

2m11.3023.1

Université de Montréal

Étude phylogénétique d' $\alpha$ -protéobactéries :  
sur les traces de l'ancêtre bactérien de l'endosymbiote mitochondrial

Par  
Delphine Marie-Egyptienne  
Département de biochimie  
Faculté de Médecine

Mémoire présenté à la Faculté des études supérieures  
en vue de l'obtention du grade de  
Maître ès sciences ( M.Sc. )  
en Biochimie

Avril 2002

©Delphine Marie-Egyptienne, 2002



W  
4

U58

2002

N. 115

Université de Montréal  
Faculté des études supérieures

Ce mémoire intitulé:

Étude phylogénétique d' $\alpha$ -protéobactéries :  
sur les traces de l'ancêtre bactérien de l'endosymbiote mitochondrial

présenté par:

Delphine Marie-Egyptienne

a été évalué par un jury composé des personnes suivantes:

Président-rapporteur  
Directeur de recherche  
Membre du jury

CHTEINBERG SERGUEI  
LANG BERND FRANZ  
ROKEACH LUIS

Mémoire accepté le :

## Résumé

Depuis la fin du siècle dernier, l'origine de la cellule eucaryote est sujet à controverse : des hypothèses ont tenté de l'expliquer, notamment en tenant compte de la mise en place des organelles spécifiques aux eucaryotes, tels que la mitochondrie ou le chloroplaste.

Une des hypothèses pour laquelle des supports ont été apportées fut celle postulant que les mitochondries auraient pour origine une bactérie qui se serait installée au sein de cellules pré-eucaryotes par endosymbiose. L'étude des questions générées par cette hypothèse, comme le nombre de fois que cet événement a eu lieu, la nature du partenaire bactérien de l'endosymbiose a amené a des évidences : par exemple, il est maintenant clairement établi que les mitochondries proviennent toutes d'une  $\alpha$ -protéobactérie et que cet événement semble n'avoir eu lieu qu'une seule fois. Actuellement, l' $\alpha$ -protéobactérie supposée être la plus proche parente des mitochondries est la bactérie causant le typhus, *Rickettsia prowazekii*. A cause de leur implantation au sein de cellules eucaryotes, les mitochondries et *R. prowazekii* ont développé des profils fonctionnels similaires. Mais si les futures mitochondries sont venues s'installer dans des cellules proto-eucaryotes, les Rickettsiales sont des parasites obligatoires de cellules animales, cellules eucaryotes d'origine relativement récente : ceci nous fait douter de la descendance directe des mitochondries à partir d'un organisme « *Rickettsia-like* ». Il semble alors raisonnable de postuler l'existence d'une  $\alpha$ -protéobactérie dont l'ancêtre serait plus proche de l'endosymbiote mitochondrial que l'ancêtre de *R. prowazekii*. Dans cette optique, nous avons examiné des  $\alpha$ -protéobactéries afin de découvrir des descendants de l'ancêtre  $\alpha$ -protéobactérien plus proches de l'endosymbiote mitochondrial.

Des expériences préliminaires de PCR à l'aide d'amorces dessinées sur les gènes mitochondriaux *cox1* et *cob* nous ayant démontré la présence fréquente d' $\alpha$ -protéobactéries dans des cultures de protistes provenant de la collection de souches ATCC, nous avons criblé des cultures de protistes. Pour augmenter le nombre de bactéries testées et la variété de protistes mis en jeu, des échantillons provenant de milieux aquatiques riches en matière organique ont également été criblés. Après avoir mis en culture ces bactéries, leur identité a été déterminée par

amplification PCR, séquençage, et analyse phylogénétique de leur gène *cox1*. Un deuxième volet a concerné des souches d' $\alpha$ -protéobactéries de collection : choisies selon leur relative proximité par rapport à l'origine mitochondriale sur un arbre déduit des séquences d'ARNr 16S d'un grand nombre de bactéries, issus de la banque de données RDP, les  $\alpha$ -protéobactéries sélectionnées ont été testées phylogénétiquement grâce à leurs gènes *cox1* obtenus par PCR, puis clonés et séquencés.

En ce qui concerne le premier volet, le clonage et le séquençage du gène *cox1* de plus de 200 bactéries, ainsi que les analyses phylogénétiques subséquentes n'ont pas suggéré la présence au sein de ces bactéries de descendants potentiels de l'ancêtre  $\alpha$ -protéobactérien de l'endosymbiote mitochondrial. Toutefois, suite à notre analyse phylogénétique des souches d' $\alpha$ -protéobactéries de collection, il nous a été donné de constater la présence de bactéries à égale distance ou plus proches de l'origine mitochondriale que *Rickettsia*. Ceci ouvre la voie à l'étude approfondie des dites bactéries, notamment grâce à la génomique bactérienne, puisque le génome de certaines de ces  $\alpha$ -protéobactéries sont en cours de séquençage.

**Mots clés :**  $\alpha$ -protéobactéries, Mitochondries, Phylogénie, Génomique

## Abstract

Since the beginning of the last century, the origin of the eukaryotic cell represents a controversial subject. Several hypotheses tried to explain this origin, many of them taking into account the setting of the specific eukaryotic organelles, such as mitochondria or chloroplasts.

The endosymbiosis theory, which postulated that a bacteria took place in a pre-eukaryotic cell by endosymbiosis, constitutes one of the hypotheses for which evidences came in support. Questions generated by this theory, such as how many times this event took place and the nature of the bacterial partner found an answer : it is now well-known that mitochondria seemed to arose from a single ancestor, an  $\alpha$ -proteobacterium.

To date, the  $\alpha$ -proteobacterium supposed to be the most close to mitochondria is the bacteria causing typhus, *Rickettsia prowazekii*. Because of their implementation into eukaryotic cells, this bacteria and mitochondria developed similar functional profiles. But, while mitochondria took place in pre-eukaryotic cells, *R. prowazekii* took place in animal cells, that are eukaryotic cells from relatively recent origin : this introduces some doubts about the origin of mitochondria directly from a *Rickettsia*-like organism. It seems reasonable to postulate the existence of an  $\alpha$ -proteobacterium whose ancestor is closer to the mitochondrial endosymbiont than the ancestor of *R. prowazekii*. With that aim in mind, we investigated within the  $\alpha$ -proteobacteria in order to find bacteria that are descended from the  $\alpha$ -proteobacterial ancestor closer to the mitochondrial endosymbiont.

Because preliminary PCR experiences with primers designed from the mitochondrial genes *cox1* and *cob* yielded on protists cultures that came from the ATCC collection demonstrated the frequent presence of  $\alpha$ -proteobacteria in these cultures, we screened some protists culture. To enhance the number of bacteria tested and the variety of protists, some samples that came from organic rich aquatic media were also screened. After their culture, the identity of these bacteria was determined by PCR amplification, sequencing and phylogenetic analysis of their *cox1* gene.

A second part was dedicated to  $\alpha$ -proteobacteria from culture collection : chosen because of their relative proximity to the origin of mitochondria on a bacterial SSU rRNA tree, they were phylogenetically tested with the help of their *cox1* gene obtained by PCR, then cloned and sequenced.

Concerning the first part, the cloning and sequencing of more than 200 bacterial *cox1* gene, as well as the subsequent phylogenetic analysis did not suggest the presence within these bacteria of a potential descended from the  $\alpha$ -proteobacterial ancestor of the mitochondrial endosymbiont. Nevertheless, the phylogenetic analysis of  $\alpha$ -proteobacteria from culture collection demonstrated the presence of bacteria as close or closer to the mitochondrial origin than *Rickettsia*. This opens the route to in-depth study on these bacteria, especially with the help of bacterial genomic, since the sequencing of the genome of these  $\alpha$ -proteobacteria.

**Key words** :  $\alpha$ -proteobacteria, Mitochondria, Phylogeny, Genomic

# Table des matières

<u>Titres</u>	<u>Pages</u>
Résumé .....	1
Abstract.....	3
Table des matières.....	5
Liste des figures.....	10
Liste des tableaux.....	12
Liste des abréviations.....	14
Remerciements.....	16
Dédicace.....	17
<u>Chapitre I . Introduction</u> .....	2
I. Les trois domaines de vie.....	2
II. Origine de la cellule eucaryote.....	4
II-1. Le génome eucaryote est une chimère.....	4
II-2. La mitochondrie et l'origine des eucaryotes.....	6
III. Les diverses hypothèses concernant l'origine des mitochondries.....	7
III-1. Les hypothèses autogènes versus les hypothèses xénogènes.....	7
III-2. La théorie de l'endosymbiose.....	9
IV. Étude phylogénétique de l'ADN mitochondrial.....	11
IV-1. Affiliation bactérienne des mitochondries.....	11
IV-1-2. L'endosymbiote mitochondrial est une $\alpha$ -protéobactérie.....	12



IV-2. Un ou plusieurs événements d'endosymbiose ?.....	13
IV-2-1. Le génome mitochondrial de <i>Reclinomonas americana</i> .....	17
V. La nature de l'endosymbiote mitochondrial.....	20
V-1. les protéobactéries.....	19
V-2. Définition des $\alpha$ -protéobactéries.....	22
V-2-1. Les $\alpha$ -protéobactéries.....	22
VI. <i>Rickettsia prowazekii</i> et l'origine des mitochondries.....	23
VI-1. Le génome de <i>Rickettsia prowazekii</i> .....	23
VI-1-1. Réduction évolutive du génome de <i>Rickettsia prowazekii</i> .....	23
VI-1-2. Profils fonctionnels de <i>Rickettsia prowazekii</i> et des mitochondries .....	25
VI-2. <i>Rickettsia prowazekii</i> , ancêtre direct des mitochondries ?.....	26
<u>Chapitre II . Matériels et Méthodes</u> .....	29
I. Principes de l'expérience.....	30
I-1. Expériences préliminaires.....	30
I-2. Mise en place de la méthode.....	31
I-2-1. Milieux de culture.....	31
I-2-2. Criblage.....	32
I-2-2-1. Les amorces sur les gènes <i>cox1</i> et <i>cob</i> .....	33
I-2-2-1-1. Les amorces sur les gènes <i>cox1</i> et <i>cob</i> .....	33
I-2-2-1-2. Les protéines Cox1 et Cob.....	33
I-2-2-2. PCR, clonage et séquençage.....	35
I-2-2-3. L'analyse phylogénétique.....	35
I-3. Recherche au sein des souches d' $\alpha$ -protéobactéries de collection.....	36
II. Matériels.....	38
II-1. Culture de protistes.....	38
II-2. Échantillons naturels.....	41
II-3. Souches d' $\alpha$ -protéobactéries de collection.....	41

III. Méthodes.....	41
III-1. Milieux de culture.....	41
III-1-1. Milieux ML, MS et CYE.....	41
III-2. Tests de croissance croisée.....	43
III-3. Récolte des bactéries.....	43
III-4. Récupération de l'ADN bactérien.....	44
III-5. Polymerase Chain Reaction ( PCR ).....	44
III-6. Clonage.....	47
III-6-1. Réparation des bouts et purification des produits PCR.....	47
III-6-2. Phosphorylation des produits PCR.....	49
III-6-3. Ligation.....	49
III-6-4. Transformation.....	49
III-6-5. Extraction rapide des plasmides des bactéries.....	50
III-7. Séquençage manuel.....	50
III-7-1. Préparation d'ADN simple brin pour le séquençage manuel.....	50
III-7-2. Réactions de séquence manuelle sur de l'ADN simple brin.....	51
III-8. Séquençage automatique.....	52
III-8-1. Préparation d'ADN plasmidique.....	52
III-8-2. Réactions de séquences Li-Cor sur l'ADN plasmidique.....	52
III-8-3. Gel Li-Cor.....	53
III-9. Obtention de la séquence complète du gène <i>coxI</i> de <i>Rhodospirillum rubrum</i> .....	53
III-9-1. Montage de la banque d'ADN génomique.....	53
III-9-2. Hybridation de la banque avec la sonde spécifique ( fragment PCR <i>coxI</i> ).....	54
III-9-3. Séquençage automatique de l'insert du clone rr2042.....	55
III-9-4. Restriction sur les enzymes SacI et KpnI sur le clone rr2042.....	55
III-9-5. Marche chromosomique sur l'ADN plasmidique du	

clone rr2042.....	55
III-10 Analyse des séquences.....	58
III-10-1. Programmes utilisés.....	58
III-10-2. Méthode.....	59
<u>Chapitre III . Résultats</u> .....	60
I. Recherche au sein des bactéries vivant en relation avec des protistes.....	61
I-1. Récolte des bactéries issues des cultures de protistes et des échantillons naturels utilisés dans nos expérimentations.....	61
I-1-1. Mise en place de la méthode.....	61
I-1-2. Tests de croissance croisée.....	61
I-1-3. Récolte des bactéries issues de la deuxième série de culture de protistes et des échantillons naturels.....	63
I-1-3-1. Cultures de protistes.....	63
I-1-3-2. Échantillons naturels.....	63
I-2. Amplification par PCR.....	65
I-4. Clonage et séquençage.....	65
I-4-1. Première série de culture de protistes.....	65
I-4-2. Deuxième série de cultures de protistes.....	68
I-4-3. Échantillons naturels.....	68
I-5. Analyse phylogénétique des séquences obtenues.....	70
I-5-1. Recherche de similarité de séquences.....	70
I-5-2. Réalisation d'arbres phylogénétiques pour mettre en évidence la position phylogénétique des bactéries de qui proviennent les fragments PCR.....	70
II. Recherche au sein des souches d' $\alpha$ -protéobactéries de collection.....	74
II-1. Obtention de la séquence complète du gène <i>coxI</i> de <i>Rhodospirillum rubrum</i> .....	74

II-1-1. montage de la banque et hybridation.....	74
II-1-2. le clone rr2042.....	75
<u>Chapitre IV . Discussion</u> .....	79
I. Méthode.....	80
I-1. La classification phylogénétique et le transfert latéral de gènes.....	80
I-1-1. La classification phylogénétique.....	80
I-1-2. Phylogénie et mitochondrie.....	81
I-2. Choix de la méthode.....	82
II. Les bactéries mises en évidence par notre méthode.....	83
II-1. Types d' $\alpha$ -protéobactéries.....	83
II-2. Relations phylogénétiques des bactéries mises en évidence par notre méthode avec des $\alpha$ -protéobactéries connues.....	85
III. <i>Rhodospirillum rubrum</i> .....	86
III-1. Relations phylogénétiques entre les $\alpha$ -protéobactéries et les mitochondries : nouvel apport de séquences $\alpha$ -protéobactériennes.....	86
III-2. Arbres phylogénétiques montrant les relations entre les mitochondries et les $\alpha$ -protéobactéries incluant <i>Rhodospirillum rubrum</i> .....	88
IV. Conclusion.....	93
Bibliographie.....	94
Appendice I.....	106

## Liste des figures

### Introduction

Figure 1	Arbre universel du vivant.....	3
Figure 2	Quelques hypothèses d'eucaryogénèse.....	5
Figure 3	Les hypothèses autogènes versus les hypothèses xénogènes.....	8
Figure 4	Théorie de l'endosymbiose mitochondriale.....	10
Figure 5	Portion d'un arbre phylogénétique des ARNr montrant les relations entre les organelles et les eubactéries.....	15
Figure 6	Arbre phylogénétique construit à l'aide des protéines de la chaîne respiratoire mitochondriale.....	16
Figure 7	Photo du jakobide flagellé <i>Reclinomonas americana</i> et carte génomique de son ADN mitochondrial.....	18
Figure 8	8(A) Contenu génomique des ADN mitochondriaux. 8(B) organisation en opéron de certains gènes.....	19
Figure 9	9(A) Photo de <i>Neisseria gonorrhoeae</i> suite à la coloration de Gram, et schéma de la structure de la paroi d'une bactérie Gram-. 9(B) photo de <i>Staphylococcus aureus</i> suite à la coloration de Gram, et schéma de la structure d'une bactérie Gram+.....	21
Figure 10	Photo de <i>Rickettsia prowazekii</i> et comparaison de la taille de son génome avec celui de génomes mitochondriaux.....	24

### Matériels et Méthodes

Figure 11	schéma de la chaîne mitochondriale de transfert d'électrons.....	34
Figure 12	Arbre phylogénétique montrant les relations entre les mitochondries et les $\alpha$ -protéobactéries réalisé à l'aide de séquences partielles de la protéine Cox1.....	37

Figure 13	galerie de photos de certains protistes utilisés lors de l'expérimentation.....	39
Figure 14	schéma des fragments amplifiables par les 5 couples d'amorces utilisés.....	46
Figure 15	Vecteurs utilisés pour les clonages.....	48
Figure 16	Schéma du processus de marche chromosomique sur l'insert du plasmide présent dans le clone rr2042 et des sites de coupure par les enzymes de restriction SacI et KpnI.....	57

## Résultats

Figures 17	Arbres phylogénétiques construits à l'aide des séquences partielles de protéines Cox1 d' $\alpha$ -protéobactéries et de mitochondries montrant la position des bactéries à partir desquelles les fragments PCR <i>cox1</i> contenus dans les clones ont été amplifiés.....	71
Figure 18	Photo des fragments de restriction obtenus suite à la digestion de l'ADN plasmidique du clone rr2042 par les enzymes de restriction SacI et KpnI.....	76
Figure 19	19(A) Séquence totale de l'insert présent dans l'ADN plasmidique du clone rr2042. 19(B) Séquence de la protéine Cox1 déduite de la séquence du gène <i>cox1</i> de <i>R. rubrum</i> .....	77

## Discussion

Figure 20	Arbre phylogénétique montrant les relations entre les mitochondries et les $\alpha$ -protéobactéries.....	87
Figures 21-22	Arbres phylogénétiques montrant les relations entre les mitochondries et les $\alpha$ -protéobactéries incluant <i>Rhodospirillum rubrum</i> .....	89-90

# Liste des tableaux

## Matériels et Méthodes

Tableau I	Liste des protistes à partir desquels des bactéries ont pu être Isolées.....	39
Tableau II	Liste des échantillons naturels testés.....	40
Tableau III	Liste des bactéries vivant librement choisies pour l'étude phylogénétique de leur gène <i>coxI</i> .....	40
Tableau IV	Milieux de culture MS, ML et MLS mis en place pour l'isolement des bactéries vivant en relation avec des protistes.....	42
Tableau V	Mélange de vitamines utilisé pour les milieux MS, ML et MLS	42
Tableau VI	Amorces utilisées pour les amplifications PCR sur les gènes <i>coxI</i> et <i>cob</i> .....	45
Tableau VII	Amorces utilisées pour la marche chromosomique sur l'insert de l'ADN plasmidique du clone rr2042.....	56

## Résultats

Tableaux VIII	Nombre de bactéries récoltées sur milieux de culture bactérienne ML, MS et CYE à partir des premières séries de cultures de protistes.....	62
Tableau IX	Tests de croissance croisée pour les bactéries issues des premières cultures de protistes isolés, sur les milieux ML, MS et CYE.....	62
Tableau X	Nombre de bactéries récoltées sur nos milieux de culture bactérienne à partir de la deuxième série de cultures de protistes testées sur milieux LB, MLS, L1pm + M et L1pm - M.....	64

Tableau XI	Nombre de bactéries récoltées sur milieu MLS à partir des échantillons naturels .....64
Tableau XII	Amplifications PCR obtenues avec le premier couple d'amorces pour les différents mélanges d'ADN bactériens, chacun étant spécifique d'un échantillon et d'un milieu de culture donné.....66
Tableau XIII	Résultats des clonages des fragments PCR amplifiés à partir des pools d'ADN de la première série de cultures de protistes isolés testés.....67
Tableau XIV	Résultats des clonages des fragments PCR amplifiés à partir des pools d'ADN de la deuxième série de cultures de protistes isolés testés.....67
Tableau XV	Résultats des clonages des fragments PCR amplifiés à partir des pools d'ADN isolés des bactéries des échantillons naturels de 1 à 7 et de 9 à 12 .....69



## Liste des abréviations

A	Adénine
ADN	Acide désoxyribonucléique
ADNmt	ADN mitochondrial
ADP	Adénosine Di-Phosphate
ARN	Acide ribonucléique
ARNase P	Ribonucléase P
ARNm	ARN messenger
ARNr	ARN ribosomique
ARNt	ARN de transfert
ATP	Adénosine Tri-Phosphate
C	Cytosine
<i>cox1</i>	Gène encodant la sous-unité 1 de la cytochrome <i>c</i> oxydase mitochondriale
Cox1	Sous-unité 1 de la cytochrome <i>c</i> oxydase
<i>cob</i>	Gène encodant l'ubiquinol :cytochrome <i>c</i> oxydoréductase
Cob	Ubiquinol : cytochrome <i>c</i> oxydoréductase
CYE	Charcoal Yeast Extract
dITP	Désoxyinosine triphosphate
dNTP	Désoxynucléotide triphosphate
ddNTP	Didésoxynucléoside Triphosphate. Appelés aussi terminateurs de chaîne dans la réaction de séquençage de Sanger
EDTA	Ethylene Diamine Tetraacetic Acid
Ef-Tu	( bacterial ) Elongation Factor Tu
g	gramme
G	Guanine
IPTG	Isopropyl- $\beta$ -D-thiogalactopyranoside
kb	Kilo paires de bases (ADN )
kDa	Kilo Daltons
LB	Luria-Bertaini ; milieu de culture bactérien

L1pm +/- M	milieu de culture L1 additionné ( + ) ou non ( - ) de méthylamine
M	Molaire
mg	Milligramme
ml	Millilitre
ML, MS	Milieus de culture contenant du lactate ( ML ) ou du succinate ( MS )
MLS	Milieu de culture contenant à la fois du lactate et du succinate
mM	Millimolaire
Mpb	Méga paires de bases
NCBI	National Center for Biotechnology Information
ng	Nanogramme
OGMP	Organelle Genome Megasequencing Program
<sup>32</sup> P	Phosphore 32
PCR	Polymerase Chain Reaction
pFBS, pBFL	vecteurs utilisés pour les clonages
RDP	Ribosomal Database Project
rpm	round per minute
<i>rnpB</i>	Gène codant la sous-unité ARN de la ARNase P
S	Constante de sédimentation ou Svedberg
<sup>35</sup> S	Soufre 35
SDS	Sodium Dodecyl Sulfate
SSPE	Sodium chloride sodium phosphate + EDTA
T	Thymine
TLG	Transfert latéral de gènes
U	Uracyle
μl	Microlitre
μM	Micromolaire
UV	Ultraviolet
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactoside

## Remerciements

Je remercie mon directeur de recherche, le Dr Franz Lang de m'avoir accueilli dans son laboratoire et m'avoir ainsi offert l'opportunité d'amorcer ma carrière de scientifique. Je le remercie de son soutien et de la confiance qu'il m'a témoigné tout au long de mon projet de recherche et lors de la rédaction de ce mémoire.

Je me dois également de remercier la perle de notre laboratoire, Lise Forget, qui avec passion et clarté, sait inculquer aux étudiants que nous sommes, ces scientifiques en devenir, l'attention toute particulière que l'on se doit d'attacher à la compréhension des choses, fut-ce le plus simple des protocoles. Je la remercie de son aide constante, de ses explications claires, de ses conseils et de nos discussions qui n'en finissaient pas sur la vie, sous bien des aspects.

Je remercie les autres membres du laboratoire pour leur compagnie rafraîchissante, Jessica Leigh, Yannick Jacob, Naiara Rodriguez, Dennis Lavrov ainsi que Elias Seif qui n'a pas son pareil pour pimenter une journée de laboratoire et Julien Lagarde, le gardien de la langue française. Je tiens également à remercier Zhang Wang, qui m'a bien souvent sauvé par ses nombreux Li-Cor .

Je remercie les membres de l'OGMP, Dr Gertraud Burger, Isabelle Plante, Yun Zhu, Dimitri Vzdornov, Jung-Hwa Seo et Shona Teijeiro, pour leur soutien technique et pour leur agréable compagnie. Je remercie les étudiants du laboratoire de Gertraud Burger, Guerline Clerzius et William Marande pour les moments passés ensemble.

Finalement, je remercie le Conseil de Recherches Médicales du Canada pour le soutien financier qui a permis de réaliser ce projet.

## Dédicace

À mes parents pour tout ce qu'il m'ont donné sans compter, à mes sœurs, vous  
tous plus précieux pour moi que de l'or

À PA & PM, grâce à qui je continue d'avancer sur la bonne route

## CHAPITRE I

# INTRODUCTION

## I. Les trois domaines de vie

L'observation des êtres vivants, au niveau macroscopique puis au niveau microscopique, avait conduit les scientifiques à séparer le monde en deux catégories : les plantes versus les animaux, et plus tard, les eucaryotes versus les procaryotes. Toutefois, ces dichotomies ont été remises en question par une accessibilité accrue à des données moléculaires qui ont conduit à une re-définition des méthodes utilisées en taxonomie. En se basant non seulement sur des données cytologiques ou morphologiques, particulièrement inadéquates pour des microorganismes, mais également sur des données moléculaires, cette science a permis une classification plus rigoureuse des organismes vivants.

Ainsi en 1977, par une étude phylogénétique basée sur des analyses de gènes ARNr d'une variété d'eucaryotes et de procaryotes, Carl Woese *et al.* (1) ont été amenés à établir l'existence de trois royaumes primaires de vie (les bactéries, les eucaryotes et les archaéobactéries) contrairement à la division en cinq royaumes alors couramment acceptée (2). Woese *et al.* ont confirmé cette trichotomie en 1990 (3) par une analyse phylogénétique incluant un nombre plus exhaustif d'espèces dans chacun des royaumes. Il en a résulté un arbre universel du vivant divisé en trois domaines (une unité taxonomique pouvant contenir plusieurs royaumes) : Bacteria (les eubactéries), Archaea (les archéobactéries) et Eucarya (les eucaryotes) (Figure 1).

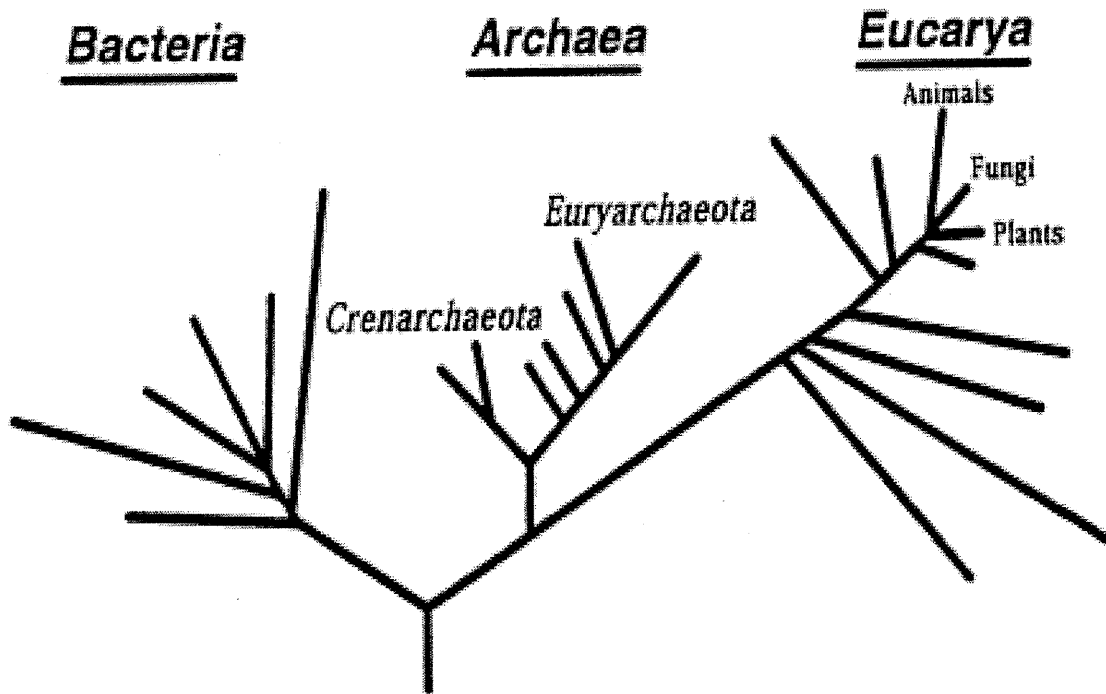


Figure 1 : Arbre universel du vivant déduit d'analyses comparatives de séquences d'ARNr (3). La racine a été déterminée en utilisant le couple de gènes paralogues EF-Tu / EFG (4). On peut observer les trois domaines primaires de vie, les Bactéries, les Eucaryotes et les Archéobactéries.

Source : Woese et al. (1990). *Proc. Natl. Acad. Sci. USA.*, **87**, 4576-4579

## II. Origine de la cellule eucaryote

### II-1. Le génome eucaryote est une chimère

Avant l'avènement des trois domaines de vie selon Woese, Iwabe *et al.* avaient enraciné l'arbre de vie au sein des bactéries en utilisant les gènes des facteurs d'élongation de la traduction présents dans tous les royaumes de vie et dupliqués avant la division Archae/Bactéries (4). Cette topologie de l'arbre a ensuite été confirmée par Woese *et al.* (2) en utilisant les ARNr et par Baldauf *et al.* (5) qui ont utilisé les facteurs d'élongation de la traduction. Ces arbres phylogénétiques universels enracinés montrent que les Archéobactéries et les Eucaryotes sont des taxons frères. En support à ces résultats, il a été démontré que des séquences de nombreux autres composants essentiels des machineries de transcription et de traduction de ces deux taxons présentent une forte similarité (6, 7).

En 1998, deux études (8, 9) postulent l'existence d'un important transfert de gènes archéobactérien ou bactériens vers le génome eucaryote: celui-ci serait une chimère. Rivera *et al.* (8) vont même jusqu'à distinguer deux classes de gènes transférés distincts: les gènes informationnels eucaryotes auraient tendance à être plus proches des gènes informationnels archaéobactériens (comme démontré en 6, 7 à une plus petite échelle), tandis que les gènes opérationnels eucaryotes auraient tendance à être plus proches des gènes opérationnels eubactériens. Même si le fait que les Archae et les Eucaryotes soient des taxons frères expliquent en partie la similarité de leur gènes informationnels, qu'en est-il de la similarité des gènes opérationnels eucaryotes avec ceux des bactéries (et pas seulement des  $\alpha$ -protéobactéries, cf IV-1-2) ? Et pourquoi cette division en deux classes de gènes : archéobactérie /informationnels et bactérie/opérationnels ?

Le débat sur la cellule eucaryote s'en est trouvé ravivé, d'autant plus qu'en cette même année (1998), de nouvelles hypothèses sont venues s'ajouter aux hypothèses d'eucaryogénèse formulées au début des années 90. Toutes ces hypothèses ont en commun l'implication de deux partenaires, un archaéobactérien et un bactérien, même si les mécanismes diffèrent (Figure 2). W. Zillig propose une fusion entre une archaéobactérie et une bactérie (10) tandis que L. Margulis propose une symbiose



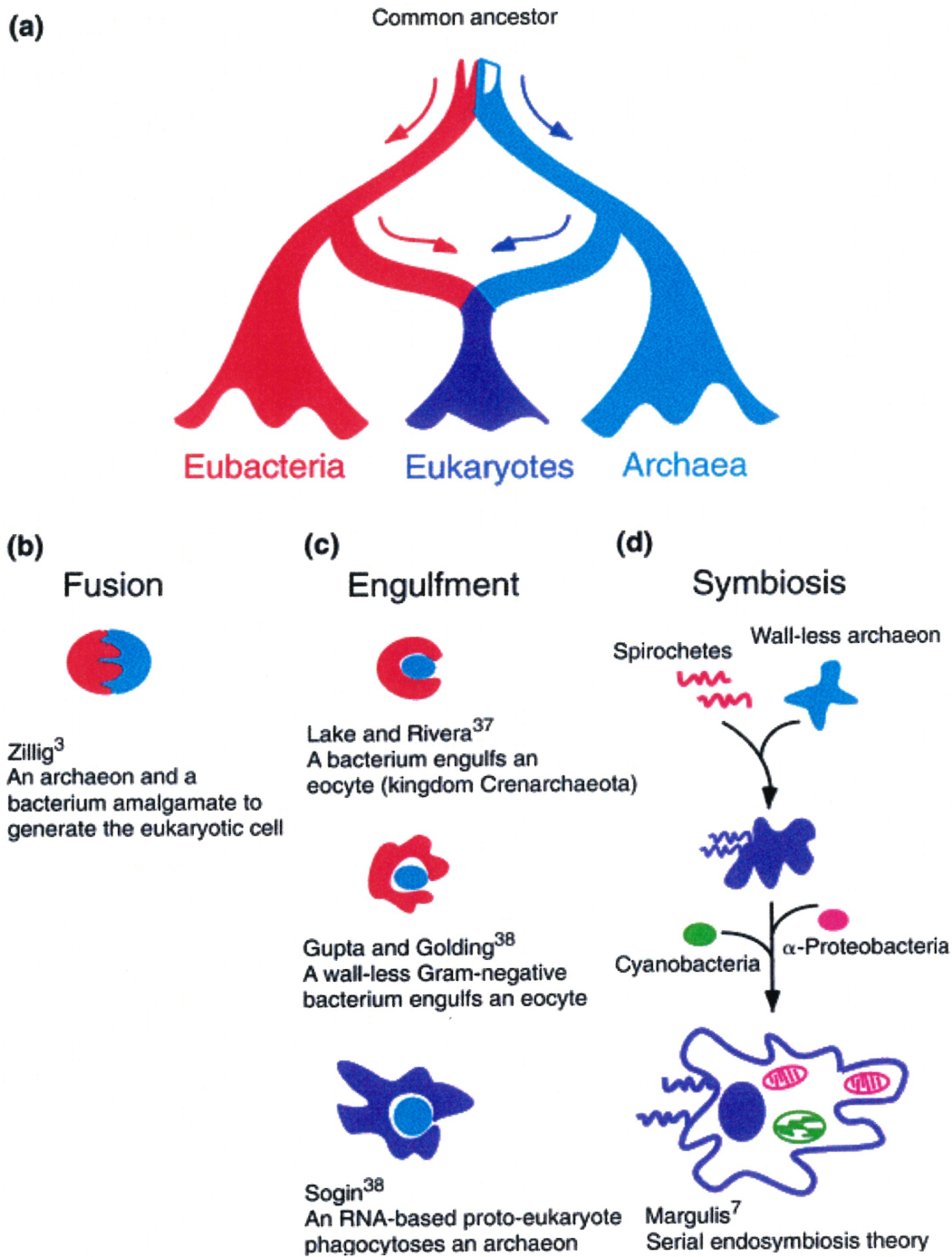


Figure 2 : Quelques hypothèses d'eucaryogenèse

(a) vue schématique de l'évolution des eubactéries, des eucaryotes et des archaea

(b)-(d) modèles chimériques de l'origine des eucaryotes

Source : Moreira et Lopez-Garcia. (1999). *TIBS*, 24 : 88-93

entre une bactérie spirochète et une archaéobactérie thermo-acidophile telle *Thermoplasma* (11-12). D'autres auteurs suggèrent l'engouffrement d'un éocyte (bactérie extrêmement thermophile métabolisant le soufre) par une bactérie (13,14), ou par un proto-eucaryote à génome ARN (15). Les plus récentes hypothèses à ce sujet (1998) sont basées sur des considérations métaboliques: l'hypothèse hydrogène (16) suppose une symbiose entre une archae strictement autotrophe dépendante de l'hydrogène et une bactérie capable de respirer mais générant de l'hydrogène comme produit de dégradation d'un métabolisme hétérotrophe anaérobie. L'hypothèse syntrophique (symbiose métabolique) (17,18), quant à elle, suppose une association symbiotique entre une archae méthanogène et une  $\delta$ -protéobactérie ou myxobactérie sulfato-réductrice.

