaatactca
210901 gttcggattg caggctgcaa ctgcctgca tgaagtgcga atcgctagta atcgctggtc
210961 agccatacag cgggtgaatac gttcccgggc cttgtacaca ccgcccgtca cactaccaga
211021 gtcggcttgg cccaaagtcg taactctaac ccgtaaggga cgagagcgcc taaggcaggg
211081 ctggtaatga gagtgaagtc gtaacaaggt agccgtactg gaaggtgctg ctggtacc
211141 tccttaatta acaaacaaag tttgttaaat aagatcttaa actttgcaaa gcatagaacc
211201 caagttttct atgtttaagt ctaaatttaa aagctacgcc taaactaccg gaggtagcgc
211261 agcgcagaaa aaaagtaaga agactttaag cgaagcttca agatcctagt ttgtgtttca
211321 cacaaacttt ggacttgttt tcttaccatg gtcttctgcy taaggttggc gaacttcggt
211381 cgcataccca ttggtttggt cattcgcaca aactttttaa tcaagtttta caaattttat
211441 ttgtaaaacc ttaaaaaaag ttcgcttgta tcttagggca aagctttaag tcttctatg
211501 tataatataa taaatttacc aattccttct taatcaaagc ctgctggatt aagatgttta
211561 caagtaacaa tctaagcaag ggtcacattc aatgtgttag agcttaggtt taaaacttat
211621 tcttgtttta atataactat ttcggtagaa tataatcaga atttaacctt aattacagcg
211681 gtttcaccct gcattgaata tataaatata ctacagtgtt aatttccaat ggagagtata
211741 agccgcacag accagactcg tcttgagaaa aataatatga ttaaactaaa agcactccga
211801 ttaaaagata aaaaactagt agatgccttt gttaccgata gaattgtgct ggaacctgct
211861 tgtactgagc gattaagtga tctttacatg agctttttca aatacagtgg caataatcct
211921 gagattgggc gacgtgcttt taaagaaata ttaatacata ttatcgtaag taagtttgaa
211981 attgaggggtg ttgttgaagt aaaaagtctt aaaggtgtat taattcaagg tattggatta
212041 cttgataatc caaaatatac gtaaaaaaat aaagcttata attaattatt ttcaatttag
212101 aaatttatat gacaaacgaa actaatgcct tatgggggggt tatatctgaa attttaggggt
212161 tatttgaaaa ggaatttgga ccaaattgtc cataccattg gacattctcg ctggattctg
212221 cttattatgc ttttgatgaa gcttcttcta atttaaattt agaacaactg tttaatcagt
212281 ggaatcaatt accttctgat gaatcttttc aggatgaaga tgaaatttta gacgttttaa
212341 aacaatctcc ggcttgggaa tggattcctg attatatttt tggttctgct ctatttttta
212401 ttctttatga taatcttgaa acgacaggaa aactctatcc aaatatttgg tccaagttgc
212461 tatatagttt aatgttcatc gatactggaa agcaaataaa attacttaat gttatagagc
212521 gtaattgggt ttttgagctt taccctgaat ttgctattaa cggagcaaaa aacttttttc
212581 ttagacaaga aagattacca gatttattct ttcagtatat acaagctcct cgacaaactt
212641 tatatactat tctcaactt ttacaactta aagaagctca aatccagctg ctagttaatg
212701 aaacttttaa aacaagaagt ttctcaaaaa taattaattt cttagaaagt tataacttac
212761 ccacacctaa tttattcaat ataaccgagt tagaagtatt attaactaaa tttggactac
212821 ctgaagaaaa acgaaaatat cttttaacta atatttctca gaatctaaat aattctggta
212881 aaaattaaaa acctgggagc cctaggttcc caggctacga tgctttttaa tttgtttct
212941 cccggtatac acagaagacc atggttaagaa cacaagtcta aaatacaatc cttccttaaa
213001 ttaccttagc gtagccccgc cattgttcat tagcaatgcc taatgaggaa tctaggcttt
213061 ataagctctg cttataaaac tttgtataaa taggtgtgtg caagcacaca ccgtacttcc
213121 agctatttat acttaggctc ccgggtaatt taagaaagta tatttcatat cagagattgt
213181 accctaaaca aactacgttt gtttagatta ggtccttgct tagatgttta ctgtaaaaca
213241 tcttagcccg gcaggctttc cccgaaggct tccccgtag gctttacgcy aagcgtaaac
213301 ctttatgttt tatcttagat aaaacattta gggaaaactt tatgttttac cgttgcttaa
213361 gtttaaaagt ttgtgcgaat gcacaaacca atgggtatgc gaacgaagtt cgccaacctt
213421 actacctcg gtagtttagg cagagctttt aaactttcta aatgttgttc aacatttaga
213481 aaagcctgcc ggcaacggca aggttctaag ataaaaacatt taggacctaa tctgaacaaa
213541 ctatgtttgt ttagggaaaa ctttatttaa caacgaagtt gttaaattta ggcaacgggc
213601 aaagcttaaa gatcctaata gaactttgcy tagccccgga tctaggcttt ataagctctg
213661 cttataaaac tttgtataaa taggtgtgtg caagcacaca ccgtacttcc agctatttat
213721 acttaggctc ccgggttcta ttttgactt ttttctgcy cttcgtacg tttgttagc

213781 ttaggtctaa gtttaacaac tatgttgtta aacaaagtaa gcaagcaaag cttgctttcc
213841 tagtttgtgt ttcacacaaa ctttggactt ccacgaagtg gaagacttta accggtgccc
213901 tgccgggcta agatgtttac aagtaaacad ctaagcaagg acctaatac aacaaacgta
213961 gtttgttttag ggtacaatct ctgattgtat tttagacttg tgttcttacc atggctctct
214021 gcgtttgttt aggaatctag ggccaagact accaaaggta gtcagggttaa tgcctttgta
214081 atgcttacia aggcatctac tcccagggtta gatatttatt cataaatatt ttaccaacg
