



الجمهورية الجزائرية الديمقراطية الشعبية
RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE

وزارة التعليم العالي و البحث العلمي
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE SCIENTIFIQUE



Université des Frères Mentouri Constantine
Faculté des Sciences de la Nature et de la Vie

جامعة الاخوة منتوري قسنطينة
كلية علوم الطبيعة و الحياة

Département Microbiologie

قسم : الميكروبيولوجيا

Mémoire présenté en vue de l'obtention du Diplôme de Master

Domaine : Sciences de la Nature et de la Vie

Filière : Sciences Biologiques

Spécialité Ecologie microbienne

Intitulé :

Analyses phénétique et bioinformatique des isolats bactériens de deux sols cultivés de blé dur : Rôle de la structure 2D des ARNr 16S dans l'identification des isolats

Présenté et soutenu par : **Djaber SAYAH**

Le : 16/06/2016

Jury d'évaluation :

Président du jury : R. Alatou (Maître de Conférences A - UFM Constantine)

Rapporteur : M. A. Hamidechi (Professeur - UFM Constantine)

Examineur : M. Gagaoua (Maître de Conférences A - INATAA UFM Constantine)

Examinatrice : I. Guergouri (Maître-Assistante A - UFM Constantine)

Année universitaire

2015 - 2016



Remerciements

Ce travail de mémoire de Master, aussi modeste soit-il, n'aurait abouti sans les efforts et les aides qui m'ont été gracieusement fournis par des personnes auxquelles il m'est impossible de trouver les mots convenables à la juste valeur des aides qu'ils m'ont démontrées.

Merci à mon encadreur, le Professeur **HAMIDECHI M. A.** pour avoir suivi ce travail et l'avoir orienté afin d'aboutir à cette quantité d'enseignements théoriques et pratiques

Je remercie infiniment les honorables membres du jury :

Mme ALATOU R., maître de conférences A pour avoir accepté de présider ce jury

Mr Gagoua M., Maître de Conférences à l'INATAA qui n'a pas manqué de nous apporter l'aide et les apprentissages indispensables à la réalisation de ce Master. Merci Dr Gagaoua M. pour votre accueil chaleureux.

Mme Guerguouri I. Maître-Assistante pour l'expertise de ce travail



Je n'oublierai jamais de remercier tous mes
enseignants d'écologie microbienne à leur tête le
Professeur Benguedouar A. qu'il trouve ici ma
meilleure reconnaissance.



Un vif remerciement pour l'équipe du laboratoire, en particulier **Mme Zoghmar S.** pour son aimable et respectueux accueil. Elle nous a offert toutes les aides possibles afin de bien avancer dans notre travail de mémoire. Nous vous disons Merci Madame pour cette aide très précieuse et riche d'enseignements.

Nous remercions également Mrs **Bennouar F., Laraba D. E.** et **Boufrioua A.** pour nous avoir bien encadrer et donné tous les réactifs et matériels indispensables à la bonne pratique durant toute la période de notre travail de mémoire.

Dédicaces

Je dédie ce travail qui est le fruit du soutien des miens

A mon père SAYAH HOCINE,

En témoigne de ma reconnaissance pour tous les efforts qu'il a consentis pour mon éducation et ma scolarisation

A ma très chère maman BAHNAS ZEHOR,

Pour m'avoir toujours prêté oreille, pour ces précieux conseils et son soutien moral et affectif

Merci mes chers parents, je voudrais que vous soyez fiers de moi.

A mes chers frères

MOHAMED, ADEM, BRAHIM et AMEUR

A mes chères sœurs

RADIA, LOUBNA et NESSRINE

A toute ma famille et tous mes amis surtout **ZEROUAL SOFIANE**

SAYAH Djaber

Liste des figures

<u>Figure 1</u> : Interactions et complexité du sol (Khelili, 2015)	3
<u>Figure 2</u> : proportions en volumes, des principaux composants du sol (Khelili, 2015)	4
<u>Figure 3</u> : Interactions Watson – Crick (G=C et A=U) et Wobble (G=U)	9
<u>Figure 4</u> : Représentation en arc de la structure 2D d'une portion d'ARN	10
<u>Figure 5</u> : Représentation des principaux motifs d'une structure 2D de l'ARN	11
<u>Figure 6</u> : Naissance des interactions entre les nucléotides pour former la structure 2d ayant le minimum d'énergie libre.	12
<u>Figure 7</u> : Séchage du sol à température ambiante	14
<u>Figure 8</u> : Tamisage à 2 mm du sol sans engrais (Ferme 1)	15
<u>Figure 9</u> : Tamisage du sol à 2 mm avec engrais (Ferme 2)	15
<u>Figure 10</u> : Etapes d'extraction et de précipitation des protéines totales des neuf isolats (Gagaoua, 2016)	20
<u>Figure 11</u> : Etapes préliminaires pour extraction de l'ADN bactérien	21
<u>Figure 12</u> :Gel d'électrophorèse de type SDS-page	26
<u>Figure 13</u> : Phylogramme rRNA 16S des 17 souches	30