## II-2. La mitochondrie et l'origine des eucaryotes

La plupart de ces théories supposent que l'événement ayant conduit à la cellule eucaryote précède l'acquisition de la mitochondrie. Seule l'hypothèse hydrogène (16) implique la possibilité d'un seul et même événement. Le débat sur le moment de l'implantation de la mitochondrie au sein de ce qui deviendra une cellule eucaryote n'est pas nouveau. En effet, en se basant sur des données morphologiques et sur leur divergence précoce dans l'arbre des eucaryotes (19, 20), on pensait que les Parabasalides (*Trichomonas vaginalis*), les Métamonades (*Giardia lamblia*) et les Microsporidies (*Varimorpha nectatrix*), des organismes eucaryotes sans mitochondries regroupées sous le terme "Archézoaires" par Cavalier-Smith (21), n'en avaient jamais possédé: ils auraient ainsi constitué des eucaryotes primitifs, vestiges des cellules eucaryotes existantes avant l'implantation des mitochondries en leur sein.

Mais la découverte de gènes probablement d'origine mitochondriale tel que *cpn60* (codant pour une protéine chaperon) chez *Entamoeba histolitica* (22) et chez *Giardia lamblia* (23) ont amené à considérer la possibilité de la perte secondaire des mitochondries chez ces organismes.

Ceci avait amené à se questionner de nouveau sur la datation de l'événement d'implantation des mitochondries qui peut difficilement être dissocié du phénomène d'eucaryogénèse. La mitochondrie

et la cellule eucaryote sont indubitablement liées, et tenter d'expliquer l'origine de l'une, apporte logiquement des éléments pour expliquer l'origine de l'autre. On peut espérer aboutir à cette clarification des données et des hypothèses pour la cellule eucaryote, comme il a été possible de vérifier et de valider la théorie de l'endosymbiose des mitochondries en ce qui concerne le partenaire eubactérien.

### III. Les diverses hypothèses concernant l'origine des mitochondries

Quelques unes des hypothèses qui vont suivre ont tenté d'expliquer l'origine des constituants typiques de la cellule eucaryote, tels que le noyau, les mitochondries et les chloroplastes. Néanmoins, la mitochondrie étant notre principal centre d'intérêt, nous ne nous intéresserons pour la suite qu'à la partie de ces hypothèses qui concerne l'origine des mitochondries.

#### III-1. Les hypothèses autogènes versus les hypothèses xénogènes

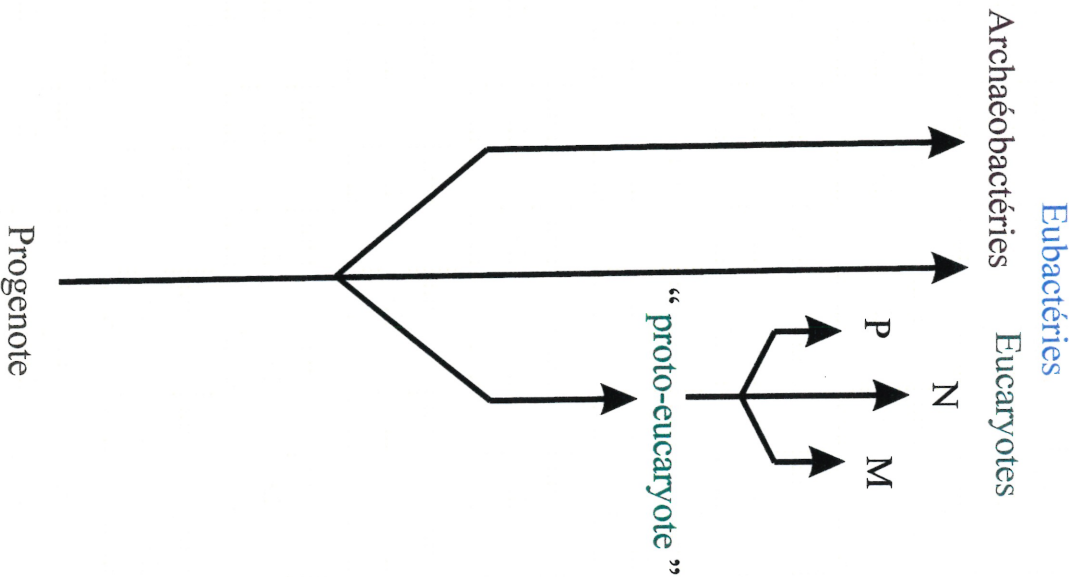
Selon la nomenclature utilisée par Gray et Doolittle (25), on peut classer les hypothèses sur l'origine des mitochondries en deux catégories (figure 3):

i. hypothèse autogène ou de filiation directe qui postule que les mitochondries proviennent d'une source intrinsèque eucaryote: la compartimentalisation physique d'une même cellule aurait été suivie d'une spécialisation fonctionnelle des compartiments ainsi créés.

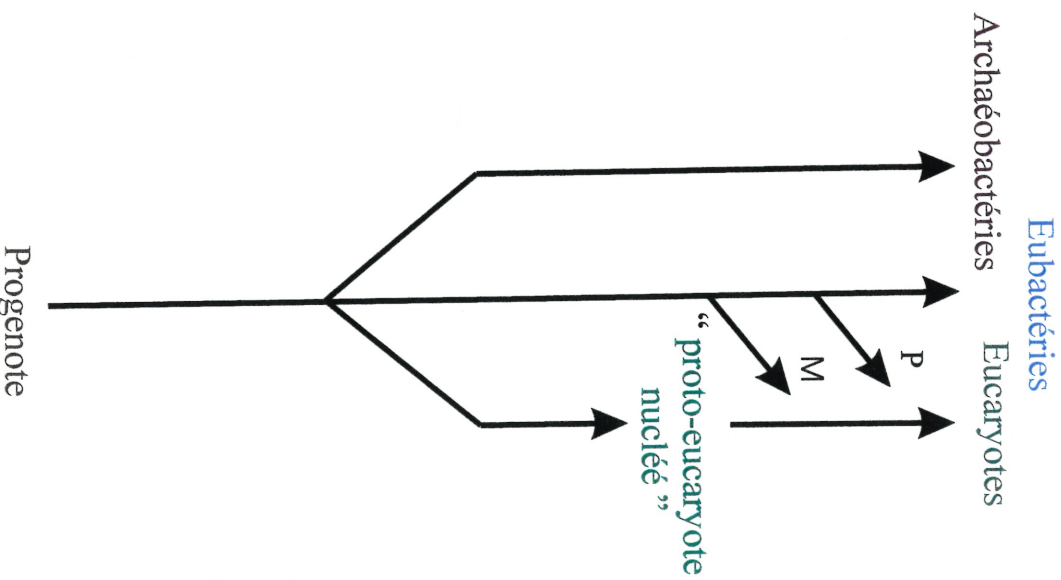
Par exemple, en 1972, Raff & Mahler (26) ont proposé l'hypothèse de l'épisome selon laquelle les mitochondries se seraient développées à partir d'une proto-mitochondrie qui serait dérivée de la membrane interne du proto-eucaryote et qui contiendrait un épisome sur lequel auraient été présents des gènes pour des composants ribosomiques mitochondriaux, ainsi que plusieurs éléments de la chaîne respiratoire. En 1975, L. Reijnders conteste cette hypothèse et en propose une autre (27) légèrement différente où chaque copie de l'ADN dupliqué d'un même procaryote se serait

Figure 3 : Origine des mitochondries

Hypothèses autogènes



Hypothèses xenogènes



P plastes  
 N noyau  
 M mitochondries

Source : Gray MW and Doolittle WF. *Microbiological reviews*, 46(1) : 1-42

retrouvée dans des invaginations de la membrane interne de l'enveloppe cellulaire, donnant ainsi naissance soit à la proto-mitochondrie, soit à la membrane nucléaire. Toujours en 1975, T. Cavalier-Smith (28) propose quant à lui une hypothèse basée sur une évolution de la phagocytose de telle sorte que des phago-lysosomes d'une même cellule dépourvue de "mur cellulaire" auraient donné les différents compartiments caractéristiques des eucaryotes.

ii. hypothèse xénogène ou d'endosymbiose qui postule que les mitochondries proviennent d'une source externe probablement bactérienne: deux types cellulaires différents seraient entrés en contact et auraient établi une symbiose si totale qu'un des deux partenaires se serait installé définitivement au sein de l'autre.

Ces hypothèses d'endosymbiose puisent leur origine à la fin du 19ème siècle, lorsque Altman propose que les mitochondries seraient des bactéries endosymbiotiques de la cellule eucaryote ancestrale (30). Bien que négligée pendant le début du siècle dernier, la théorie de l'endosymbiose a connu un regain d'intérêt au cours des années 60-70, durant lesquelles de nombreuses versions de cette hypothèse ont été publiées ( 31-32 ). En 1970, L. Margulis, la plus persévérante à défendre cette hypothèse publia un livre (33) qui fait aujourd'hui figure de référence pour cette hypothèse ayant acquis depuis le statut de « théorie de l'endosymbiose mitochondriale ».

### III-2. La théorie de l'endosymbiose

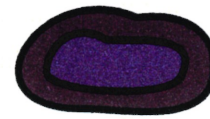
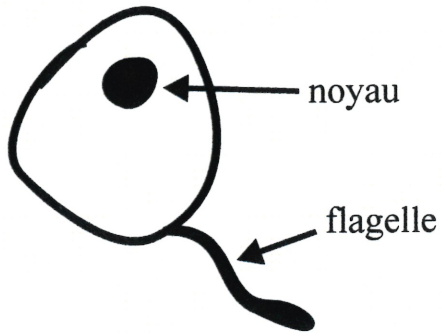
Cette théorie postule que la cellule eucaryote tire son origine d'interactions entre des partenaires qui se sont ensuite joints en une endosymbiose (figure 4). Mais de quel type aurait été les différents partenaires?

Le futur symbiote aurait été un procaryote vivant librement, qui aurait développé la capacité d'utiliser l'oxygène présent dans l'atmosphère pour respirer. Cette bactérie aérobique aurait ensuite été phagocytée par le futur hôte, le proto-eucaryote, une cellule de type amoeboïde possédant d'ores et déjà un noyau mais dépourvu de la faculté de réaliser la phosphorylation oxydative. Les partenaires auraient alors mis en place une symbiose, soit une relation durable et réciproquement

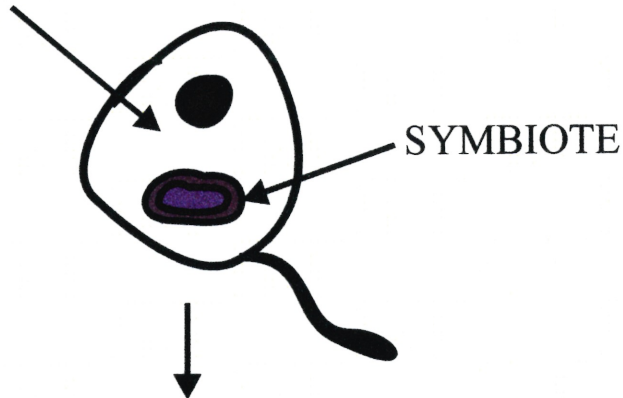
Figure 4 : schéma de la théorie de l'endosymbiose mitochondriale

Eucaryote primitif capable de phagocytose

Bactérie aérobie  
apparentée aux  $\alpha$ -protéobactéries

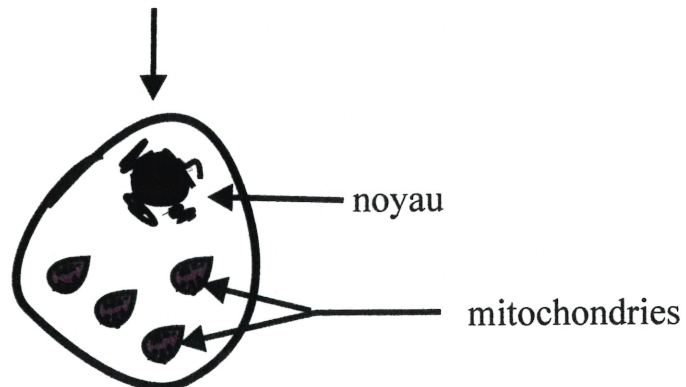


CELLULE HÔTE



ENDOSYMBIOSE

Transformation du symbiote en mitochondrie :  
réorganisation structurale, réduction du génome  
( transfert de gènes vers le noyau )



CELLULE EUCARYOTE MODERNE

favorable au cours de laquelle la bactérie symbiote aurait reçu une protection physique et des substrats métabolisables du proto-eucaryote, tandis que ce dernier aurait reçu de l'ATP et un avantage métabolique notable, la phosphorylation oxydative, de son symbiote. Au cours du temps, la symbiose serait devenue totale, le symbiote perdant progressivement son autonomie, tant au niveau de son métabolisme qu'au niveau de l'expression de son génome, au profit de l'hôte. Il y aurait ainsi eu transfert partiel de gènes du symbiote à la cellule hôte. Le symbiote, privé de la possibilité d'indépendance et d'autonomie serait devenu une organelle.

De manière intuitive, de par leur structure cytologique globale et parce que les mitochondries possèdent la capacité de produire de l'ATP suivant un système similaire à celui présent chez certains procaryotes, il est facile de les rapprocher des bactéries. Un support métabolique pour la théorie de l'endosymbiose et l'affiliation bactérienne des mitochondries a été fourni en 1975 par John et Whatley (34) qui ont démontré la forte ressemblance de *Paracoccus denitrificans* et des mitochondries au niveau de leur chaîne respiratoire. Mais une simple ressemblance, fut-elle cytologique ou biochimique, ne saurait constituer un élément de preuve. Toutes ces hypothèses, tant autogènes que xénogènes, tiennent compte de la présence d'un ADN dans les mitochondries. A l'époque où ces hypothèses ont été formulées, les outils pour exploiter les données présentes dans cet ADN n'étaient pas encore suffisamment développés. La révolution apportée par la phylogénie moléculaire a permis d'utiliser ces clés phylogénétiques, les gènes mitochondriaux, pour retracer l'histoire des mitochondries et tenter de valider la théorie de l'endosymbiose.

#### IV. Étude phylogénétique de l'ADN mitochondrial

##### IV-1. Affiliation bactérienne des mitochondries

Tout comme pour la classification à grande échelle des êtres vivants réalisée par Woese *et al.* en 1977 (1), ce fut l'ARNr des génomes mitochondriaux qui constitua la molécule de choix pour leur étude phylogénétique. Ceci permettait de comparer les données mitochondriales à celles déjà

obtenues pour des génomes nucléaires d'eucaryotes et des génomes bactériens.

Dès 1982, Gray et Doolittle (26) sont en mesure de faire une revue de la littérature qui regroupe tous les éléments connus à ce jour sur la biologie moléculaire des mitochondries et qui met en avant les éléments en faveur de l'hypothèse endosymbiotique. Ils notaient toutefois le manque de données suffisantes pour établir sans l'ombre d'un doute la véracité de l'hypothèse.

De nouvelles données sur la provenance eubactérienne des mitochondries s'accumulèrent encore. Par exemple, en 1984, Spencer *et al.* (35) ont démontré que l'ARNr de la mitochondrie de blé partageait une forte similarité tant sur le plan de la séquence primaire que des structures secondaires potentielles, avec son homologue eubactérien chez *E. Coli*. Ils en déduisirent qu'il s'agissait là d'une évidence supplémentaire en faveur d'une provenance eubactérienne des mitochondries. Mais les ARNr des mitochondries d'autres eucaryotes que le blé sont très différents des autres ARNr eucaryotes et différents également entre eux (36, 37). Tout ce que l'on pouvait conclure des données sur les ARNr étudiés à ce jour (qui n'étaient certes pas exhaustives), c'était que les mitochondries avaient une provenance eubactérienne, qu'elles branchaient précocement dans l'arbre de vie, et qu'elles provenaient peut être de plusieurs événements d'endosymbiose (37, 38).

#### IV-1-1. L'endosymbiote bactérien est une protéobactérie

En 1985, au vu de toutes les évidences citées précédemment, on en était rendu au point où la communauté scientifique acceptait le fait que les mitochondries soient le résultat d'une (ou plusieurs) endosymbiose (s) bactérienne(s) au sein d'une cellule eucaryote ancestrale, mais se questionnait sur la nature de l'endosymbiote.

Des analyses phylogénétiques de protéines présentes chez les mitochondries et les bactéries, les cytochromes *c* avaient pavé la voie vers un groupe particulier d'eubactéries, les alpha-protéobactéries (40, 4). Mais comme les cytochromes *c* étudiées à cette époque étaient codées par le noyau, on pensait que la généalogie des mitochondries déduite de l'étude de ces protéines aurait pu être sujet à controverse. C'est pour cela que cette même année, Yang *et al.* (39) ont étudié des ARNr provenant de protéobactéries des divisions alpha et bêta et en ont déduit que l'endosymbiote ayant



donné naissance aux mitochondries appartenait effectivement à la sous-division alpha des protéobactéries. Cette conclusion a été confirmée par la suite par d'autres études utilisant la comparaison d'ARNr mitochondriaux, eubactériens et eucaryotes.

Toutefois certains chercheurs, au vu en particulier des taux d'évolution radicalement distincts des ARNr entre les différentes lignées eucaryotes (26, 42), ont choisi d'utiliser d'autres « chronomètres » moléculaires, les chaperonnes mitochondriales (Hsp60). Ces protéines codées par le noyau, remarquablement conservées tant au niveau structural que fonctionnel, sont distribuées de manière ubiquitaire au sein des eubactéries et des organelles des cellules eucaryotes (43). La conclusion fut la même : les mitochondries partagent un ancêtre commun avec les alpha-protéobactéries, en particulier avec la sous division contenant des parasites obligatoires de cellules eucaryotes tels que *Rickettsia*, *Ehrlichia* et *Anaplasma* (44, 45).

D'autres gènes, codés par les mitochondries cette fois ont été utilisé dans cette optique : les gènes codant le cytochrome *b* et la sous-unité I de la cytochrome *c* oxydase, qui constituent deux des protéines les plus conservées de la chaîne de transport des électrons (46). L'analyse phylogénétique de ces gènes a confirmé une nouvelle fois les résultats précédents. Ainsi, à la fois par l'étude de gènes codés par le noyau et de gènes codés par les mitochondries, il a été montré que les mitochondries proviennent des alpha-protéobactéries et en particulier du groupe des Rickettsiales.

#### IV-2. Un ou plusieurs événements d'endosymbiose ?

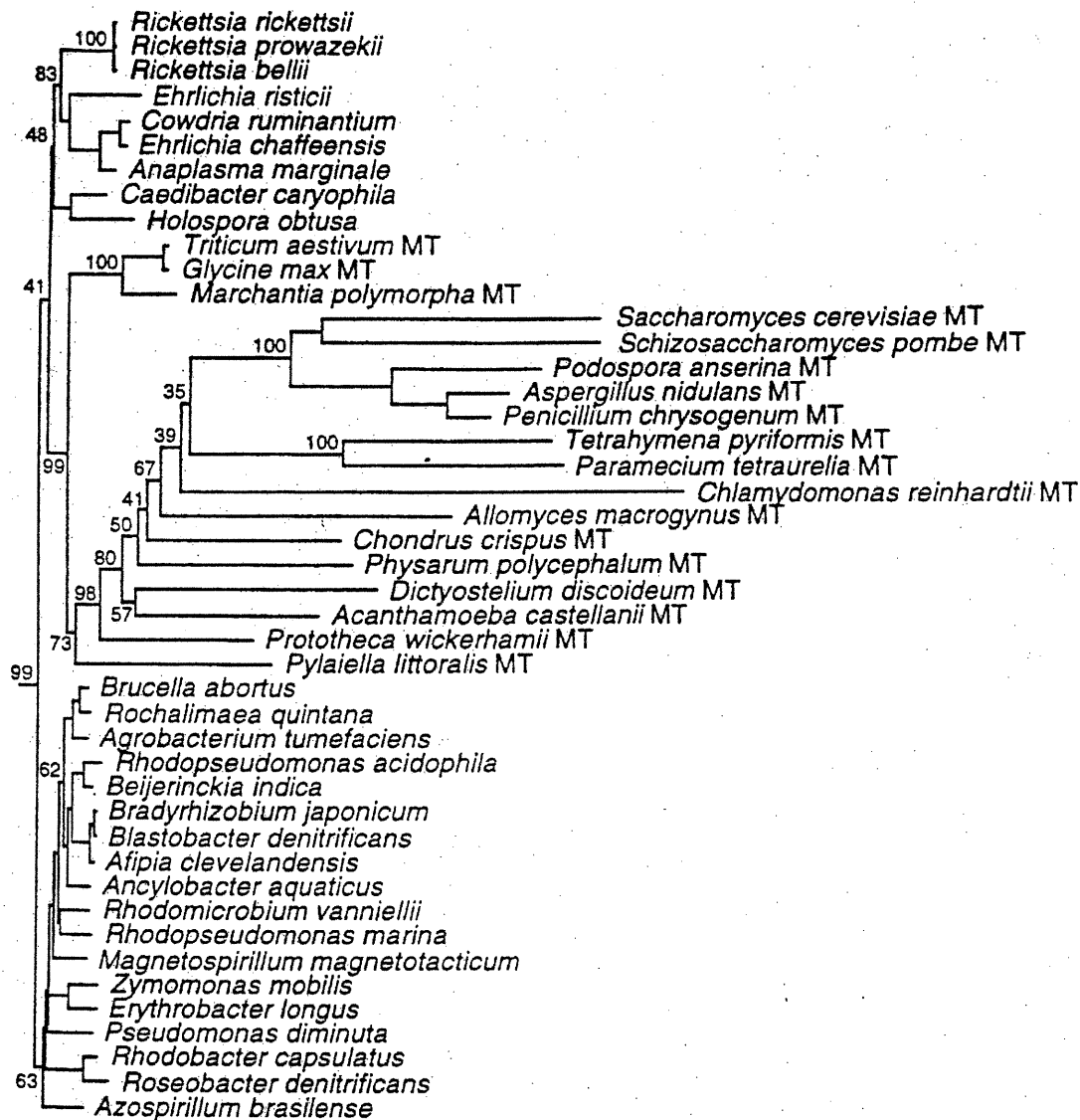
Tout au long de la recherche sur la nature de l'endosymbiote mitochondrial, les chercheurs se sont également questionnés sur la possibilité d'une origine monophylétique (soit un seul événement d'endosymbiose, donc un ancêtre communs) ou polyphylétique (plusieurs événements d'endosymbiose donc plusieurs ancêtres communs) des mitochondries.

De fait, bien avant que l'accumulation de données nous incite à la prudence dans l'utilisation des ARNr mitochondriaux, la découverte d'une déviation au code génétique dans de nombreuses mitochondries avait amené à ce même questionnement sur la possibilité d'une origine polyphylétique des mitochondries (26).

Dès l'utilisation des gènes ARNr de mitochondries de diverses espèces, telles les mitochondries de plantes versus les mitochondries d'animaux, la question s'est posée de savoir si toutes ces mitochondries si différentes entre elles pouvaient provenir du même événement d'endosymbiose (26). Comme cité précédemment (26, 42), le fait que les ARNr des mitochondries montrent des taux extrêmement variés d'évolution et des particularités dans les différentes lignées eucaryotes empêche une reconstruction phylogénétique fiable de l'origine des mitochondries. Seules les mitochondries de plantes échappaient quelque peu à cette règle (35). Il est cependant possible de construire une phylogénie en laquelle on puisse avoir confiance en utilisant des ARNr mitochondriaux : en 1996, Gray et Spencer (48, Figure 5), ont construit un arbre phylogénétique basé sur des séquences d'ARNr qui démontre la monophylie des mitochondries et leur affiliation  $\alpha$ -protéobactérienne. L'interprétation d'autres arbres utilisant des séquences protéiques et non plus nucléotidiques (44, 45) avaient conduit à cette même conclusion.

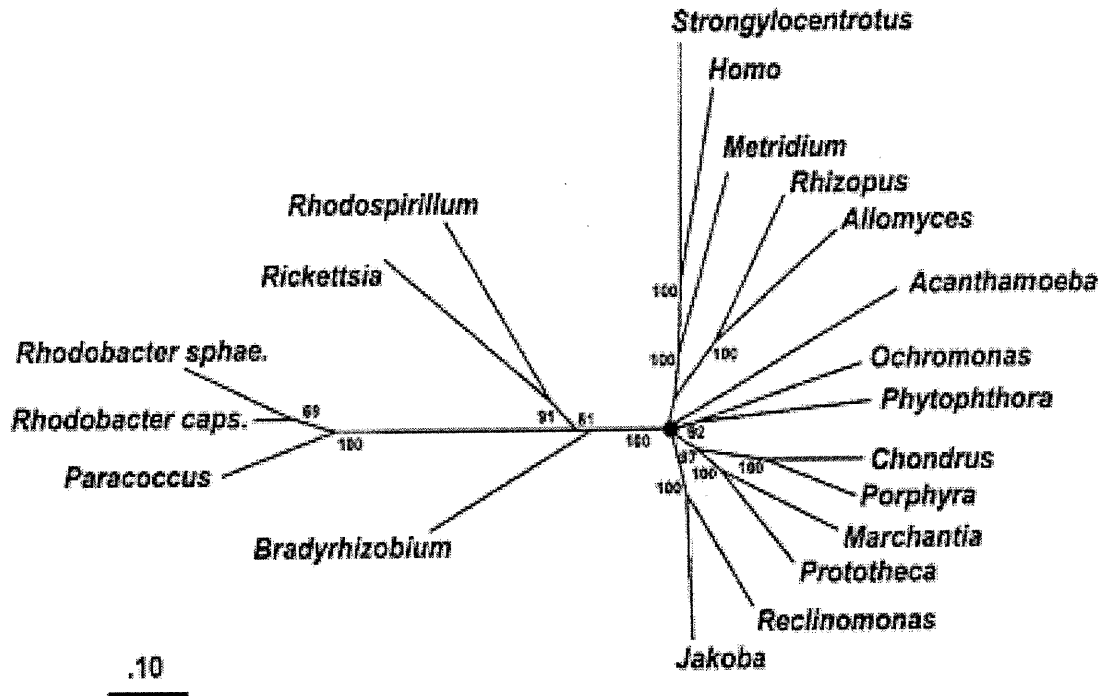
La révolution génomique est venue au secours des chercheurs en apportant un flot de données sous forme notamment de génomes mitochondriaux entièrement séquencés et d'un nombre de gènes séquencés plus important dans les trois domaines de vie. En intégrant des données sur plusieurs gènes protéiques dont la séquence est connue à la fois chez les mitochondries, les bactéries et les eucaryotes, on peut maximiser la reconstruction phylogénétique : en effet, en concaténant des séquences protéiques (c'est à dire en les alignant les unes à la suite des autres) dans une même analyse phylogénétique, on augmente le nombre de sites informatifs. C'est par cette méthode que Gray, Lang et Burger (49,50) démontrent la radiation rapide de lignées distinctes d'eucaryotes, qui constitue un support pour la monophylie des mitochondries (Figure 6).

D'autre part, l'étude du génome mitochondrial le plus eubactérien séquencé à ce jour a apporté de nouveaux éléments en faveur de la probable monophylie des mitochondries.



**Figure 5** : Portion d'un arbre phylogénétique des ARNr montrant les relations entre les organelles et les eubactéries. Cette portion concerne les relations entre les mitochondries (MT) et les alpha-protéobactéries. L'analyse utilise un ensemble de données alignées de 236 séquences de la petite sous-unité ribosomique d'eubactéries et de mitochondries (1009 caractères). L'arbre a été généré en utilisant le programme DNADIST de PHYLIP version 3.5 (90) avec l'option ML et un ratio de transition-transversion de 1.0. La matrice de dissimilarité de ce programme a été utilisé comme entrée (« input ») dans le programme NEIGHBOR (le programme neighbor-joining de PHYLIP). Les analyses de bootstrap (100 répliqués) ont été réalisés en utilisant les programmes SEQBOOT, DNADIST, NEIGHBOR et CONSENSE.

Source : Gray et Spencer (1996). Organellar evolution (Cambridge Univ. Press, Cambridge) pp 107-126



**Figure 6** : Relations phylogénétiques entre les mitochondries et les alpha-proteobacteries. Un ensemble de séquences d'acides aminés alignées et concaténées, correspondant aux protéines de la chaîne respiratoire apocytochrome *b* (Cob) et les sous-unités 1 à 3 de la cytochrome *c* oxydase (Cox1-3) a été utilisée pour l'analyse.

Les taxa comprennent des représentants des groupes eucaryotes majeurs incluant des jakobides et toutes les  $\alpha$ -protéobactéries pour lesquelles les données étaient disponibles à l'époque. L'analyse phylogénétique a été réalisée avec la plus récente implémentation de PROTDIST/FITCH (90), qui permet une correction Jin/Nei pour les taux de changements inégaux aux différentes positions d'acides aminés. Le coefficient de variation utilisé était de 0,5. Le cercle noir illustre les courtes branches ayant été collapsées pour refléter l'ordre de branchement non résolu. (> 60% de support de bootstrap). La barre horizontale illustre le nombre moyen de substitution par site. Cette topologie de l'arbre est aussi supportée par des analyses de maximum likelihood. Organismes et séquences (numéros d'accèsion à GenBank entre parenthèses) sont *Strongylocentrotus purpuratus* (" sea urchin" ; X12631), *Paracoccus denitrificans* (X05829, M17522, X05934, X05828), *Rickettsia prowazekii* (AJ235270 a AJ 235273), *Rhodobacter sphaeroides* [ X56157, X62645, M57680, C454164 (protéine) ], *Rhodospirillum rubrum* (X55387), *Rhodobacter capsulatus* (X05630), *Bradyrhizobium japonicum* (J03176, X68547).

Source : Lang et al. (1999). *Annu. Rev. Genet.*, **33** : 351-397

#### IV-2-1. Le génome mitochondrial de *Reclinomonas americana*

*Reclinomonas americana*, un jakobide, protozoaire hétérotrophe d'eau douce (Figure 7), possède le génome mitochondrial le plus eubactérien séquencé à ce jour. Ses 69,034 paires de bases contiennent la plus grande collection de gènes jamais identifiés dans un ADN mitochondrial : il possède les gènes pour l'ARNr 5S, la composante ARN de la ARNase P (*rnpB*) et 18 protéines qui n'avaient jamais été identifiées dans un autre génome mitochondrial (51, 52, Figure 7).

Entre autres faits marquants, ce génome code pour une ARN polymérase de type eubactérienne ( $\alpha 2 \beta \beta' \sigma$ ) alors que dans les autres eucaryotes, l'ARN polymérase mitochondriale est un polypeptide codé par le noyau et de même type que celle du bactériophage T3-T7. Il code également pour une autre protéine typiquement eubactérienne, SecY, une composante universelle de la machinerie de translocation des protéines à travers les membranes des eubactéries et des chloroplastes. On peut aussi observer dans ce génome des vestiges d'une organisation sous forme d'opéron pour plusieurs regroupements de gènes ainsi que des interactions de type Shine-Dalgarno, typiquement eubactériennes (51, 52, Figure 8).

Au vu de toutes ces caractéristiques eubactériennes, Lang *et al.* concluent que ce génome mitochondrial est celui qui ressemble le plus au génome proto-mitochondrial putatif (51, 52). Il est vrai que tous les gènes codant des protéines présents dans les génomes mitochondriaux séquencés constituent des sous-ensembles des gènes codant des protéines présents chez *Reclinomonas americana* (Figure 8). Ceci serait difficilement expliqué par une origine polyphylétique des mitochondries où chaque génome mitochondrial dans les différentes lignées eucaryotes serait le résultat de réductions indépendantes de génomes eubactériens plus grands (49). D'autre part, comme le génome mitochondrial de *Reclinomonas* code pour moins de 2% des gènes protéiques trouvés chez *E. coli*, une réduction convergente au même ensemble de gènes est improbable (49).