214141 ggatttccct aaatgtttta tctaggataa aacataaagt tttacgcttt gcccttggg
214201 ccttaccag gggttaaggac gtaaagcctt tgggtaaaac tttattaaaa aacttcggtt
214261 ttttaatttag acaacggggg ctataatcaa ttaatttaat ttactaatta tgtctatctt
214321 taacgaagat tttctaccag atttcctgaa atatctaaaa gcaaatttta ctaaaaattt
214381 aggtaaaagc attttttata aaaataaaaa taacaagaa aaaaaaagc aaaaggcgat
214441 tcttcaggat ctttaaggata agcaagaaaa taacgattgt aattcgtgta ataagctgag
214501 ctgtaattac tgtaggacaa aatttacgaa ggatcttgac cagagcgttt ttcgtccaat
214561 tttaaacata atgtgggaac tccagaaaaga aacacttcac aaagcccggc aaacattaca
214621 attaaagctc gtaacagcag aagattattc tacagtttca aaagaaataa aaaaaaata
214681 ttataaggag catggatttt cagaaaacta taaattgtca caggaacaaa catcagattt
214741 tgatcagttc tgcgttgaca gactaaaaga atacgagaaa agcgataacg ctaaaattgc
214801 gagatttcat cgttgtgcac gatataaata caaagaagat agcagattag atccagtcgg
214861 gaatttataa ttggcttaca cattattaaa tctctcatat gtactacaag gaaaaacaa
214921 cgaaaaataa cttattgatt ttttcatgtc aggagctttc caactcttcg caccgatga
214981 tttcagctca ggcagcttt ttaataaaca ttttgagaga ctagggcact acctaaccgg
215041 aaataaggct actcaatcta ctgcgcctc taaccaacc tctacccta accaacctc
215101 taccttttgg tctttgccag agacataagc cgaataaaaa tcagctagcg caccctgaa
215161 cgataaacag tatctagatt tacatatggg taaatttcta gcagcatttg gtactgataa
215221 aatagcagaa ctagataaaa ttgtagaagt ttacggtaag ctgcgtaaaa agaagttta
215281 ttagttgaag atcctatctt atgctatgca taagatttgg actttaagca aagcttaag
215341 atcctaataa aagctgtgat tttatttggg acttatatat aataggtatt attctaattc
215401 tatcttatgc tacgcaaaaag attttgattt aaagctatgc ctaaactacc tttgtagta
215461 tttattattt ttttagttaa aataagtcta gtaagtttat cttataaaaa tctttttaa
215521 tctatatttt ttttctggg atattataac aatgtttcaa agtttctaata attggttta
215581 ttactaccct tctctattag accttcttta taaacaataa agtttataaa taaaagtagt
215641 atttattaat gtttatggta aaataaagtt agcaatttta tttataaact tctttta
215701 tcttacttta aactaaaata caacatcact tcccttggg ctttactcat tagccgttgc
215761 cctttgggtc ttactcatta gccgttgcta atgagcaacg gccaaagggt aaagactaaa
215821 cttttactat caattttata atatgaatca aataattagc aatgtcttgc taactgtcag
215881 ttatcaaat ttttacgca taaaagttaa tttattagaa ttgcctttc cagcgtctaa
215941 aagctatagc aattttttaa ttaataaat tgcgccaagc tactttaaaa acccttttt
216001 ttatttatac ctccggagtga aactaaacct tttttattgt cgaaattttt taaaaatat
216061 ttcttattat cgagaactta tcattagttt tatgcaggat gaatacgttg atggatcat
216121 aatttctaatt tccaagaatt attttgaaag aaaacctatt catctttttc ctaaattgcc
216181 gaaacgcaat atagaaattt taaaacaaaa ccctgagtta caaaataaac aaattgcaca
216241 agatatacac tatattactc aagttgagta tacagtccaa cgtctggaaa acctggaagc
216301 ttacttatta tgggtgatga tgcctatttg gttcgggttgc gctttttgtt taaactcgaa
216361 agcttttttt taatattatt tttgacattt tacaatttaa aatattatg caaaatttaa
216421 ataattccac aggatcatct tctcgtctctg aatctttacc aattaataca gacgatgcta
216481 ctctgttctgc tcttctaaca gcatatttga acgaagctca aaaagtttaa attccgctg
216541 ccgctgagtc ttcgctagaa aaaatgattg aagcttattt cattaggttc ttggagctta
216601 aaaaaaataa ttttctaat tattatacca atttccatga ttttctgca gtcaatatac
216661 tggatctacc attattcact tataatttcc gacattgggc taccatgaag aatcgtgcag
216721 taatacagag cttacattct aggcgtttac aactactaac taatgaaca tggcgagaac
216781 ttctcgacgt tttaaaaaac aacggtttat ctctacctga tgacgatttt aattcaagt
216841 gtaatctaaa cagacgaaga cgggggtggca cagcagttgc tccccgaga aggatcccaa
216901 caccttggag taccctgtat acagatttta aaatttctga ttctgataat gaaaattgag
216961 ctcgagatgt tagacttcaa acgtattatg ctaaggcaca atcgcatttt atagatcagg
217021 gacggcgagg agaactcgta tccacaatag actgcacgtg taacctacc ccatgcgttg
217081 gccacccaaa gtcattgaaa gaacattgtt taaccattg cctaaaagcc aacgcaatga
217141 aaattgagga cgtggttgaa aaactttcaa aacatactaa cggatatatac gacaaaagaa

217201 gacaagatta tatgtttgga atctattcga cagcaatact tttcaaacta ttatttgact
217261 ttcatacaact gtttttacca gaagatggta ccgcatttat ggatgaactt tcgaatgaag
217321 aattagttga tgaacaaggc atgaaagctc aaattcaaaa ctccgctact tttttacaaa