Liste des tableaux

<u>Tableau 1</u> : Énergie d'empilement de paires de base (kcal/mole à 37°C) (Turner <i>et al.</i> , 1980)	12
<u>Tableau 2</u> : Énergie de déstabilisation pour les autres boucles (kcal/mole à 37°C) (Serra et Turner, 1995)	12
<u>Tableau 3</u> : Volumes des composants ajoutés pour préparer le gel de séparation	17
<u>Tableau 4</u> : Volumes des composants ajoutés pour préparer le gel de concentration	18
<u>Tableau 5</u> : Résultats de la caractérisation morphologique des colonies des isolats du sol de la ferme 1 (sans engrais)	23
<u>Tableau 6</u> : Résultats de la caractérisation morphologique des colonies des isolats du sol de la ferme 2 (avec engrais)	23
<u>Tableau 7</u> : Caractères biochimiques des isolats de la ferme 1 (sans engrais)	24
<u>Tableau 8</u> : Caractères biochimiques des isolats de la ferme 2 (avec engrais)	24
<u>Tableau 9</u> : Caractères physiologiques des isolats de la ferme 1 (sans engrais)	25
<u>Tableau 10</u> : Caractères physiologiques des isolats de la ferme 2 (avec engrais)	25
<u>Tableau 11</u> : Fantôme du gel d'échantillons bactériens représente l'ordre de dépôt des isolats dans les puits du gel.	26
<u>Tableau 12</u> : Résultats du BLAST pour les isolats des deux fermes	27
<u>Tableau 13</u> : Synthèse de la place systématique des différents isolats	29
<u>Tableau 14</u> : Energies libres ΔG° des différentes structures 2D des rRNA 16S pour les isolats de la ferme 1 (sans engrais)	31
<u>Tableau 15</u> : Energies libres ΔG° des différentes structures 2D des rRNA 16S pour les isolats de la ferme 2 (avec engrais)	33

Liste des abréviations et acronymes

AAF : Aérobie Anaérobie Facultatif

ADH: Arginine DiHydrolase

BLAST: Basic Local Aligement Search Tools

CTL: Catalase

CTR: Citrate

EMBL: European Molecular Biology Laboratory

LCT : Lactose

LDC: Lysine décarboxylase

MBT : Mobilité

MEVAG H : Milieu d'Etude de la Voie d'Attaque des Glucides avec l'huile

MEVAG SH : Milieu d'Etude de la Voie d'Attaque des Glucides sans l'huile

MTL : Mannitol

NCBI: National Centre for Biotechnology Information

NDL : Indole

NR : Nitrate Réductase

ONPG:Orthonitrophényl- β -galactoside

PCR: Polymérase Chain Réaction

RM: Rouge de Méthyle

RRS: Ribosomal *RNA Small* subunit

S: Svedberg

TDA: Tryptophane Désaminase

URS : Uréase

VP:Voges-Proskauer

VF : Viande-Foie

XDS : Oxydase

SOMMAIRE

Liste des figures	i
Liste des tableaux	ii
Liste des abréviations et acronymes	iii
INTRODUCTION	1
REVUE BIBLIOGRAPHIQUE	
1- Structure du sol	3
2- Les bactéries du sol	4
3- Eléments de la systématique bactérienne	5
4- Structure des ARN	8
MATERIEL ET METHODES	
A- MATERIEL	14
1- Description des sites de prélèvement	14
2- Echantillonnage	14
3- Matériels bioinformatiques	14
B- METHODES	14
1- Analyse microbiologique du sol	14
2- Analyse protéomique des isolats	16
3- Analyse biomoléculaire par séquençage rRNA 16S	21
4- Traitement bioinformatique des séquences 16S rRNA	22
RESULTATS ET DISCUSSION	23
1- Caractérisation phénotypique des isolats	23
2- Caractérisation métabolique des isolats	24
3- Caractérisation physiologique des isolats	25
4- Analyse du profil électrophorétique des protéines totales pour neuf isolats	26
5- Traitement bioinformatique des séquences rRNA 16S	27
CONCLUSION	37
REFERENCES BIBLIOGRAPHIQUES	38
Annexes	
Résumés	