La parcimonie peut expliquer une autre caractéristique des génomes mitochondriaux. Dans de nombreux cas, les regroupements de gènes codant des protéines sont similaires à ceux présents dans les génomes de leurs homologues bactériens. Toutefois, chez les mitochondries ces regroupements sont affectés par des délétions ou réarrangements spécifiques dont la présence s'explique le plus

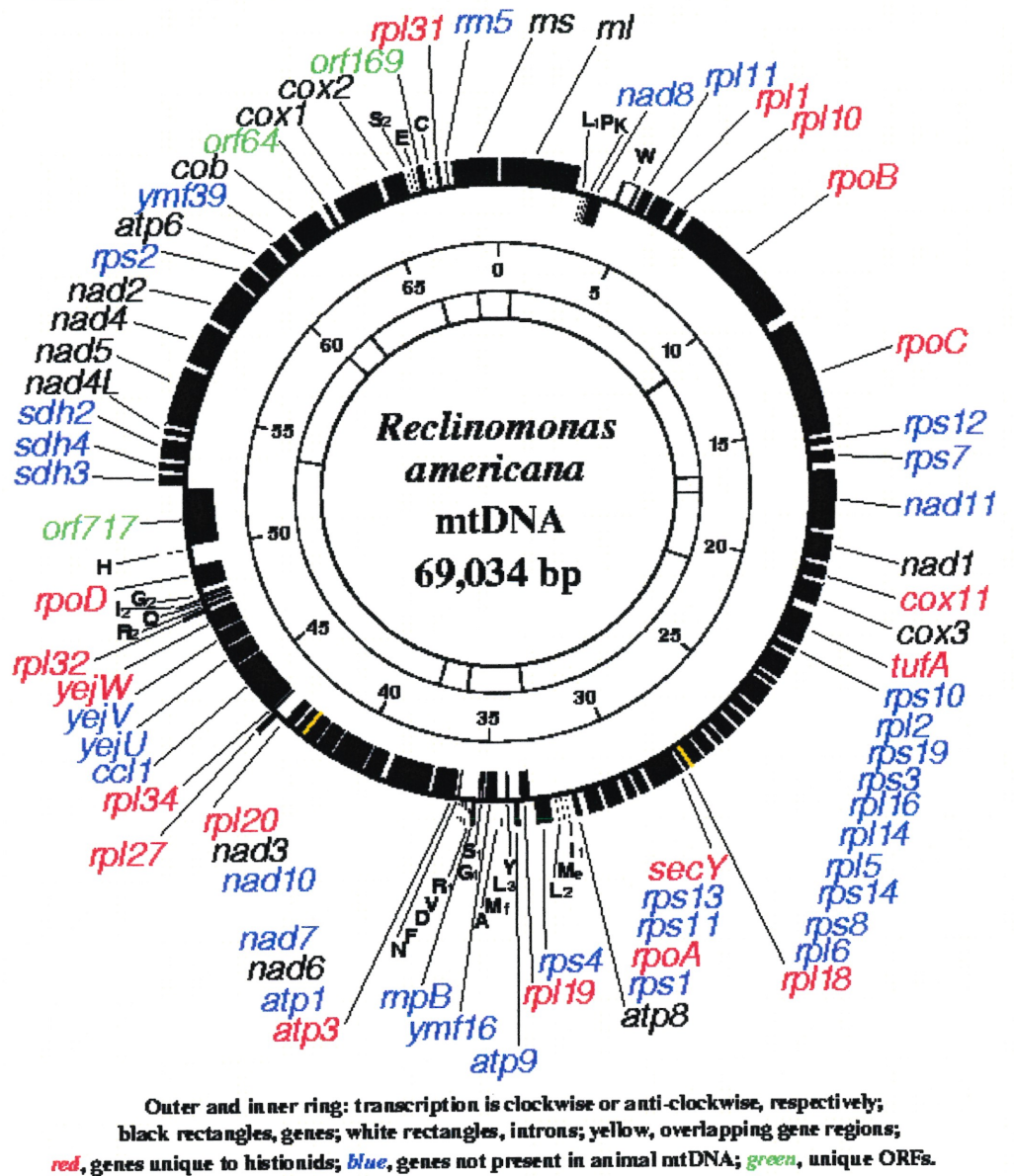


Figure 7 : Photo du jakobide flagellé *Reclinomonas americana* et carte génomique de son ADN mitochondrial. Source : OGMP <http://megasun.bch.umontreal.ca/protists/gallery.html>  
<http://megasun.bch.umontreal.ca/maps/ramermap.gif>

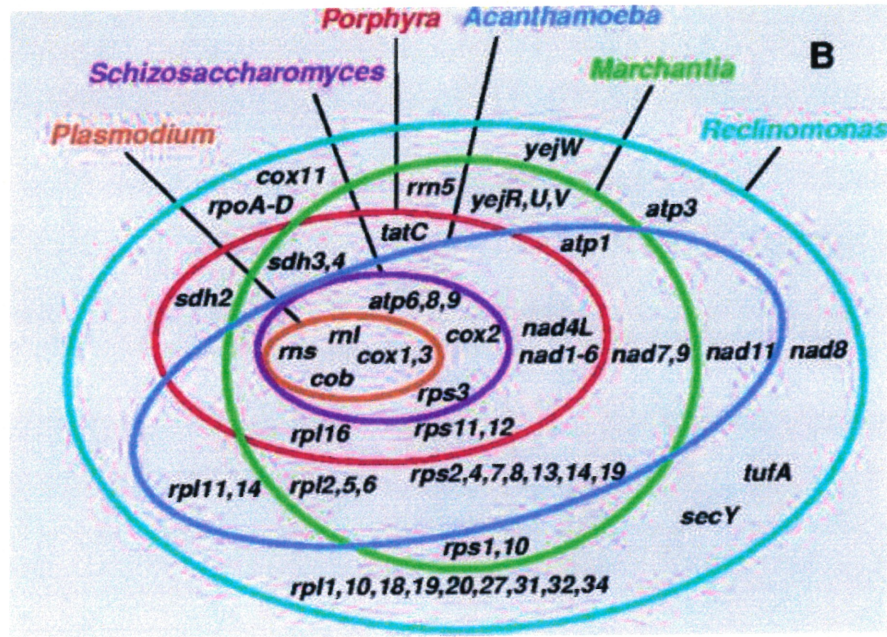


Figure 8(A). Contenu génomique des génomes mitochondriaux. Chaque ovale correspond à un organisme ; les gènes compris dans un ovale sont présents dans l'ADNmt de l'organisme. Seuls les gènes rRNA (*rnl*, *rns*, *rrn5*) et les gènes codant des protéines sont montrés.

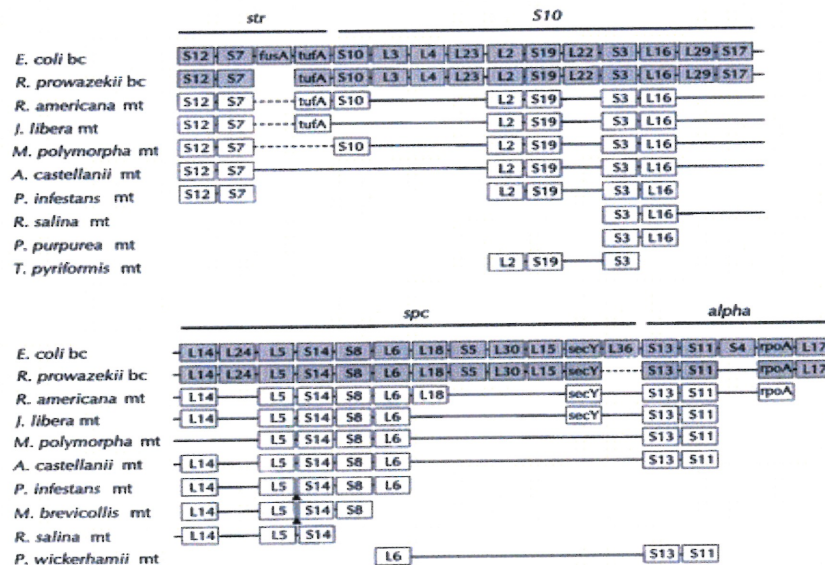


Figure 8(B). Organisation en opéron de certains gènes (codant pour les protéines de la petite et de la grande sous-unité ribosomique) comparée entre des génomes mitochondriaux et les génomes de protéobactéries, d'*E. coli* et de *R. prowazekii*.

Sources : (A) Gray et al. (1999). *Science*, **283** : 1476-1481  
 (B) Lang et al. (1997). *Nature*, **387** : 493-497

parcimonieusement par un événement unique ayant eu lieu dans le génome de l'ancêtre commun aux mitochondries, après sa divergence de son ancêtre bactérien (53).

Ces observations conduisent Lang et *al.* (outre l'utilisation des arbres phylogénétiques construits à l'aide de séquences protéiques concaténées qui démontre que les Jakobides font partie de la radiation rapide des lignées eucaryotes (Figure 6)) à établir la possibilité de la monophylie des mitochondries. L'étude phylogénétique des ADN mitochondriaux nous a permis de valider la théorie de l'endosymbiose par de multiples études dont les résultats convergent vers un même scénario.

Néanmoins, il reste encore matière à exploration en ce qui concerne la nature de l'endosymbiote  $\alpha$ -protéobactérien ayant donné naissance à l'ensemble des mitochondries.

## V. La nature de l'endosymbiote bactérien

### V-1. Les protéobactéries

Il a précédemment été démontré que les mitochondries tire leur origine au sein des alpha-protéobactéries (cf. IV-1-1). Les protéobactéries font partie d'un des groupes bactériens majeurs mis en place avant l'avènement de la phylogénie moléculaire, les bactéries Gram négatives (un autre groupe étant les bactéries Gram positives). Sous cette appellation sont regroupées des bactéries de toutes sortes de forme, d'habitat, de mode de nutrition, qui ont une chose en commun : elles colorent rouge suite à l'application de la coloration de Gram, autrement dit elles ne retiennent pas la coloration violette de Gram, contrairement aux bactéries Gram<sup>+</sup>. Ces différences de coloration correspondent à une différence de structure de leur paroi bactérienne, les bactéries Gram<sup>+</sup> possédant une couche de peptidoglycane à l'extérieur de leur membrane cellulaire, tandis que les bactéries Gram<sup>-</sup> possèdent une membrane externe en plus de leur peptidoglycane (Figure 9).



Figure 9(A) : Photo de *Neisseria gonorrhoeae*, suite à la coloration de Gram (bactérie Gram- qui colore rouge) et schéma de la structure de la paroi d'une bactérie Gram-

source : <http://www.meddean.luc.edu/lumen/DeptWebs/microbio/med/gram/slides/slide2.html>

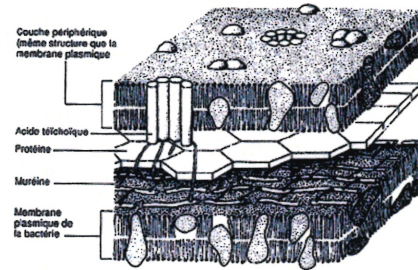
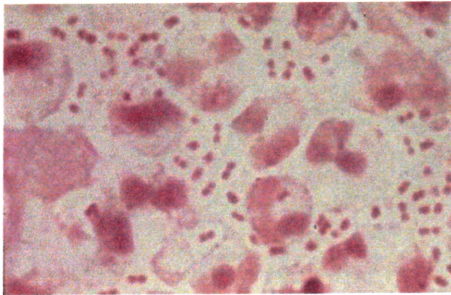
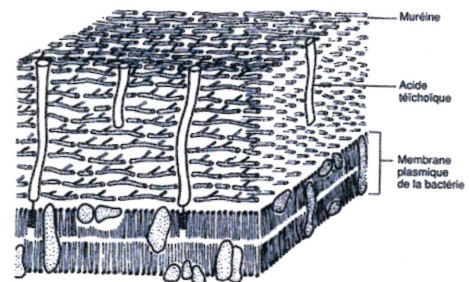
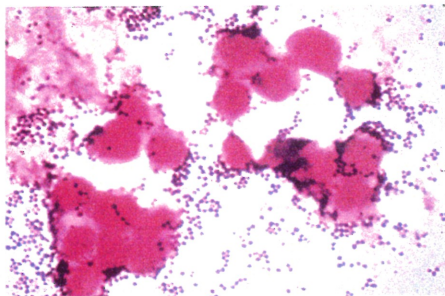


Figure 9(B) : Photo de *Staphylococcus aureus*, suite à la coloration de Gram (bactérie Gram+ qui colore violet) et schéma de la structure d'une bactérie Gram+

source : <http://medic.med.uth.tmc.edu/path/00001456.htm>



Les phylogénies moléculaires construites à partir des ARNr 16S des bactéries Gram- ont démontré que celles-ci ne forment pas dans leur ensemble un groupe monophylétique : malgré tout, un groupe populeux, les protéobactéries (54) forment un groupe monophylétique qui comprend la plupart des organismes Gram- connus.

## V-2. Définition des protéobactéries

Avant d'être nommées protéobactéries, ces bactéries Gram- étaient regroupées sous le terme générique, « bactéries pourpres photosynthétiques et parentes », qui n'était pas approprié puisque la plupart de ces bactéries n'étaient ni pourpres, ni photosynthétiques. Ce groupe bactérien se distingue par la grande diversité des styles de vie de ses membres. Ces bactéries peuvent être photosynthétiques, non photosynthétiques, aérobies, anaérobies ou microaérophiles, certaines encore pathogènes ou commensales d'organismes multicellulaires. Au vu de toute cette diversité, les taxinomistes ont désigné ces bactéries par le terme Protéobactéries, du nom du dieu grec de la mer *Proteus*, capable d'adopter différentes formes, et du nom *bakterion*, petit bâton : ainsi, *Proteobacteria* désigne un groupe de bactéries de diverses propriétés ayant une ancestralité commune (54).

### V-2-1. Les $\alpha$ -protéobactéries

Les  $\alpha$ -protéobactéries incluent la plupart des protéobactéries oligotrophes, ainsi que les bactéries pourpres non sulfureuses (55). Ces dernières possèdent les mêmes caractéristiques que les bactéries pourpres, soit l'utilisation d'une photosynthèse non-oxygénée, des bactériochlorophylles a et b ainsi que des photosystèmes dans leurs membranes en continu avec la membrane plasmique. Les  $\alpha$ -protéobactéries sont flexibles dans le choix de leur source d'énergie : certaines de ces bactéries sont des pathogènes (*Rickettsia*, *Holospora*, *Ehrlichia*), certaines ne tolèrent que de faibles concentrations d'oxygène et sont hétérotrophes (*Magnetospirillum*) alors que d'autres sont hétérotrophes mais ne croissent qu'à de hautes quantités d'oxygène (*Rhizobium*, *Aquaspirillum*) (50).

Au sein des  $\alpha$ -protéobactéries, on remarque une tendance de certains de ses membres à vivre en

interaction plus ou moins forte, dommageable ou bénéfique, avec des cellules eucaryotes. En effet, certaines familles d' $\alpha$ -protéobactéries sont des symbiotes, des commensales ou des parasites d'autres cellules. Par exemple, la famille des Rhizobiacées comprend des bactéries telles que *Rhizobium*, qui est un symbiote fixateur d'azote bénéfique à la plante au niveau de ses racines, et des bactéries telle que *Agrobacterium*, qui provoque des tumeurs dans les cellules végétales où elle s'installe. Le règne animal n'est pas épargné par ce type d'interactions : la famille des Rickettsiales, dont tous les membres sont soit des bactéries parasites, soit des bactéries mutualistes, comprend de nombreux pathogènes responsables de maladies chez l'humain, telles *Rickettsia prowazekii* responsable du typhus, et *Rickettsia rickettsii*, responsable de la « fièvre boutonneuse des Montagnes Rocheuses ».

La génomique est venue accroître notre connaissance des  $\alpha$ -protéobactéries avec le séquençage en 1998 de l' $\alpha$ -protéobactérie responsable du typhus, *Rickettsia prowazekii* (57) et plus récemment en 2001 de *Rickettsia conorii* (57). Au delà de la connaissance d'importantes bactéries pathogènes de l'espèce humaine, ces nouvelles données ont permis d'approfondir nos connaissances sur un groupe de bactéries, les Rickettsiales, alors le plus proche de l'origine des mitochondries (46, 48).

## VI. *Rickettsia prowazekii* et l'origine des mitochondries

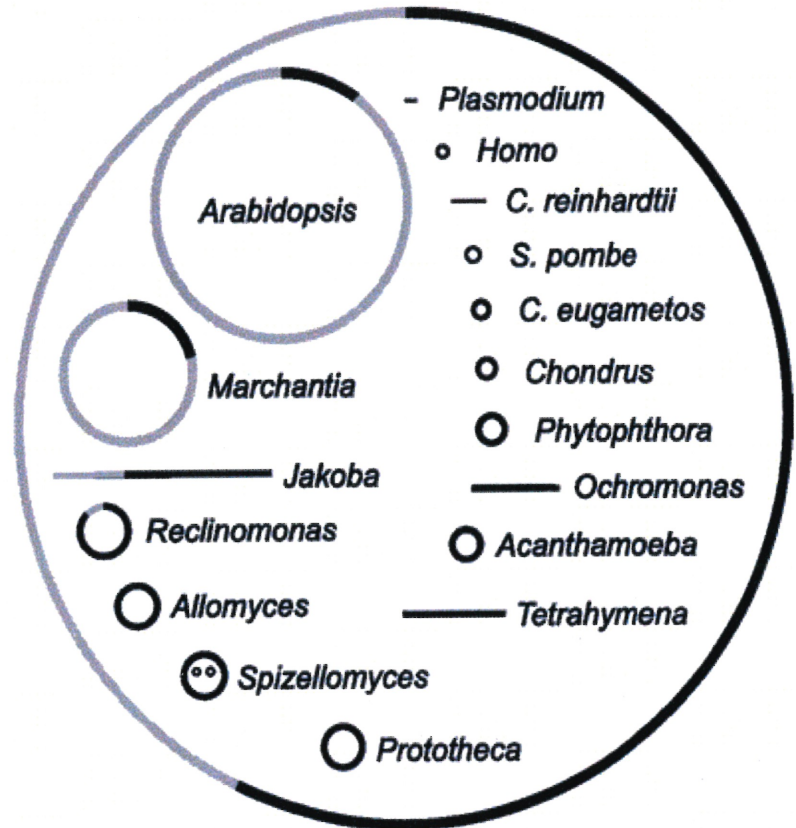
### VI-1. Le génome de *Rickettsia prowazekii*

#### VI-1-1. Réduction évolutive du génome de *Rickettsia prowazekii*

*Rickettsia prowazekii*, bactérie responsable du typhus est un parasite intracellulaire obligatoire : c'est à dire qu'elle ne peut vivre qu'à l'intérieur de cellules eucaryotes. Comme d'autres bactéries vivant dans cette niche écologique particulière, la transition à son nouvel environnement s'est accompagnée d'une réduction significative de son génome qui se traduit par la perte caractéristique des gènes codant pour les composantes des voies de biosynthèse d'acides aminés et de nucléotides (56,58). Ces processus de biosynthèse, parmi les plus lourds à maintenir pour la



## Rickettsia



**Figure 10** : Photo de *Rickettsia prowazekii* et comparaison de la taille de son génome avec celui de génomes mitochondriaux. Les cercles et les lignes représentent les formes des génomes circulaires et linéaires, respectivement. Pour les génomes dont la taille est supérieure à 60 kpb, l'ADN codant pour des gènes de fonction connue (noir) est différencié de celui codant pour des cadres ouverts de lecture non identifiés et les régions intergéniques (gris). Les noms d'espèces sont *C. (Chlamydomonas) reinhardtii*, *C. (Chlamydomonas) eugametos*, et *S. (Schizosaccharomyces) pombe*.

Source : Lang et al. (1999). *Annu. Rev. Genet.*, **33** : 351-397 (figure)  
[http://www.nature.com/genomics/papers/r\\_prowazekii.html](http://www.nature.com/genomics/papers/r_prowazekii.html) (photo)

cellule bactérienne, sont remplacés par des systèmes permettant de puiser les éléments manquants directement au cœur du cytoplasme de l'hôte, ce qui conduit finalement à une dépendance totale de la bactérie vis-à-vis de son hôte. La perte de gènes essentiels tels que les gènes de biosynthèse d'acides aminés et de nucléotides explique en partie la taille relativement petite (1,11 Mpb pour *R. prowazekii*) de ce type de génome (59) comparativement à celles d' $\alpha$ -protéobactérie vivant librement (~3.7 Mpb pour *Rhodobacter capsulatus*).

Cette réduction des génomes de micro-organismes résidents obligatoires d'un environnement conditionné par le génome de l'hôte peut s'expliquer par la théorie du « cliquet de Muller » (60). Au sein du cytoplasme de l'hôte, ces micro-organismes résidents forment de petites populations clonales

de bactéries avec de faibles possibilités de recombinaison entre les variants. Ceci peut entraîner des effets irréversibles sur les génomes de ces organismes : par exemple, des délétions et des mutations délétères non létales pourraient s'accumuler avec le temps dans ces génomes en absence d'une contre-sélection efficace (61). Cette augmentation de la charge évolutive sur ces génomes entraîne l'apparition de pseudogènes, et la perte totale de gènes.

Ce phénomène explique la petitesse du génome chez *R. prowazekii* (61, 62) ainsi que la présence de nombreux pseudogènes ainsi que le haut contenu en régions non codantes (24%).

#### VI-1-2. Profils fonctionnels de *R. prowazekii* et des mitochondries

Les phylogénies moléculaires démontrent clairement que les mitochondries tirent leur origine des  $\alpha$ -protéobactéries, et plus particulièrement des Rickettsiales (cf. IV). Le séquençage d'une  $\alpha$ -protéobactérie offrait l'opportunité de comparer les génomes mitochondriaux, en particulier le génome ancestral de *Reclinomonas americana* (67 protéines) avec celui d'une de leur plus proche parente bactérienne *Rickettsia prowazekii* (834 protéines) (figure 10).

On observe que (63-64) :

- i. Le génome de *R. americana* est dépourvu de gènes opérationnels tels que ceux impliqués dans la biosynthèse des cofacteurs, le métabolisme des acides gras et des phospholipides, le métabolisme intermédiaire et énergétique, la synthèse de l'enveloppe cellulaire et la division cellulaire. Il est dépourvu aussi de la plupart des gènes informationnels, soit ceux impliqués dans la réplication, la transcription et la traduction. Pour la plupart, ces gènes sont maintenant codés dans le noyau suite à un transfert de gènes (49).
- ii. Les deux génomes sont dépourvus de gènes impliqués dans la biosynthèse d'acides aminés et de nucléotides.
- iii. Tous deux sont dépourvus de gènes impliqués dans le métabolisme des sucres en absence d'oxygène, soit la glycolyse anaérobie.
- iv. Tous deux possèdent un système de production et de transport d'ATP similaire : des protéines pour le cycle de Krebs (qui métabolise le pyruvate), la chaîne de transport d'électrons, une  $F_1F_0$  ATP synthétase et une ADP/ATP translocase.

On peut conclure que les profils fonctionnels de ces deux génomes hautement dérivés sont similaires (63) : ceci peut s'expliquer par le mode de vie particulier de leur propriétaire (64). La réduction de ces deux génomes suite à leur installation au sein de cellules hôtes a certainement eu lieu de manière indépendante (49, 56, 63) : mais à quel moment cette installation a-t-elle eu lieu ? Si les futures mitochondries sont venues s'installer au sein de cellules proto-eucaryotes, les Rickettsiales sont des parasites de cellules animales, cellules eucaryotes d'origine relativement récente par rapport aux présumées cellules proto-eucaryotes. Dans ces conditions, serait-il possible qu'une Rickettsiale ait directement donné naissance aux mitochondries ? Il semblerait que non, et d'autres évidences viennent étayer cette hypothèse.

#### VI-2. *Rickettsia prowazekii*, ancêtre direct des mitochondries ?

L'organisation des gènes en opéron (soit le regroupement de gènes codant pour un même

processus) est strictement caractéristique des bactéries. Ce type d'organisation est retrouvée dans le génome mitochondrial de *Reclinomonas americana*, ainsi que dans celui de *Rickettsia prowazekii* ou d'autre bactérie telle que *Escherichia coli*. L'organisation de certains opérons, comme celle des protéines ribosomiques (51) est même conservée entre le génome mitochondrial de *R. americana* et le génome d'*E. coli* (avec la délétion de certains gènes chez *R. americana*).

Mais, les génomes de *Rickettsia* et de *Reclinomonas* ne sont liés spécifiquement par aucune caractéristique dérivée partagée (ou synapomorphie), notamment au niveau de l'organisation des gènes (63, 53). Autrement dit, les différences observées dans l'organisation en opéron (par rapport à *E. coli*) chez *Rickettsia* et chez *Reclinomonas* ne sont pas identiques : *R. prowazekii* en possède même certaines qui lui sont propres.

De plus, une interaction putative de type Shine-Dalgarno, typiquement eubactérienne, entre des ARNm et l'ARN de la petite sous-unité ribosomique a été découverte dans le génome mitochondrial de *Reclinomonas* (51), mais semble avoir été perdue dans les bactéries reliées aux Rickettsiales (52). Si l'on continue de parcourir le génome de *Rickettsia*, on voit que les régions intergéniques sont habituellement longues et extrêmement riche en A+T (cf. VI-1-1), ce qui ressemble plus au génome mitochondrial de levure, très dérivé, qu'au génome ancestral de protistes tels que *Reclinomonas americana* (52). Ainsi *Rickettsia* manque de caractéristiques présentes chez les mitochondries considérées comme primitives : par conséquent, il semble peu probable qu'une telle bactérie ait donné naissance aux mitochondries.

En ce qui concerne le profil fonctionnel similaire des mitochondries et de *Rickettsia* pour la production d'ATP, les aspects de l'utilisation de ce même ATP sont relativement différents. Alors que les mitochondries exportent leur ATP vers le cytoplasme, *Rickettsia* utilise l'ATP qu'il produit, et, dans les phases précoces de son développement, profite de l'environnement riche en ATP qu'est le cytoplasme de son hôte pour échanger l'ADP qu'il produit avec l'ATP de son hôte à l'aide de son ADP/ATP translocase (58, 63). Cette protéine n'est pas reliée spécifiquement aux ADP/ATP translocases mitochondriales : les translocases seraient apparues de manière indépendante au cours de l'adaptation intracellulaire du parasite et de l'organelle, après leur divergence d'un ancêtre commun (63).

Du point de vue phylogénétique, il a clairement été démontré que les mitochondries proviennent des  $\alpha$ -protéobactéries, et qu'au sein de ces  $\alpha$ -protéobactéries, les Rickettsiales sont particulièrement proches des mitochondries. Mais si l'on regarde de plus près les arbres phylogénétiques, construits à partir de séquences nucléotidiques (48), comme à partir de séquences protéiques (52), on remarque que les distances entre *Rhodobacter*, *Rhodospirillum* et *Paracoccus* ( $\alpha$ -protéobactéries vivant librement), ainsi que *Agrobacterium*, *Bradyrhizobium* ( $\alpha$ -protéobactéries parasites/symbiotes) et les mitochondries sont seulement légèrement plus grandes que la distance entre *R. prowazekii* et les mitochondries (52). D'autres arbres phylogénétiques construits avec de plus grandes collections de protéines (49-50) indiquent même que certaines  $\alpha$ -protéobactéries telles que *Rhodospirillum* et *Bradyrhizobium*, branchent exactement au même endroit voire même avant *Rickettsia* avec de bons supports statistiques (Figure 7).

À la suite de toutes ces évidences, il semble raisonnable de postuler l'existence d'une  $\alpha$ -protéobactérie dont l'ancêtre serait sensiblement plus proche de l'endosymbiote mitochondrial que l'ancêtre de *Rickettsia*, autant en terme de séquences, qu'en terme d'organisation et d'expression du génome.

Dans cette optique, mon projet de recherche a visé à découvrir une  $\alpha$ -protéobactérie entre autre phylogénétiquement plus proche des mitochondries, soit plus susceptible d'avoir donné naissance aux mitochondries que *Rickettsia*.



## CHAPITRE II

# MATÉRIELS & MÉTHODES

## I. Principes de l'expérience

### I-1. Expériences préliminaires

Notre laboratoire a séquencé le génome mitochondrial de l'algue verte *Klebsormidium flaccidum* à partir d'une culture de la souche ATCC du protiste en question (Tableau I). Lors de la purification de l'ADN total de cette algue sur un gradient de chlorure de césium, le Dr Lang a constaté la présence de bandes d'ADN bactérien difficilement dissociable de la bande d'ADN mitochondrial. Pour tenter d'identifier la bactérie en question, des PCR ont été effectuées sur l'ADN mitochondrial contaminé par de l'ADN bactérien à partir d'amorces dessinées sur des gènes mitochondriaux *cox1* et *cob* ayant d'ores et déjà permis l'amplification de gènes *cox1* d' $\alpha$ -protéobactéries. Le séquençage des fragments PCR obtenus a révélé la présence d'un fragment du gène *cox1* d'une  $\alpha$ -protéobactérie : un arbre phylogénétique partiel réalisé à partir du fragment PCR codant pour la protéine Cox1 a révélé que la distance séparant cette  $\alpha$ -protéobactérie de l'origine mitochondriale était courte.

Des bactéries contaminant d'autres cultures de protistes dont notre laboratoire a séquencé le génome mitochondrial ont été découvertes. Suite à une PCR avec les amorces sur les gènes *cox1* et *cob* puis à l'analyse phylogénétique des protéines partielles, la distance séparant ces bactéries (qui se sont révélées être des  $\alpha$ -protéobactéries) de l'origine mitochondriale semblait courte comparée à la distance séparant les  $\alpha$ -protéobactéries connues de l'origine mitochondriale. Suite à ces observations, nous en avons déduit qu'il était possible de trouver fréquemment des  $\alpha$ -protéobactéries associées à des eucaryotes, notamment dans les souches de protistes ATCC, et en couvrant un plus large spectre, dans des échantillons provenant de boues et d'eau riches en matière organique, milieux où sont également trouvées bon nombre d' $\alpha$ -protéobactéries.

## I-2. Mise en place de la méthode

### I-2-1. Milieux de culture

Les bactéries contaminantes mises en évidence par les études préliminaires (cf. I-1) encodent les protéines Cox1 et Cob, ce qui suggère la présence d'une chaîne respiratoire. En outre, comme ces bactéries vivent dans un milieu de culture aérobie, elles seront tout comme les protistes, aérobiques. C'est à ce type de bactéries aérobiques, démontrées comme étant proches de l'origine mitochondriale que nous nous intéresseront.

Nous voulons isoler ces bactéries des protistes en les faisant croître sur un milieu de culture solide. Comme nous nous intéressons à des bactéries ayant la faculté de respirer, nous avons décidé d'ajouter de succinate et de lactate, tous deux des éléments intervenant dans la chaîne respiratoire, soit directement après le cycle de Krebs via le complexe II de la chaîne respiratoire (Succinate déshydrogénase) pour le succinate, soit après fermentation pour le lactate (65). Nous supposons que la bactérie présente dans le milieu de culture du protiste peut utiliser certains produits de dégradation du protiste, tels que des acides aminés, des nucléotides. Nous supposons également que la bactérie peut utiliser des éléments traces qui composent, entre autres ingrédients, les milieux de cultures des protistes. Nous avons donc élaboré deux milieux de culture (Tableau IV) contenant pour chacun des acides aminés, des nucléotides, des éléments traces puis du succinate (milieu MS) ou du lactate (milieu ML). Nous espérons que ces milieux seraient suffisamment riches pour la croissance des bactéries, mais trop peu riches en éléments pour la croissance des protistes.

Nous envisageons également la possibilité de la présence dans les milieux de culture de protistes de bactéries possédant la faculté de respirer, capable de vivre en relation étroite avec un protiste et de manière autonome sur un milieu de culture solide. Un exemple de ce type de bactérie est la bactérie *Afipia felis*, qui est capable de vivre à l'intérieur d'*Acanthamoeba*, et de vivre seule sur un milieu de culture (66). Le milieu de culture sur lequel peut croître *Afipia felis* étant le milieu CYE (Charcoal Yeast Extract, 67), nous avons décidé de l'utiliser comme milieu de croissance dans nos expériences.

Ces trois milieux de culture, MS, ML et CYE ont constitué nos milieux test pour la croissance des bactéries. Suite à des tests de croissance croisée sur les trois milieux (CYE, MLS et MS) qui nous ont révélé que les bactéries se développant sur un des trois milieux pouvaient pour la plupart se développer tout aussi bien sur les deux autres milieux, nous avons décidé de ne plus utiliser qu'un seul milieu de culture, contenant le succinate et le lactate, le milieu MLS (Tableau IV).

Seules trois cultures de protistes ont été utilisées au départ pour tester notre méthode. Lorsque nous avons étendu notre collection de protistes et que nous avons travaillé avec des échantillons naturels, nous avons également utilisé le milieu de culture riche LB (Appendice 1) et le milieu de culture à base d'eau de mer, L1 (Appendice 1), parfois additionné de méthylamine, cette molécule oxydée en carbone et azote par la méthylamine déshydrogénase dans le périplasme de nombreuses bactéries méthylophiles. Par l'utilisation de ces milieux de culture, nous pensions accroître la variété métabolique des bactéries récoltées par rapport au seul milieu MLS.

Pour tous les milieux de culture utilisés, la température et le pH du milieu où vit le protiste, soit température de la pièce et pH neutre sont utilisés pour les manipulations.

### I-2-2. Criblage

Des bactéries issues des milieux de culture de protistes ou des échantillons naturels se sont développés sur nos milieux de culture synthétiques. Il faut alors cribler cette population de bactéries pour y retrouver des  $\alpha$ -protéobactéries telles que celles découvertes dans les expériences préliminaires. Nous avons pour ce faire utilisé l'outil nous ayant permis de faire cette découverte, soit les amorces sur les gènes mitochondriaux *cox1* et *cob* qui ont été conçues pour amplifier spécifiquement les gènes *cox1* et *cob* mitochondriaux et  $\alpha$ -protéobactériens.

### I-2-2-1. Les amorces sur les gènes *cox1* et *cob*

#### I-2-2-1-1. Les amorces sur les gènes *cox1* et *cob*

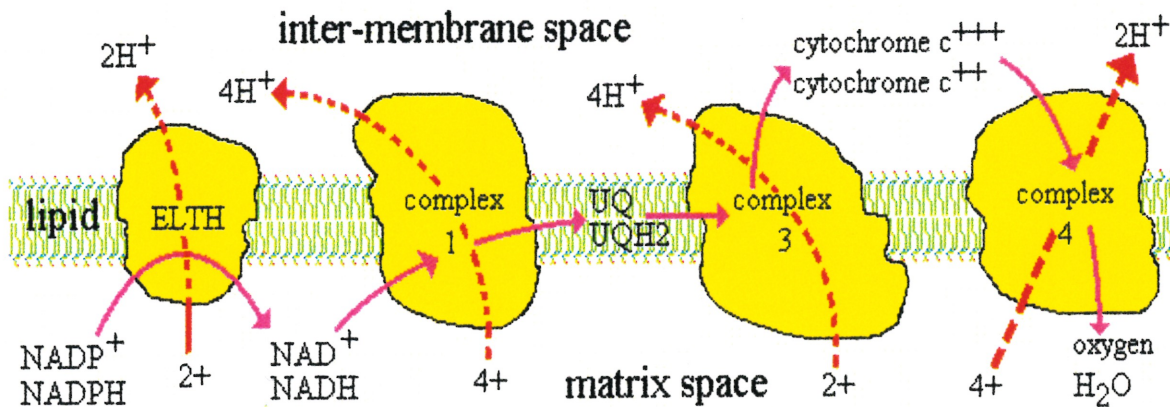
Les amorces ont été construites à partir de l'alignement de protéines Cox1 et Cob de mitochondries. Leur utilisation nous a permis d'amplifier des  $\alpha$ -protéobactéries.