217381 atgctgtcaa aaatccttca ggctctacag tattcatgct tgtgggtattt ttagcacgaa
217441 ataggtgggc ttttgcacat tttttgtaag aaaaaagcag gtttattaac tttctatfff
217501 attagccttt gtaaaacatt taggtaaagc cttactcatt aggcgttgct aatgagcaac
217561 ggccggacac ggtaaggcct tataagccga aggccttagc gcagaaaaaa agtccaaaat
217621 agaacagagt tctattagga tctttaagct tcgcccgttg cctaagtfta aaagtttggtg
217681 cgaatgcaca aaccaatggg tatgcgaacg aagtccgcca accttactac gtaatcaagt
217741 aaagcttgat taagtagttt aggcagagct tttaaacttt ctaaagtfta ttcaacattt
217801 agaaaagcct gccgggcacg gcaagggfta aagtcttctg cgttataaaa cttgaacaaa
217861 cttggcgaag catagaacca tagaaaactt gggttctatg cttcgccaag tttctatgt
217921 ttgttttaga ggcacggcaa gggftacaatc ttcgatatga aatataactt cttaaattac
217981 ccgggagccc tagattcccg ggctacgcta aggttaattta aggaaggatt gtattttaga
218041 cttgtattct taccatggtc ttctgctgtt gtttaggatc cccttgccct ttggctttta
218101 ctccctctac tacgtaggag tgtaggctct tagagtaaag accaaaaggc aagggttaat
218161 gagggttgac agactgatta gttttagttt accatagttt tataagggct attagctcag
218221 ttggtttagag cgcacccttg ataagggfta ggtcgtctgat tcgaattcag catagcccat
218281 aataaagcca catagaaaac ttggcgaagc atagaacca agttttctat gtataatttc
218341 agctatftat acaaattftat gggcagagct tataaagcct aagttcctca ttaggcattc
218401 aaaagttggt ttttcccctt atcagttagt tgtaatttag tataataaag aggaagaaga
218461 aagctftftat gaattgtcgt ttcaattacc cagttttctt ctttttttag tgctgatata
218521 tcttgtatct cttctaaatt cgtttctfta tcgagtaact ggcttaatcc taattgtcta
218581 aatactacag atacgttatt atttgataag actftftfta tcatacctac ttgaccagaa
218641 aaaattacta tgttctcgtt ctcttgtftt atagcaaaaat cataaacaat ggttactaaa
218701 tttccagggt caaatctaga agacataaaa ttatggftaa taataatgft tcttctfta
218761 ttaattaaagt ctaaactctt ttgatattct tctaataacg ttgtaccttc ttctaaagct
218821 ttttttaccg aattaacggg ataactattt gctatctgct gtttaaaatg aagtttaagt
218881 atctcatcta agtgaagaat tggagttgca aaagcgtcat aagaattaac attcttcagt
218941 ttattftaatt catcataaaa atttataaaa tatgtttctg attcctgctt atctaactft
219001 acttttggtc caactcgacc tatctgcctt ataaagaaac aagaaagggft ttttttatca
219061 attaaactctg tacgcaatgc aacataatcc ccagggtgaga taagacgttc tctgctatat
219121 atttcaaagg caatagacgt attctcctft gctttattat tttttgattt cttactttcg
219181 taacctaaat tttctgfttg tttttctatt ctaaattgat ttagtftaatg gtttacatgt
219241 ctaaagaagt taaaatagcc ttaaaatcag tatcattaaa cggatctftc ccgttaaaata
219301 ctttatataa agtatttgta acctcctftt tttgttgftc acttaaaataa gtacactftg
219361 ataatacatc attaaaatta taaaactftt gtaaaatttc aagaattftt ctttaattctt
219421 ggtaagaatt ttggttccac gagcctactt gtttatatcg gtttttatta gactgattgg
219481 ttccaccctt aacaaaatcg gttacgggct cgatccacca caactcgfta aaaagacaaa
219541 tgggtataaat caatgftaat gaacctccat atagttftft atccacactt atttcacgat
219601 gaacgtcatg gcgaaatctc gcagtaagaa tctgaggagt tfttaaatfta ttattftftg
219661 agtaagcaat aaagtcgaga tcattcgcag ttaaatftftg aaaattgtcg taattftgtac
219721 ttaagttagt ttctgftaatt gggftftctat aaattftcgtt gaaaatgfaat accataaaaa
219781 acaacgftaa taaactctft aattftftct cttftftcgt caattgagat gttctcgtat
219841 tataaaaggft cgtggattfta taaatattftg gtaggftgca aattftcaatc aaatctaaaa
219901 tatgctftftg ttctagacaa tgttcaaaaa ttctattftaa ccaggctgga aaatcatatt
219961 gaaagtcag ggaaggftaaa aagttaccta cagtcaatat taatctatct aggttaattfta
220021 tagaacgacc tggtagctftt tcaaccgaag tataaccattg ctctgctgta aattftggat
220081 atggagatgc tggaggccat ttgggccttg aagagagagc tgctaattgg ttgaaaggat
220141 tggattgttg atctfttagcc aaagttggfta tagcccgcca aatcggftfta gggatctftt
220201 ttattftgtcc agttfttaagt atattftcfaa actcataggg tctcctftaaa ggcattgctcag
220261 gtctatgfta agggattgct gcgftftaaag atgatgftaa agtatacata gttftftgftt
220321 aaaataaaca aagftftaaat tgaaaaataa atfttatat fttaattftac acaaaaaaat
220381 ttgacaaatt ctcgattftta gttftggcggft aaagfttcaa acaactftta gctgcaattt
220441 atatttagtca gtagccaatg ttgaaatac aactftftaaag taagaagact tgtgttctta
220501 ccatggctctt tftftftftfta atfttatagaa gttftaaaaca aacaaagftt gttftgattfta
220561 aatcaaaatt ttgggtatta gcaaaataa aaaaattftt atftattftta tctaagctta