INTRODUCTION

Les techniques de la biologie moléculaire sont en nette avancée technologique. Les laboratoires sont capables, en une seule journée, de séquencer la totalité d'un génome bactérien et les méthodes traditionnelles d'identification bactérienne qui se basent sur quelques caractéristiques phénotypiques, physiologiques et biochimiques paraissent limitées, surtout dans le cas des microorganismes non cultivables ou même ceux difficilement cultivables (Dauga *et al.*, 2005).

L'analyse microbiologique, telle que nous l'appliquons encore aujourd'hui ne conduit donc pas à une bonne description des différentes communautés microbiennes existantes dans les échantillons analysés et l'interprétation des interactions intermicrobiennes se voient sous-estimées ; voire non aboutissantes. C'est l'avènement des méthodes biomoléculaires, sur la base d'analyses des gènes 16S rRNA (en tant que marqueur moléculaire), des gènes d'expression (ARNm), et la phylogénie moléculaire qui apporte une valeur ajoutée à l'analyse microbiologique classique (Woese, 1987).

Dans le cas de notre étude, nous nous intéressons à la description d'une communauté bactérienne isolées de deux types de sols destinés aux cultures céréalières (blé dur : *Triticum durum*) afin de mettre en exergue l'importance des communautés bactériennes en fonction des conditions de leurs environnements (telles que les engrais) et les interactions microbiennes comparativement aux sols non additionnés d'engrais.

Il ressort que les objectifs tracés pour cette analyse sont :

1. Identifier par méthodes phénétique (caractéristiques phénotypiques, physiologiques et biochimiques : dont le profil électrophorétique des protéines totales en SDS-PAGE) et biomoléculaire (basée sur le séquençage rRNA 16S)
2. Proposer une prédiction des structures 2D des rRNA 16S basée sur l'estimation d'énergie libre afin de confirmer les espèces isolées.

Ce travail comporte une partie bibliographique destinée à un état d'art sur la prédiction des structures 2D des ARN et une partie expérimentale décrivant le procédé et les résultats obtenus ; pour finir avec une conclusion générale et des perspectives.

En fin, il est à noter, que notre travail est une partie d'une thèse (LMD en Ecologie microbienne) inscrite dans le cadre des activités de recherche du laboratoire de Biologie cellulaire et moléculaire, dans l'équipe dirigée par Le Professeur BENHIZIA Y. et encadrée par le Directeur de thèse Pr HAMIDECHI M. A.

1- STRUCTURE DU SOL

1-1 La composition du sol

Le sol est l'environnement le plus complexe de la biosphère et est considéré comme le réservoir majeur de la diversité microbienne (Gobat, 2003). Ce milieu est une union de la lithosphère, l'atmosphère et l'hydrosphère et englobe cinq composants fondamentaux (figures 1 et 2) : la matière organique, eau, air, fraction minérale et organismes vivants. La porosité créée par l'organisation entre les matières organique et minérale va servir de flux d'eau et d'air dans cet habitat (Alexander, 1977). C'est une matrice vivante, complexe et structurée.