Comme ces amorces ont permis la mise en évidence d' $\alpha$ -protéobactéries, nous pouvons les utiliser afin de mettre en évidence d'autres  $\alpha$ -protéobactéries qui pourraient se trouver dans de nouvelles cultures de protistes et des échantillons naturels.

#### I-2-2-1-2. Les protéines Cox1 et Cob

Les gènes *cox1* et *cob* codent respectivement pour la sous-unité 1 de la cytochrome *c* oxydase (complexe IV) et pour l'ubiquinol:cytochrome *c* oxydoréductase (complexe III), protéines transmembranaires de la chaîne respiratoire mitochondriale de transfert d'électrons (Figure 11). Comme nous l'avons vu précédemment (cf. chapitre I), l'étude de la chaîne respiratoire des mitochondries et des bactéries avait montré la similarité fonctionnelle des chaînes respiratoires de *Paracoccus* et des mitochondries (34). D'autres études sont venues compléter cette étude préliminaire et démontrer la forte conservation des protéines de la chaîne respiratoire entre les  $\alpha$ -protéobactéries et les mitochondries, tant au niveau de la séquence primaire (68-70) que de la structure secondaire et tertiaire. Par exemple, les structures tertiaires des protéines Cox de *Paracoccus* et de la mitochondrie de cœur de bœuf (71-72) publiées au courant des années 1990 ont permis de constater à quel point ces protéines sont similaires. Ces protéines ne sont pas seulement similaires, mais certainement aussi homologues étant donné la provenance reconnue des mitochondries du sein des  $\alpha$ -protéobactéries (cf. chapitre I). Dans les cellules eucaryotes, seules quelques sous-unités sont codées par le génome mitochondrial : pour la cytochrome *c* oxydase, seuls 3 sous-unités sur 13, soit Cox 1-3 sont encodées dans le génome mitochondrial. Quant à l'ubiquinol:cytochrome *c* oxydoréductase, seule une sous-unité sur 11, celle de l'apocytochrome *b*,

Figure 11 : Schéma de la chaîne mitochondriale de transfert d'électrons



- ELTH : « Energy Linked TransHydrogenase »
- Complexe 1 : NADH déshydrogénase
- Complexe 3 : Ubiquinol:cytochrome *c* oxydoréductase
- Complexe 4 : Cytochrome *c* oxydase

Le chemin de transfert des électrons est indiqué en rose, et celui des protons chargés positivement (H<sup>+</sup>) en rouge. Les électrons sont transférés entre les complexes 1 et 3 par la protéine soluble dans la membrane, l'ubiquinone (UQ), et entre les complexes 3 et 4 par le cytochrome *c*, situé à la périphérie de la membrane. Le complexe 2, non montré, est constitué par la Succinate déshydrogénase qui transfère les électrons du succinate à l'ubiquinone. La transhydrogénase liée à l'énergie (ELTH) est une enzyme membranaire qui joue un rôle auxiliaire de pompe à protons qui échange 2H<sup>+</sup> contre 2 électrons.

Sources :

Voet and Voet (1995). Biochemistry, second edition, John Wiley & sons, Inc. Chapter 20, Electron transport and oxidative phosphorylation, figure 20-13, pp 573

Respiratory chain components : <http://www.bmb.leeds.ac.uk/illingworth/oxphos/carriers.htm>

Cob est encodée dans le génome mitochondrial. Les autres sous-unités sont encodées par le génome nucléaire.

Si l'on compare tous les génomes mitochondriaux séquencés entièrement, bien que le nombre de gènes codant pour les protéines de la chaîne respiratoire et les protéines ribosomales varient largement entre les espèces, présentement tous ces génomes contiennent les gènes *cox1* et *cob*.

#### I-2-2-2. PCR, clonage et séquençage

Pour chaque culture de protiste et pour chaque échantillon naturel étalés sur un milieu de culture donné, nous avons purifié en même temps l'ADN génomique de toutes les bactéries présentes sur le milieu de culture : ce qui signifie que nous avons mélangé tous les ADN génomiques de toutes les bactéries s'étant développées sur le milieu de culture, formant ainsi des « pools » d'ADN. Puis, avec les amorces spécifiques des gènes mitochondriaux *cox1* et *cob*, nous avons fait des PCR sur chacun de ces mélanges d'ADN spécifiques à une culture de protistes (ou un échantillon naturel) sur un milieu de culture donné. On obtient ainsi une population de fragments amplifiés pour chaque mélange d'ADN bactérien.

Ces populations de fragments PCR sont alors clonés et séquencés. Les séquences partielles des protéines Cox1 et Cob sont analysées phylogénétiquement afin de savoir si on a affaire à une ou plusieurs  $\alpha$ -protéobactérie(s) dans la population de bactéries, et si oui, si elle est proche de l'origine mitochondriale.

#### I-2-2-3. L'analyse phylogénétique

L'obtention de fragments des protéines Cox1 et Cob est avantageux pour l'analyse phylogénétique. En effet, l'ARN ribosomique a longtemps été la molécule de prédilection pour retracer l'histoire des organismes vivants. Pourtant, l'arbre universel du vivant de Woese *et al.* (1) illustrant cette histoire a été remis en question de nombreuses fois par de nouvelles données provenant notamment de protéines. Un des exemples le plus connu est celui des « Archézoaires », ces eucaryotes que l'on

pensait originellement dépourvus de mitochondries, et qui se sont révélés être des eucaryotes ayant certainement perdus secondairement les dites mitochondries (21). Ou encore l'exemple des microsporidies, ces eucaryotes que l'on pensait primitifs et qui se sont révélés être des champignons (73).

Les ARNr ont d'autre part le désavantage de présenter un taux de mutation de certaines positions phylogénétiques constant, alors que d'autres positions ont un taux de mutation élevé, ce qui cause une sursaturation mutationnelle de ces positions qui résulte par un signal phylogénétique trop faible pour être fiable. En particulier les ARNr mitochondriaux présentent une très forte accélération de leur taux de mutation, ce qui se traduit lors des analyses phylogénétiques montrant les relations entre les mitochondries et les  $\alpha$ -protéobactéries par des branches extrêmement longues pour les mitochondries, comparées aux branches des  $\alpha$ -protéobactéries (cf. figure 3). Une des conséquences de la sursaturation des sites phylogénétiques pourrait être le phénomène d'attraction des longues branches qui résulterait au regroupement des génomes bactériens avec les génomes mitochondriaux évoluant rapidement.

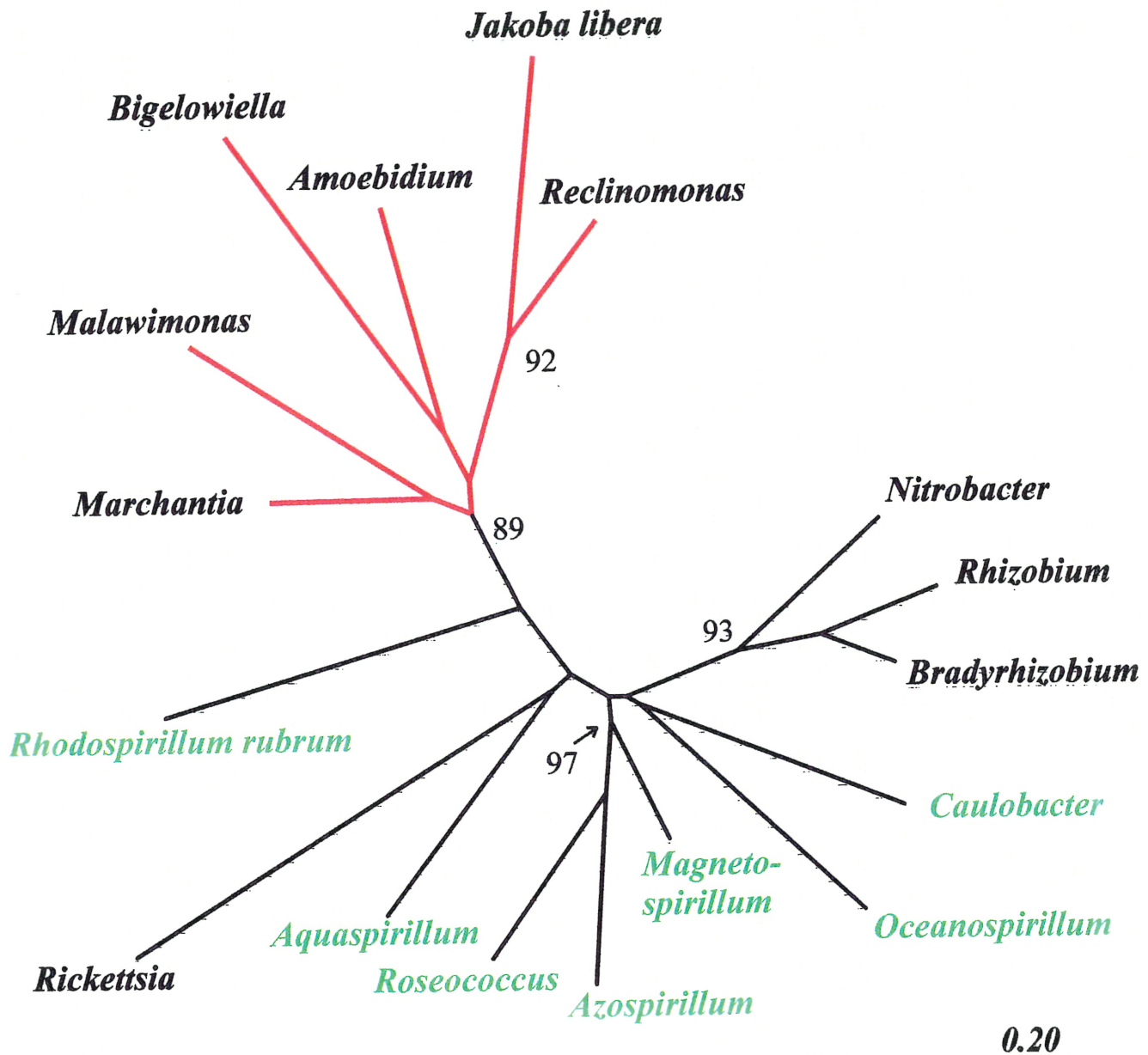
Il est donc préférable d'utiliser des protéines hautement conservées présentes dans la majorité des génomes mitochondriaux telles que les protéines Cox 1-3 (des sous-unités de la cytochrome *c* oxydase) et Cob (apocytochrome *b*) (50). Ces protéines homologues remarquablement bien conservées au cours de l'évolution sont des marqueurs tout désignés pour une analyse phylogénétique.

### I-3. Recherche au sein de souches d' $\alpha$ -protéobactéries de collection

Au sein des  $\alpha$ -protéobactéries, deux grands groupes se distinguent : les  $\alpha$ -protéobactéries vivant librement soit isolément des cellules eucaryotes, et les  $\alpha$ -protéobactéries vivant en relation plus ou moins étroite, voire obligatoire avec des cellules eucaryotes. Du point de vue phylogénétique, Lang *et al.* (50-52) avaient amorcé l'étude des  $\alpha$ -protéobactéries isolées et démontré pour certaines leur



Figure 12 : Arbre phylogénétique montrant les relations entre les  $\alpha$ -protéobactéries et les mitochondries réalisé à l'aide de séquences partielles de la protéine Cox1



L'analyse a été réalisée avec un ensemble de données d'acides aminés correspondant à la portion de la protéine Cox1 encodée par les fragments PCR alignée avec les protéines homologues Cox1 mitochondriales et  $\alpha$ -protéobactériennes. Les taxons comprennent des représentants de groupes eucaryotes et d' $\alpha$ -protéobactéries, ainsi que les bactéries vivant librement dont nous souhaitons connaître la position phylogénétique. L'analyse phylogénétique a été réalisée par une approche de distance, avec la plus récente implémentation de PROTDIST (option matrice JTT) et WEIGHBOR (87, 88) qui permet une correction Jin/Nei pour des taux de changements inégaux à des positions d'acides aminés différentes. Le coefficient de variation utilisé était de 0.5. La barre horizontale indique le nombre moyen de substitution par site. Les valeurs de bootstrap supérieures à 80% sont indiquées au dessus des noeuds. La clade des eucaryotes est représentée en rouge et les  $\alpha$ -protéobactéries vivants librement dont nous avons obtenu une portion du gène *cox1* sont représentés en vert. Les organismes et les séquences (code d'accèsion à GenBank entre parenthèses) sont *Jakoba libera* ("core jakobid", 112), *Reclinomonas americana* ("core jakobid", AF007261), *Marchantia polymorpha* ("liverwort", M68929), *Bigelowiella natans* (112), *Amoebidium parasiticum* (choanoflagellé, 112), *Malawimonas jakobiformis* (jakobide, 112), *Rickettsia prowazekii* (AJ235270 à AJ235273), *Nitrobacter winogradskyi* (X89566), *Rhizobium leguminosarum* (X74341), *Bradyrhizobium japonicum* (J03176).

placement quasi identique à *Rickettsia prowazekii* quant à l'origine des mitochondries.

J'ai poursuivi l'étude de ce domaine des  $\alpha$ -protéobactéries en utilisant des  $\alpha$ -protéobactéries sélectionnées suivant leur relative proximité de l'origine mitochondriale sur un arbre phylogénétique déduit des séquences d'ARNr 16S d'un grand nombre de bactéries, issus de la banque de données RDP (qui réunit l'ensemble des séquences des petites sous unités ribosomiques connues à ce jour (74)). Comparativement aux bactéries étudiées dans le premier volet du projet (cf. I-2), ces bactéries sont d'ores et déjà isolées de leur habitat naturel : elles sont disponibles dans les collections de souche ATCC ou DSMZ et peuvent être cultivées dans un milieu synthétique adéquat connu.

Par PCR sur l'ADN génomique de ces bactéries, je suis allée chercher des fragments des gènes *cox1* et *cob* grâce aux amorces citées précédemment. Les fragments PCR obtenus ont été clonés et séquencés, ce qui a permis d'analyser phylogénétiquement les séquences partielles quant à l'origine mitochondriale. Nous avons alors déterminé que la bactérie *Rhodospirillum rubrum* branchait entre l'ensemble des mitochondries et la bactérie *Rickettsia prowazekii*, ce qui n'avait pas été observé auparavant (cf. Figure 12). Nous avons voulu obtenir la totalité de ce gène *cox1* au cours de ma maîtrise afin de confirmer ou d'infirmer la position phylogénétique observée pour cette bactérie car l'étude que nous avons faite n'était basée que sur un fragment de la protéine Cox1.

Pour ce faire, une banque génomique a été construite sur cette bactérie, puis les colonies de la banque ont été hybridées avec le fragment *cox1* ayant permis l'analyse phylogénétique, qui a constitué notre sonde marquée radioactivement au  $^{32}\text{P}$ . Les clones hybridants avec la sonde, susceptibles de contenir des fragments du gène *cox1* ont été séquencés. La séquence totale de la protéine Cox1 de *Rhodospirillum rubrum* a ensuite été utilisée (avec d'autres séquences Cox1 d' $\alpha$ -protéobactéries) pour réaliser une analyse phylogénétique montrant les relations entre des  $\alpha$ -protéobactéries et les mitochondries.

## II. Matériel

### II-1. Cultures de protistes

Nous avons commencé nos expérimentations sur des cultures de protistes : *Malawimonas*,



Figure 13 : Galerie de photos de certains protistes utilisés lors de l'expérimentation.

Source : pour *Malawimonas*, *Cafeteria*, et *Salpingoeca*, site du PID (Protist Image Database, <http://megasun.bch.umontreal.ca/protists/gallery.html>) et pour *Klebsormidium*, site du Protist Information Server, ([http://mac2031.fujimi.hosei.ac.jp/Protist\\_menuE.html](http://mac2031.fujimi.hosei.ac.jp/Protist_menuE.html))

Tableau I : Liste des protistes mis en culture à partir desquels des bactéries ont pu être isolées.

Protistes	Classification taxonomique	Souche ATCC
<i>Achnophrys</i>	Bicosoécides	
<i>Cafeteria roenbergensis</i>	Straménopiles	
<i>Cercomonas longicauda</i>	Cercozoaires	BER-1 50344
<i>Globalomonas rotunda</i>	indéterminé	CA-1 50729
<i>Goniomonas sp.</i>	Cryptomonades	9/16/81 NB(1) 50108
<i>Jakoba bahamiensis</i>	Jakobides	
<i>Klebsormidium</i>	Viridiplantées	UTEX 176 [ CCAP-335/3 ] 30444
<i>Malawimonas californica</i>	Malawimonades	
<i>Macropharyngomonas</i>	Hétérolobosides	
<i>Monosiga brevicollis</i>	Choanoflagellés	ATCC:07:86:1 50154
<i>Neoramoeba aesturina</i>	Paramoebidées *	Patuxent 50806
<i>Paraphysomonas imperforata</i>	Chrysomonades	TX-5 50309
<i>Salpingoeca orceolata</i>	Choanoflagellés	
<i>Stachyamoeba sp.</i>	Hétérolobosides	1974-1 50324
<i>Stygamoeba regulata</i>	Rhizopodes	BSH-02190019 50892

\* Sujet à controverse

Tableau II : Liste des échantillons naturels testés

Numéro d'échantillons	Provenance de l'échantillon
1	étang à Rigaud (Québec)
2	eau de l'Oust (France)
3	eau du Val d'amour (France)
4	eau de la Manche (France)
5	vase de la Manche (France)
6	étang à Gavray (France)
7	étang en foret de Lanouée (France)
8	bactéries isolées des milieux de culture de protistes (voir Tableau I)
9	quai en bois immergé dans l'Océan atlantique à Bigelow (Maine, ÉU)
10	algues d'une plage de Boothbay Harbor (Maine, ÉU)
11	boue d'eau douce Wiscasset (Maine, ÉU)
12	sable d'une plage à Boothbay Harbor (Maine, EU)
13	terre de Bigelow (Maine, ÉU)

Tableau III

Liste des cultures d' $\alpha$ -protéobactéries isolées choisies pour l'étude phylogénétique de leur gène *cox1*

Bactéries	Classification taxonomique
<i>Aquaspirillum peregrinum</i>	Rhodospirillacées
<i>Azospirillum brasilense</i>	Rhodospirillacées
<i>Holospora obtusa</i>	Rickettsiales
<i>Magnetospirillum gryphiswaldense</i>	Rhodospirillacées
<i>Rhodospirillum molischianum</i>	Rhodospirillacées
<i>Rhodospirillum rubrum</i>	Rhodospirillacées
<i>Roseococcus thiosulfatophilus</i>	Acétobacteracées
<i>Oceanospirillum pusillum</i>	groupe des <i>Oceanospirillum</i>

*Salpingoeca* et *Klebsormidium* (Figure 13, Tableau I). Par la suite, nous avons étendu notre collection de cultures de protistes (Tableau I).

## II-2. Échantillons naturels

Nous avons augmenté le nombre de bactéries testées et la variété de protistes mis en jeu grâce à des échantillons naturels. Ce sont presque tous des échantillons d'eau, prélevés soit dans des étangs, soit près d'une rivière, soit près de la mer (Tableau II). Il est à noter que les échantillons 6 et 13 n'ont pas été utilisés pour la suite des manipulations, car ils présentaient tous deux de trop grandes proportions d'algues et de champignons, respectivement. En effet, il nous aurait été impossible de repiquer les bactéries individuellement en évitant les cellules eucaryotes : lors de la purification des mélanges d'ADN bactérien, nous risquions ainsi de contaminer nos ADN bactériens avec de l'ADN eucaryote, ce que nous voulions éviter.

## II-3. Souches d' $\alpha$ -protéobactéries de collection

Ces bactéries sont disponibles dans la collection de souches ATCC (Tableau III).

## III. Méthode

### III-1. Milieux de culture

#### III-1-1. Milieux MS, ML et MLS

Nous avons élaboré les milieux MS (milieu où le succinate est l'élément du cycle de Krebs suppléé), ML (le lactate est suppléé) et MLS (Tableau IV).

Tableau IV : Milieux de culture mis en place pour l'isolement de bactéries

Ingrédients pour le milieu de base		
	peptone <sup>1</sup> extrait de levure <sup>2</sup> ADN dégradé <sup>3</sup> éléments traces SL-6 <sup>4</sup> mélange de vitamines <sup>5</sup>	
ML	MS	MLS
lactate <sup>6</sup>	succinate <sup>5</sup>	succinate & lactate <sup>7</sup>
	6 < pH < 7	

<sup>1</sup> 0.5g/500ml, <sup>2</sup> 0.5g/500ml, <sup>3</sup> ADN de sperme de saumon dégradé Fluka biochemica 31149, <sup>4</sup> solutions d'éléments traces (ref), <sup>5</sup> 2.5 ml/500 ml (Tableau VI), <sup>6</sup> : acide lactique DL sigma L-1375, 2ml/500ml, <sup>5</sup> acide succinique sigma S-9512, 5ml/500ml d'une solution stock à 1g/10ml d'eau, pH 6.5., <sup>7</sup> éléments ajoutés dans les mêmes proportions que pour les milieux MS et ML.  
 Du bacto-agar est ajouté pour solidifier ces milieux si besoin est à hauteur de 15g/1000ml.

Tableau V

Mélange de vitamines utilisé pour les milieux MS, ML et MLS

Vitamines	Concentration finale dans la solution stock* ( mg/l )
Thiamine HCl	8,5
Biotine	2,5
B12	25
Pyridoxine	30
Calcium panthotenate	5,5
Acide p-amino benzoïque	8,5
Acide folique	2

\* la solution stock est conservée à 4°C pour manipulations ultérieures.

### III-2. Tests de croissance croisée

Les bactéries issues des premières cultures de protistes ont été testées pour leur capacité à croître sur chacun des milieux MS, ML et CYE.

### III-3. Récolte des bactéries

Les échantillons naturels de 1 à 7 ainsi que les milieux de culture des protistes ont été dilués en série (de  $10^{-1}$  à  $10^{-4}$ ), et chaque dilution a été étalée sur un milieu solide. Les autres échantillons naturels (de 9 à 12) ont été traités par la méthode de l'« embedded agar » qui présente l'avantage de figer les organismes présents dans le milieu gélifié, de manière à pouvoir les distinguer les uns des autres. 250ml des différents milieux de culture liquides (LB, MLS, L1pm + M, L1pm – M) sont additionnés d'agar en concentration suffisante pour la solidification et amenés à ébullition pour permettre à l'agar de fondre. Les milieux sont alors refroidis jusqu'à une température d'environ  $37^{\circ}\text{C}$  : à ce moment, 100ul d'échantillon naturel sont ajoutés au milieu, puis des pétris sont coulés. Les milieux de culture solidifiés contiennent ainsi les différents organismes présents dans l'échantillon. Après croissance, les bactéries seront par la suite isolées sur boîte de pétri contenant le milieu approprié. Pour ce faire, on chauffe à la flamme du bec bunsen une pipette pasteur stérile dont on étire l'extrémité. Cette très fine extrémité (de l'ordre d'un millimètre) permet d'aller chercher séparément chaque bactérie enfouie dans le milieu solide par succion, puis de les déposer sur une boîte de pétri contenant le milieu de culture adéquat.

Les bactéries isolées sur milieux ML, MS ou CYE à partir d'un échantillon naturel ou d'un milieu de culture de protistes donné, sont ensuiteensemencées en milieu MLS liquide dans des « microtitres » (Corning) à 96 puits. Les bactéries isolées sur milieu LB, L1pm +/- méthylamine sont repiquées de la même manière dans des « microtitres » contenant le milieu liquide correspondant. Afin de les conserver par congélation, du glycérol à 25% final est ajouté à chaque « microtitre » qui est ensuite déposée à  $-20^{\circ}\text{C}$  jusqu'à la prochaine utilisation.

#### III-4. Récupération de l'ADN bactérien

Les bactéries isolées dans les « microtitres » sont remises en culture sur milieu solide MLS, LB ou L1pm (suivant le milieu liquide utilisé) . Toutes les bactéries issues d'une même microtitre sont mélangées afin de récupérer leur ADN génomique : cet ADN constituera ce que nous nommons un « pool » , soit l'ensemble des ADN génomiques bactériens constitué à partir d'une seule microtitre.

A 9,75 ml de tampon T<sub>100E5</sub> (Tris 100mM, pH 8; EDTA 5mM) sont ajoutés un peu de chaque colonie bactérienne fraîchement repiquée. 100µl de lysozyme (10mg/ml) sont ajoutés au mélange qui est ensuite laissé 10 minutes à température de la pièce. 100µl de protéinase K (10mg/ml) et 50µl de SDS 10% sont alors ajoutés, puis le mélange est mis à incuber 30 minutes à 37°C. Une précipitation des protéines au chlorure de sodium (NaCl) est réalisée : du NaCl à 1M final est ajouté au mélange, qui est ensuite laissé sur glace pour au moins une heure. Deux centrifugations de 15 minutes à 15000 rpm (à 4°C) permettent de récupérer un surnageant qui sera traité pour récupérer l'ADN bactérien. Pour ce faire, 2,5 volumes d'éthanol 95% additionné de 0,5M d'acétate d'ammonium froid sont ajoutés au surnageant puis le tout est mis sur glace pour au moins 15 minutes. Une centrifugation au froid pendant 15 minutes à 15000 rpm permet de récupérer un culot d'ADN. Ce culot est lavé avec de l'éthanol 80% froid puis séché à 50°C pendant 10 minutes. Finalement, le culot est dissous dans 1,5 ml de tampon TE (Tris 1mM pH8, EDTA 0,1mM).

#### III-5. Polymerase Chain Reaction ( PCR )

Chaque « pool » d'ADN est ensuite soumis à un test d'amplification PCR avec des couples d'amorces spécifiques dessinées à partir de séquences de gènes *cox1* et *cob* de mitochondries (Figure 14 et Tableau VI). Pour le test, une quantité variable d'ADN bactérien (entre 1 et 10 ng) a été ajouté à un mélange contenant des déoxyribonucléotides à 200 µM final, des oligonucléotides à 0,5 µM final, 1X final de tampon PCR (Boehringer Mannheim), 2 mM final de chlorure de magnésium,



## Tableau VI

Amorces utilisées pour les amplifications par PCR sur les gènes *cox1* et *cob*

Nom de l'amorce	Séquence en acides aminés	Séquence en nucléotides
Cox1-101 →	FVPIMIGAPDMAFDP	5'-TTGTICCIATATIGGIGCICCIGATATGGCITTTCC-3'
Cox1-247 →	GGDPVLYQHFWFFGHP L F	5'-GGIGGIGATCCIIITITITITCAICATTTITTTGGTTTTTGGICATCC-3'
Cox1-236 ←	QHFWFFGHPEVYI	5'-ATGTAIACCTTCIGGGTGICCGAAGAACCAGAAIAGGTGCTG-3'
Cox1-441 ←	GMPRRIPDYPDFAFA	5'-GCGAAIGCGTCIGGGTAGTCIGGIATICTICTIGGCATICC-3'
Cox1-294 →	HHMYTVGLDVDT F	5'-CATCATATGTATACIGTIGGITTIGATGTIGATAC-3'
Cob-086 →	NYGWLLRYXHANGA	5'-AATTATGGITGGTTITTIIGITATIHCATGCIAATGGIGC-3'
Cob-266 ←	HIVPEWYFLPFYAI	5'-ATIGCGTAGAAIGGIAGGAAGTACCATTTCIGGIACIATGTG-3'

Toutes les amorces sont écrites de 5' vers 3'. Les amorces suivies du signe → iront hybrider sur le brin sens (5' → 3') lors de l'amplification, tandis que les amorces suivies du signe ← iront hybrider sur le brin complémentaire.

Les I présents dans les séquences nucléotidiques correspondent à des inosines : ces nucléotides peuvent se lier à des cytosines aussi bien que les guanines, mais forment une double liaison au lieu d'une liaison triple avec les cytosines.

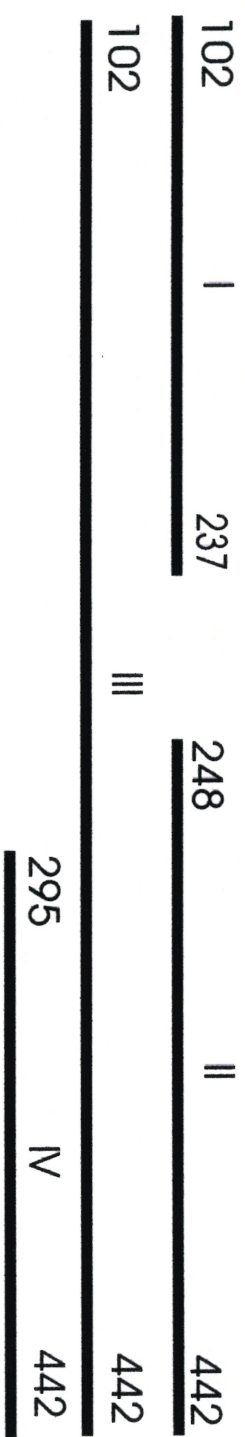
Pour chaque amorce, le chiffre qui suit le nom du gène correspond à la position du dernier acide aminé codé par l'amorce. (cf. figure 14)

Seuls cinq couples d'amorces sont possibles :

- Couple 01 : Cox1-101 + Cox1-236
- Couple 02 : Cox1-247 + Cox1-441
- Couple 03 : Cox1-101 + Cox1-441
- Couple 04 : Cox1-294 + Cox1-441
- Couple 05 : Cob-086 + Cob-266

**Figure 13 : schéma des fragments amplifiables sur les gènes *cox1* et *cob***

**Gène *cox1***



**Gène *cob***



Les schémas ne sont pas à l'échelle : les fragments I, II, III, IV et V correspondent aux cinq fragments qu'il est possible d'amplifier sur ces deux gènes par nos cinq couples d'amorce ( cf. Tableau VI )s. Les chiffres aux deux extrémités des fragments I, II, III, IV et V correspondent au premier et dernier acide aminé de la protéine codée par le fragment amplifié, à l'exclusion de l'amorce.

ainsi que de l'eau HPLC pour atteindre un volume total de 50 µl. Un « hot start » est réalisé en dénaturant les « pools » à 95°C pendant 5 minutes puis incubés 5 minutes à 88°C : c'est durant ces dernières 5 minutes que l'on ajoute 1,25 unités d'enzyme, l'Expand High Fidelity (mélange de Taq polymérase et de Pfu ; Roche). Par la suite, les échantillons sont soumis à 30 cycles d'amplification qui se déroulent comme suit : une dénaturation à 95°C pendant 30 secondes, un annelage à 45°C pendant 30 secondes puis une élongation à 72°C pendant 45 secondes. Quelquefois, pour obtenir une amplification, des températures d'annelage inférieures, telles que 40°C ont du être utilisées. Finalement, les échantillons sont incubés pendant 5 minutes à 72°C. Ces réactions PCR ont toutes été faites dans l'appareil Perkin Elmer Cetus 9600.

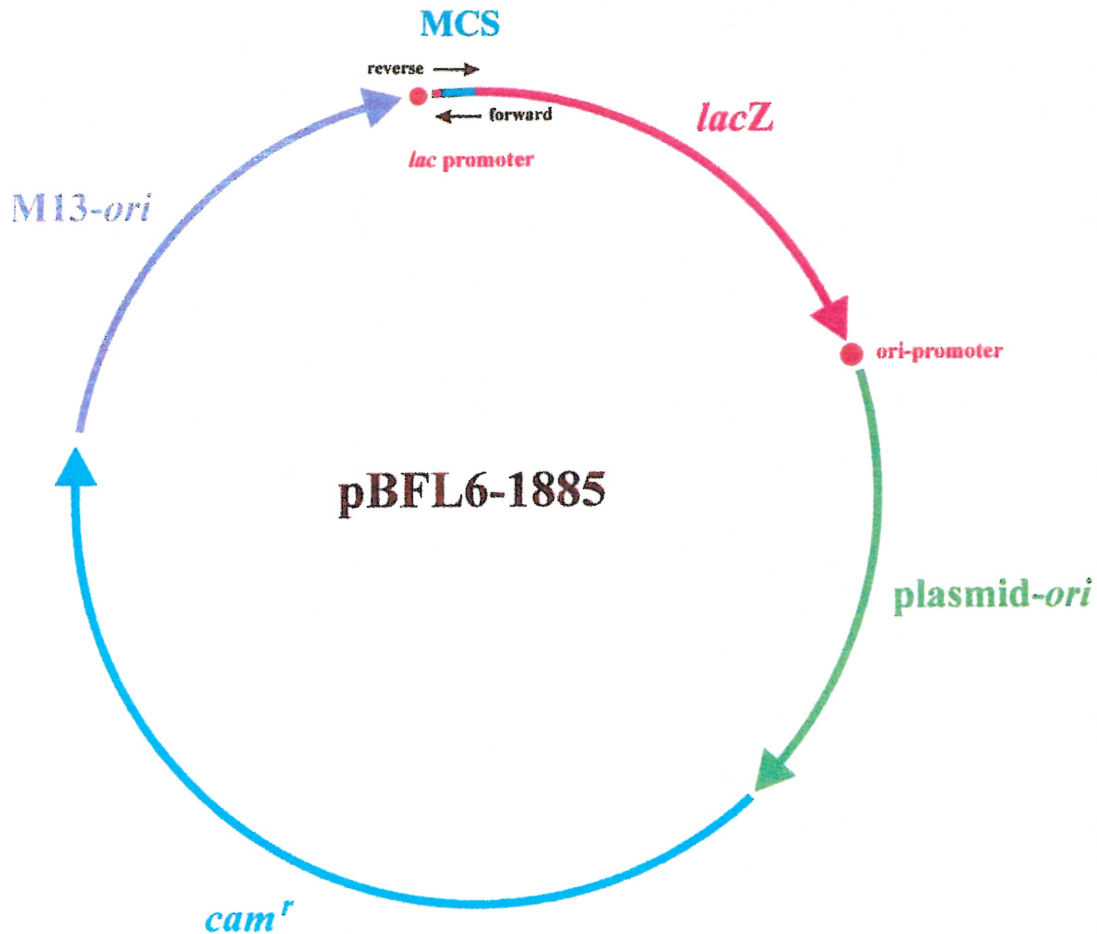
### III-6. Clonage

Les produits d'amplification PCR de taille attendue purifiés et phosphorylés ont été clonés dans le site *EcoRV* du Site de Clonage Multiple ou MCS des vecteur pFBS (pour les clonages des fragments de *Magnetospirillum gryphiswaldense*), pBFLV (pour les clonages des autres fragments obtenus suite à l'amplification sur les autres cultures de bactéries isolées) et pBFLVI (pour les bactéries issues d'échantillons naturels ou de cultures de protistes isolés), qui sont des versions modifiées du vecteur " bluescript " (Figure 15).