220621 gatttaaagg ctaataacca ttataagaa aaaacaaaa aagttttact cggttagaat
220681 ttacaattgt ttaattcgcg ggcttttctc tagtttattt agtagctcta atgcaagata
220741 atttttacct aaaaccggtg ccccgttggc taagatattt accgttgcct aagtttaaaa
220801 gtttgtgcca atgcacaaac caatgggtat gcgaacgaag ttcgccaacc ttactacgta
220861 atcaagtaaa gcttgattaa gtagtttagg cagagctttt aaactttcta aatattattc
220921 aacattttaga aaagcctgcc gccgttgtct aaattaaaaa acgaagtttt ttaataaagt
220981 tttacccaaa ggctttacgt ccttaccctt tggtaaggac caaggggcaa agcgtaaaac
221041 tttatgtttt accgttgcct aagtttaaaa gtttgtgcca atgcacaaac caatgggtat
221101 gcgaacgaag ttcgccaacc ttactacgta atcaagtaaa gcttgattaa gtagtttagg
221161 cagagctttt aaactttcta aatgttgctc aacattttaga aaagcctgcc gccgttatcc
221221 cgttgggtaa aatatttatg aataaatatc taacctggga gccctagatt cccaggctac
221281 gcccaacagg caaaggctaa atttaacaac gaagttgtta aatcaagttt gacccaagg
221341 tttacgtccc ttacccttgc gccgttgctc attagcaacg gctaagtagt aaggccttcc
221401 accattcggg ggaagatcct ggttatcaaa caaagtttgg taaccttggg ccaaggggca
221461 aagcgtaaaa ctttatgttt taccgttgcc ttgccggcaa cggcaagggt ctacgataaa
221521 acatttaggt aaagccttac tcattaggcg ttgctaata gcaacggccg gacacggtaa
221581 ggctttataa gccgaaggct tataaaactt gaacaaactt tgtttgtttt aggtcagaga
221641 taaaacattt agatcttaaa gctgcgcttt aagtccaaac cttatgcgta gcataaggta
221701 ggatcttcta ttatatagaa gtcttctgcg ttcgccaacc ttactactta attagcttt
221761 gcttgattaa gtagtttagg caaagctttt aaacttaggt atagctaaag gttttaattt
221821 aactaaaatc ttaagttttt ggcaaagctt cttttgtttt atttataaaa acaccaagcc
221881 ctcccctta ctgctgagag ctaaccaaat taaatcttta gaaattgggc tattctcaaa
221941 ttctaatttc tgaaattcaa ctgaagctaa gtttgctgat aattttttct taactttacc
222001 aatttgtcct aattttattt ttggatgtag cttacacata agtatgggtg tatttttttc
222061 aaacggaaaa aggacagtca attctagtac ttcagttttt cgtttttttg ttttttttga
222121 atcgttatga cgagttagat gaaaataaat agcaactccg actaaactta ttaaacttag
222181 taaagaaaca aaagtcaaaa cattgccgta tcgagtatca ttaggcttta cttccggact
222241 aggagggtaa cgctgctgtc cgagttcggg attaacaatt ccccttgag tgttgaactc
222301 aacaaaagca gagcttattg gggattgagc ctcaactatt gtgtgttctg cctctgagc
222361 ggaacgggat ttggcagaat ttctcgatcc cccttgaatc ttgggctcac ccaaacagg
222421 ggttattcgg acttcagctc cagctattgc ctgttctgtt ttctcagag gacgaaattc
222481 gtttagattt cccggtgcct cttgaggagt gggatcacca ctaatattgg gttttgggac
222541 tttatgctga gctattgtgt cttcggcagt ttcagatgaa cgggatttgt tagaattttc
222601 cgaattaatt ttagcgaata aatttctctt tgtcatttaa attatcctta attacactaa
222661 ttttatatac gggcacttag taatacgcca tattaataaa aaattattcg atgattctaa
222721 aaaattaatt aaataattac caatttctat ttaatatttt aatataaaca aattaatatt
222781 ttatacttat ttaagcttat ctttctttta gaaagatata aattattgat atgagttcag
222841 tgtagttctt ctaattaaca acaaaatgat tgtctctaca atttgacaaa ttttcaattt
222901 tagtttggcg gtaaagttcc aaacaacttt aagctgtaat ttaaattagt aaatagctaa
222961 tgttacaaat acaactttaa agtaagaaga cttgtgttct taccatggcc ttttctttt
223021 tgatttagag aagtttaatc atttaaattt aaaattttgg gtataggaag aagactttaa
223081 gcgaagctcg tagcctggcc cattgggtta aacactttat gtttaaactt taataagcga
223141 agcttattaa aacttgatat tgttcttcca acaatattta ggaacctagg aagaagactt
223201 ccaccttac catgtccggt aaggacaatg gcgaagtgga agatcctaaa caaacatagc
223261 gaagcgcaga aaaaaagtcc aaagtataca aactttgttt gtatactagg atcttaaagc
223321 tatgctttaa gtcttctgcg tttgttttaga ttaggtctaa acttaacaac ttcgttgtta
223381 agtaaagttt tgcctaaatg ttttatctta gacccttgcc gtgccgggca gggcaacggg
223441 aaaacataaa gttttacgct tcgccccttg gtccctacca aggggtaagg acgtaaagcc
223501 tacgggcaaaa tcctagtgtg tgcaagcaca cactttggac ttcccctttag ggaagacttt
223561 aatacaatct tcgattgtat tttagacttg tgttcttccc atgggtaaac aagcgaagct
223621 tgtcaacctt actacttgaa cttgttcaag tagtttaggg ctcccaggta aagatcctag
223681 tgtgtgcaag cacacacttt ggccattgct ccttagcaat ggctaaggag actgtcaaaa
223741 agtgacagac ctgagccctt gcttagatgt ttacaagtaa acatctaagc aagggcaaaa
223801 tacaatcaga gattgtattt agatctgcca ttacatggaa ggcattgggt agcaagcaaa