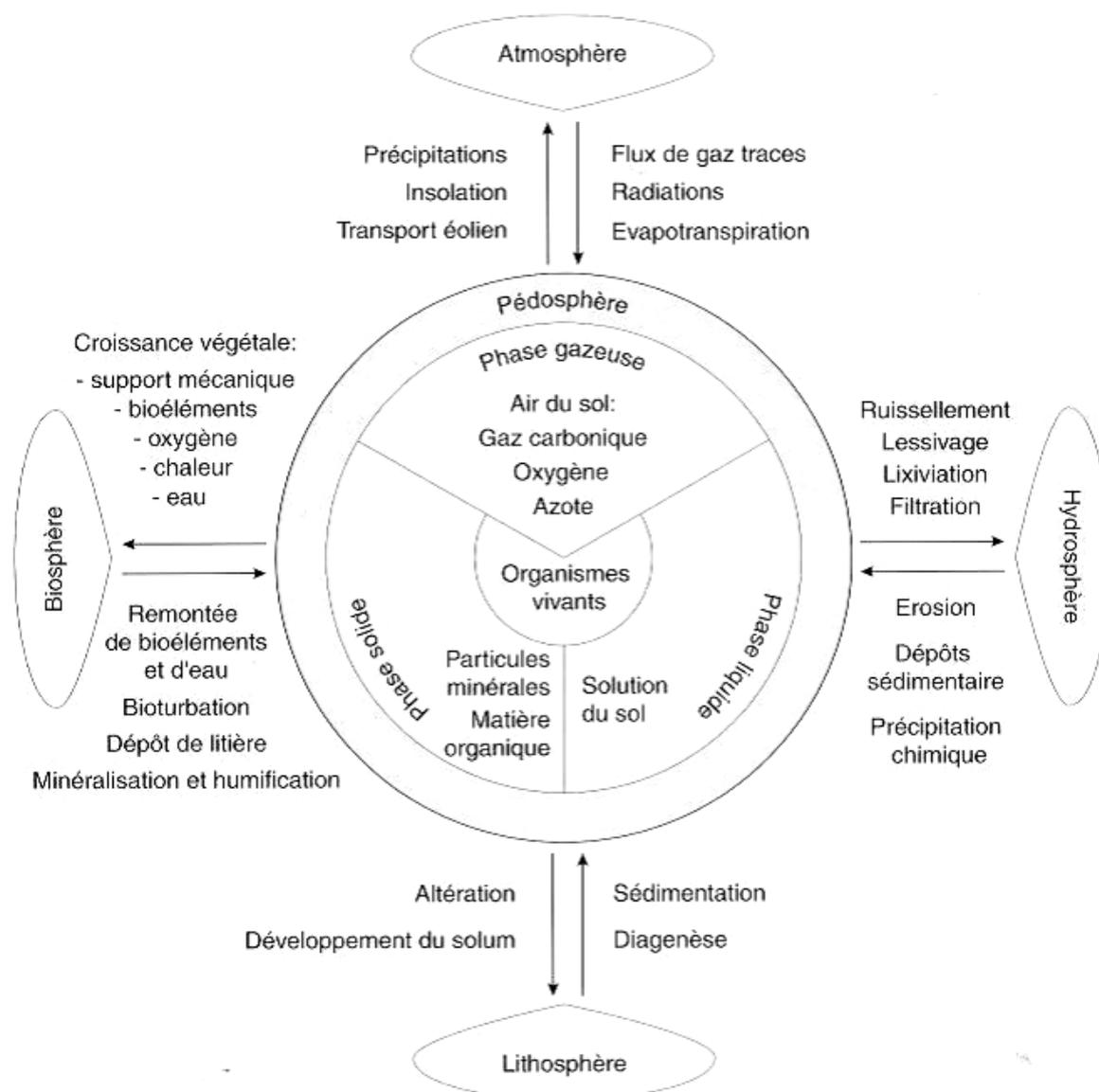


Figure 1 : Interactions et complexité du sol (Khelili, 2015)

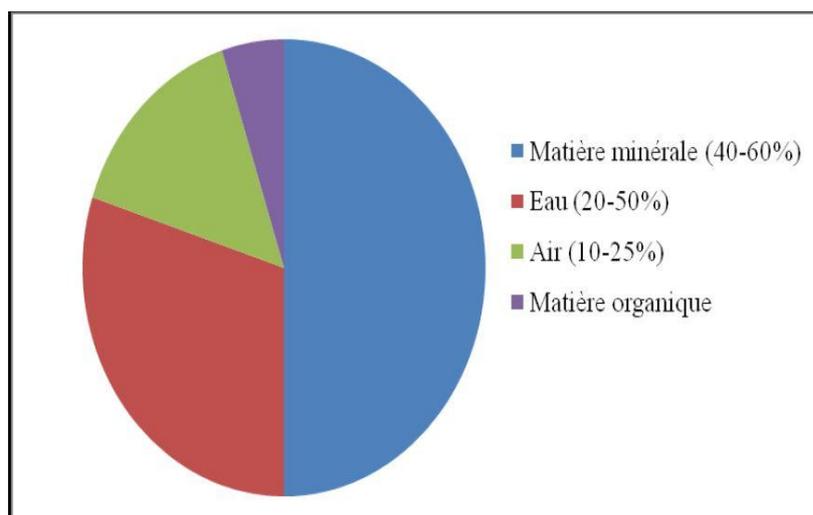


Figure 2 : proportions en volumes, des principaux composants du sol (Khelili, 2015)

2-LES BACTERIES DU SOL

Composent une partie de l'écosystème biotique du sol, se distinguent entre elles par rapport à leurs fonctions, aspects, types respiratoires, types trophiques, etc.