#### III-6-1. Réparation des bouts et purification des fragments PCR

Les produits issus de l'amplification par PCR de taille attendue sont traités pendant 30 minutes à 12°C par la T7 DNA polymérase combinée au Fragment de Klenow afin d'obtenir des extrémités franches, nécessaires pour le clonage au site *EcoRV*. La réaction enzymatique est arrêtée par de l'EDTA (5mM final), puis les enzymes sont inactivées par un traitement à 65°C pendant 10 minutes puis l'échantillon se refroidit lentement à la température de la pièce.

Le produit PCR est ensuite débarrassé des amorces et nucléotides dus à la PCR par une électrophorèse par gel « low-melting » à 0,8%. Le fragment PCR est soumis à une électrophorèse.



<i>lacZ</i>	séquence codante du gène <i>lacZ</i> précédé de sa région promotrice ( <i>lac promoter</i> )
plasmid-ori	origine de réplication du plasmide
ori-promoter	région promotrice permettant l'expression du gène de résistance au chloramphénicol (dans le cas du vecteur pBFLVI ou du gène de résistance à l'ampicilline pour le vecteur pFBS)
<i>cam<sup>r</sup></i>	gène de résistance au chloramphénicol
M13-ori	origine de réplication du phage M13
MCS	Multiple Cloning Sites : sites de clonage multiples de pBFLVI comportant des sites de coupure unique pour de nombreuses enzymes de restriction
reverse/forward	zone d'hybridation des amorces utilisées pour le séquençage

Les deux vecteurs ont les caractéristiques ci-dessus en commun, à l'exception du gène de résistance à un antibiotique et de la taille (pBFLVI est de 1000pb plus petit que pFBS).

et la bande d'agarose qui contient le fragment est coupée du gel au moyen d'une lame de scalpel stérile : le morceau d'agarose résultant est ensuite traité par la méthode du « freeze-squeeze » afin d'être récupéré (Appendice1).

### III-6-2. Phosphorylation des produits PCR

Les amorces utilisées n'étant pas phosphorylées, il est nécessaire de phosphoryler les fragments PCR pour pouvoir les insérer dans un vecteur afin de les cloner. Pour ce faire, entre 250 et 450 ng de produits PCR sont dénaturés à 90-95°C pendant 2 minutes puis refroidis sur glace. Du tampon kinase (0.3 µl du tampon 10X initial), de l'ATP (0.15 µl de la solution d'ATP 10mM initial) et 2 unités de T4 kinase sont ajoutés aux produits PCR. Le mélange est ensuite placé à 37°C pendant 30 minutes. Au bout de ces 30 minutes, l'expérience est répétée en ajoutant uniquement les mêmes

Figure 15 : Schéma du vecteur pBFLVI utilisé lors des clonages des fragments PCR

quantités d'ATP et de T4 kinase. Afin de dénaturer l'enzyme, les produits ainsi traités sont chauffés à 95°C pendant 2 minutes puis refroidis lentement à la température de la pièce.

### III-6-3. Ligation

Les produits PCR ainsi traités sont ensuite ligués au vecteur pBFL VI dans une proportion de un mole de vecteur pour un mole de produits PCR. Les quantités de vecteur et de produits PCR sont ajustées de telle sorte que la concentration finale d'ADN total (soit vecteur + insert) soit de 5 ng/µl. La réaction de ligation se fait avec une concentration finale de 1mM pour l'ATP, 1X pour le tampon de ligation et 0.1 unité/µl de volume final de T4 ligase. Le mélange ainsi constitué est incubé pour au moins 14 heures à 14°C.

### III-6-4. Transformation

Les plasmides ainsi obtenus sont utilisés pour transformer des cellules compétentes de

*E.coli*, de la souche XL-1 Blue dont le génotype est le suivant : recA1 endA1 gyr A96 thi-1 hsd R17 sup E44 rel A1 lac [F' pro AB lac Iq ZΔM15 Tn 10 ( Tet )<sup>r</sup>]. La délétion de cette portion spécifique du gène *lac Z* permet un clonage par  $\alpha$ -complémentation.

Les marqueurs de sélection utilisés sont le chloramphénicol ou l'ampicilline, dont les gènes de résistance sont présent sur le vecteur pBFLVI, pBFLV ou pFBS respectivement (Figure 15).

La transformation s'effectue comme suit : à 2 ng de produit de ligation sont ajoutées 200  $\mu$ l de cellules compétentes, puis le mélange est laissé sur glace pendant 20 minutes. Un choc thermique est appliqué au mélange pendant 30 secondes à 42°C, puis l'échantillon est remis sur glace. Un ml de LB préalablement chauffé à 37°C est ajouté au mélange qui est ensuite incubé pendant 40 minutes à 37°C. Après une courte centrifugation de 3000 rpm pendant 3 minutes, les cellules sont étalées sur des boites de pétri contenant du milieu LB additionné de 5  $\mu$ g/ml final de tétracycline, de 10  $\mu$ g/ml final de chloramphénicol, 95,2  $\mu$ g d'IPTG et de 0,8 mg de X-Gal. Ces pétris sont ensuite placés dans l'incubateur à 37°C pour la nuit.

Seules les colonies blanches qui sont censées contenir un plasmide avec insert sont repiquées sur LB + 5  $\mu$ g/ml de tétracycline + 10  $\mu$ g/ml de chloramphénicol.

### III-6-5. Extraction rapide des plasmides des bactéries

Cette étape permet de savoir quels clones ont un insert ou non donc lesquels seront traités afin d'obtenir la séquence (75).

### III-7. Séquençage manuel

#### III-7-1. Préparation d'ADN simple brin pour le séquençage manuel

Suite à l'extraction rapide des plasmides bactériens, on choisit spécifiquement les clones

possédant un plasmide ayant un insert de taille attendue. Ces bactéries d'intérêt sont repiquées sur une boîte de pétri, puis le protocole de préparation de l'ADN simple brin leur est appliqué (76).

### III-7-2. Réactions de séquence manuelle sur de l'ADN simple brin

Le séquençage a été réalisé par la méthode de terminaison de chaîne (77) en utilisant de l'ADN simple brin comme matrice et du  $^{35}\text{S}$ -dATP comme marqueur radioactif. Des réactions déaza-dITP ont été préférées aux réactions classiques car cette méthode nous permettait d'obtenir des réactions plus facilement lisibles et comportant moins de compressions. En effet, le dITP est de l'inosine triphosphate : ces nucléotides peuvent se lier aux cytosines par des doubles liaisons, plus faciles à rompre lors du séquençage que les guanines.

Dans un tube sont mélangés  $n \mu\text{l}$  ( $n$  étant le nombre de simples brins à séquencer) de tampon de séquençage 5X (0,2 M de Tris pH 7,5; 100mM de  $\text{MgCl}_2$ ; 25 mM de NaCl) et  $n \mu\text{l}$  d'amorces (50uM). 2 $\mu\text{l}$  de ce tube sont prélevés et déposés dans  $n$  tubes auxquels on ajoute 3,5 $\mu\text{l}$  d'ADN simple brin. Ce mélange est ensuite incubé à 65°C pendant 2 minutes puis refroidi lentement jusqu'à 37°C. Dans 4 tubes identifiés G, A, T ou C, sont ajoutés 1,25 $\mu\text{l}$  de terminateurs didéoxynucléotides (80  $\mu\text{M}$  de dNTP, 8 $\mu\text{M}$  de ddATP, ddCTP, ddGTP ou ddTTP et 50 mM NaCl) : ces tubes sont conservés sur glace puis incubés à 37°C peu avant de commencer les réactions de séquençage en elles-mêmes. Parallèlement, un tube contenant  $n$  fois 0,5  $\mu\text{l}$  de DTT (100 mM),  $n$  fois 1  $\mu\text{l}$  de mélange de marquage,  $n$  fois 0,75 $\mu\text{l}$  de tampon de dilution de l'enzyme (20 mM de Tris pH 7,5; 5% glycérol et 0,1 mg/ml BSA),  $n$  fois 0,25  $\mu\text{l}$  de T7 DNA polymérase (3u/ $\mu\text{l}$ ),  $n$  fois 0,25  $\mu\text{l}$  de Klenow (2u/ $\mu\text{l}$ ) et  $n$  fois 0,25  $\mu\text{l}$  de dATP-S $^{35}$  (12,5  $\mu\text{Ci}/\mu\text{l}$ ). On se retrouve avec  $n$  fois 3  $\mu\text{l}$  du mélange. Ensuite, pour chaque simple brin, on mélange les 5,5  $\mu\text{l}$  contenant l'ADN simple brin et les amorces avec les 3  $\mu\text{l}$  du mélange contenant les enzymes et le marqueur radioactif. Ce mélange global est incubé 1 minute à température de la pièce pour que les réactions de polymérisation et de marquage de l'ADN se fassent. Puis la réaction de terminaison est amorcée en ajoutant 1,75  $\mu\text{l}$  de d'ADN marqué dans chacun des tubes contenant les terminateurs didéoxynucléotides. Les réactions de

terminaison sont incubées à 37°C pendant 3 minutes. Pour arrêter les réactions de terminaison, 2,5 µl de solution stop ( 95% formamide, 20 mM EDTA, 50% de bleu de bromophénol et 50% de xylène cyanol ) sont ajoutés à chacun des tubes identifiés G,A,T, ou C. Les réactions de séquence ainsi obtenus sont traitées sur gel de polyacrylamide selon la méthode de Lang et Burger (78). Les gels sont ensuite autoradiographiés , et les séquences, lisibles sur environ 400 à 500 nucléotides, sont entrées manuellement dans des fichiers informatiques.

### III-8. Séquençage automatique

Certaines séquences ont été séquencées automatiquement sur un appareil Li-Cor (Lincoln, NE) 4000L qui utilise des amorces marquées aux extrémités par fluorescence et un protocole de séquençage par PCR (« cycle sequencing », Amersham). Cette méthode permet également d'augmenter le nombre de nucléotides lues pour un même clone, soit de l'ordre de 1000 paires de bases pour une direction de lecture (« forward » ou « reverse »; cf. ci-dessous, § III-8-1-2).

#### III-8-1. Préparation d'ADN plasmidique

Les clones contenant ces séquences d'intérêt sont repiquées sur LB + 5 µg/ml de tétracycline + 10µg/ml de chloramphénicol puis sont inoculées dans 1ml de milieu (0,5X LB, 1X M9, 5 mM MgCl<sub>2</sub>, 3% glycérol, 5 µg/ml de chloramphénicol) et incubées avec agitation pour une durée de 16 à 18 heures à 37°C. A ces cultures est appliqué le protocole de récolte et de purification des ADN plasmidiques doubles brins sur colonnes Quiagen (Qiaprep8 Miniprep Kit). Suite à l'utilisation de ce kit, nous obtenons les ADN plasmidiques double brins dans du tampon TE.

#### III-8-1-2. Réactions de séquence Li-Cor sur l'ADN double brin plasmidique

Ces réactions de séquence sont réalisées selon le protocole et avec les produits de « cycle sequencing » d'Amersham LIFE SCIENCE.



Les réactions sont faites dans des plaques de PCR à 96 puits (Diamed). Les gels Li-Cor sont conçus pour 16 réactions de séquence qui seront lues dans les 2 directions soit « forward » (5' → 3') et « reverse » (3' → 5') grâce à deux amorces marquées aux extrémités par des molécules qui fluorescent à 2 longueurs d'onde différentes. Les amorces suivantes ont été utilisées : l'amorce « forward » Franz2/iRD800 (5'-CGTAATACGACTCACTATAGGGCGAATTGG-3') qui fluoresce à 800nm et l'amorce « reverse » M13reverse/iRD700

(5'-AACAATTTACACAGGAAACAGCTATGACCATG-3') qui fluoresce à 700nm.

Les puits de la plaque sont identifiés avec des couleurs spécifiques pour chaque nucléotide, G, A, T ou C : 3 µl d'ADN plasmidique double brins sont ajoutés dans chaque puits identifié. Dans chacun de ces puits, on ajoute encore 1µl de mélange de réaction (qui contient une ADN polymérase séquenase thermostable) ainsi que 0,2 µl d'amorce Franz2/iRD800 et 0,15 µl d'amorce M13reverse/iRD700. La plaque est ensuite déposée dans l'appareil PCR pour être soumise au traitement suivant : 2 incubations à 95°C pendant 5 minutes afin de chauffer l'appareil puis de dénaturer l'ADN, puis 15 cycles décomposés en 30 secondes à 95°C (dénaturation) et 1 minute 15 secondes à 70°C (polymérisation). A la fin de ces cycles, les échantillons sont amenés à 4°C, puis 1,35 µl de solution « stop » sont déposés afin d'arrêter complètement les réactions.

### III-8-1-3. Gel Li-Cor

Les réactions Li-Cor sont migrées sur un gel dénaturant de polyacrylamide de 3,5% (Appendice 1) durant environ 16 heures. La lecture du gel se fait par le programme *Base Image IR version 4000* fourni par la compagnie.

## III-9. Obtention de la séquence complète du gène *cox1* de *Rhodospirillum rubrum*

### III-9-1. Montage de la banque d'ADN génomique

Le protocole de construction d'une banque génomique d'organelle selon l'OGMP (Organelle Genome Megasequencing Program) a été adapté pour la banque génomique sur *Rhodospirillum rubrum* afin d'obtenir après nébulisation des fragments de 4 à 8 kb (78). Ce sont des fragments de cette taille qui ont été privilégiés pour le clonage dans le même vecteur que précédemment soit pBFLVI. Par la suite, le protocole a été rigoureusement suivi.

### III-9-2. Hybridation de la banque avec la sonde spécifique ( fragment PCR *cox1*)

Suite à la transformation, les colonies blanches sont repiquées dans 100µl de milieu (LB 1X, glycérol 3%, M9 0,5X, tétracycline 5µg/ml et 5µg/ml de chloramphénicol) dans des microtitres à 96 puits qui sont ensuite placées dans un incubateur à 37°C pendant 16 à 18 heures.

10µl de chaque puits (pour chaque « microtitre ») sont déposés sur des membranes de Nylon qui sont lavées par la suite dans différentes solutions de lyse, de dénaturation et de neutralisation (Annexe 1), et dans 10mM de Tris pH8 additionné de 50µg/ml de protéinase K. Les membranes sont séchées sur un papier filtre puis irradiées aux UV afin de fixer l'ADN : finalement, elles sont lavées brièvement dans 2X SSPE (3,6M NaCl, 0,2M NaPO<sub>4</sub>, 20 mM EDTA pH 7,5-7,7). Les membranes sont ensuite plongées dans une solution de préhybridation (SDS 1%, SSPE 1,5X, Denhardt 2,5X de l'eau et de l'ADN dégradé à 0,1mg/ml préalablement dénaturé pendant 10 minutes à 95-100°C). La préhybridation dure au moins une heure à 60°C. Pendant ce temps, la sonde est préparée : 0,5ng du produit PCR *cox1* de *Rhodospirillum rubrum* est dénaturé pendant 10 minutes à 95°C, puis refroidi sur glace pendant 2 minutes. Puis, à cet ADN sont ajoutés 4 µl de mélange de réaction à 1X final, 3,5 µl de <sup>32</sup>P, 2 unités de Klenow et de l'eau en quantité suffisante pour atteindre un volume final de 10,5 µl : le mélange est ensuite incubé à 37°C pendant au moins 30 minutes. L'hybridation se fait alors sur les membranes (dans une solution identique à la solution de préhybridation) avec la sonde radioactive pendant au moins 4 heures à 60°C. Les membranes sont ensuite lavées successivement dans trois solutions de 2X SSPE, 0,1% SDS, puis 1X SSPE, 0,1% SDS et finalement 0,1X SSPE, 0,1% SDS. Les membranes sont alors exposées à -70°C pendant 24 heures, puis les films sont

autoradiographiés. Finalement les membranes sont déshybridées pour se débarrasser de la sonde : elles sont incubées dans une solution de 0,4N NaOH pendant 30 minutes à 45°C, puis incubées dans 0,1X SSPE, 0,1% SDS, 0,2M Tris pH 7,5 pendant 15 minutes à 45°C. Les colonies ayant hybridées sont repiquées sur boîte de pétri et traitées afin d'obtenir la séquence de l'insert présent sur leur ADN plasmidique (cf. III-4-2).

### III-9-3. Séquençage automatique de l'insert du clone rr2042

Le clone rr2042 pour lequel un signal a été obtenu après hybridation de la banque par la sonde a été séquencé par la méthode décrite précédemment (cf. ci-dessus, § III-8).

### III-9-3. Restriction par les enzymes *SacI* et *KpnI* sur le clone rr2042

Nous avons soumis L'ADN plasmidique du clone rr2042 à une digestion par les enzymes *SacI*, *SmaI* et *KpnI* (Roche) afin d'isoler l'insert du plasmide, et d'obtenir des fragments de cet insert de taille réduite (de l'ordre de 2 à 2.5kpb). Ces trois enzymes coupent toutes une fois dans le MCS du plasmide : nous avons testé les enzymes *SacI* et *SmaI* en combinaison avec *KpnI* pour identifier avec quelles enzymes nous obtenions des fragments de taille adéquate pour un séquençage manuel ou automatique par nos propres moyens. Pour couper le fragment par ces enzymes, à 5 µl d'ADN plasmidique à 100 ng/µl, sont ajoutés 1 µl de tampon de restriction ( tampon Roche L pour *SacI* et *KpnI*, tampon Roche A pour *SmaI*), 1 µl de BSA (Bovine Serum Albumine) à 1mg/ml initial, 2,8 µl d'eau et 0,1 µl de chaque enzyme. Dans le cas des restrictions par les enzymes *SacI* et *KpnI*, nous avons incubé chaque enzyme successivement pendant 85 minutes pour une digestion complète à 37°C. Pour les enzymes *SmaI* et *KpnI*, nous avons traités les échantillons respectivement à 25 et 37°C pour la même durée.

### III-9-4. Marche chromosomique sur l'ADN plasmidique du clone rr2042

Ayant eu de la difficulté à obtenir la séquence des fragments de restriction (contamination

## Tableau VII

Amorces utilisées pour la marche chromosomique sur l'insert du plasmide présent dans le clone rr2042

---

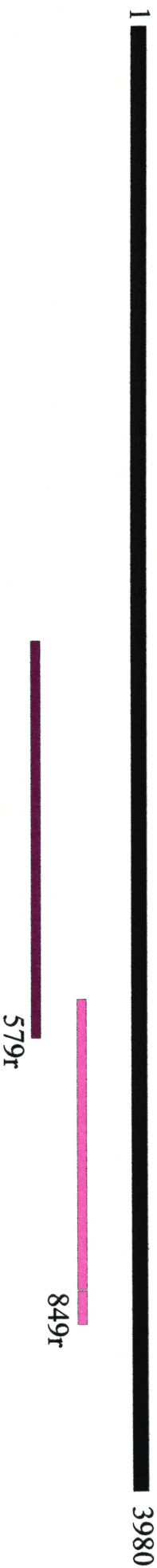
Numéro de l'amorce	Séquences en nucléotides
692f →	5'-CTGGGTGATCGCCACTCTGGCCGAG-3'
624f →	5'-GCCTATGTCATGCTCGATCTCGACC-3'
255f →	5'-TTTCCCGACTGACGCCGC-3'
849r ←	5' GCGTGTCCGGCATGTTTCGTTCCAGTA 3'
579r ←	5'-CCACCATCGCATAGACCA-3'

---

Toutes les amorces sont écrites de 5' vers 3'. Les amorces suivies du signe → iront hybrider sur le brin sens (5' → 3') lors de l'amplification, tandis que les amorces suivies du signe ← iront hybrider sur le brin complémentaire.

Pour chaque amorce, le chiffre qui suit le numéro de l'amorce correspond à la position du premier nucléotide de l'amorce dans la séquence d'où elle est issue (cf. figure 16). Le «f» qui suit le numéro de l'amorce correspond à une amorce s'hybridant sur le brin sens, tandis que le «r» correspond à une amorce s'hybridant sur le brin antisens.

**Figure 16 : Schéma du processus de marche chromosomique sur l'insert du plasmide présent dans le clone rr2042**



Les amorces permettant le séquençage à partir du brin sens sont indiquées en jaune, et les amorces permettant le séquençage à partir du brin antisens sont indiqués en violet-rose. Les chiffres au dessus des rectangles de couleur schématisant les amorces indiquent le numéro de l'amorce (cf. Tableau VII). L'insert du plasmide présent dans le clone rr2042 est schématisé par un rectangle noir : les chiffres à ses 2 extrémités indiquent le premier et le dernier nucléotide de l'insert. Les indications Sacl et KpnI représentent les sites de coupure par ces enzymes. Cette représentation schématique n'est pas à l'échelle.

bactérienne et problème de séquençage de tous les fragments de restriction), nous avons procédé à une marche chromosomique sur l'insert du plasmide présent dans le clone rr2042 à l'aide de premières amorces dessinées sur les portions de séquences obtenues lors du séquençage automatique (Tableau VII). Ces amorces, 849r et 674f, ont été envoyés ainsi que l'ADN double brin plasmidique du clone rr2042 au service de séquençage de la compagnie BIO S&T (79). Comme le séquençage à partir de l'amorce 674f n'a pas donné de résultats, une deuxième amorce dans la même région a été dessinée, l'amorce 624f. Cette amorce, ainsi que l'ADN double brin plasmidique du clone rr2042 a été envoyé au service de séquençage de la compagnie BIO S&T (79). Sur cette séquence, l'amorce 255f a été dessinée et sur la séquence obtenue suite au séquençage à partir de l'amorce 849r, l'amorce 579r a été dessinée. Ces deux amorces ont été envoyés au service de séquençage de l'Université Laval

(80) ainsi que l'ADN double brin plasmidique du clone rr2042 : les séquences obtenues alors nous ont permis de reconstituer la totalité de l'insert du clone rr2042. Ce processus est résumé dans la figure 16.

### III-10. Analyse des séquences

#### III-10-1. Programmes utilisés

Les séquences issues du Li-Cor ont été assemblées par le système XGAP (81). Le programme FASTA (82) a été utilisé pour des recherches de similarités de séquences avec des bases de données locales (83), alors que le service BLAST (84) a été utilisé pour des recherches dans la base de données GenBank sur le site du NCBI (National Center for Biotechnology Information; 85). De nombreux programmes additionnels intervenants au cours de l'analyse, par exemple pour la traduction de séquences nucléotidiques ont été développés par l'OGMP. Ces programmes sont décrits plus en détail et disponible sur Internet (86). Les alignements multiples de protéines ont été réalisés avec le programme CLUSTALW (87), inclus dans le système GDE (Genetic Data Environment; 88) : un contrôle manuel des alignements a été effectué.

### III-10-2. Méthode

Chaque séquence d'un fragment PCR est traduite par le programme flip dans les six cadres de lecture possible. Une recherche de similarité de séquence a été effectuée pour chacune de ces protéines par le biais des programmes FASTA et BLAST. Les protéines similaires à des protéines Cox1 d' $\alpha$ -protéobactéries ont alors été alignées avec un ensemble de protéines Cox1 de mitochondries et d' $\alpha$ -protéobactéries. Pour chacune de ces protéines similaires à des protéines Cox1 d' $\alpha$ -protéobactéries, un arbre phylogénétique est construit afin de déterminer la position phylogénétique de la bactérie dont l'ADN a été amplifié. Cet arbre est construit par une méthode de distance à l'aide de la plus récente implémentation du programme PROTDIST/WEIGHBOR (89) de PHYLIP (90) en utilisant la portion conservée de la protéine alignée avec les autres protéines Cox1 d' $\alpha$ -protéobactéries et de mitochondries. Pour vérifier la solidité de la topologie de l'arbre obtenu, des « bootstraps » (1000) ont été effectués à l'aide des programmes SEQBOOT, PROTDIST et CONSENSE. Des arbres ont été construits avec le même ensemble de données à l'aide d'une méthode de reconstruction PROML incluse dans PHYLIP (90).

### **CHAPITRE III**



# RÉSULTATS

## I. Recherche au sein des bactéries vivant en relation avec des protistes

### I-1. Récolte de bactéries issues des cultures de protistes et des échantillons naturels utilisés dans l'étude

#### I-1-1. Mise en place de la méthode

Nous avons mis en place notre méthode en utilisant les cultures des protistes, *Klebsormidium*, *Malawimonas* et *Salpingoeca*. Le nombre de bactéries récoltées sur les milieux de culture bactérienne que nous avons mis en place (tels MS et ML) ou sélectionné (le milieu CYE ; cf. Chapitre II, § I-2-1) est présenté dans le tableau VIII. Le milieu ML est celui qui présentait le plus petit nombre de bactéries, quelque soit la culture de protiste utilisée. A l'inverse, le milieu MS présentait en moyenne le nombre le plus élevé de bactéries, le milieu CYE présentant quant à lui des nombres de bactéries plus élevés que ceux du milieu ML mais plus bas que ceux du milieu MS (sauf pour le protiste *Malawimonas*).

#### I-1-2. Tests de croissance croisée

Suite à cette récolte, nous avons testé la spécificité de croissance de chacune des bactéries se développant sur un milieu donné en la transférant sur chacun des deux autres milieux de culture. Les tests de croissance croisée présentés dans le tableau IX nous ont montré que de manière générale, les bactéries se développant sur un des trois milieux pouvaient pour la plupart se développer tout aussi bien sur les deux autres milieux. De plus, le milieu CYE n'apportant pas de variété supplémentaire, nous avons décidé de ne plus utiliser par la suite qu'un seul milieu de culture contenant à la fois du succinate et du lactate, le milieu MLS.

Tableau VIII : Nombre de bactéries récoltées sur nos milieux de culture bactérienne ML, MS et CYE à partir des premières séries de cultures de protistes

	ML	MS	CYE
<i>Klebsormidium</i>	4	24	11
<i>Malawimonas</i>	4	8	9
<i>Salpingoeca</i>	-	5	2

Tableau IX : Tests de croissance croisée pour les bactéries issues des premières cultures de protistes isolés, sur les milieux ML, MS et CYE

Culture de protistes isolés	Nombre de bactéries se développant sur les milieux de départ		Nombre de bactéries pouvant se développer sur d'autres milieux de culture	
	ML	MS	ML	CYE
<i>Klebsormidium</i>	ML	4	MS 4	CYE 2
	MS	24	ML 10	CYE 8
	CYE	11	ML 5	MS 10
<i>Malawimonas</i>	ML	4	MS 4	CYE 1
	MS	8	ML 8	CYE 6
	CYE	9	ML 4	MS 7
<i>Salpingoeca</i>	ML	-	-	-
	MS	5	ML 4	CYE 2
	CYE	2	ML -	MS 1

### I-1-3. Récolte des bactéries issues de la deuxième série de cultures de protistes et des échantillons naturels

L'étude des cultures des protistes *Klebsormidium*, *Malawimonas* et *Salpingoeca* a constitué un test pour la faisabilité du projet : ce test s'est révélé concluant par la suite, puisque nous avons réussi à amplifier spécifiquement de l'ADN d' $\alpha$ -protéobactéries (cf. I-3, I-4 et I-5). Nous avons donc étendu notre collection de protistes et testé également des échantillons naturels, afin d'accroître si possible la variété et le nombre de bactéries.

#### I-1-3-1. Cultures de protistes

Pour la deuxième série de cultures de protistes, un total de 66, 34, 27 et 30 bactéries respectivement sur milieux LB, MLS, L1pm + M et L1pm – M ont été récoltées. Les détails sont présentés dans le Tableau X.

Un total de 224 bactéries ont été traitées par la suite afin d'obtenir la séquence de leurs gènes *cox1* et *cob*.

#### I-1-3-1. Échantillons naturels

Tous les échantillons de 1 à 7 n'ont été testés que sur milieu MLS (cf. I-1-2). Un total de 275 bactéries, tous échantillons confondus, a été récolté (cf. Tableau XI pour les détails). Pour la deuxième série d'échantillons naturels (de 9 à 12), un total de 138, 135, 190 et 106 bactéries ont été récoltées respectivement sur milieux LB, MLS, L1pm + M et L1pm – M (cf. Tableau XI).

A partir des échantillons naturels, quelques 844 bactéries ont été récoltées, soit quasiment 4 fois plus que pour les cultures de protistes isolés.

Tableau X : Nombre de bactéries récoltées sur nos milieux de culture bactérienne à partir de la deuxième série de cultures de protistes testées sur milieux LB, MLS, L1pm + M et L1pm – M

	LB	MLS	L1pm + M	L1pm – M
<i>Achnophrys</i>	4	-	5	-
<i>Cafeteria</i>	4	7	6	-
<i>Cercomonas</i>	6	-	-	-
<i>Globalomonas</i>	2	5	4	-
<i>Goniomonas</i>	5	10	-	9
<i>Jakoba bahamiensis</i>	4	-	3	-
<i>Jakoba libera</i> *	16	7	5	-
<i>Malawimonas californica</i>	4	2	-	7
<i>Macropharyngomonas</i>	2	-	-	-
<i>Monosiga</i>	3	-	-	-
<i>Neoparamoeba</i>	3	-	-	11
<i>Paraphysomonas</i>	5	3	-	-
<i>Stachyamoeba</i>	3	-	4	-
<i>Stygamoeba</i>	5	-	3	-

\* les valeurs tiennent compte de plusieurs cultures de *Jakoba libera*

Tableau XI : Nombre de bactéries récoltées sur milieux LB, MLS, L1pm + M et L1pm – M à partir des échantillons naturels

Échantillon	LB	MLS	L1pm + M	L1pm – M
1	-	53	-	-
2	-	48	-	-
3	-	48	-	-
4	-	30	-	-
5	-	60	-	-
7	-	36	-	-
9	26	33	21	36
10	68	35	40	38
11	-	25	48	32
12	44	42	81	-

## I-2. Amplifications par PCR

Nous avons réalisé des amplifications par PCR sur les différents pools d'ADN génomique bactérien. Nous rappelons qu'un pool est constitué par toutes les bactéries d'une origine donnée s'étant développées sur un milieu de culture donné et pour lesquelles les ADN génomiques bactériens ont été purifiés ensemble.

Tous les couples d'amorces (5 couples d'amorces, cf. Tableau VI et figure 13) ont été testés sur les 26 pools d'ADN formés. Seul le couple d'amorces 01, qui amplifie le début du gène *cox1* (soit la portion en nucléotides de 101 à 236) a permis l'amplification d'un fragment du gène *cox1*. Toutefois, pour certains pools d'ADN, même avec le couple d'amorces 01 et même à des températures d'anneau plus basses, l'amplification n'a pas donné de résultats : par exemple, on voit ce phénomène pour l'échantillon 10 sur le milieu MLS où l'amplification avec le couple d'amorces 01 n'a pas pu être obtenue. Le couple 05, le seul contenant les amorces sur le gène *cob* n'a lui non plus jamais permis d'amplification.

Tous ces résultats sont résumés dans le tableau XII.

## I-3. Clonage et séquençage

### I-3-1. Première série de cultures de protistes

Les fragments amplifiés suite à la PCR ont été clonés. Les résultats du clonage sont résumés dans le tableau XIII.

D'une manière générale, les clonages de ces fragments PCR ont été difficiles : certains fragments ont eu plus de difficulté que d'autres à se cloner (comme *Salpingoeca*, tableau XIII), et lorsque certains présentaient des taux d'insertion dans les vecteurs satisfaisants, la concentration d'ADN des simples brins obtenus étaient faible, ce qui rendait le séquençage difficile. Néanmoins, les séquences obtenues qui ont été analysées étaient suffisantes pour la démonstration que par notre méthode nous avons réussi à sélectionner des  $\alpha$ -protéobactéries, ce qui nous a conduit à continuer sur une plus grande échelle.

Tableau XII

Amplifications PCR obtenues avec le premier couple d'amorces pour les différents mélanges d'ADN bactériens, chacun étant spécifique d'un échantillon et d'un milieu de culture donné.

Échantillon	Mélange d'ADN provenant des bactéries s'étant développées sur le milieu LB	Mélange d'ADN provenant des bactéries s'étant développées sur le milieu MLS	Mélange d'ADN provenant des bactéries s'étant développées sur le milieu L1pm + M	Mélange d'ADN provenant des bactéries s'étant développées sur le milieu L1pm - M
1	-	+	-	-
2	-	+	-	-
3	-	+	-	-
4	-	+	-	-
5	-	+	-	-
6	-	-	-	-
7	-	+	-	-
8	+	+	+	+
9	+	+	-	+
10	+	-	+	+
11	-	+	+ *	+
12	+	+ *	+ *	-

« - » absence d'amplification avec le couple d'amorces 01

« + » amplification avec le couple d'amorces 01

\* amplification à une température d'anneau de 40°C au lieu de 45°C

« échantillon 8 » deuxième série de cultures de protistes ( cf. Chapitre II, Tableau II )

Tableau XIII

Résultats des clonages des fragments PCR amplifiés à partir des pools d'ADN de la première série de cultures de protistes isolés testés

	Nombre de colonies blanches / nombre total de colonies	% de colonies blanches	Nombre de colonies blanches avec insert / nombre total de colonies testées	Nombre de clones séquences
<i>Klebsormidium</i>	121 / 548	22	18 / 28	5
<i>Malawimonas</i>	159 / 633	25,1	25 / 28	7
<i>Salpingoeca</i>	165 / 708	23,3	12 / 28	3

Tableau XIV

Résultats des clonages des fragments PCR amplifiés à partir des pools d'ADN de la deuxième série de cultures de protistes isolés testés

	Nombre de colonies blanches / nombre total de colonies	% de colonies blanches	Nombre de colonies blanches avec insert / nombre total de colonies testées	Nombre de clones séquences
LB*	23 / 109 et 20 / 95	21,1	9 / 23 et 5 / 20	14
MLS*	22 / 78 et 18 / 73	26,5	2 / 22 et 1 / 20	3
L1pm + M	40 / 197	20,3	10 / 20	9
L1pm - M	21 / 80	26,2	10 / 20	10

\* deux clonages successifs ont été effectués ( cf. I-4-2 )

### I-3-2. Deuxième série de cultures de protistes isolés testée

Cette deuxième série de cultures de protistes isolés a été nommée échantillon 8 (cf. tableau II). Les ADN ont été réunis à partir des bactéries d'un même milieu de culture. Les fragments amplifiés ont été clonés, les résultats de ces clonages sont résumés dans le tableau XIV.