223861 gcttgccaac cttgctacct tcggtagctt agggctcctg ggcgaagcat agaaccacag
223921 ttttctatgg ttctatgctt cgctatgttg ttaaacttag acctaactta acaaaactac
223981 gtttgtttag ggtaaagact aagatgttta caagtaaaaa tctaagcaag ggcaaacctt

224041 tatgggggcta cgggggtttta cttataatta taccatgaca gctctatgta atgtatcctg
224101 acatcaagtt atctttccaa agccccacgg cttaagtatt gtcactcttt ctacgtttca
224161 caaccaagtt cgagatgggt tgggtgtggt ccatagaaac aaggacacca ggattatcca
224221 acccttgcca ttgccaaggg caatggcgaa gttggataag gtttatctta aactacctgg
224281 tagtcaagca aagcttgact actcattagc ccttgccctt tggcttttac tcattagccg
224341 ttgctaata gaacggcca aagggtaaag acttccactt tgtggaagat cctagtttgt
224401 gttttacaca aactttggac taatgagcaa gggccagata atctaagact tgaagcttgg
224461 cttcaagatc ctaatagact ttttaagtct attttggaaa aatthagatc aaaatcatac
224521 ggtctgttag tctatctcag ctgcaactat tactagcctt acaacttgata gctatcgaac
224581 ggttagtctt accgtgacct ttgtctagaa aactagacga gagcactcat cttaaaggtgg
224641 gcttctact tagatgcttt cagcagttat cctctccata cttggctacc cagcgtatcc
224701 ccttggcagg agaactggta tactagtggt atgttcttcc aggtcctctc gtacaatgga
224761 agactccttt caatgctctt acgcctacac cggatatgga ccgaactgtc tcacgactt
224821 ctgaaccag ctcacgtacc gctttcatgg gcgaacagcc caacccttgg aacgtactac
224881 cgctccaggt tgcgatgagc cgacatcgag gtgccaaacc ttcccgtcga tgtgaactct
224941 tggggaagat cagcctgtta tccctagagt aactttttatc cgttgagcga cggcccttcc
225001 actcggtagc gtcggatcac taaggccggc tttcgccctt gctcgaactt taggtcttgc
225061 agtcaagctc ccttctgctt ttacactctt cggctgattt ccgtccagcc tgagggaaacc
225121 tttgcaagcc tccgttactg tttaggaggc gaccgcccc a gtcaaactgc ccacctgaaa
225181 ctgtcaaagc tcttgattca aggatcgcca ttagaattct agcttctcca gagggtctc
225241 tcaactgatg ctccaacct cccggaagag aattatcaac gcctcccacc taggctgcgc
225301 aagaaaagcc cgaattcaat tccaggctac agtcaagctt catagggctt ttctgtccag
225361 gtgcaggaag tccgcatctt cacggacatg tctatctcac cgagcctctc tccgagacag
225421 taccagatc gttacgcctt tctgtcgggt cggaaacttac ccgacaagga atttcgctac
225481 cttaggaccg ttatagttac ggccgcccgtt caccggggct tccggttgtca gcttttctta
225541 ttccgaagaa taaaataacc aacttccctc accttccggc actgggcagg cgtcagcccc
225601 catatattgt cttacgactt agcggagacc tgtgtttttg gtaaacagtc gcctgggctt
225661 ggtcactgcg gcctctatct cttagaggac cccttctccc gaagttacgg ggccattttg
225721 ccgagttcct tagagagagt tatctcgcgc ccctaggtat actctaccca cctacctgtg
225781 tccggtttggg gtacaggtaa ttcttaatta ttagtagttc gagcttttct aggaagttta
225841 acaactagtt aagtttctact atgaaactag cttacactaa ttccagggaa tgaatccgta
225901 tcacgtctca gctcaagggt agtttttatt ctacctctcg acaccttgaa cgtttccacc
225961 aaaaatccaa ttttgattaa cctttgcttt cttcgtccct cggtaactaat caagaattag
226021 tacaggaata ttaacctgtt ttccatcgac tacgccttcc ggccttgcct taggtcctga
226081 ctcaccctcc atggacgagc ctagtggagg aacccttagg ttttcggggc attggattct
226141 caccaatggt tgcgttactc aagccgacat tctcacttcc gctttgtcca cattaacttg
226201 cgtttatgct tcaccctaga gcggaacgct cccctaccga tttttaacat aggtcgtagc
226261 gcagaaaaaa agtaagaaga ctttaagctt cgcttaaaga tcctaattgtg tgcaagcaca
226321 ctttttgacc ttccacgaag tggaaagatc tagtttgtgt ttcacacaaa ctttggactt
226381 gtgttcttcc taaactactt gaacaagttc aagcagtaag gttggcgagc tacgctcgca
226441 taccatggt cttctgcgta tataatttga tggaaatcaa ttatggttgt ttagttttaa
226501 aatcccacag cttcggcaga ccgcttagtc ccgttcattt tccggcgaag aacgctccac
226561 cagtgaagta ttacgcactc ttttaagggt ggctgcttct aagcaaacct cctggttgtt
226621 tcagcattct tacctccttt gccacttagc ggtcatttag gggccttagc tggatgatctg
226681 ggctgtttcc ctctcgacga tgaagcttat cccccaccgt ctcaactggt gatggaaagc
226741 atatctagta ttcttagttt gctacgattt ggtaccgctc tcgcagcccg caccgaaaca
226801 gtgctttacc cctagatagc ccaagtcac aaccgctgcy cctcaacgca tttcggggag
226861 aaccagctag ctccgagttc gattggcatt tcaccctaa ccacagctca tccgcccatt
226921 tttcaacatc ggtcggttcg gacctccact tggttttacc caagcttcat cctggcctag
226981 gttagatcac ccgggttcgg gtccataaga agtgacatta gcgccctatt aagacttgc
227041 ttcgctttgg cttcagatgg ttttctttaa ccacgccact ccctataagt cgcggctca