2-1 le type respiratoire

- Les bactéries aérobies strictes et les microaérophiles (niveau d'oxygène moléculaire : inférieur à celui de l'atmosphère) qui exigent l'oxygène comme accepteur final d'électrons, participant aux réactions des métabolismes d'oxydation de la matière organique ($\text{CO}_2 + \text{H}_2\text{O}$).
- Les bactéries anaérobies strictes pour lesquelles l'oxygène est toxique, réduisant la matière organique au cours de processus de fermentation mais n'aboutissant jamais à une oxydation complète de la matière organique.
- Les bactéries aérobies anaérobies facultatives (AAF) dont le métabolisme énergétique est intermédiaire entre les deux précédents.

2-2 le type trophique :

On distingue plusieurs types selon leurs rapports avec la nature de la matière (organique ou minérale) et la source d'énergie métabolique.

- Bactéries hétérotrophes : utilisent les constituants organiques comme source de carbone et d'énergie.
- Bactéries autotrophes : Capables de générer leur propre matière organique à partir des constituants minéraux.

3- ELEMENTS DE LA SYSTEMATIQUE BACTERIENNE

3-1 Systématique à l'échelle phénétique

L'analyse préliminaire dans l'identification microbienne se fait souvent avec une ou plusieurs méthodes phénotypiques. Les méthodes phénotypiques conviennent aux micro-organismes qui sont cultivables, ayant des profils physiologique et biochimique bien établis.

3-1-1 Analyse des caractères morphologiques

Ce sont des méthodes qui reposent sur la description des critères morphologiques. L'étude de la morphologie bactérienne est le premier acte effectué au laboratoire et est réalisée plusieurs fois au cours de l'identification présomptive d'une bactérie car elle permet une orientation préliminaire quant à l'identification de la souche à analyser. C'est une approche préalable et obligatoire à toute identification bactérienne.

3-1-2 Analyses des caractéristiques biochimiques

Les tests biochimiques s'appuient sur l'utilisation des milieux de cultures et les réactifs spécifiques. Ils permettent une orientation plus précise dans le diagnostic bactérien car ils reflètent le comportement et les signatures biomoléculaires des isolats.

3-1-3 Caractérisation physiologique

Fondée sur la réaction des microorganismes vis-à-vis des changements *in vivo* ou *in vitro* dans lequel elles vivent. Cette caractérisation peut être réalisée par :

a- étude de l'influence des facteurs abiotiques sur la croissance bactérienne : Certains facteurs abiotiques peuvent influencer et intervenir, de façon primordiale dans la croissance des bactéries et par conséquent dans l'obtention d'une culture optimale. Parmi ces facteurs on peut citer : La température de croissance, le pH, la pression osmotique (différentes concentrations de NaCl), la pression partielle d'oxygène (mode respiratoire aérobie/anaérobie).

b- Recherche d'enzymes spécifiques : Certaines bactéries possèdent des enzymes qui jouent un rôle déterminant dans l'identification bactérienne. L'on cite :

- L'oxydase : enzyme présente dans certaines chaînes respiratoires bactériennes et qui catalyse une réaction d'oxydo-réduction en impliquant une molécule d'oxygène comme accepteur final d'électrons,
- la nitrate-réductase dont le rôle est la réduction des nitrates en nitrites,

- la catalase dont la fonction est de prévenir l'accumulation de niveaux toxiques de peroxyde d'hydrogène (H₂O₂),
- la β-lactamase (résistance aux antibiotiques β-lactames).

3-2 Systématique à l'échelle biomoléculaire (Données et outils)

L'utilisation de marqueurs phénotypiques pour l'identification d'un germe présente des limites. L'avènement de la biologie moléculaire permet une approche génotypique rapide et complémentaire pour l'identification bactérienne. Ces méthodes reposent toutes sur l'acquisition de séquences nucléotidiques spécifiques. L'identification repose sur la comparaison (alignement) de séquences d'acides nucléiques (ADN, ARN) ou des protéines d'un microorganisme avec les données documentées d'organismes connus et hébergés dans des bases et banques de données universelles libres de droits.