Les clonages des fragments 8(LB) et 8(MLS) ont été réalisés 2 fois notamment pour essayer d'améliorer le rendement de ces clonages. Mais, comme on peut le constater avec les clonages des fragments 8(L1pm + M) et 8(L1pm – M), ces rendements ont été ceux avec lesquels il nous a fallu travailler pour quasiment tous les clonages de fragments PCR de ce projet. Tous les clones testés possédant un insert dans leur plasmide et dont le simple brin présentait une qualité satisfaisante ont été séquencés. Ces résultats sont rapportés dans le tableau XIV.

### I-3-3. Échantillons naturels

Ceci correspond aux échantillons 1 à 12, à partir desquels des bactéries ont été récoltées sur milieux LB, MLS, L1pm + M, L1pm – M. Les amplifications PCR ont été clonées : les résultats sont résumés dans le tableau XV.

Ces clonages ont eu un rendement satisfaisant pour la plupart, le nombre de clones avec insert augmentant sensiblement par rapport aux clonages précédents.

Certains clones ont été séquencés directement en Li-Cor : 9, 6, 12 et 7 clones respectivement pour les échantillons 2, 4, 5 et 7. Trois clones issus du clonage du fragment PCR amplifié à partir du pool d'ADN de l'échantillon 3, séquencés manuellement et présentant une similarité de séquence intéressante par rapport aux mitochondries ont également été séquencés en Li-Cor afin d'obtenir plus de données, soit une lecture plus longue du fragment PCR.

Tous les autres clones testés possédant un insert dans leur plasmide et dont le simple brin présentait une qualité satisfaisante ont été séquencés manuellement. Ces résultats sont rapportés dans le tableau XV.



Tableau XV

Résultats des clonages des fragments PCR amplifiés à partir des pools d'ADN isolés des bactéries des échantillons 1 à 7 et 9 à 12

Échantillon	Milieux de culture	Nombre de colonies blanches / nombre total de colonies	% de colonies blanches	Nombre de colonies blanches avec insert / nombre total de colonies testées	Nombre de clones séquencés
1	MLS	27 / 212	12,7	10 / 27	10
2	MLS	51 / 162	31,5	46 / 51	43
3	MLS	36 / 136	26,5	34 / 36	32
4	MLS	42 / 94	44,6	41 / 42	40
5	MLS	41 / 141	29	40 / 41	35
7	MLS	42 / 165	30,9	30 / 42	25
9	LB	63 / 258	23,3	9 / 63	9
	MLS	23 / 146	15,7	6 / 23	4
	L1pm – M	37 / 201	18,4	22 / 37	19
10	LB	38 / 202	18,8	1 / 38	1
	L1pm + M	30 / 176	17	2 / 30	2
	L1pm – M	42 / 235	18,3	4 / 36	2
11	MLS	14 / 21	66,7	14 / 14	14
	L1pm + M	7 / 10	70	6 / 7	6
	L1pm – M	7 / 12	58,3	6 / 7	6
12	LB	7 / 14	50	7 / 7	7
	MLS	4 / 12	33,3	4 / 4	4
	L1pm + M	2 / 4	50	2 / 2	2

## I-5. Analyse phylogénétique des séquences obtenues

### I-5-1. Recherche de similarité de séquence

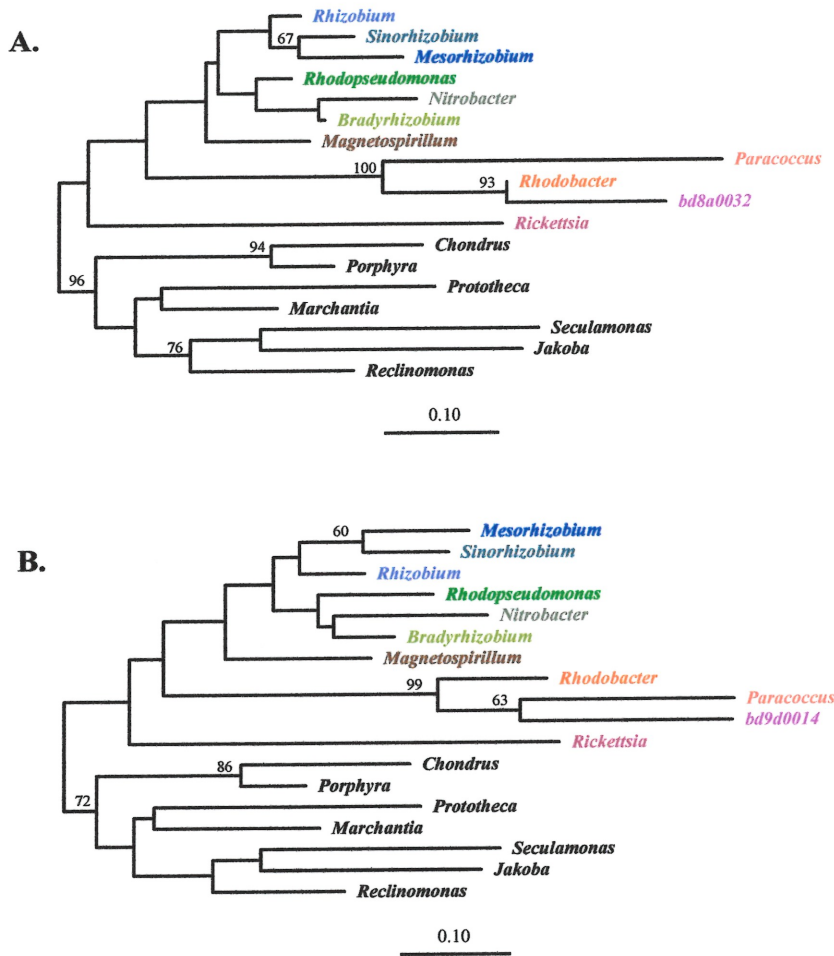
Un total de 177 séquences ont été traitées, soit traduites et comparées à des séquences protéiques connues afin de savoir à quelle séquence connue elles étaient similaires. Les séquences ont été testées à la fois pour leur similarité avec des séquences protéiques de notre base de données (par FASTA), et pour leur similarité avec des séquences protéiques de la base de données GenBank (par BLAST), beaucoup plus exhaustive. Il est important de noter que les observations faites sur notre base de données restreinte ont toujours été confirmées suite à l'utilisation de la base de données Genbank.

De manière générale, 79,4% des séquences traitées sont similaires à des fragments de protéines Cox1 d' $\alpha$ -protéobactéries, les séquences restantes se séparant en 19,3% de séquences ne démontrant aucune similarité avec des séquences connues (avec notre base de données locales FASTA comme avec la base de données GenBank BLAST), et en 1,3% de séquences provenant certainement de conataminations démontrant des similarités avec des protéine d'*E. coli* (protéine d'insertion ou protéine Cox1 dans deux clones).

### I-5-2. Réalisation d'arbres phylogénétiques pour mettre en évidence la position phylogénétique des bactéries de qui proviennent les fragments PCR

Les protéines partielles obtenues suite à nos expérimentations qui étaient similaires à des protéines Cox1  $\alpha$ -protéobactérienne ont toutes été alignées avec un ensemble de protéines Cox1 mitochondriales et  $\alpha$ -protéobactériennes, et un arbre phylogénétique a été généré grâce au programme PROTDIST/WEIGHBOR avec un coefficient de variation de 0,5 pour chaque protéine. Ces bactéries se retrouvaient toujours au sein des  $\alpha$ -protéobactéries avec une position variable. Nous avons déterminé que seules dix positions phylogénétiques étaient obtenues pour les 140 séquences protéiques testées, soient des bactéries parentes des  $\alpha$ -protéobactéries *Nitrobacter*, *Rhodobacter*,

Figure 16 : Arbres phylogénétiques montrant les relations entre les mitochondries, les  $\alpha$ -protéobactéries, et les bactéries issues de nos expérimentations.

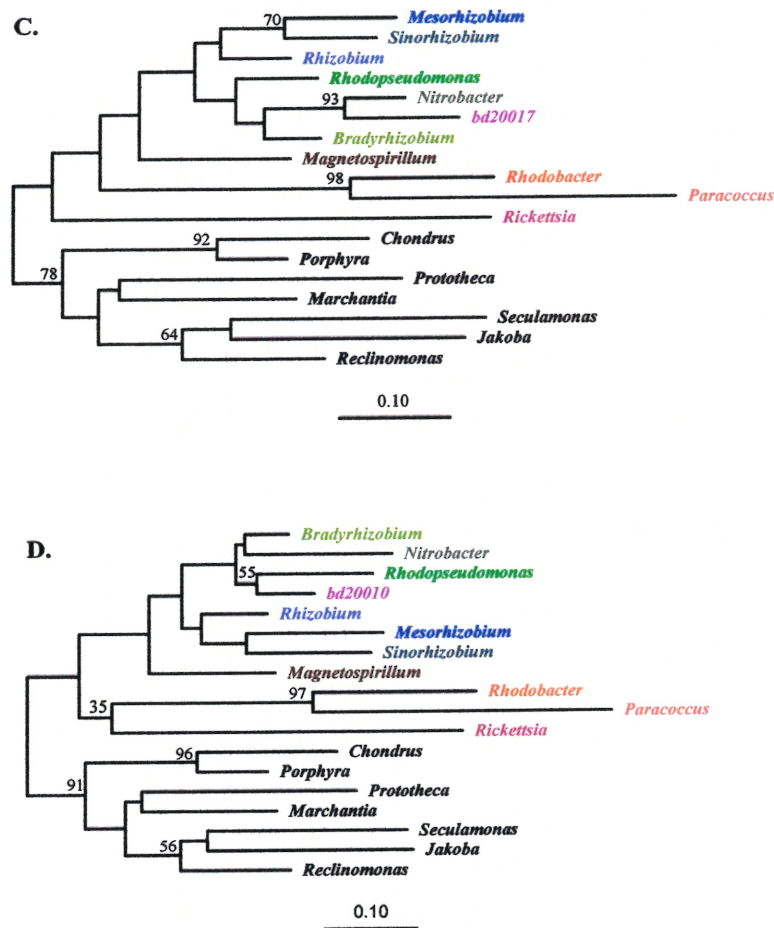


L'analyse a été réalisée avec un ensemble de données d'acides aminés correspondant à la portion de la protéine Cox1 encodée par l'insert des clones d'intérêt alignée avec les protéines homologues Cox1 mitochondriales et  $\alpha$ -protéobactériennes. Les taxons comprennent des représentants de groupes eucaryotes et d' $\alpha$ -protéobactéries, ainsi que les bactéries issues de nos expérimentations dont nous souhaitons connaître la/les plus proche(s) parente(s). L'analyse phylogénétique a été réalisée par une approche de distance, avec la plus récente implémentation de PROTDIST (option matrice JTT) et WEIGHBOR (89, 90) qui permet une correction Jin/Nei pour des taux de changements inégaux à des positions d'acides aminés différentes. Le coefficient de variation utilisé était de 0.5. La barre horizontale indique le nombre moyen de substitution par site. Les valeurs de bootstrap supérieures à 50% sont indiquées au dessus des noeuds correspondants.

Les eucaryotes sont représentés en noir, *Rickettsia* en rose, *Rhodobacter* en orange léger, *Paracoccus* couleur pêche, *Magnetospirillum* en marron, *Rhodopseudomonas*, *Bradyrhizobium* et *Nitrobacter* en vert, vert martien et vert armée respectivement, ainsi que *Rhizobium*, *Sinorhizobium*, et *Mesorhizobium* en bleu électrique, bleu du désert et bleu respectivement.

Les organismes et les séquences (code d'accèsion à GenBank entre parenthèses) sont *Chondrus crispus* (algue rouge, Z47547), *Porphyra purpurea* (algue rouge, AF114794), *Jakoba libera* (jakobide, 112), *Seculamonas ecuadoriensis* (jakobide, 112), *Reclinomonas americana* (jakobide, AF007261), *Marchantia polymorpha* (liverwort, M68929), *Prototheca wickerhamii* (algue verte, U02970), *Rickettsia prowazekii* (AJ235270 à AJ235273), *Magnetospirillum magnetotacticum* (NC\_002725), *Rhodobacter sphaeroides* (NC\_002718), *Paracoccus denitrificans* (X05829), *Nitrobacter winogradskyi* (X89566), *Rhizobium leguminosarum* (X74341), *Bradyrhizobium japonicum* (J03176), *Rhodopseudomonas palustris* (NC\_002804), *Mesorhizobium loti* (NC\_002678), et *Sinorhizobium meliloti* (NC\_002678).

**Figure 16** : Arbres phylogénétiques montrant les relations entre les mitochondries, les  $\alpha$ -protéobactéries, et les bactéries issues de nos expérimentations. ( suite )

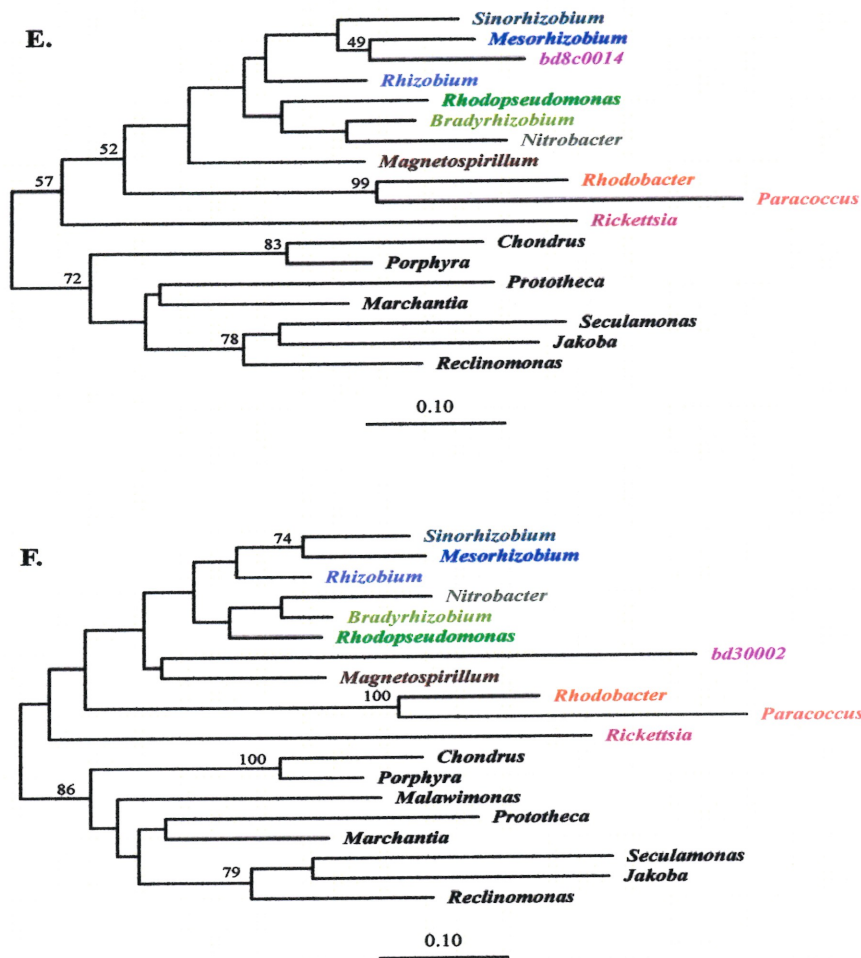


L'analyse a été réalisée avec un ensemble de données d'acides aminés correspondant à la portion de la protéine Cox1 encodée par l'insert des clones d'intérêt alignée avec les protéines homologues Cox1 mitochondriales et  $\alpha$ -protéobactériennes. Les taxons comprennent des représentants de groupes eucaryotes et d' $\alpha$ -protéobactéries, ainsi que les bactéries issues de nos expérimentations dont nous souhaitons connaître la/les plus proche(s) parente(s). L'analyse phylogénétique a été réalisée par une approche de distance, avec la plus récente implémentation de PROTDIST (option matrice JTT) et WEIGHBOR (89, 90,) qui permet une correction Jin/Nei pour des taux de changements inégaux à des positions d'acides aminés différentes. Le coefficient de variation utilisé était de 0.5. La barre horizontale indique le nombre moyen de substitution par site. Les valeurs de bootstrap supérieures à 50% sont indiquées au dessus des noeuds correspondants.

Les eucaryotes sont représentés en noir, *Rickettsia* en rose, *Rhodobacter* en orange léger, *Paracoccus* couleur pêche, *Magnetospirillum* en marron, *Rhodopseudomonas*, *Bradyrhizobium* et *Nitrobacter* en vert, vert martien et vert armée respectivement, ainsi que *Rhizobium*, *Sinorhizobium*, et *Mesorhizobium* en bleu électrique, bleu du désert et bleu respectivement.

Les organismes et les séquences (code d'accèsion à GenBank entre parenthèses) sont *Chondrus crispus* (algue rouge, Z47547), *Porphyra purpurea* (algue rouge, AF114794), *Jakoba libera* (jakobide, 112), *Seculamonas ecuadoriensis* (jakobide, 112), *Reclinomonas americana* (jakobide, AF007261), *Marchantia polymorpha* (liverwort, M68929), *Prototheca wickerhamii* (algue verte, U02970), *Rickettsia prowazekii* (AJ235270 à AJ235273), *Magnetospirillum magnetotacticum* (NC\_002725), *Rhodobacter sphaeroides* (NC\_002718), *Paracoccus denitrificans* (X05829), *Nitrobacter winogradskyi* (X89566), *Rhizobium leguminosarum* (X74341) *Bradyrhizobium japonicum* (J03176), *Rhodopseudomonas palustris* (NC\_002804), *Mesorhizobium loti* (NC\_002678), et *Sinorhizobium meliloti* (NC\_002678).

**Figure 16** : Arbres phylogénétiques montrant les relations entre les mitochondries, les  $\alpha$ -protéobactéries, et les bactéries issues de nos expérimentations. ( suite )



L'analyse a été réalisée avec un ensemble de données d'acides aminés correspondant à la portion de la protéine Cox1 encodée par l'insert des clones d'intérêt alignée avec les protéines homologues Cox1 mitochondriales et  $\alpha$ -protéobactériennes. Les taxons comprennent des représentants de groupes eucaryotes et d' $\alpha$ -protéobactéries, ainsi que les bactéries issues de nos expérimentations dont nous souhaitons connaître la/les plus proche(s) parente(s). L'analyse phylogénétique a été réalisée par une approche de distance, avec la plus récente implémentation de PROTDIST (option matrice JTT) et WEIGHBOR (87, 88) qui permet une correction Jin/Nei pour des taux de changements inégaux à des positions d'acides aminés différentes. Le coefficient de variation utilisé était de 0.5. La barre horizontale indique le nombre moyen de substitution par site. Les valeurs de bootstrap supérieures à 50% sont indiquées au dessus des noeuds correspondants.

Les eucaryotes sont représentés en noir, *Rickettsia* en rose, *Rhodobacter* en orange léger, *Paracoccus* couleur pêche, *Magnetospirillum* en marron, *Rhodopseudomonas*, *Bradyrhizobium* et *Nitrobacter* en vert, vert martien et vert armée respectivement, ainsi que *Rhizobium*, *Sinorhizobium*, et *Mesorhizobium* en bleu électrique, bleu du désert et bleu respectivement.

Les organismes et les séquences (code d'accèsion à GenBank entre parenthèses) sont *Chondrus crispus* (algue rouge, Z47547), *Porphyra purpurea* (algue rouge, AF114794), *Jakoba libera* (jakobide, 112), *Seculamonas ecuadoriensis* (jakobide, 112), *Reclinomonas americana* (jakobide, AF007261), *Marchantia polymorpha* (liverwort, M68929), *Prototheca wickerhamii* (algue verte, U02970), *Rickettsia prowazekii* (AJ235270 à AJ235273), *Magnetospirillum magnetotacticum* (NC\_002725), *Rhodobacter sphaeroides* (NC\_002718), *Paracoccus denitrificans* (X05829), *Nitrobacter winogradskyi* (X89566), *Rhizobium leguminosarum* (X74341) *Bradyrhizobium japonicum* (J03176), *Rhodopseudomonas palustris* (NC\_002804), *Mesorhizobium loti* (NC\_002678), et *Sinorhizobium meliloti* (NC\_002678).

*Paracoccus*, *Rhodopseudomonas*, *Magnetospirillum*. Ces résultats sont résumés dans la figure 17, A-B-C-D-E-F. Comme nous ne disposions que de courts fragments, il nous était difficile d'obtenir du support pour les topologies des arbres. Néanmoins, pour tous les arbres obtenus les clades formées par les algues rouges *Chondrus* / *Porphyra* d'une part et par les bactéries *Rhodobacter* / *Paracoccus* d'autre part sont supportées par des valeurs de bootstrap de plus de 80%. Les autres nœuds sont peu voire pas du tout supportés, même si les regroupements bactériens et eucaryotes respectent des relations phylogénétiques connues : par exemple, on peut voir que *Nitrobacter winogradsky* est phylogénétiquement proche de *Rhodopseudomonas sp* et de *Bradyrhizobium* (91), que *Sinorhizobium* forme une clade avec *Mesorhizobium* et que les mitochondries forment une clade. Quant aux clades formées par les bactéries issues de nos expérimentations avec des bactéries connues, elles sont en général supportées par les valeurs de bootstrap, et une topologie identique est obtenue avec une méthode de maximum likelihood. Pour certaines bactéries issues de nos expérimentations, non seulement la topologie des arbres obtenus ne respectaient pas les relations phylogénétiques connues, mais ces bactéries branchaient avec une clade : avec la clade formée par les bactéries *Rhodobacter* / *Paracoccus*, avec la clade formée par les bactéries *Mesorhizobium* / *Sinorhizobium*. Aucune de ces topologies n'était supportée par des valeurs de bootstrap, et les arbres obtenus avec la méthode de maximum likelihood avaient une topologie différente.

## II. Recherche au sein des souches d' $\alpha$ -protéobactéries de collection

### II-1. Obtention de la totalité du gène *cox1* de *Rhodospirillum rubrum*

#### II-1-1. Montage de la banque et hybridation

Comme *Rhodospirillum rubrum* présente une position phylogénétique privilégiée entre les mitochondries d'une part et les  $\alpha$ -protéobactéries plus *Rickettsia prowazekii* d'autre part, nous avons décidé de confirmer cette position en obtenant la totalité du gène *cox1* de *R. rubrum*.

Nous avons effectué 14 transformations simultanées qui ont donné environ 15000 colonies blanches

avec un ratio de 8 colonies blanches pour 2 colonies bleues. Nous avons repiqué 4896 colonies blanches réparties en 51 « microtitres » à 96 puits. Comme le génome est estimé à environ 4 mégabases (soit 4000 kpb) et que les inserts ont en moyenne un insert de 4 à 7 kpb, ces 4896 clones sont censés couvrir environ 6,7 fois le génome. Nous avons ensuite hybridé la banque avec le fragment du gène *cox1* de *R. rubrum* obtenu par amplification PCR, et nous n'avons trouvé qu'un seul clone ayant hybridé suite à une exposition de 40 heures, le clone rr2042. Nous avons ré-exposé les membranes pendant une semaine, ce qui a conduit à l'apparition de 14 autres clones qui se sont révélés être de faux positifs, correspondant vraisemblablement aux bruit de fond d'hybridation. Seul le clone rr2042 contenait bien un insert possédant le gène *cox1*.

#### II-1-2. Le clone rr2042

Ce clone possède un insert d'environ 3500 pb. Par séquençage automatique via le Li-Cor, nous avons déterminé la présence de la protéine Cox1 dans le clone. Ceci constituait une étape préliminaire, étant donné que nous savions que nous n'aurions pu obtenir la totalité de la séquence par cette méthode, le clone ayant une taille de l'ordre de 3,5 à 4 kpb, et le séquençage automatique ne permettant de lire au mieux que jusqu'à 2 kpb.

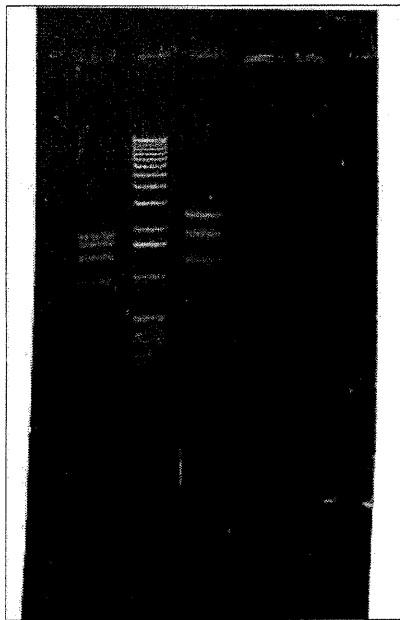
Nous avons cherché à isoler l'insert du plasmide par des coupures endonucléotidiques avec les enzymes de restriction combinées Sac I et Kpn I (figure 18), et nous avons parfait le séquençage par une marche chromosomique (cf. chapitre II, § III-9-4). Les séquences du gène *cox1* et de la protéine déduite sont présentées figure 19.

Figure 18 :

Photo des fragments de restriction obtenus suite à la digestion de l'ADN plasmidique du clone rr2042 par des enzymes de restriction

Puits :

1 2 3



Puits1 : Digestion de l'ADN plasmidique du clone rr2042 par les enzymes de restriction SacI et KpnI

Puits 2 : Marqueur de taille en pb

Puits3 : Digestion de l'ADN plasmidique du clone rr2042 par les enzymes de restriction SmaI et KpnI

A droite de la photo, le schéma indique les tailles en pb des fragments du marqueur.



Figure 19-A : Séquence de l'insert contenu dans l'ADN plasmidique du clone rr2042

```

1   CCCCCACCGGATAGCCGAAGTAGTTGAGGGCCGCTGTCACCTCCTTGTTCTGCTCCCCGT
61  GGGCGCTGGACATGGTGGCCTTGGCCGGGGCCCGGCGATAGACCTGCTGCTGCTGGGCGC
121 GGGCGGGTCTGGTTCGTTGGCGATCGCATGECGATGATGCCGCCACGAGCGCCCCCG
181 CGATGGCGTCGCCCCGATCCGCCTTCACAGGCTCGGCCGGGGCCAGGCCGATGAGGGCAG
241 CGAGGCTGGCGGTGACTAGATATTTCCCATACATCGGACATCTCCCGGAAGTCGAAGGAC
301 TCGGGAAGCGTAACACGGCCATGTGAACTGCCGGCATGTTTTTTGCGGTGGAGAAGGGGA
361 TCATAAGGCGATGTCCGCACTGGAATGGTCGATCCTGCCCGGCCTGAGCCCCCTATCGGGA
421 GACGCTCGAGGCGATGGAGAACC GGTTGGCCGCCATCCGGGCCGGAGAGGCCGCGGAGGC
481 GATCTGGCTGCTCGAGCACCCCCGCTCTACACCGCCGGCACCTCGGCCCGGCCGAGGA
541 TCTGGTTCGAGCCCAGCGCTTTCCTGTTCCATGTGGCCGGTTCGCGGCGGACAATATACT
601 ACCACGGGCCGGGGCAGCGGGTGGCCTATGTGCTGATCTCGACCGGCGCGGCCGGG
661 ACGTGCGCCGCTTCGTGACGGCGCTCGAGGACTGGGTGATCGCCACTCTGGCCGAGTTCA
721 ACGTGCGGGGCGAGCGGCGCACGGGGCGCGTTCGGGGTCTGGGTGCTGCGGCCCGACCGG
781 CCCGCGGGCCTCGACGGCAGCCCGCGGAGGACAAGATCGCGGCCATCGGCGTCAAGCTC
841 CGCCGCTGGGTGAGCTTTCACGGGCTCTCGATCAATGTGAGCCGAACCTGACGCATTTT
901 GAGGGGATCGTGCCCTGCGGCATCCGGGAACACGGGGTGACGAGCCTCGTGGATCTGGGG
961 CTTCCCCTCACGATGCAGGACCTCGACCGGCCCTGCTGCGCACCTTCCCAGCAGCATTTT
1021 CCCGACTGACGCCGAGGTCCAGGCTTGGCCGTTCGCCTGCCGGGCCGAACCGGCCCTCC
1081 CGCCACAGGCCCGCACGGCAGCGCCCTGTTGCGGAGGAGCCGTTCTCGTGATCCGGGC
1141 CGCATGGCTCAACGCATGAGGGTGCAGCCGGTGGCACGCAGGCGCTCGGNACGGCCTGC
1201 CGACTATGCCTGCCGCCGGGCGAGTGTCCCGCCGGGCCGGCCCGCCGCGGGTCTCGGG
1261 TCGGCTTCGTTTCCACAGACCAAATACCACCCCAAGGAAGGAGTGGACGATTTTAT
1321 TGATCCAGATCAAAGTGCGGCGTTCGCCGAACAAGACCGGCGCTCTACAGATATCCGGGA
1381 GATGCTGGTACGGATCCATCCGACGCCGCTTGATCGGAAACAGCACGAGGGGAGGCAAG
1441 CCA
;
G-COX1 ==> START
1444 ATGGCCGACGCAGCCATCCATGGCCACGACCACGACCGGAGGGGGTTCTTACCCGCTGG
1504 TTCATGTGACGAACCACAAGGACATCGGAGTTCTCTATCTCTTACCCGGGGCCCTCGTC
1564 GGGCTGATCTCGGTGGCCTTACCCTTACATGCGGATGGAGCTCATGGCGCCGGGCGTG
1624 CAATCATGTGCCCGAACATCTGGAATCGGGCCCTGAAGGGCTTCTTCCAGTTCGCTCTGG
1684 CCCTCGGCGGTGGAAAACATGCACCCCGAACGGCCATCTGTGGAACGTGATGATCACCGGC
1744 CACGGGATCCTGATGATGTTCTTCGTGGTCAATCCCGCGCTCTTCGGCGGCTTCGGCAAC
1804 TATTTTCATGCCGCTGCACATCGGCGCGCCGACATGGCCTTCCCAGCGGATGAACAACCTC
1864 TCCTACTGGCTCTATGTGCGGGCACCTCGTTCGCGGTCGCTCTTCGGCCGGGGC
1924 GGCAACGGCCAGCTGGGCTCTGGCATCGGCTGGGTGCTCTATCCGCCGCTCTCCACCTCC
1984 GAATCGGGCTATTCCACCGACCTCGCGATCTTCGCGGTGCACCTGTGCGGCGCCTCCTCG
2044 ATCCTCGGCGGATCAACATGATCAGGACCTTCTGAACATGCGCGCACCCGGCATGACC
2104 ATGCACAAGGTGCCGCTCTTCGCTGGTTCGATCTTCGTCACCGCCTGGCTGATCCTGCTG
2164 GCGCTGCCCGTGTCTCGCCGGCGCCATTACCATGCTGCTGACCGACCGGAACTTCGGCAC
2224 ACCTTCTTCCAGCCCTCGGGCGGGCGACCCGGTGTCTACCAGCACATCCTGTGGTTT
2284 TTCGGCCACCCGGAGGTCTACATCATCGTGTGCTGCCGGCCTTCGGCATCGTCAGCCACGTC
2344 ATCGCGACCTTCGCCAAAAGCGACTCTTCGGCTATTTGCCGATGGTCTATGCGATGGTG
2404 GGAATCCTTCTCCTGGATCGCGACCATGTGGGGCGGCTCGATCGAGCTCAAGACCGGATG
2464 CTCACCCAGCAGACTACTTTATGATGGCGACCATGGTGTGATCGCAGTGGCCACCGGCATC
2524 AAGATCTTCTCCTGGATCGCGACCATGTGGGGCGGCTCGATCGAGCTCAAGACCGGATG
2584 CTCTGGGCGCTCGGGTTCTCTTCTTCTTACCCTGGGCGGCGTCAACCGGCATCGTGCTG
2644 AGCCAGGCGAGCGTTCGACCGCTATTATCACGACACCTACTATGTGCTGGCGCACTTCCAT
2704 TATGTGATGAGCCTCGGCGCGGTCTTCGGCATCTTCGCAGGGATCTACTTCTGGATCGGC
2764 AAGATGTGCGGGCCGGCAATATCCGGAATGGGCCGGGAAGCTGCATTTCTGGATGATGTTT
2824 GTGGGCGCAACCTCACCTTCTTCCCAGCACTTCTCGGCCCGCCAGGGCATGCCGCGG
2884 CGCAACACCGATTATCCCGGAGCCTTCCCACCTGGGAATTCGTCTCGTCCGTGGGCCCC
2944 TTCTGTCTTCCCTTCTTCTTTTTTCTCGGCGTGTCTTCTAACACTGACCGGGGGC

```

```

3004  CGGGTGACGGCGAACAACTACTGGAACGAACATGCCGACACGCTGGAGTGGACGCTGACC
3064  TCTCCGCCCGCCGGAGCATAACGTTTCGAGCAGCTTCCCAAGCGGGAAGACTGGGAACGCGCG
3124  CCGGCCCATGATGCCGCCTTCTGA
;      G-COX1 ==> END
3148  TGCCCTACGAATGGTCATTGCCCCGAAGGCCCGGAGACATCCGGGGCCTTTCTCGTTC
3208  GGGGGAGGAAGATCCTCTCGGCCGGCTTCATCTGTGGCCCTACCGCTCGCTGCCACGGCG
3268  CGGTTTCGCGGCCTTCATCGGGCGTGACCTGCCTGCTGCTCGCCCTGCCGCTGATCGGGCTG
3328  CTGGGCTCGCCCGTGCTCTGGGCGCTGCTGCCCTTCTCGTAGCGGTCTGACCCGGCATG
3388  TGGTGGCGCTGAACCGCAGCTACGGCGACGGGGAGATCCTCGAGGAGTGACGCTCTGGCC
3448  CGACAGGGTGACGCTGGATCGGCACGGGCCGCGCGGCCGGCGGCAGAGCTGGCAGGCCAA
3508  TCCGCACTGGGTGCGGGTGACGCTGCACCGAAGCGGGCGGGCCGGTGCCCCATTACCTCAC
3568  GCTCGGCGGGCGGACCGCGCGAGGTGGAGATCGGCGCCTTTCTGGGCGAGGAGGAGCGTCT
3628  GCGGCTCTGGCCCGAACTGGAACGCGCTTTCTCCCGGGCGCGCTGACGGGCGGGAGGAGC
3688  GTCGACGGGGCCATCGAGGCCCTCGCGCCGCGTTGCCACGGACCGGGAGCCGGAGGC
3748  CCGCAGCTTGAACCGGCGCGCCGCGGGAGAAGAGTCTGTCGGGAGCCCCAGGGGCGCGAA
3808  GGGCATGACTGCGCTCGCCCGGACGGGCAGCGTCCCCATNGCTGCGAGCCGCCCTGCC
3868  CGCGCATCGCCCGNCCCCGGNCCCC
;;    3892

```

Figure 19-B : Séquence de la protéine Cox1 déduite de la séquence du gène *cox1* de *R. rubrum*

```

1  MADAATHGHDHRRRGGFFTRWFMSSTNHKDIGVLYLFTGGLVGLISVAFTVYM
   RMELMAPGVQSCAPMIWNRALKGFFQSLWPSAVENCTPNGHLWNVMITGHG
   ILMFFVVI PALFGGFGNYFMPLHIGAPDMAFPRMNNLSYWLYVAGTSLAV
   ASLFAPGGNGQLGSGIGWVLYPPLSTSESGYSTDLAIFAVHLSGASSILGA
   INMITTFLNMRAPGMTMHKVPLFAWSIFVTAWLILLALPVLGAIITMLLTD
   RNFGTFFFQPSGGGDPVLYQHILWFFGHPEVYIIVLPAFGIVSHVIATFAK
   KRLFGLPMVYAMVGLGVLGFWVAHMYTAGLSLTQQSYFMMATMVIAP
   TGIKIFSWIATMWGGSIELKTPMLWALGFLFLFTVGGVTGIVLSQASVDY
   YHDTYYVVAHFHYVMSLGAVFAGIFAGIYFWIGKMSGRQYPEWAGKLFWMM
   FVGANLTFPQHFLGRQGMRRNTDYPGAFPTWEFVSSVGPFLSSLLLLFP
   RRDLLTLTRARVTANNYWNEHADTLEWTLTSPPEHTFEQLPKREDWERAP
   AHDAAF

```

534

Dans la figure 19-A, les chiffres à gauche de la séquence indiquent la position du premier nucléotide de la ligne qui suit dans la séquence. Les indications G-COX1 début ou fin se réfèrent au début et à la fin respectivement de la région codant la protéine Cox1. Les régions soulignées dans l'insert correspondent à des sites de coupure potentiels par l'enzyme de restriction KpnI. Dans la figure 19-B, la région soulignée de la protéine indique la région que nous avons amplifiée au départ grâce au couple d'amorces 01 sur l'ADN génomique de *Rhodospirillum rubrum* (cf. chapitre II). La protéine contient 534 acides aminés.