227101 ttcttcaaca ggcacgagg gaggcttttc gccctcccac tgettgaacag cttacggtt
227161 catgttctat ttactcccc tacaggggtt ctttttccac tttccctcgc ggtactatct
227221 cactatcggc cacctaggag tatttagcct tacgaggtgg tctctcgtga ttcacaggg
227281 atttcacgtg ccccatgcta ctccgggataa gaaatttagt tgatccgtag cccggccctt
227341 gctcattagt ctttactcat tagccgttgc taatgagcaa cggccaaagg gtaaagacta
227401 atgagcaagg gcccaacaa tcaaggctga ccttattaat aagggttttta gttggatata

227461 tcaactgaact ttcgggetact ggacttttcac catctatggt gcaggattca gctgcttcac
227521 ctagttaagt gaattcttca taaaatttta cctggccctt gccttttgggt ttttaccaaa
227581 aggtaaagac cgtggccggc aggctttccc cgtaggcttt acgcaaagcg taaaacttta
227641 tgttttaccg ttgccctgcc gtcaacggca agggctctag ataaaacatt tagggaaaac
227701 tttatttaac aactttgttg ttaaatttag gcaacgggta atcaagcttt gcttgataac
227761 ccgtagcctg ggaacctagg gctcccaggt tgttttctag gtaaaacata gactttcttc
227821 ccacaacccc gactcgacag atcgagttgg tttaggctgt tcccatttcg ctcccgcta
227881 ctaagggat cgcgtttget ttcttttctt ccagctacta agatgtttca gttcactggg
227941 ttgtctcttt cctatctatg aattcaatag gtagtataaa aggttgcctt attcgggaat
228001 ctccgaatca tagcttgttt ccagctcctc gaagcatttc gttggtttcc acgcccctca
228061 tcgtctctag gtgcctaggt atccaccgta agcttttctt catttgatct ttactaaaac
228121 aagcaaagct tgttcaactt ttctaagttt aaaaacgaag tttttaaaca aagatcgtag
228181 cgcagaaaaa aagtccaaaa tagaacagag ttctattagg atcttccact ttgtggaagt
228241 ccaaaatgtg tgcttgcaca cattaggatc ttttaagcttc gcccgttgcc ctggcgggca
228301 cggcaagggt taaagtcttc tgcgttagaa agtttgtgca gggcacaac caatggacca
228361 gctattaatt aataagctgg tggaaagtaag cggactcgaa ccgctgacat ctgccttgca
228421 aaggcagcgc tctaccaact gagctatact cccttacctg attaggattt gtccaaaggc
228481 tttacacttg acgtcaaact aggtttcacc taaactacgg taaaacattt agatcttgaa
228541 gcttcgcttc aagtccaaac cttatgcgta gcataaggta ggatcttcta ttatatagaa
228601 gtcttctgtg tttaaattat tttttaactt ttctcttctt caagaatata ctacagagct
228661 tatttacgtc aatccagtc gcctctctct aatcctactt tatttaaaga ggctttatac
228721 ggtgtctcta aaacaaactt tgcgtagtct ttacccttgg ggcgtagtct caccctcac
228781 ctgggagccc taaactactc cttagctatt gctaagagc aatggccgga ggtagcgcag
228841 cgcagaaaaa aagtaagaag acttccatcc ttaccatgct cggtaaggac aatggcgaag
228901 tggaaagatc tagtgtgtgc aagcacacac tttggacttg ttttcttacc atggtcttct
228961 gcgtaagggt gacaagcttc gcttgtttac ccatgccttc cactacgtgg aagacctaa
229021 tacaatctcc gattgtattt tgccttggct tagatgttta cttgtaaaaca tcttagcccg
229081 gcaggctttt ctaaagtgtg aacaacattt agaaaaactt tgttaaaaag ctctgcctaa
229141 actaccggag gtagtaagggt tggcgaactt cgttcgcata cccttgggtt tgtgcattcg
229201 cacaaacttt taaacttagg caacggggga agaacacaag tctaaaatac aatcagagat
229261 tgtacccttg cttagatggt tacttgtaaa catcttagcc cggcaaggca acggttaaat
229321 cactatgtga aagtccaaag tgtgtgcttg cacacactag gatgttccac ttcgccattg
229381 tccttaccgg acatggtaag gatggaagtc ttcttacttg tgttctgtgc tacgctgcgc
229441 ttcgctatgt ttgttttagga tcttctaaa tgttttaccg ttgccctgcc ggttacggca
229501 agggtaaaag ataaaacata aagttttacg ctctgcttaa agcctttagg gaagtcttct
229561 tcctagggtt ctaaataattg ttgccaacgg caacaatata aagttttaac aagcgtagct
229621 tgccattgtc ttaccgcccgt tgtctaagtt taaaagctgc gcttttaaat ttaggtaagc
229681 aagcaaagct caccaacctt acgggtatgc gaacaaagtt cgccaacctt acccgggagc
229741 cctagattcc tcattaggca ttgctaatga gcaatggccc ccttgccgtg cccggcaggg
229801 caacgggggc tacgctactt aatcaagctt tgcttgatta agtagtttag gcaaggcttt
229861 atgttttctc tccgaccctt gccgtaccgg gcaggctttt ctaaagtgtg aacaacattt
229921 agaaaaactt tgtttaaaag tttgtgcgaa tgcacaaacc aatgggtatg cgaacgaagt
229981 tcgccaacct tactacctcc ggtagtttag gcagagcttt taaacttagg caacggtaaa
230041 acatttagat ctggaagctt cgcttcaagt ccaaacctta tgcgtagcat aaggtaggat
230101 cttctattat atagaagtct tctgggcaa agctcaattt agttacaagt acggagttag
230161 agggattcga accctcggta caggtagctg tacactcgat tagcaatcga gctctttcgg
230221 ccaactcagac accactccaa ttttagtttt acgttttaca tagtgatatt aaatttcaca
230281 ctgaataaat agtaggatac attttttagt ctcaataatc aagaccccac gagattttat