3-2-1 Méthodes génotypiques :

Méthodes telles que le GC% de l'ADN et l'hybridation ADN/ADN. Mais cette dernière requière des régions génomiques hautement conservées au sein des séquences cibles telles que l'ARNr 16S, ou les séquences codantes (ARNm) des certains enzymes spécifiques (Gyrase, enzymes allostériques de régulation comme la Phospho-Fructo-Kinase, Cytochromes, ...).

3-2-2 Approche moléculaire :

Les techniques de biologie moléculaire ont bouleversé les méthodes d'identification bactérienne et ont mis en exergue les insuffisances et les erreurs d'identifications phénotypiques qui étaient jusque-là seules disponibles (Ronsin, 2005 ; Denis *et al.*, 2007). L'identification moléculaire des bactéries repose essentiellement sur l'analyse de la séquence du gène ARNr 16S. La séquence est comparée via le réseau Internet, à l'aide de logiciels spécialisés (blast), avec les banques de séquences comme "Nucleotide" du portail NCBI, ou EMBL en Europe.

L'identification va reposer sur le pourcentage d'identité des séquences. Nous admettons généralement, qu'une similarité supérieure ou égale à 97% permet l'identification au niveau espèce. Une identité de 95% indique un genre commun des deux séquences. Enfin, en dessous de 95%, il y a une possibilité que l'espèce soit nouvelle car non décrite dans les banques de données. Par ailleurs, certains logiciels permettent de positionner la séquence obtenue au laboratoire parmi l'ensemble des séquences ARNr 16S disponibles sous forme d'arbres phylogénétiques (Sanger *et al.*, 1977 ; Thomas, 1980 ; Smith *et al.*, 1986; Swerdlow *et al.*, 1991; Zhang *et al.*, 2003). Cette possibilité est particulièrement intéressante lorsque nous déterminons au laboratoire une séquence

ARNr 16S de similarité inférieure à 97% ; cela permet alors d'avoir une idée du groupe (clade) bactérien dans lequel se situe la bactérie nouvellement identifiée (Staden, 1979).

3-3 Les marqueurs moléculaires (cas de l'ARN 16s) :

Le gène codant l'ARN ribosomal 16S a été le premier et le plus complètement exploré. Les ARNr étant essentiels et ayant un rôle fonctionnel identique pour tous les ribosomes, leur structure évolue très lentement au cours du temps ; vraisemblablement en raison de leur rôle critique et constant dans la physiologie bactérienne. Le gène codant l'ARN ribosomal 16S présente à la fois des régions hautement conservées tout en contenant de courtes séquences signatures spécifiques de genre et parfois d'espèces. Ces caractéristiques particulières font de ce gène une cible intéressante. La généralisation de l'analyse des séquences d'ADNr 16S a incité les auteurs à corréliser les pourcentages d'identité entre séquences d'ADNr 16S et les pourcentages de réassociations ADN/ADN. Il a ainsi été mis en évidence que des séquences d'ADNr 16S ayant des similarités inférieures à 97% ne correspondaient jamais à des pourcentages de réassociation ADN/ADN supérieurs à 60% (Stackebrandt et Goebel, 1994). Par extension, il est généralement admis que des séquences partageant moins de 97% de similarité correspondent à des espèces différentes. Le pouvoir résolutif de l'ADNr 16S, a priori suffisant pour séparer des espèces différentes, ne permet pas toutefois de regrouper des espèces identiques. Si le pourcentage de similarité des ADNr 16S de deux souches est égal ou supérieur à 97, le placement de deux souches dans une unique espèce ou dans deux espèces différentes doit reposer sur les résultats des hybridations ADN/ADN. Dès lors, la définition d'espèce ne peut être basée uniquement sur la base de similarité de séquences d'ADNr 16S ou même d'autres gènes (Rossello Mora et Amann, 2001). Cette technique a largement contribué dans la classification et l'identification des bactéries (Chun et Goodfellow, 1995 ; Embley et Stackebrandt, 1994 ; Kataoka *et al.*, 1997 ; Laurent *et al.*, 1999).

L'ARNr 16S possède de nombreuses fonctions parmi lesquelles :

- Rôle structural : définissant les positions des protéines ribosomales,
- l'extrémité 3' contient une séquence "anti-Shine-Dalgarno" (AGGAGG) qui se lie en amont du codon AUG (Start codon) de l'ARNm. L'extrémité 3' de l'ARNr 16S se lie aux protéines S1 et S21 connues pour être impliquées dans l'initiation de la synthèse des protéines,
- interagit avec le 23S, contribuant ainsi à la liaison des deux sous-unités ribosomiques (30S et 50S),

- stabilise l'appariement codon-anticodon sur le site A du ribosome par l'intermédiaire de la formation d'une liaison hydrogène entre l'atome N1 des résidus Adénine 1492 et 1493 et le groupe 2'OH de l'épine de l'ARNm.