## CHAPITRE IV

# DISCUSSION

## I. Méthode

### I-1. La classification phylogénétique et le transfert latéral de gènes

#### I-1-1. La classification phylogénétique

L'ère moléculaire de la biologie a aidé à la classification des micro-organismes pour qui les critères morphologiques ou comportementaux sont de piètres supports : ainsi, c'est en utilisant l'ARNr que Woese *et al.* (1) construisent une classification hiérarchisée des organismes vivants, un arbre universel de la vie généralement accepté par la communauté scientifique depuis sa publication il y a maintenant 25 ans. Cet arbre a depuis été remis en question par les données génomiques sans cesse en croissance : ce sont surtout des reconstructions phylogénétiques avec des gènes codant des protéines qui viennent remettre en cause les arbres ARNr. Par exemple, dans le génome de l'archéobactérie *Archaeoglobus fulgidus*, WF. Doolittle a mis en évidence la présence d'une réductase qui ressemblait plus aux réductases bactériennes qu'aux réductases des autres archéobactéries ou des eucaryotes, supposément les plus proches des archéobactéries selon l'arbre universel du vivant (92). Un autre exemple concerne la bactérie vivant à haute température *Aquifex aeolicus*, que l'on peut placer proche de la bactérie du sol *Bacillus subtilis* (alors que ces deux bactéries proviennent de deux branches différentes de l'arbre ARNr bactérien) en utilisant le gène codant pour la protéine FtsY qui intervient dans le contrôle du cycle cellulaire, ou encore proche des archéobactéries en utilisant un gène codant une enzyme nécessaire à la synthèse de l'acide aminé tryptophane (92). Une explication pour ce dernier résultat serait que *Aquifex* ait substitué sa version du gène au gène archéobactérien par transfert latéral de gènes. Chez les eucaryotes, des transferts latéraux de gènes ont été détectés non seulement pour des gènes, mais aussi à l'échelle de catégories entières de gènes (8), les gènes informationnels seraient proches des gènes informationnels archéobactériens tandis que les gènes opérationnels seraient proches des gènes opérationnels eubactériens. Ce phénomène serait-il si répandu dans les organismes vivants qu'il nous obligerait à

reconsidérer la classification hiérarchisée que nous avons mis en place pour les êtres vivants ? Doolittle l'affirme (93) : en tenant compte du transfert latéral de gènes (TLG) l'arbre universel du vivant serait alors représenté par un réseau réticulé de branches qui conduirait aux domaines modernes que nous connaissons. Malgré tout, en tenant compte du TLG, certains chercheurs comme Carl Woese espèrent qu'une image cohérente de l'évolution microbienne naîtra, par exemple sous forme d'un arbre réticulé (92).

### I-1-2. Phylogénie et mitochondrie

Si il nous est difficile de retracer l'histoire des organismes vivants, il nous est néanmoins possible d'éclaircir certains points de la mise en place des cellules modernes, comme la présence de mitochondries dans les cellules eucaryotes. Grâce aux études réalisées depuis les années 1970, certains points de la théorie de l'endosymbiose mitochondriale ont été éclaircis : la nature de l'endosymbiote bactérien, le nombre de fois que l'événement d'endosymbiose a pu avoir lieu et l'absence de découverte d'endosymbiose secondaire chez les mitochondries, contrairement à ce qui a été découvert chez les chloroplastes. Ce que la théorie supposait, i.e que la plupart des gènes de l'endosymbiote aient été perdus et que certains aient été transférés au génome nucléaire de l'hôte, s'est vérifié par la suite : pour certains génomes mitochondriaux, des transferts récents (200 millions d'années) ont été mis en évidence, suggérant que le processus serait toujours en cours (94-95). Ainsi, pour la phylogénie des mitochondries, à l'inverse des situations où l'on fait des études phylogénétiques sans idée de la possibilité d'un TLG, on présume qu'un TLG a eu lieu entre la bactérie symbiote et son hôte, et on peut reconstruire des phylogénies dans ce contexte. En effet des gènes provenant de l'endosymbiote  $\alpha$ -protéobactérien ayant été transférés au noyau tels que les protéines HSP60 ont été utilisés pour des études phylogénétiques qui ont amené à la conclusion que les mitochondries partagent un ancêtre commun avec les  $\alpha$ -protéobactéries (44-45).

Pour reconstituer l'histoire des mitochondries, il est possible à l'inverse d'utiliser des gènes restés dans le génome de l'endosymbiote devenu une organelle. Pour les raisons citées précédemment (cf. chapitre II, § I-2-2-3), il est préférable d'utiliser des gènes codant des protéines : comme nous avons

choisi cette approche pour notre étude, nous avons choisi d'utiliser les gènes *cox1* et *cob*, les rares gènes protéiques présents dans tous les ADNmt séquencés à ce jour ce qui laisse présager qu'ils proviennent de l'endosymbiote qui a donné naissance aux mitochondries.

Dans ce point précis de l'étude des organismes vivants que constitue l'avènement des mitochondries, la phylogénie peut nous aider à avoir une idée des événements ayant conduit à la cellule eucaryote moderne.

## I-2. Choix de la méthode

Par la méthode mise en place, nous voulions mettre en évidence une  $\alpha$ -protéobactérie vivant librement qui provienne d'une  $\alpha$ -protéobactérie qui soit plus proche de l'origine des mitochondries que *R. prowazekii* (cf. figure 6, chapitre I). Étant donné qu'il est estimé que seul un infime pourcentage, soit 0,4% des espèces bactériennes aient été identifiées et formellement décrites (96), il est fort probable que cette bactérie ait appartenu à un groupe d' $\alpha$ -protéobactéries jamais encore décrites auparavant. De plus, puisqu'il est également estimé que seuls 0,1% des espèces présentes dans la nature peuvent être cultivées sur milieu synthétique (97), il est possible que ce groupe de bactéries soit difficilement cultivable en dehors de son milieu naturel.

Pour éviter cet écueil, nous aurions pu effectuer notre étude en réalisant des PCR directes sur les ADN issus d'échantillons naturels en évitant le passage sur milieux synthétiques, comme N. Pace *et al.* l'ont fait (98). Mais les motivations ne sont pas les mêmes : en plus de démontrer la richesse de la diversité biologique comme N. Pace *et al.*, nous voulions identifier un type particulier de bactérie, une bactérie cultivable sur des milieux synthétiques, plus proche de l'origine mitochondriale donc plus apte à servir d'organisme modèle que *Rickettsia prowazekii* dans l'étude de l'origine et de l'évolution des mitochondries. Pour constituer un modèle adéquat, il était préférable que la(le groupe de) bactérie(s) réalisât la phosphorylation oxydative, une fonction centrale de la mitochondrie qui est probablement une des principales forces de sélection ayant conduit à l'établissement définitif des mitochondries au sein d'une grande majorité d'eucaryotes : le passage sur les milieux de culture synthétique ML, MS et MLS qui contiennent du succinate et du lactate (cf. chapitre II) nous a permis de faire une sélection de telles bactéries. Un autre avantage de l'utilisation des milieux de culture

synthétiques était l'évitement de la contamination par les cellules eucaryotes (en général, sauf pour les échantillons 6 et 13) : en effet, si par hasard de l'ADNmt s'était retrouvé dans nos mélanges d'ADN bactérien, les gènes *cox1* et *cob* mitochondriaux auraient été le plus souvent amplifiés en lieu et place des ADN  $\alpha$ -protéobactériens en présence.

Par conséquent, notre méthode présentait les avantages de sélectionner des bactéries qui respirent, qui réalisent la phosphorylation oxydative, d'éviter les contaminations eucaryotes en passant sur milieux synthétiques et de sélectionner l'amplification de gènes d' $\alpha$ -protéobactéries grâce au crible des amorces conçues pour cette spécificité. Néanmoins, la possibilité d' $\alpha$ -protéobactéries ne pouvant se développer sur les milieux de culture synthétiques utilisés ne peut être écarté : nous pourrions rééditer les expériences sur nos échantillons et nos cultures de protistes en réalisant des PCR directes sur les ADN issus de ces échantillons tout en gardant à l'esprit l'éventualité de l'amplification préférentielle des gènes mitochondriaux, qui pourrait masquer l'amplification potentielle d'ADN  $\alpha$ -protéobactérien. L'ampleur de l'effort nécessaire à l'identification de séquences sous ces conditions dépasserait cependant le temps et les ressources disponibles pour le projet pressenti.

## II. Les bactéries mises en évidence par notre méthode

### II-1. Types d' $\alpha$ -protéobactéries

Quelque soit la provenance des échantillons que nous avons traité, soit cultures de protistes isolés ou échantillons naturels, nous avons obtenu une grande uniformité dans les  $\alpha$ -protéobactéries que nous avons mises en évidence : nous avons déterminé la présence de bactéries affiliées à *Rhodopseudomonas*, *Rhodobacter*, *Nitrobacter*, *Paracoccus*, *Mesorhizobium* et *Magnetospirillum*. Ces bactéries possèdent des modes de vie différents : par exemple, tandis que *Mesorhizobium* est une symbiote de cellules végétales, *Rhodobacter* et *Rhodospseudomonas* sont des bactéries photosynthétiques et *Nitrobacter* une bactérie oxydant le nitrite (99). Pourtant, elles ont toutes en commun d'être des bactéries aérobies et de se développer dans les sols, les milieux aquatiques, les

milieux riches en matière organique telle que la boue. Il est donc logique de retrouver de telles bactéries dans notre étude puisque les milieux de départ, nos échantillons naturels proviennent justement de tels environnements (cf. Tableau II). Le même type de bactérie a été mis en évidence à partir des cultures de protistes : comme l'avaient démontré les expériences préliminaires, il n'est pas rare de trouver des  $\alpha$ -protéobactéries dans des cultures de protistes ATCC. Il est à noter que les bactéries que nous avons mises en évidence à partir de ces cultures ne sont pas des bactéries du genre *Enterobacter*, bactérie utilisée pour nourrir des protistes tels que *Jakoba bahamiensis*. Par conséquent, les bactéries mises en évidence ne sont pas des contaminations mais semblent avoir été présentes dès la réception de ces souches. On peut toutefois se questionner sur la présence d' $\alpha$ -protéobactéries dans des cultures de protistes isolés : étant donné que les habitats naturels des protistes utilisés sont majoritairement des environnements aquatiques (100), il se peut que l'isolement des protistes n'ait pas éliminé les  $\alpha$ -protéobactéries présentes, et qu'elles se retrouvent ainsi dans le milieu de culture du protiste. Suivant cette hypothèse, nous aurions alors pu mettre en évidence des bactéries vivant en relation étroite avec des cellules eucaryotes, telle que *Afipia felis* (66), ce qui aurait pu expliquer la présence de ces bactéries dans des cultures de protistes isolés. Pourtant, tel n'a pas été le cas : par conséquent, une des explications reste la présence de ces  $\alpha$ -protéobactéries vivant en milieux aquatiques tout comme les protistes, et qui ont été isolées en même temps que les protistes.

Comme expliqué précédemment (cf. chapitre II), nous envisagions la possibilité de la présence dans les milieux de culture de protistes de bactéries possédant la faculté de respirer, capable de vivre en relation étroite avec un protiste et de manière autonome sur un milieu de culture solide, tout comme la bactérie *Afipia felis* (66). L'utilisation du milieu de culture d'*Afipia felis*, le milieu CYE (67) sur lequel des bactéries du même type aurait pu se développer n'a pas donné de résultats supplémentaires ou différents des deux autres milieux (cf. test de croissance croisée, Tableau IX). Nous avons néanmoins mis en évidence des bactéries affiliées à des  $\alpha$ -protéobactéries symbiotes de plantes telles que *Mesorhizobium*, notre milieu présentant la possibilité de suppléer aux éléments présents dans les milieux de culture synthétiques des bactéries du genre Rhizobiales (101). En ce qui concerne des bactéries parasites de cellules eucaryotes telles que les Rickettsiales, nous ne nous attendions pas à



en mettre en évidence par notre étude, ce type de bactéries se révélant difficile à mettre en culture en dehors de leurs hôtes : les chercheurs ont la possibilité de cultiver certaines Rickettsiales dans des œufs, ou dans des cultures de cellules humaines (102, 103, 104), des «milieux de culture» beaucoup plus élaborés que les milieux de culture synthétiques utilisés lors de cette étude.

## II-2. Relations phylogénétiques des bactéries mises en évidence par notre méthode avec des $\alpha$ -protéobactéries connues

La figure 17 (chapitre III) nous montre les relations phylogénétiques entre les  $\alpha$ -protéobactéries mises en évidence par notre méthode et des  $\alpha$ -protéobactéries connues. Les topologies de ces arbres ne sont pas fortement soutenues par des valeurs de bootstrap à tous les nœuds : seuls certains le sont, parmi lesquels en général ceux concernant la bactérie dont on veut établir la parenté. Ceci est sûrement dû au manque de signal phylogénétique employé dans cette analyse, puisque seuls 100 à 150 acides aminés (soit la longueur du fragment protéique présent dans le fragment PCR) ont pu être analysés. Néanmoins, dans la majorité des cas la topologie obtenue était confirmée par une autre méthode de reconstruction phylogénétique, le maximum de vraisemblance (maximum likelihood). Cette topologie était également consistante avec des arbres publiés (50, 91, 105, 106) et avec l'analyse phylogénétique réalisée avec la majorité des séquences disponibles pour les  $\alpha$ -protéobactéries (figure 20) : par conséquent, les résultats montrés suffisent à une classification préliminaire des  $\alpha$ -protéobactéries mises en évidence par notre méthode.

Nous n'avons pas identifié de bactéries appartenant à un groupe d' $\alpha$ -protéobactéries jamais encore décrites auparavant (chapitre IV, § I-2). Il serait intéressant de refaire cette étude avec la méthode de PCR directes sur les ADN des échantillons, sans passage sur milieu de culture synthétique : ceci démontrerait certainement la présence de bactéries aérobies réalisant la phosphorylation oxydative mais ne s'étant pas développées sur nos milieux de culture. Il est également possible que l'absence de bactéries jamais encore décrites dans notre étude soit dû à notre échantillonnage: en effet, nous nous sommes confinés à des environnements aquatiques «classiques» : des lacs, des étangs, de l'eau de mer, qui aurait pu révéler des organismes non encore décrits. Nous ne nous sommes pas intéressés

(et nous n'avions pas les ressources pour ce faire), à des environnements «extrêmes», tels que des sources d'eau chaude ou les fonds profonds des océans. Pourtant, ce sont dans de tels environnements que des chercheurs tels que N. Pace pour des bactéries (98) et plus récemment D. Moreira et P. López-García pour des eucaryotes (107), ont mis en évidence l'ampleur de la diversité biologique, toujours toutefois en utilisant la méthode de PCR directe. Il serait particulièrement intéressant d'utiliser nos amorces sur des échantillons provenant de tels environnements, afin de tester si elles peuvent permettre d'élargir la diversité des  $\alpha$ -protéobactéries en identifiant des  $\alpha$ -protéobactéries jamais encore décrites, qui pourraient se révéler plus proche de l'origine mitochondriale que les espèces connues. Toutefois, comme nous l'avons dit précédemment, ceci dépasse le temps et les ressources disponibles pour ce projet.

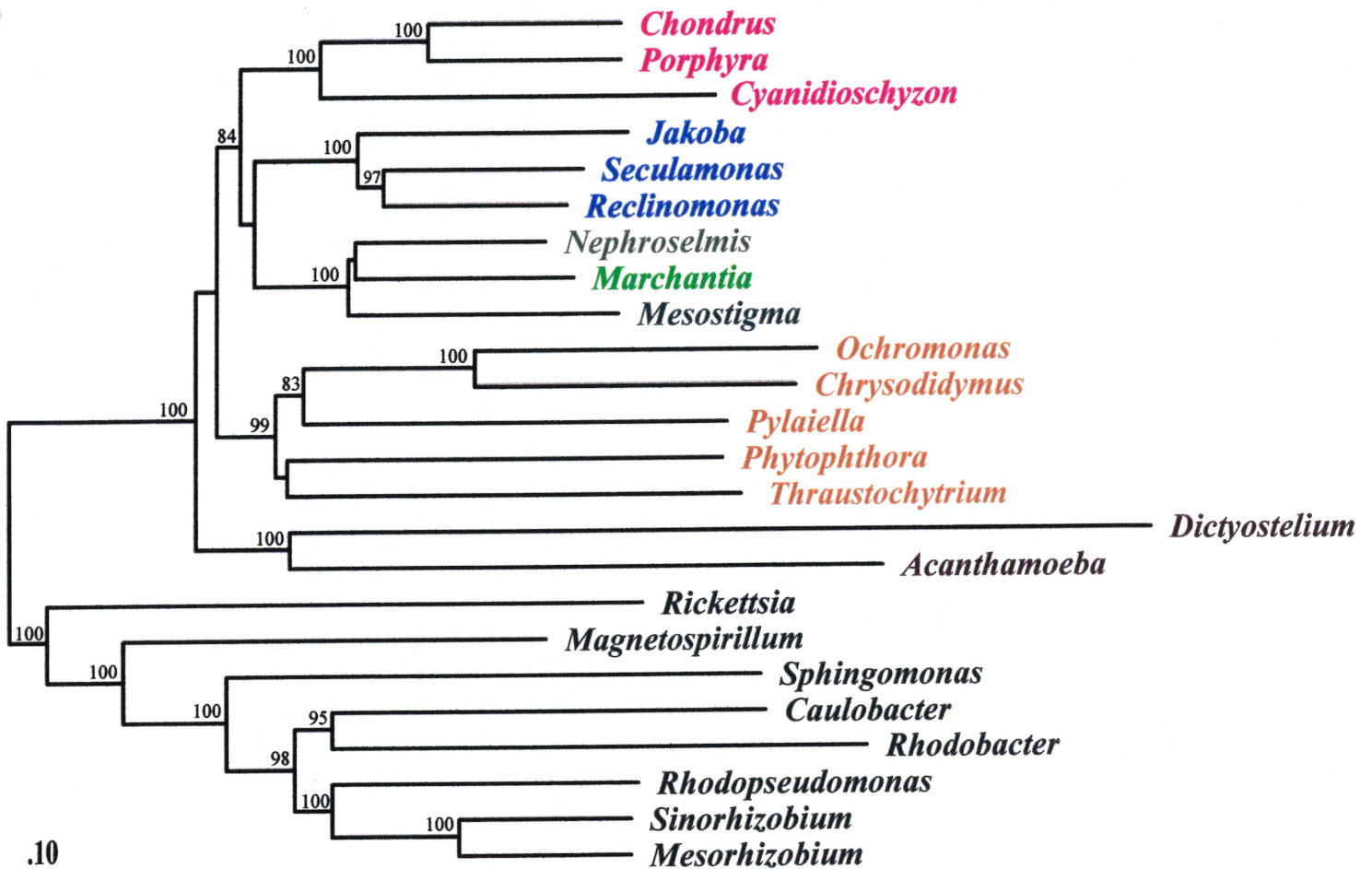
Des centaines de bactéries examinées, aucune ne s'est révélée dérivée d'une espèce plus proche de l'origine mitochondriale au contraire de notre hypothèse de départ. Mais, bien que cette étude n'ait pas permis de confirmer la possibilité d'un changement de la position de *R. prowazekii* au profit d'un autre type d' $\alpha$ -protéobactérie, de nouvelles données viennent étayer l'hypothèse selon laquelle d'autres types d' $\alpha$ -protéobactéries seraient plus ou tout aussi susceptibles d'avoir donné naissance aux mitochondries que *R. prowazekii*, comme discuté ci-dessous.

### III. *Rhodospirillum rubrum*

#### III-1. Relations phylogénétiques entre les $\alpha$ -protéobactéries et les mitochondries : nouvel apport de séquences $\alpha$ -protéobactériennes

Depuis le séquençage complet en 1998, du génome de l' $\alpha$ -protéobactérie *Rickettsia prowazekii*, de nombreuses autres  $\alpha$ -protéobactéries ont été entièrement séquencées ou sont en voie de l'être. L'année 2001 a vu la publication des génomes complets des  $\alpha$ -protéobactéries *Caulobacter crescentus*, *Mesorhizobium loti*, *Rickettsia conorii*, *Agrobacterium tumefaciens* et *Sinorhizobium meliloti* (108-111). Les génomes des cinq  $\alpha$ -protéobactéries suivantes sont quant à elles en cours

**Figure 20** : Relations phylogénétiques entre les mitochondries et les  $\alpha$ -protéobactéries.



L'analyse a été réalisée avec un ensemble de données d'acides aminés alignées et concaténées correspondant aux protéines de la chaîne respiratoire apocytochrome b (Cob) et aux sous-unités 1 à 3 de la cytochrome c oxydase (Cox1-3). Les taxons comprennent des représentants de groupes eucaryotes et de la plupart des  $\alpha$ -protéobactéries dont le génome entier est accessible. L'analyse phylogénétique a été réalisée par une approche de distance, avec la plus récente implémentation de PROTDIST (option matrice JTT) et WEIGHBOR (89, 90) qui permet une correction Jin/Nei pour des taux de changements inégaux à des positions d'acides aminés différentes. Le coefficient de variation utilisé était de 1. La barre horizontale indique le nombre moyen de substitution par site. Les valeurs de bootstrap supérieures à 80% sont indiquées au dessus des noeuds correspondants. La topologie de l'arbre est également supportée par des analyses de maximum-likelihood.

Les  $\alpha$ -protéobactéries sont indiquées en noir, les algues rouges en fuschia, les jakobides en bleu, les algues vertes, les plantes dans divers tons de vert, les straménopiles en jaune et les "slime-molds" en marron. Les organismes et les sequences (code d'accésion à GenBank entre parenthèses) sont *Chondrus crispus* (algue rouge, Z47547), *Porphyra purpurea* (algue rouge, AF114794), *Cyanidioschyzon merolae* (algue rouge, D89861), *Jakoba libera* ( jakobide, 112), *Seculamonas ecuadoriensis* (jakobide, 112), *Reclinomonas americana* (jakobide, AF007261), *Nephroselmis olivacea* (algue verte, AF137379), *Marchantia polymorpha* ("liverwort", M68929), *Mesostigma viride* (plante verte, AF353999), *Ochromonas danica* (algue brune, 112), *Chrysodidymus synuroideus* (stramenopile, AF222718), *Pylaiella littoralis* (stramenopile, AJ277126), *Phytophthora infestans* (stramenopile, U17009), *Thraustochytrium aureum* (stramenopile, AF288091), *Dictyostelium discoideum* ("slime-mold", AB000109), *Acanthamoeba castellanii* (amibe rhizopode, U12386), *Rickettsia prowazekii* (AJ235270-AJ235273), *Magnetospirillum magnetotacticum* (NC\_002725), *Sphingomonas aromaticivorans* (NC\_02719), *Caulobacter crescentus* (NC\_002696), *Rhodobacter sphaeroides* (NC\_002718), *Rhodopseudomonas palustris* (NC\_002804), *Mesorhizobium loti* (NC\_002678), et *Sinorhizobium meliloti* (NC\_002678).

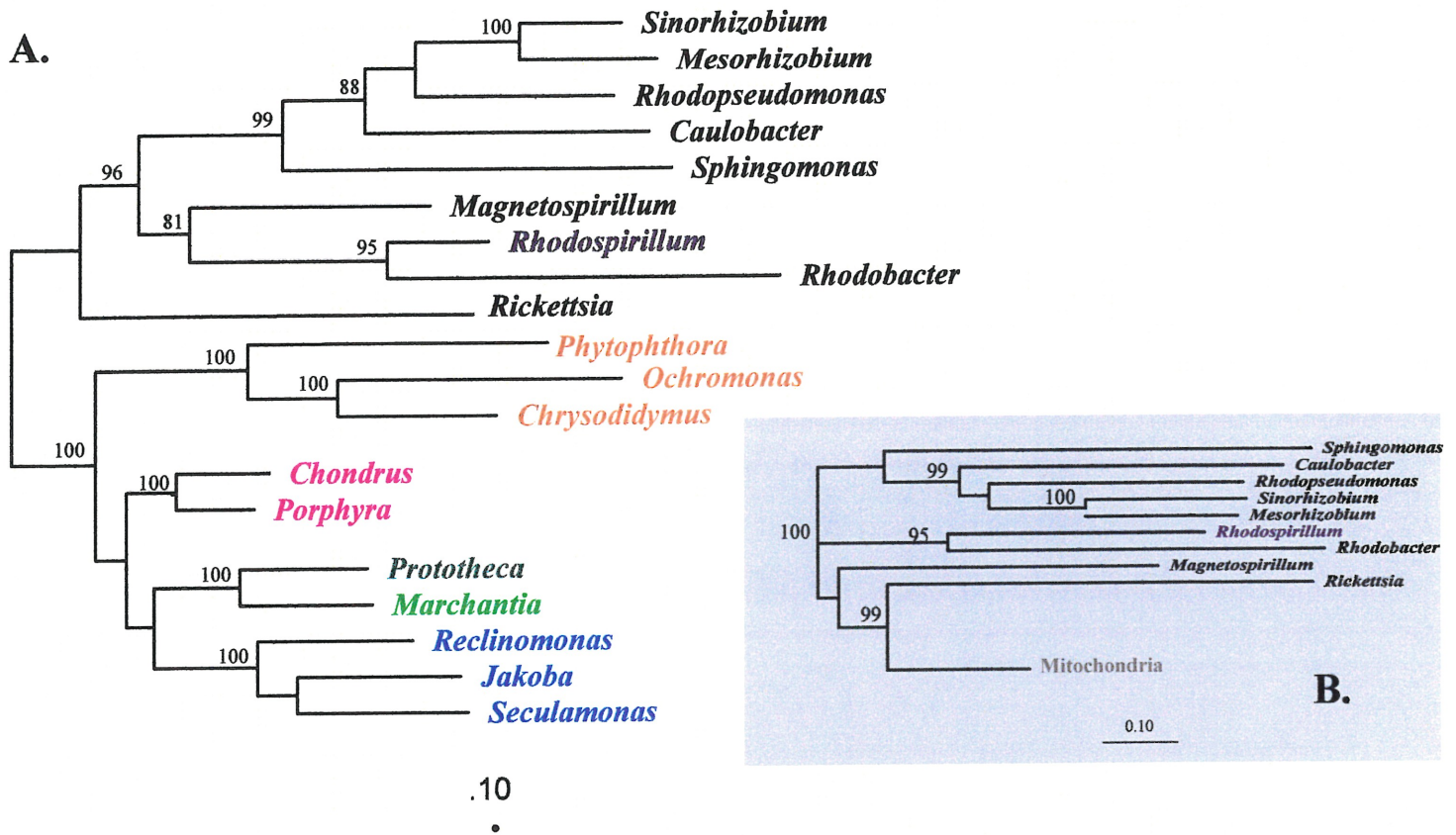
d'annotation : *Rhodobacter capsulatus*, *Magnetospirillum magnetotacticum*, *Rhodobacter sphaeroides*, *Rhodospseudomonas palustris* et *Sphingomonas aromaticivorans* (112). L'accès à ces nouvelles données nous avait amené à réaliser un arbre phylogénétique montrant les relations entre les  $\alpha$ -protéobactéries et les mitochondries (Figure 20) avec les séquences concaténées des gènes *cox1-3* et *cob*, bien avant l'obtention de la séquence du gène *cox1* de *R. rubrum*. Cet arbre nous montre que *Rickettsia* n'est plus l' $\alpha$ -protéobactérie la plus proche de l'origine des mitochondries en terme de distance évolutive même si sa divergence est plus proche de l'origine mitochondriale que les autres  $\alpha$ -protéobactéries : elle est remplacée par l' $\alpha$ -protéobactérie *Magnetospirillum magnetotacticum* avec un fort soutien par les valeurs de bootstrap (100).

### III-2. Arbres phylogénétiques montrant les relations entre les mitochondries et les $\alpha$ -protéobactéries, incluant la bactérie *Rhodospirillum rubrum*

Avant d'obtenir la séquence complète du gène *cox1* de *Rhodospirillum rubrum*, nous avons fait une reconstruction phylogénétique à l'aide des 50 premiers acides aminés de la protéine Cox1 concaténés avec la totalité des protéines Cox2-3, Cob, Atp6 et Atp9 (Figure 21-A). On voit que la topologie de l'arbre est semblable à celle obtenue pour l'arbre de la Figure 20, avec un aussi fort soutien par les valeurs de bootstrap : cependant, des exceptions majeures se distinguent. En effet, on observe que *Magnetospirillum* est toujours la bactérie la plus proche en distance de l'origine des mitochondries, mais les bactéries *Rhodospirillum* et *Rhodobacter* forment une clade et *Magnetospirillum* partage un ancêtre commun avec cette clade, deux résultats qui n'avaient jamais été observés auparavant (cf. Figure 6, Figure 20).

Pour confirmer ce résultat, nous avons utilisé deux autres méthodes de reconstruction phylogénétique qui ont conduit à des résultats similaires : une méthode de maximum likelihood, Puzzle, et une méthode de distance, PROTDIST / WEIGHBOR. La méthode de maximum likelihood donnaient la même topologie que sur la figure 20 (soit la monophylie des bactéries *Magnetospirillum*, *Rhodospirillum* et *Rhodobacter*) et la méthode de distance donne ces trois bactéries comme branchant au même point mais n'étant pas monophylétiques (cf. figure 21-B). Le regroupement de

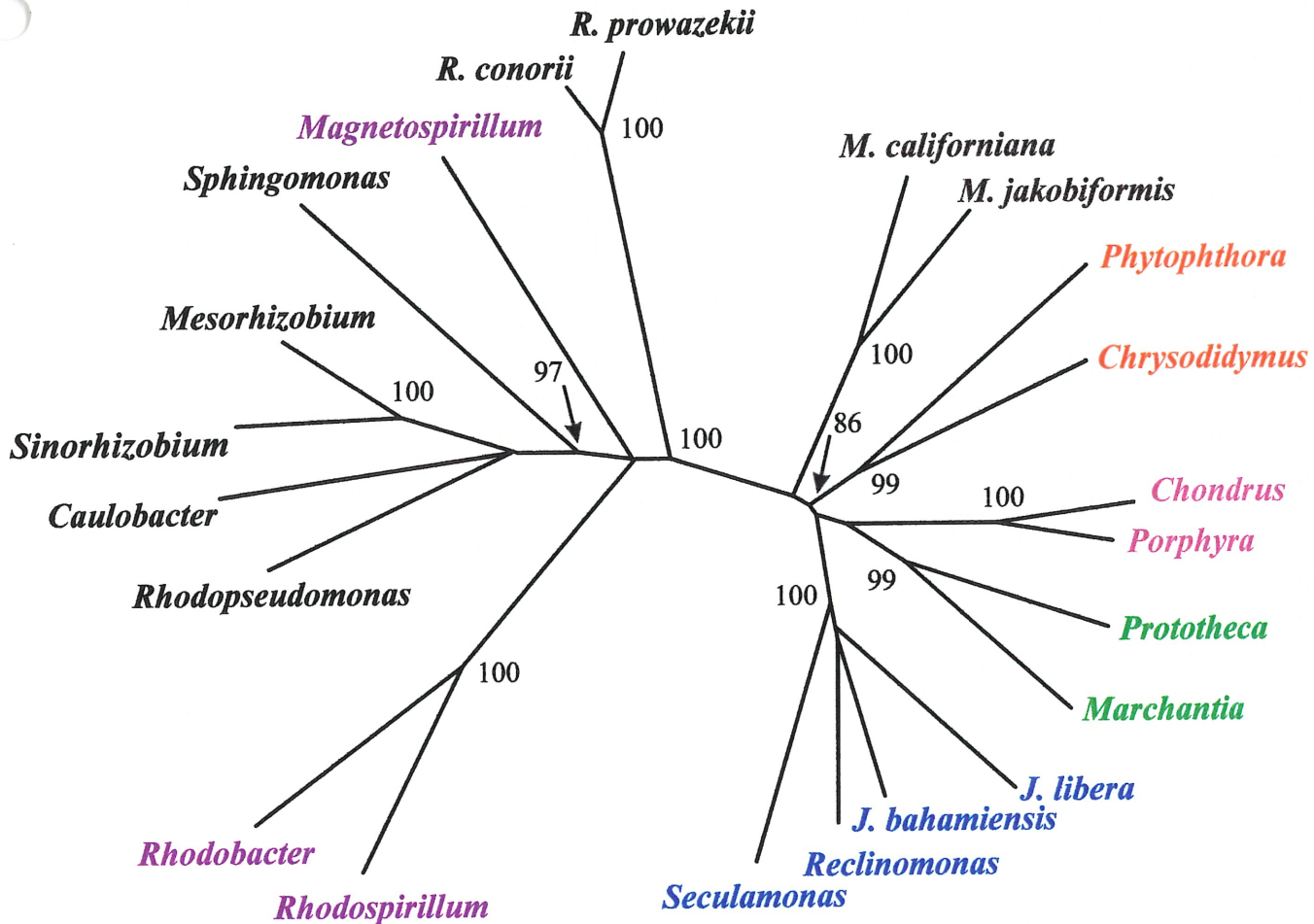
**Figure 21** : Arbre phylogénétique montrant les relations phylogénétiques entre les mitochondries, les  $\alpha$ -protéobactéries, incluant *Rhodospirillum rubrum*



L'analyse a été réalisée avec un ensemble de données d'acides aminés alignées et concaténées correspondant aux protéines de la chaîne respiratoire apocytochrome b (Cob) et aux sous-unités 1 à 3 de la cytochrome c oxydase (Cox1-3). Les taxons comprennent des représentants de groupes eucaryotes et de la plupart des  $\alpha$ -protéobactéries dont le génome entier est accessible. L'analyse phylogénétique a été réalisée par une approche de distance, avec la plus récente implémentation de PROTDIST (option matrice JTT) et WEIGHBOR (89, 90) qui permet une correction Jin/Nei pour des taux de changements inégaux à des positions d'acides aminés différentes. Le coefficient de variation utilisé était de 1. La barre horizontale indique le nombre moyen de substitution par site. Les valeurs de bootstrap supérieures à 80% sont indiquées au dessus des noeuds correspondants. La topologie de l'arbre est également supportée par des analyses de maximum-likelihood.