230341 gccgaagtata cttcgcataa agcaaggata tttttatttg acccttggg cgtagcccca
230401 cccctcacct gggagcccta aactactcct tagctattgc taatgaacaa tggccggagg
230461 ttacgtagcg cagaaaaaaa gtctatagtc tataatagag cgtagctcta ttaagatctt
230521 tcttttaaga aagtcttctg cgtaagggtg acaagcttgc cttgtttacc catgccttc
230581 attacatgga atacctaaat gttttatcgt agacccttgc cgttgatcct taccgtgtcc
230641 ggccgttgct cattagcaac gcctaattag taaggcttta cctaaatggt ttatcttaga
230701 cccttgccgt tgatccttac cgtgtccggc cgttgctcat tagcaacgcc taatgagtaa
230761 ggctttacct aaatgtttta tcttagacct ttgccgttga tccttaccgt gtccggccgt
230821 tgctcattag caacgcctaa tgagtaaggc tttacccttt gggtaaaact tgagttaaca

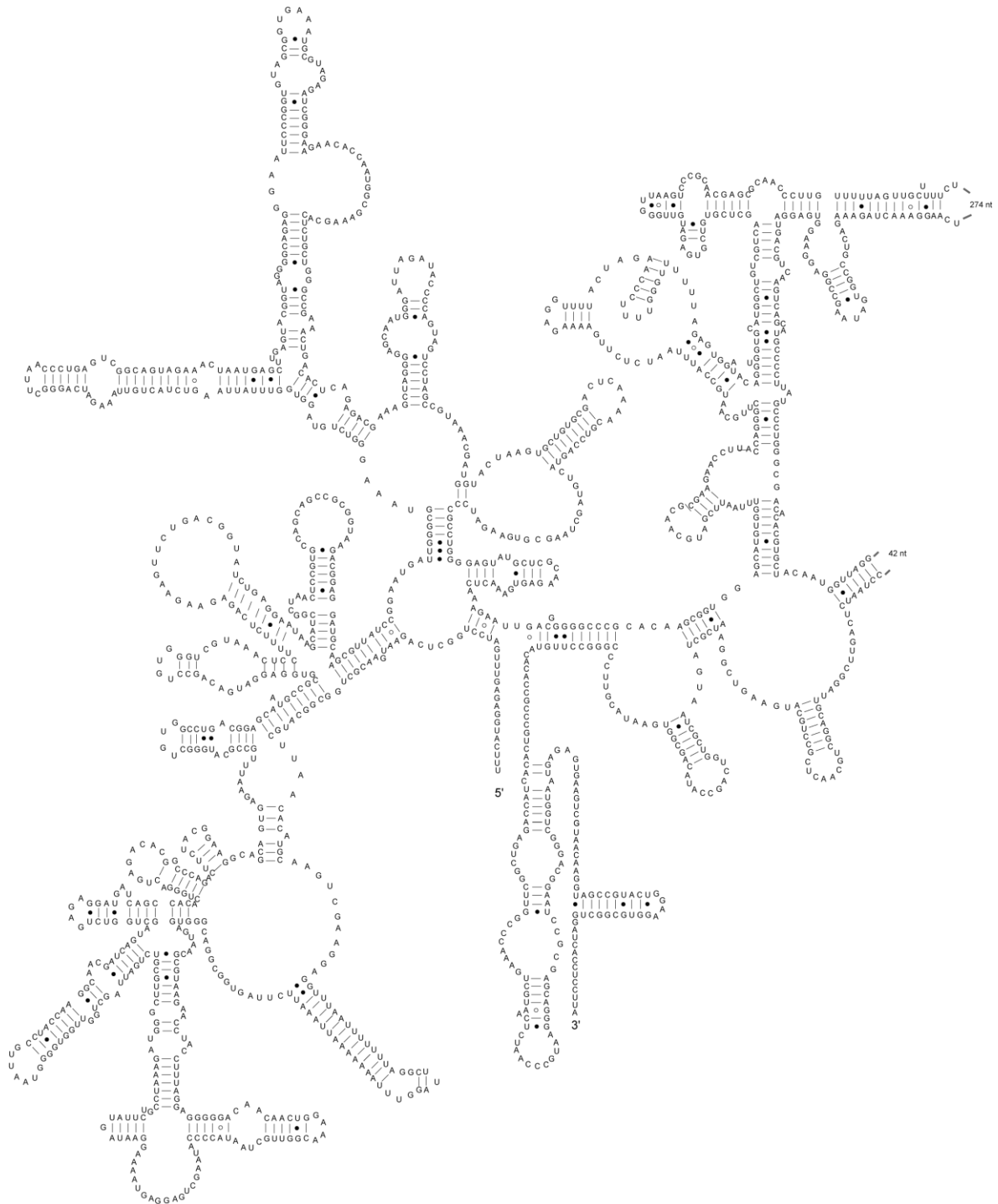
230881 acgaagttgt taaatthtaga caacggcggc aggcctttccc cgaaggcttt ccccgtaggg
230941 gaaaacttta tgthttaccg ttgccttgcc gtcaacggca aggggtctaag ataaaacatt
231001 tagggaaaac thttatthaac aacgaagttg thaaatthtag gcaacggtaa aacataaagt
231061 thttacgcttg gcgtaaagcc thcggggaaa gcctcccggg ctaagatgth tacaagthaa
231121 catctaagca aggggtacaat cgaagattgt atthttgcctt tgcccgttg g tcaaaggcca
231181 acgggatcct aaacaaacat agcgaagcgc agcgtagcac agaacacaag thaagaagact
231241 tccatcctta ccatgtccgg taaggacaat ggcggaagtg aacatcctag tgtgtgcaag
231301 cacacactth ggacttgtgt thttaccatg gtcttctgtg aaaaaaagtc caaagtatac
231361 aaactthgtt tgtatactag gatctthaaag ctatgctthta agtcttctgc gthttatthtag
231421 gggagaaga caagtctaaa gtgaaattcg thgthaaatgt thaaaaaatt agaagthata
231481 aaaatthttag thaatgtthcaa aaaatthaaaa gtatggaaaa thtttagthaa gacgaattac
231541 aaatthttat atthtcgatat thctththgaa thtcaaataca thttthacaaa aatthttgtc
231601 ataaaaaaa aaataaaact thctcagact thgtattcaa thtttagthatt thcattgggc
231661 thggthttagt thctthtaact caatctaaat cggcatatgc thtgcaacaa ccagatccga
231721 thttgggcaac aggcgctgaa gcgtcctccc caaaccaagc thgattatcac thctcatctat
231781 cagagthtctc ggggtgcacc acttcaggga thcccgcgat ggaaccaaca actgagacaa
231841 thcctgthgtc gtgtththaac thggatgactt gctgctthtc thcgcctata gctccgccac
231901 cggatgacca gccgccacgt acacacgcath thcctthttc gatgtctggc ccatthcccc
231961 thtagatcttc thttgctagt thcattagcag aaaatcccc thctthcttca thctccgagag
232021 caaggtctgt aggagctgcc gctactagcg thcaaacacg thagaggagag agthaacgtat
232081 thagccactc aagtcggagt gaaactccac gagcacaag gatgctatca thtaagthactc
232141 aagctgtthca atthcaagctha aaaaaaagca cathaaacc gaaagthaaaa thtaatgaac
232201 thcaatathca cggthgagatt ccaactctag thgtctctag thcaagathact thggthtaact
232261 thgcaatathca aataththaac thcatcggatg atcgtgaaat agthagctthc atccgcgatg
232321 cagatthgtc thgcattccga aacacgactha thaatggthaa gctthgatggg aaactthggthg
232381 thactatthtg thagcthaaga aaagatgctha atththgaaac agcagacaaa thgtthggcga
232441 aagatcaaaa agaaagaaag aactacacag atgagthatt thaacacacath thctgaaagath
232501 gthgggathat thacggaagth agaaaaaaa thattthcgtth thaatgaaact cgaagcgc
232561 thgaacccaa gcaattcagag thtcacgaga gagcggthagg gthagaaaa agcactgaca
232621 aagthaaag thactcctggg gaaacacath thggthcaaac ththaaagacaa thattthcgc
232681 thaaagactgt thactthgtac gaaaatthcag thaccaacth aaagthcagag acathcaatth
232741 ataaaaaagg thaacccggc attgctccaa aatagactcc thcctthaaat thactagggcg
232801 thagcccgcc attgctcath agcaatgcct aatgaggaath ctaggctthta thagctctgc
232861 thataaaaact thgtataaath aggtgtgtgc aagcacacac cgtactthca gctatthata
232921 thtacgcaga agaccatggt aagaaaaacaa gtccaaagth thgtgctthgca cacactaggc
232981 thctthcacth cgcctatgtc thtacccggac atggthaaag thggaagthct thctactthth
233041 thctgctgctg cggthaaacaa gcgaagctth thcaaccthac thaccggaggt agththagggc
233101 thcccgggthaa ththaaagaa thtatththcat atcgtagagth ctatthaggath thctthacaaa
233161 gthagaagthca thtagcaatgc cthaatgagga gccctagath cctcathagg cathgtathat
233221 gagcaatggc cgggctacgc acggththgt thctththgt thatatgtata thattgathata
233281 thaaaagagath thaaaaaath gaththcaath thctgaththg thththagaath thcgtgathth
233341 atthctcctath cthggccgth ththaaaaaag cthaggagath acgathcggc thactatthgt
233401 aaagaaaact gthattgaaag aatctgggag thctthactc agccccctth agctacaaaa
233461 thgcaaaactha cctthththac thcagcaath thggthaaagth thgctggaaath thattththaga
233521 agccacgag caththththc ggatgthaaa aathcaacth acgaaathth thgtththaccg
233581 aaaaataath aathththcth cgagththatt thataaaaaa aathctthth aaataagccg
233641 ggagccctag atcctgggct acgathctag thcccgtthgc thaccthaac gcaaggagaa
233701 aatccggaac agaatthatt thcaaththct caththctgath thctthctgath aathagthctg
233761 atgcagath thggaaaagc thcctththgat thgtctgth thaatththgath aaacaaagaa
233821 gthaatthth gthagthgath thgtcgagth thataaaaa aathcgagga thhaagcaaaa
233881 caaggathaaa gthccaaacc ggccathgct ccaaaathaga thcctthctth aaathactag
233941 ggcgtagcth thagathcctha gthatacaaac aagththgth thactthggac thccthaathg
234001 thttaccgth gccthaagth thaaagththgt gcgaathgcac aaaccaathg gthagcgaac
234061 gaagthctgcc aacctthacth cctccggth ththaggcagath gctththaaac thcthaathg
234121 thgtthcaaca ththagaaag cctgcccgtth thagathgth thcaagthaac atcthaagca
234181 gggthcaagath thaaacath agthththacgc thcgcctth ggtccaaggt thaccaacaa
234241 agththggthaa ccaggathcth thcaccgaath gthggaaaggc thactcath gcccgtthgth

234301 atgagcaacg gccaaaggggt aaggacgtaa agccttttagg gaagatccta aacaaacata
 234361 gcgaagcgca gcgtagcaca gaacacaagt ccaaagtgtg tgcttgca cactaggatc
 234421 ttccacttcg ccattgtcct taccggacat ggtaaggatg gaagtcttct tacttttttt
 234481 ctgcgctgcg ctacctccgg ccattgctca ttagcaatag ctaaggagta atttagggct
 234541 cccgggttta atacaatctc tgatatgaaa tatactttct taaattaccg gggagcccta
 234601 gattcccggg ctacgctaag gtaatttaag gaaggattgt attttagact tgtgttctta
 234661 ccatggtctt ctgcttcgcc attgcccttg gcaatggcgg attatcacat ttatttttat
 234721 tatccagata ttcttttagtc tttacctttt ggctattgct cattagcaat agctaagtag
 234781 taaagaccaa aaggaccgta gcccggggaa ctaggcttta taagccgaag gccgtagcgc
 234841 agaaaaaaaa taagaagact tccatcctta ccatgtccgg taaggacaat ggcgaagtgg
 234901 aagatcctag tgtgtgcaag cacacacttt ggacttgttt tcttaccatg gtcttctgcg
 234961 ttataaaaact tgaccaaaact tcggttgttt taggctcccg gaaccggtag attaagcata
 235021 attagaatta tgcgttgatt gaaggagctt caacagaagc taagtctaga gggagttgt
 235081 gcgcgttacg ttcgtgcatt acttcatac ctaagttagc acggttaata atgtcagccc
 235141 aagtgttgat tacacgacct tgagagtcta caacagattg gttgaagttg aaaccattta
 235201 gggttgatgc catagttgaa atacctaaag cagtgaacca gataccaact acaggccaag
 235261 cagctaggaa gaagtgtaga gaacgagagt tgttgaaaga agcatattgg aagattaaac
 235321 gaccaaagta accgtgagca gctacgatgt tgtaagtttc ttcttcttga ccgaatttgt
 235381 aaccagcgtt agcagattca ttctcagtag ttccacggat taaagaagaa gttactagag
 235441 aaccgtgcat agcagagaat aaagaaccac cgaaaacacc agcaacacca agcatgtgga
 235501 atgggtgcat taagatgttg tgctcagctt ggaatacaat catgaagttg aaagtaccag
 235561 aaatacctaa aggcataccg tcagagaaaag aaccttgacc gatagggtaa atgataaata
 235621 cagcagtagc agcagcaact ggagcagagt aagctacagc aatccaagga cgcataccta
 235681 aacggaaaga aagttcccac tcacgaccca thtagcagca gatacctaag aagaaatggc
 235741 aaacgataag ttggtaagga ccaccgttgt ataaccattc gtctaaagaa gcagcttccc
 235801 aaattgggta gaagtgaaga ccaattgctg ttgaagttgg aacaacagca ccagaaatga
 235861 tgttgtttcc gtaaagtaga gaaccagaaa caggctcacg aataccatcg atatctactg
 235921 gaggtgcagc gatgaatgcy atgataaata cagaagttgc agttaataaa gttgggatca
 235981 ttagaacacc aaaccaaccg atgtataaac ggttttcagt gctagtaacc cattcacaga
 236041 agcagcccca taggctagcy ctttcacgct tttctaaaat agcagtcata ataattttta
 236101 attcttaaaa aaattttatt tctcaaaacc ggatactctt acataatcta agcttccgta
 236161 gcctggcttt aacaagcata gcttgttaaa actttatatt gttgccactc gtagcgaagc
 236221 gcagaaaaaa aaatcttctg cgaaatttat aagttttttt attccgactt ggtaaatagt
 236281 agtattacat attaaacacc agttatcaag tgattataac aaagttaatt aaataaatat
 236341 atagagcaat ttggaaagtc aaaagataaa gttgaagaaa gtttttaatt gaacacgggg
 236401 cccttgacgt gctccttagc ctcaccccgga agaccgtagc gcagaaaaaa agtcttctgc
 236461 gtt

//

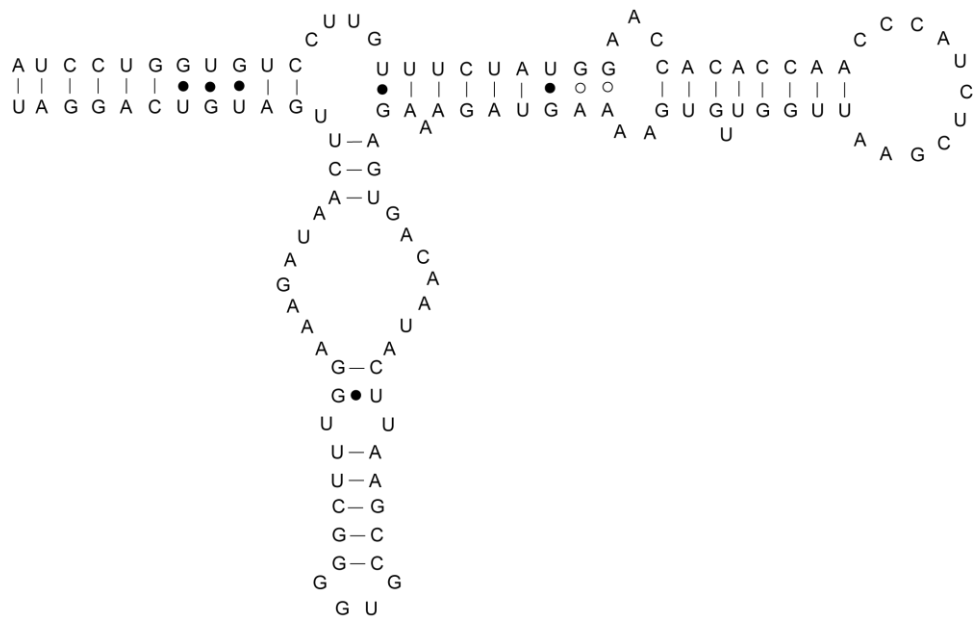
Annexe II

La structure secondaire du gène rrs.



Annexe III

La structure secondaire du gène rrf.



Annexe IV

Les structures secondaires du gène rrl.