Les  $\alpha$ -protéobactéries sont indiquées en noir, les algues rouges en fuschia, les jakobides en bleu, les algues vertes, les plantes dans divers tons de vert, les stramenopiles en jaune et les "slime-molds" en marron. Les organismes et les sequences (code d'accèsion à GenBank entre parenthèses) sont *Chondrus crispus* (algue rouge, Z47547), *Porphyra purpurea* (algue rouge, AF114794), *Cyanidioschyzon merolae* (algue rouge, D89861), *Jakoba libera* (jakobide, 112), *Seculamonas ecuadoriensis* (jakobide, 112), *Reclinomonas americana* (jakobide, AF007261), *Nephroselmis olivacea* (algue verte, AF137379), *Marchantia polymorpha* (liverwort, M68929), *Mesostigma viride* (plante verte, AF353999), *Ochromonas danica* (algue brune, 112), *Chrysodidymus synuroideus* (stramenopile, AF222718), *Pylaiella littoralis* (stramenopile, AJ277126), *Phytophthora infestans* (stramenopile, U17009), *Thraustochytrium aureum* (stramenopile, AF288091), *Dictyostelium discoideum* ("slime-mold", AB000109), *Acanthamoeba castellanii* (amibe rhizopode, U12386), *Rickettsia prowazekii* (AJ235270 a AJ235273), *Magnetospirillum magnetotacticum* (NC\_002725), *Sphingomonas aromaticivorans* (NC\_02719), *Caulobacter crescentus* (NC\_002696), *Rhodobacter sphaeroides* (NC\_002718), *Rhodopseudomonas palustris* (NC\_002804), *Mesorhizobium loti* (NC\_002678), et *Sinorhizobium meliloti* (NC\_00267).

Figure 22 : Arbre phylogénétique montrant les relations phylogénétiques entre les mitochondries, les  $\alpha$ -protéobactéries, incluant *Rhodospirillum rubrum*



.10

L'analyse a été réalisée avec un ensemble de données d'acides aminés alignées et concaténées correspondant aux protéines de la chaîne respiratoire apocytochrome b (Cob) et à la sous-unité 1 de la cytochrome c oxydase Cox1. Les taxons comprennent des représentants de groupes eucaryotes et de la plupart des  $\alpha$ -protéobactéries dont le génome entier est accessible. L'analyse phylogénétique a été réalisée par une approche de distance avec la plus récente implémentation de PROTDIST/WEIGHBOR, (option matrice JTT) (88, 90). Le coefficient de variation utilisé était de 0.5. La barre horizontale indique le nombre moyen de substitution par site. Les valeurs de bootstrap supérieures à 80% sont indiquées au dessus des noeuds correspondants.

Les  $\alpha$ -protéobactéries sont indiquées en noir (les bactéries *Rhodospirillum*, *Rhodobacter* et *Magnetospirillum* sont indiquées en violet), les algues rouges en fuschia, les jakobides en bleu, les algues vertes, les plantes dans divers tons de vert et les straménopiles en jaune et *Malawimonas* en brun. Les abréviations sont les suivantes : *J.*, pour *Jakoba*, *M.* pour *Malawimonas* et *R.* pour *Rickettsia*. Les organismes et les séquences (code d'accèsion à GenBank entre parenthèses) sont *Chondrus crispus* (algue rouge, Z47547), *Porphyra purpurea* (algue rouge, AF114794), *Jakoba libera* (jakobide, 112), *Jakoba bahamiensis* (jakobide, 112), *Malawimonas jakobiformis* (NC\_002553), *Malawimonas californiana* (112), *Seculamonas ecuadoriensis* (jakobide, 112), *Reclinomonas americana* (jakobide, AF007261), *Marchantia polymorpha* (liverwort, M68929), *Prototheca wickeramii* (algue verte, 112), *Chrysodidymus synuroideus* (straménopile, AF222718), *Phytophthora infestans* (straménopile, U17009), *Rickettsia prowazekii* (AJ235270 à AJ235273), *Rickettsia conorii* (NC\_003103), *Magnetospirillum magnetotacticum* (NC\_002725), *Sphingomonas aromaticivorans* (NC\_02719), *Rhodobacter sphaeroides* (NC\_002718), *Caulobacter virioides* (NC\_002696), *Rhodopseudomonas palustris* (NC\_002804), *Mesorhizobium loti* (NC\_002678), et *Sinorhizobium meliloti* (NC\_002678).

ces bactéries semble être un résultat constant. Malgré tout, comme nous avons concaténé les séquences des protéines afin de réaliser notre reconstruction phylogénétique, il était possible que ce résultat soit dû à un conflit dans les topologies obtenues avec les protéines individuelles : nous avons testé cette hypothèse en réalisant des reconstructions phylogénétiques pour chacune de ces protéines, ce qui a conduit au même résultat. Par conséquent, avec l'obtention de la totalité de la séquence du gène *cox1*, nous n'avons pas confirmé la position observée dans la figure 12, mais nous avons obtenu un résultat qui demande une étude approfondie des bactéries *Magnetospirillum* et *Rhodospirillum*. Une fois la séquence complète de la protéine Cox1 de *Rhodospirillum rubrum* obtenue, nous avons concaténé les protéines Cox1 et Cob de nombreuses  $\alpha$ -protéobactéries, incluant *R. rubrum* et *Rickettsia conorii* et réaliser une reconstruction phylogénétique en utilisant la méthode PROTDIST/WEIGHBOR. L'arbre résultant, présenté dans la figure 22 confirme ce qui a déjà été observé pour les arbres précédents : *Rhodobacter* et *Rhodospirillum* forme une clade, mais la position de *Magnetospirillum* vis à vis de cette clade est mal définie (cf. les bootstraps dans les figures 21-B, 21A et 22). Avec ces mêmes séquences, nous avons réalisé une reconstruction à l'aide de la méthode de maximum likelihood : un résultat similaire à celui obtenu dans la figure 21-A a pu être observé. Pour ces deux méthodes de reconstruction, l'ajout de la bactérie *R. conorii* ne modifie en rien la position de *R. prowazekii* : la clade formée par deux bactéries reste la première divergence par rapport à l'origine des mitochondries avec de fortes valeurs de bootstraps. Par conséquent, nous pouvons déduire de ces reconstructions préliminaires (une reconstruction à l'aide des protéines complètes concaténées Cox1-3, Cob, Atp6 et Atp9 devant être réalisé avec les deux méthodes de reconstruction phylogénétiques utilisées ici pour confirmer ce résultat) que les bactéries *Rhodobacter* et *Rhodospirillum* forment une clade, et que la divergence de *Magnetospirillum*, bien que proche de l'origine des mitochondries en terme de distance est mal définie par rapport notamment à la clade *Rhodobacter/Rhodospirillum* et ne précède pas celle des Rickettsiales.

Ni par la méthode que nous avons mise en place pour les bactéries issues d'échantillons naturels, ni pour les bactéries provenant de collection de souches isolées pour lesquelles les séquences accessibles sont en croissance, nous ne pouvons démontrer la présence d'une  $\alpha$ -protéobactérie dont la divergence pourrait se situer entre les mitochondries et l'ensemble de toutes les autres  $\alpha$ -

protéobactéries, comme c'était le cas pour *Rhodospirillum* dans l'arbre préliminaire (figure 12). Comme nous l'avons expliqué dans le § II-2, il se peut qu'en allant chercher dans de nouvelles niches écologiques, nous réussissions à mettre en évidence l'existence d'une  $\alpha$ -protéobactérie qui soit plus proche de l'origine mitochondriale. Il est également possible que tout comme pour *Rhodospirillum* et *Magnetospirillum*, des reconstructions phylogénétiques comportant le maximum de signal phylogénétique (ceci s'avère possible par l'accessibilité à un plus grand nombre de séquences  $\alpha$ -protéobactériennes) nous permette de repositionner certaines  $\alpha$ -protéobactéries sur un arbre retraçant les relations phylogénétiques entre les mitochondries et les  $\alpha$ -protéobactéries. L'une d'elle pourrait alors se révéler être la première  $\alpha$ -protéobactérie divergente par rapport à l'origine des mitochondries. Une autre hypothèse serait que les plus proches descendants de l'ancêtre  $\alpha$ -protéobactérien de l'endosymbiote mitochondrial aient disparu, ou soient particulièrement difficiles à isoler et mettre en évidence. Une telle situation est envisagée pour les eucaryotes : il n'a pas encore été déterminé de descendants des cellules eucaryotes ancestrales, et le débat est vif en ce qui concerne ce que peut constituer une cellule eucaryote ancestrale : est ce une cellule eucaryote n'ayant jamais possédé de mitochondries ou encore un organisme d'ores et déjà complexe ? (50, 113-115) . On ne peut toutefois pas écarter la possibilité de la disparition de ces eucaryotes, dont la découverte nous aiderait à répondre aux questions fondamentales soulevées par l'origine de la cellule eucaryote. Ainsi, pour les  $\alpha$ -protéobactéries également, nous ne pouvons écarter la possibilité de la disparition des plus proches descendants de l'ancêtre  $\alpha$ -protéobactérien de l'endosymbiote mitochondrial.



#### IV. Conclusion

Bien que nous n'ayons pas réussi par notre méthode à découvrir une  $\alpha$ -protéobactérie qui soit plus proche de l'origine mitochondriale que *Rickettsia prowazekii*, notre incursion dans le monde des  $\alpha$ -protéobactéries vivant librement nous a confirmé dans notre hypothèse de départ, que *Rickettsia prowazekii* peut être supplantée par une autre  $\alpha$ -protéobactérie quant à l'origine des mitochondries. Il ne s'agit ici que de résultats préliminaires qu'il serait rentable d'approfondir à l'aube de la révolution génomique qui nous donne d'ores et déjà accès à un nombre appréciable de génomes d' $\alpha$ -protéobactéries. Il nous est déjà possible d'utiliser ces informations pour faire, par exemple de la génomique comparative entre les  $\alpha$ -protéobactéries proches des mitochondries et le génome mitochondrial ancestral du protiste *Reclinomonas americana*, afin d'étudier l'ordre des gènes, ou encore de comparer le mode d'expression de ces génomes. Toutes ces données nous permettront également de baser nos argumentations concernant la proche parenté des mitochondries et de certaines  $\alpha$ -protéobactéries sur l'utilisation d'autres génomes que celui extrêmement dérivé du parasite obligatoire *Rickettsia prowazekii*.

## BIBLIOGRAPHIE

1. Woese CR. et Fox GE. (1977). Phylogenetic structure of the prokaryotic domain: The primary kingdoms. *Proc. Natl. Acad. Sci. USA.*, **74**, (11), 5088-5090
2. Whittaker RH. (1959) *Q. Rev. Biol.* **34**, 210-226
3. Woese CR., Kandler O., et Wheelis ML. (1990). Towards a natural system of organisms: Proposal for the domains Archaea, Bacteria and Eucarya. *Proc. Natl. Acad. Sci. USA.*, **87**, 4576-4579
4. Iwabe N., Kuma K., Hasegawa M., Osawa S. and Miyata T. (1989). Evolutionary relationship of archaebacteria, eubacteria, and eukaryotes inferred from phylogenetic trees of duplicated genes. *Proc. Natl. Acad. Sci. USA.*, **86**, 9355-9359
5. Baldauf SL., Palmer JD. and Doolittle WF. (1996). The root of the universal tree and the origin of eukaryotes based on elongation factor phylogeny. *Proc. Natl. Acad. Sci. USA.*, **93**, 7749-7754
6. Koonin EV, Mushegian ER., Galperin MY and Walker DR. (1997). Comparison of archaeal and bacterial genomes : computer analysis of protein sequences predicts novel function and suggests a chimeric origin for the archaea. *Mol Microbiol*, **25**: 619-637
7. Podani J., Oltvai ZN., Jeong H., Tombor B., Barabasi AL. and Szathmary E. Comparable system level organization of Archaea and Eukaryotes. (2001) *Nature Genet.* **29**(1) : 54-56
8. Ribeiro S. et Golding GB. (1998). The mosaic nature of the eukaryotic nucleus *Mol. Biol. Evol.*, **15** (7), 779-788

9. Rivera MC., Jain R., Moore JE. et Lake JA. (1998). Genomic evidence for two functionally distinct gene classes. *Proc. Natl. Acad. Sci. USA.*, **95**, 6239-6244
10. Zillig W. (1991). Comparative biochemistry of Archaea and bacteria. *Curr. Opin. Genet. Dev.*, **1**, 544-551
11. Margulis L. (1993). *Symbiosis in Cell Evolution*, W. H. Freeman
12. Margulis L, Dolan FL, Guerrero R.(2000). The chimeric eucaryote origin of the nucleus from the karyomastigont in amitochondriate protists. *Proc. Natl. Acad. Sci. USA*, **97**(13): 6954-6959
13. Lake JA, Rivera MC. (1994). Was the nucleus the first endosymbiont ? *Proc. Natl. Acad. Sci. USA*, **91**: 2880-2881
14. Gupta R, Golding GB. (1996). The origin of the eukaryotic cell. *Trends. Biochem. Sci*, **21**: 166-171
15. Sogin ML. (1991). Early evolution and the origin of eukaryotes. *Curr. Opin. Genet. Dev*, **1**: 457-463
16. Martin W, Muller M. (1998). The hydrogen hypothesis for the first eucaryote. *Nature*, **5**, 392(6671): 37-41
17. Moreira D, Lòpez-García P. (1998). Symbiosis between methanogenic archaea and  $\delta$ -proteobacteria as the origin of eukaryotes: the Syntrophic Hypothesis. *J. Mol. Evol*, **47**: 517-530
18. Moreira D, Lòpez-García P. (1999). Metabolic symbiosis at the origin of eukaryotes. *TIBS*, **24**: 88-93

19. Vossbrinck CR, Maddox JV, Friedman S, Debrunner-Vossbrinck BA, Woese CR. (1987). Ribosomal RNA sequence suggests microsporidia are extremely ancient eukaryotes. *Nature*, **362**: 411-414
20. Leipe DD, Gunderson JH, Nerad TA, Sogin ML. (1993). Small subunit ribosomal RNA of *Hexamita inflata* and the quest for the first branch in the eukaryotic tree. *Mol. Biochem. Parasitol.* **59**: 41-48
21. Cavalier-Smith T. (1983). A six-kingdom classification and a unified phylogeny in *Endocytobiology II*, ed. W Schwemmler, HEA Schenk, pp1027-34. Berlin: De Gruyter
22. Clark CG, Roger AJ. (1995). Direct evidence for secondary loss of mitochondria in *Entamoeba histolytica*. *Proc. Natl. Acad. Sci. USA*, **92**: 6518-21
23. Roger AJ, Staffan GS, Tovar J, Clark CG, Smith MW, Gillin FD, Sogin ML. (1998) A mitochondrial-like chaperonin 60 gene in *Giardia Lamblia*: evidence that diplomonads once harbored an endosymbiont related to the progenitor of mitochondria. *Proc. Natl. Acad. Sci. USA*, **95**: 229-234
24. Hashimoto T, Sanchez LB, Shirakura T, Muller M, Hasegawa M. (1998). Secondary absence of mitochondria in *Giardia lamblia* and *Trichomonas vaginalis* revealed by valyl-tRNA synthetase phylogeny. *Proc. Natl. Acad. Sci. USA*, **95**: 6860-6865
25. Gray MW, Doolittle WF (1982). Has the endosymbiont hypothesis been proven ? *Microbiological reviews*, **46** (1): 1-42
26. Raff RA, Mahler HR (1972). The nonsymbiotic origin of mitochondria. *Science*, **177**: 575-582
27. Reijnders L (1975). The origin of mitochondria. *J. Mol. Evol.*, **5**: 167-176

28. Cavalier-Smith T. (1975). The origin of nuclei and of eukaryotic cells. *Nature*, **256**: 463-468
29. Sagan L. (1967). On the origin of mitosing cells. *J. Theoret. Biol.* **14**(3) : 255-74
30. Nass S. (1969). The significance of structural and functional similarities of bacteria and mitochondria. *Int. Rev. Cytol.* **25** : 55-129
31. Raven PH. (1970). A multiple origin for plastids and mitochondria. *Science* **16**(946): 641-646
32. Stanier RY. (1970). *Organisation and control in prokaryotic and eukaryotic cells*, Charles H.P., Knight B.C.J.G., Eds., Cambridge University press.
33. Margulis L. *Origin of Eukaryotic cells* (1970) Yale Univ. Press, New Haven, CT
34. John P., Whatley FR. (1975) *Paracoccus denitrificans* and the evolutionary origin of the mitochondrion. *Nature*, **254** : 495-498
35. Spencer DF, Schnare MN. And Gray MW. (1984) Pronounced structural similarities between the small subunit ribosomal RNA genes of wheat mitochondria and *Escherichia coli*. *Proc. Natl. Acad. Sci. USA*, **81** : 493-497
36. McCarroll R., Olsen GJ., Stahl YD., Woese CR. and Sogin ML. (1983) *Biochemistry*, **22** : 5858-5868
37. Kuntzel H., Kochel HG. (1981) Evolution of rRNA and origin of mitochondria. *Nature (London)* **293** : 751-753
38. Gray MW., Sankoff D., Cedergren RJ. (1984) On the evolutionary descent of organisms and

organelles : a global phylogeny based on a highly conserved structural core in small subunit ribosomal RNA. *Nucleic Acids Research*. **12** : 5837-5852

39. Yang D., Oyaizu Y., Oyaizu H, Olsen GJ., Woese CR. (1985). Mitochondrial Origins. *Proc. Natl. Acad. Sci. USA*, **82** : 4443-4447

40. Dickerson RE. (1980) Evolution and gene transfer in purple photosynthetic bacteria *Nature (London)* **283** : 210-212

41. Woese CR., Stackebrandt E., Weisburg WG, Paster BJ., Madigan MT., Fowler VJ., Hahn CM., Blanz P., Gupta R., Nealson KH., Fox GE. (1984). The phylogeny of purple bacteria : the alpha subdivision. *Syst. Appl. Microbiol.* **5** : 315-326

42. Gray MW. (1982) Comparative sequence analysis as an approach to evaluating structure, function, and evolution of 5S and 5.8S ribosomal RNAs. *Can. J. Biochem.*, **60** : 157-171

43. Ferreyra RG, Soncini FC and Viale AM. (1993) Cloning, characterization, and functional expression in *Escherichia coli* of chaperonin ( groESL ) genes from the phototrophic sulfur bacterium *Chromatium vinosum*. *J. Bacteriol.* **175** : 1514-1523

44. Viale AM, Araki AK. (1994) The chaperone connection to the origins of the eukaryotic organelles. *FEBS letters* **341** : 146-151

45. Gupta RS. (1995). Evolution of the chaperonin families (Hsp60, Hsp10 and Tcp-1) of proteins and the origin of the eukaryotic cells. *Molecular Microbiology* **15** : 1-11.

46. Sicheritz-Ponten T, Kurlan CG et Andersson SGE. (1998) A phylogenetic analysis of the cytochrome *b* and cytochrome *c* oxidase I genes support an origin of mitochondria from within the Rickettsiaceae . *Biochimica et Biophysica acta* , **1365** : 545-551

47. Cousineau B, Cerpa C, Lefebvre J et Cedergren R. (1992). The sequence of the gene encoding elongation factor Tu from *Chlamydia trachomatis* compared with those of other organisms. *Gene*, **120** : 33-41
48. Gray MW et Spencer DF. (1996). Organellar evolution., in *Evolution of Microbial life*, DM Roberts, P Sharp G. Alderson M. Collins, Eds. (Cambridge Univ. Press, Cambridge) pp 107-126.
49. Gray MW, Burger G, Lang BF (1999). Mitochondrial evolution. *Science*, **283** : 1476-1481
50. Lang BF, Gray MW, Burger G (1999). Mitochondrial genome evolution and the origin of eukaryotes. *Annu. Rev. Genet.* **33** : 351-397
51. Lang BF, Burger G, O'Kelly CJ, Cedergren R, Golding GB, Lemieux C, Sankoff D, Turmel M et Gray MW (1997). An ancestral mitochondrial DNA resembling a eubacterial genome in miniature. *Nature*, **387** : 493-497
52. Lang BF., O'Kelly CJ. and Burger G. (1998). Mitochondrial genomics in protists, an approach to probing eukaryotic evolution. *Protist* **149** : 313-322
53. Gray MW., Burger G. and Lang BF. (2001). The origin and early evolution of mitochondria. *Genome Biology*, **2(6)** : reviews1018.1-1018.5
54. Stackebrandt E., Murray RGE. and Truper HG. (1988). Proteobacteria classis nov., a name for the phylogenetic taxon that includes the "purple bacteria and their relatives" . *Int. J. Syst. Bact.*, **38** : 321-325
55. Systematics and proteobacteria : <http://www.cs.uoregon.edu/~kglass/microbio/additional/sytems.html>

56. Andersson SGE., Zomorodipour A., Andersson JO., Sicheritz-Ponten T., Alsmark CM., Podowski RM., Naslund K., Eriksson A-S, Winkler HH. & Kurland CG. (1998). The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. *Nature*, **396** : 133-140

57. *Rickettsia conorii* génome : <http://igs-server.cnrs-mrs.fr/>

58. Andersson SGE. (1998). Bioenergetics of the obligate intracellular parasite *Rickettsia prowazekii*. *Biochimica et Biophysica Acta*, **1365** : 105-111

59. Birkelund S., Stephens RS. (1992). Construction of physical and genetic maps of *Chlamydia trachomatis* serovar L2 by pulse-field electrophoresis gel. *J. Bacteriol.*, **174**(9) : 2742-2747

60. Muller HJ. (1964) *Mutat.Res.*, **1** : 2-9

61. Andersson SGE. and Kurland CG. (1998). Reductive evolution of residents genomes. *Trends Microbiol.*, **6** (71) : 263-268

62. Andersson JO. and Andersson SG. (1999). Genome degradation is an ongoing process in *Rickettsia*. *Mol. Biol. Evol.*, **16**(9) : 1178-1191

63. Gray MW. (1998). *Rickettsia*, typhus and the mitochondrial connection. *Nature*, **396** : 109-110

64. Müller M. and Martin W. (1999). The genome of *Rickettsia prowazekii* and some thoughts on the origin of mitochondria and hydrogenosomes. *Bioessays*, **21** : 377-381

65. Voet and Voet (1995). Biochemistry, second edition, John Wiley & sons, Inc. Chapter 19-20, pp538-598

66. La Scola B. and Raoult O. (1999). *Afipia felis* in hospital water supply in association with free-living amoebae. *The Lancet*, **353** : 1330



67. Protocole du milieu de culture CYE : <http://www.dsmz.de/media/med585.htm>
68. Steffens GCM., Buse G., Oppliger W. and Ludwig B. (1983). Sequence homology of bacterial and mitochondrial cytochrome *c* oxidases, partial sequence data of cytochrome *c* oxidase from *Paracoccus denitrificans*. *Bioch. Biophys. Res Comm.* **116**(1) : 335-340
69. Bott M., Bolliger M. and Hennecke H. (1990). Genetic analysis of the cytochrome *c-aa3* branch of the *Bradyrhizobium japonicum* respiratory chain. *Mol. Microbiol.* **4**(12) : 2147-2157
70. Westenberg DJ and Guerinot ML. (1999). Succinate deshydrogenase (Sdh ) from *Bradyrhizobium japonicum* is closely related to mitochondrial Sdh. *J. Bacteriol.* **181**(15) : 4676-4679
71. Iwata et al. (1995). Cytochrome oxidase from *Paracoccus denitrificans*. *Nature.* **376** : 660-669
72. Tsukihara T., Aoyama H., Yamashita E., Tomizaki T., Yamaguchi H., Shinzawa-Itoh K., Nakashima R., Yaono R. and Yoshikawa S. (1996). The whole structure of the 13-subunit oxidized cytochrome *c* Oxidase at 2,5 Å. *Science.* **272** : 1136-1144
73. Keeling PJ., Luker MA. and Palmer JD. (2000). Evidence from beat-tubulin phylogeny that microsporidia evolved from within the fungi. *Mol. Biol. Evol.* **17**(1) : 23-31
74. Adresse internet du site RDP (Ribosomal Database Project) : <http://rdp.cme.msu.edu/html>
75. Sambrook J, Fritsch EF and Maniatis T. (1989). *Molecular biology, a laboratory manual* , 2<sup>nd</sup> edition, Cold Spring Harbor Laboratory Press , 1.25-1.28
76. Single-stranded DNA preparation for sequencing : <http://megasun/ogmp/ssdna.html>
77. Sanger F., Nicklen S. and Coulson AR. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA*, **74** : 5463-5467.

78. Lang B.F. and Burger G. (1990). A rapid, high resolution DNA sequencing gel system. *Anal. Biochem.* **188** : 176-180.

<http://megasun/ogmp/gels.html>

79. Compagnie BIO S&T : <http://www.biost.com>

80. Université Laval, Service de séquençage, local 3233, pavillon C.-E.-Marchand, Québec (Québec), G1K 7P4 : <http://nucleus.rsvs.ulaval.ca:8000/>

81. Dear S. and Staden R. (1991). A sequence assembly and editing program for efficient management of large projects. *Nucleic Acid Res.* **19** : 3907-3911

82. Pearson WR. (1990). Rapid and sensitive sequence comparison with FASTP and FASTA. *Methods Enzymol.* **183** : 63-98

83. Bases de données locales pour FASTA disponible sur requête à [maried@bch.umontreal.ca](mailto:maried@bch.umontreal.ca) .

84. Altschul SF., Gish W., Miller W, Myers EW. And Lipman DJ. (1990). Basic local alignment search tool. *J. Mol. Biol.* **215** : 403-410

85. Adresse internet du NCBI : <http://www.ncbi.nlm.nih.gov/BLAST>

86. Programmes informatiques OGMP pour l'analyse de séquences :

<http://megasun.bch.umontreal.ca/ogmp/ogmpid.html>

87. Thompson JD., Higgins DG. and Gibson TJ. (1994). CLUSTALW : improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22** : 4673-4680

88. Smith SW, Overbeek R., Woese CR, Gilbert W. and Gillevet PM. (1994). The genetic data environment, an expandable GUI for multiple sequence analysis. *Comput. Appl. Biosci.* **10** : 671-675
89. Bruno WJ, Socci ND and Halpern AL. (2000). Weighted neighbor joining : a likelihood-based approach to distance-based phylogeny reconstruction. *Mol Biol Evol.* **17** (1) : 189-197
90. Felsenstein J. Phylip (phylogeny inference package) Versions 3.6 a2.1 distributed by the author. Dep. Genet., Univ. Wash., Seattle.
91. Woese CR. (1987). Bacterial evolution. *Microbiol. Rev.* **51**(2) : 221-271
92. Pennisi E. (1998). Genome data shake tree of life. *Science* 280(5364) : 672-674
93. Doolittle WF. (1999). Phylogenetic classification and the universal tree. *Science.* **284**(5423) : 2124-2129
94. Covello PS. and Gray M. (1992). Silent mitochondrial and active nuclear genes for subunit 2 of cytochrome c oxidase (cox2) in soybean : evidence for RNA-mediated gene transfer. *Embo J.* **11**(11): 3815-3820
95. Nugent JM. and Palmer JD. (1991). RNA mediated transfer of the gene coxII from the mitochondrion to the nucleus during flowering plant evolution. *Cell.* **66** : 473-481
- 96.” The microbial world “ : foundation of the biosphere. based on a report of the American Academy of Microbiology colloquium held January 19-21, 1996 in Palm Coast, Florida. <http://www.asmta.org/acasrc/acal.html>
97. Staley JT. and Konopka AE. (1985). Measurement of in situ activities of nonphotosynthetic in aquatic and terrestrial habitats. *Annu. Rev. Microbiol.* **39** : 321-346
98. Hugenholtz P., Pitulle C., Hershberger KL. and NR. Pace. (1998). Novel division level bacterial diversity in a Yellowstone hot spring. *J. Bacteriol.* **180**(2) : 366-376

99. Holt JG., Krieg NR, Sneath PHA, Staley JT. and Williams ST. (1994). Bergey's manual of determinative bacteriology, 9<sup>th</sup> edition, edition Williams & Wilkins
100. Informations sur les protistes : <http://megasun.bch.umontreal.ca/protists/protists/html>,  
<http://megasun.bch.umontreal.ca/protists/jak.html> et <http://www.ucmp.berkeley.edu>
101. Milieu de culture des Rhizobiales : <http://www.dsmz.media/med098.htm>
102. Tarasevich IV., Makarova VA., Fehsova NF., Stepanov AV., Miskarova ED., Balayeva N. *et al.* (1991). Astrakhan fever, a spotted fever rickettsiosis. *The Lancet*, **377**: 172-173
103. Raoult D., La Scola B., Enea M., Fournier PE., Roux V., Fenoller F. *et al.* (2001) Isolation and characterization of a flea associated rickettsia pathogenic for humans. *Emerg. Infect. Dis.* **7**: 73-81
104. Houpiikian P. and Raoult D. (2002). Traditional and molecular techniques for the study of emerging bacterial diseases : one laboratory's perspective. <http://www.cdc.gov/ncidod/eid/vol8no2/01-0141.html#1>
105. Ambler RP., Meyer TE., Bartsch RG. and Cusanovitch MA. (2001). An alternative to the accepted phylogeny of purple bacteria based on 126S RNA: analyses of the amino acid sequences of cytochrome c2 and c556 from *Rhodoabcter (Rhodovulum) sulfidophilus*. *Arch. Biochem. Biophys.* **388**(1) : 25-33
106. Gupta RS. (2000). The phylogeny of proteobacteria : relationships to other eubacterial phyla and eukaryotes. *FEMS Microbiology Reviews* **24**: 367-402
107. Moreira D, Lòpez-García P. (2002). The molecular ecology of microbial eukaryotes unveils a hidden world. *Trends Microbiol.* **10**(1) : 31-8
108. Nierman WC. *et al.* (2001). Complete genome of *Caulobacter crescentus*. *Proc. Natl Acad. Sci. USA.* **98** : 4136-4141
109. *Mesorhizobium loti* génome : [ftp://ncbi.nlm.nih.gov/genomes/Bacteria/Mesorhizobium\\_loti/NC\\_002678.gbk](ftp://ncbi.nlm.nih.gov/genomes/Bacteria/Mesorhizobium_loti/NC_002678.gbk)

110. Goodner B. et al. (2001). Genome sequence of the plant pathogen and biotechnology agent *Agrobacterium tumefaciens* C58. *Science*. **294**(5550) : 2323-2328
111. Capela D. (2001). Analysis of the chromosome sequence of the legume symbiont *Sinorhizobium meliloti* strain 1021. *Proc. Natl Acad. Sci. USA*. **98**(17) : 9877-9882
112. génomes bactériens en cours d'annotation : [http://www.ncbi.nlm.nih.gov/PMGfs/Genomes/eub\\_p.html](http://www.ncbi.nlm.nih.gov/PMGfs/Genomes/eub_p.html)
113. Sogin ML. (1997). History assignment : when was the mitochondrion founded ? *Curr. Opin. Gen. Dev.* **7**: 792-799
114. Embley TM. and Hirt RP. (1998). Early branching eukaryotes ? *Curr. Opin. Gen. Dev.* **8**: 624-629
115. Forterre P. and Philippe H. (1999). Where is the root of the universal tree of life ? *Bioessays* **21**(10): 871-879

**APPENDICE 1**

### **Milieu de culture LB**

- 10g de tryptone
- 5g d'extrait de levure
- 10g de NaCl
- 1% de glycérol

Dissoudre dans 1000ml d'eau distillée et amener le pH de la solution à 7,5. Autoclaver.

### **Milieu de culture L1pm**

- 1 litre de milieu contenant de l'eau de mer L1 ( Guillard and Hargraves (1993). *Phycologia* 32(3) : 234-236 )
- 1g de bacto-peptone
- 0.675g de Methylamine-HCl

### **Méthode du freeze-squeeze**

- Préparer un gel d'agarose low-melting 0,8-1% (selon la taille du fragment).
- Prendre 0,25-1 $\mu$ g du produit PCR et le faire migrer dans le gel low-melting à maximum 70V
- Couper la bande du gel d'agarose, transférer le morceau d'agar dans un tube eppendorf
- Centrifuger rapidement à 2000rpm pour faire culoter le morceau d'agarose
- Mettre à  $-20^{\circ}\text{C}$  pendant 30 minutes, puis centrifuger 15 minutes à 13000rpm à la température de la pièce

- Transférer le liquide qui contient l'ADN dans un nouveau tube. Ajouter 5  $\mu$ l d'eau au culot d'agarose, mettre à  $-20^{\circ}\text{C}$  pendant 30 minutes, centrifuger de nouveau et combiner le surnageant à celui déjà récupéré.