Dans les années 1980, il a été montré que les relations phylogéniques entre les êtres vivants pouvaient être déterminées en comparant leurs séquences nucléiques (Clarridge, 2004). Des travaux préliminaires ont montré que le 16S rRNA gene code pour un ARN ribosomal de fonction constante dans l'évolution ; ce gène peut donc servir d'horloge moléculaire pour suivre les changements dans l'évolution bactérienne. Les séquences du gène 16S ont ainsi joué un rôle majeur dans l'étude de la phylogénie et de la taxinomie bactérienne (Janda and Abbott, 2007; Olsen and Woese, 1993). Il a d'abord été utilisé à la fin des années 1980 par Carl Woese comme outil d'étude de l'évolution bactérienne (Woese, 1987). En 1991, Weisburg *et al.* ont décrit des amorces dites «universelles» permettant d'amplifier l'intégralité du gène codant pour l'ARNr 16S de la plupart des bactéries (Weisburg *et al.*, 1991).

Les séquences types de l'ARNr 16S de la plupart des souches bactériennes sont disponibles sur les bases de données publiques telles que Genbank du portail NCBI (www.ncbi.nlm.nih.gov/). Cependant, la qualité de ces séquences n'est pas validée pour la plupart d'entre elles. Par conséquent, les bases de données spécifiques qui n'hébergent que les séquences d'ARNr 16S, sont plus fiables. Les bases de données les plus utilisées en bioinformatique sont :

- EzTaxon-e (<http://eztaxon-e.ezbiocloud.net/>),
- RDP (<http://rdp.cme.msu.edu/>),
- SILVA (www.arb-silva.de),
- Greengenes (<http://greengenes.secondgenome.com>).

4- Structure des ARN

L'ARN est une chaîne polynucléotidique pouvant être composée de 20 à 3000 résidus mais reste plus courte l'ADN. La demi-vie de l'ARN ne peut dépasser trois minutes. Il peut être répété plusieurs milliers de fois (Kaplan *et al.*, 2007) L'ordre des nucléotides et leurs interactions dues aux liaisons Hydrogène donnent naissance à plusieurs structures (primaire, secondaire et tertiaire) :

La structure primaire :

C'est simplement la séquence (orientée de 5' vers 3') des nucléotides qui composent la molécule d'ARN. La structure primaire est une succession de nucléotides liés entre-eux de manière covalente (liaison phosphodiester). Elle renseigne principalement sur le %GC.

La structure secondaire :

L'ARN est une molécule simple brin qui a la capacité de se replier sur lui-même permettant ainsi à ses nucléotides éloignés dans la structure primaire de s'apparier entre eux par des interactions hydrogènes. C'est un processus stochastique hiérarchique entraîné par le couplage d'un sous-ensemble des nucléotides qui composent la molécule d'ARN. La structure 2d met en évidence les régions appariées dues à des interactions entre nucléotides. Ces interactions sont de deux types (figure 3) : Type Watson-Crick : concernent les appariements $G \equiv C$ et $A = U$ et type Wobble pour le cas rare des interactions $G = U$:

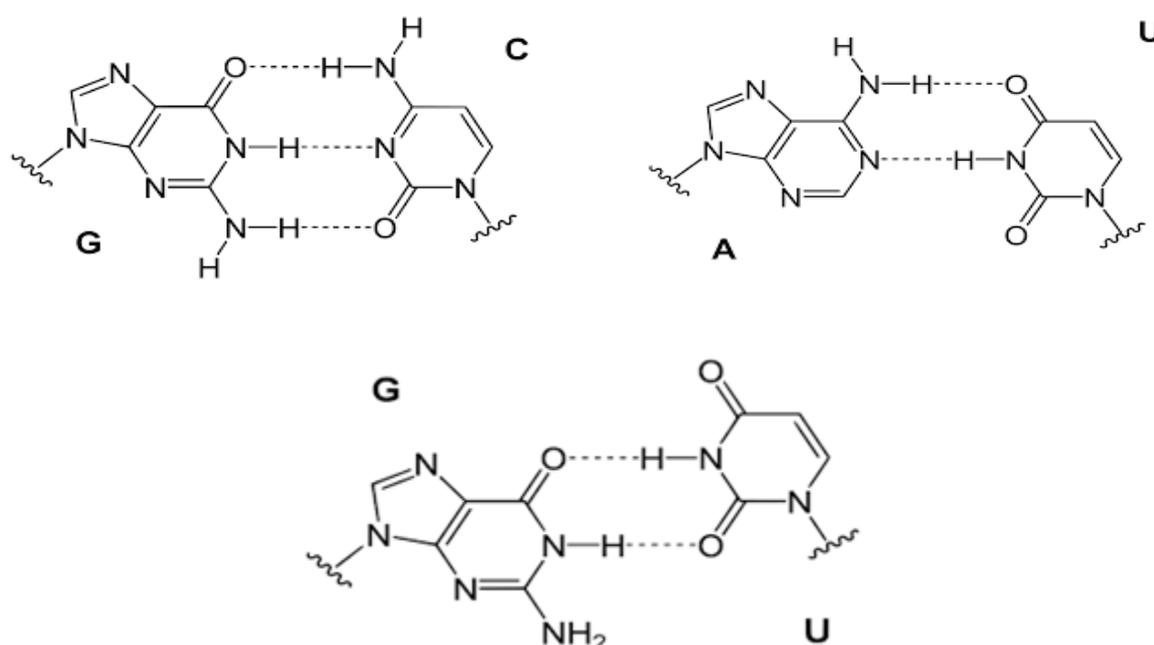


Figure 3 : Interactions Watson – Crick ($G \equiv C$ et $A = U$) et Wobble ($G = U$)

4-1- Représentation des structures secondaires

La représentation graphique des structures 2d des ARN peut se faire de plus d'une manière. Elle peut être représentée par :

a- des diagrammes en arcs (figure 4) où la séquence est écrite sur une ligne et où les paires de bases sont matérialisées par des arcs tracés au-dessus :

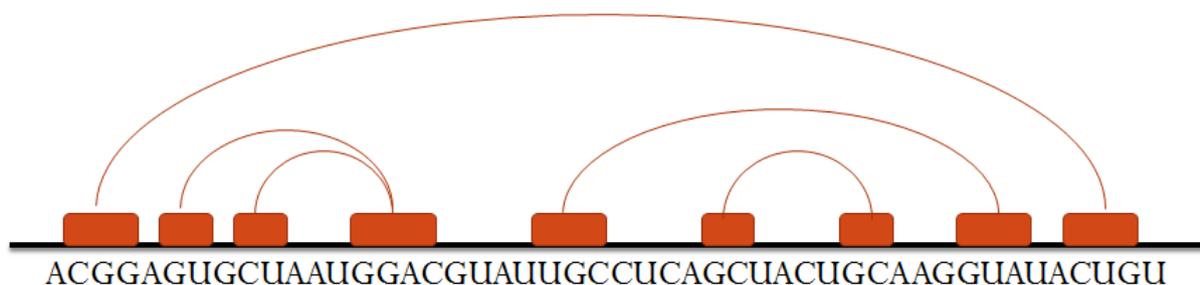
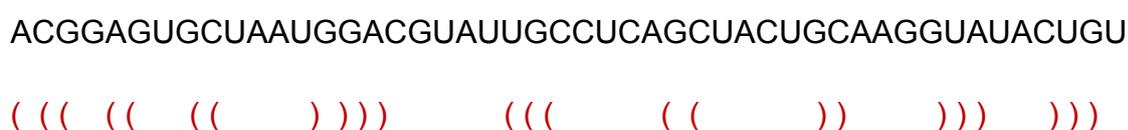


Figure 4 : Représentation en arc de la structure 2d d'une portion d'ARN

b- annotée à l'aide de parenthèses ouvrantes et fermantes, chacune représentant une des deux bases d'une même paire. Les bases libres sont représentées par des points. C'est le format Vienna :



c- la représentation type des structures 2d est une structure primaire représentée dans un espace à deux dimensions (figure 5). Dans cette représentation, plusieurs motifs (éléments) apparaissent. Les plus rencontrés sont :

- La tige : segment bicaténaire dans lequel plusieurs nucléotides s'apparient,
- La boucle terminale est une structure circulaire faite d'une succession de nucléotides comme dans la structure primaire. Il n'y a pas d'appariements possibles entre les nucléotides de cette boucle. La boucle vient toujours à la fin d'une tige.
- Les hernies (bulge) sont sous la forme d'une demi-boucle. Ce sont des renflements latéraux sur un seul brin.

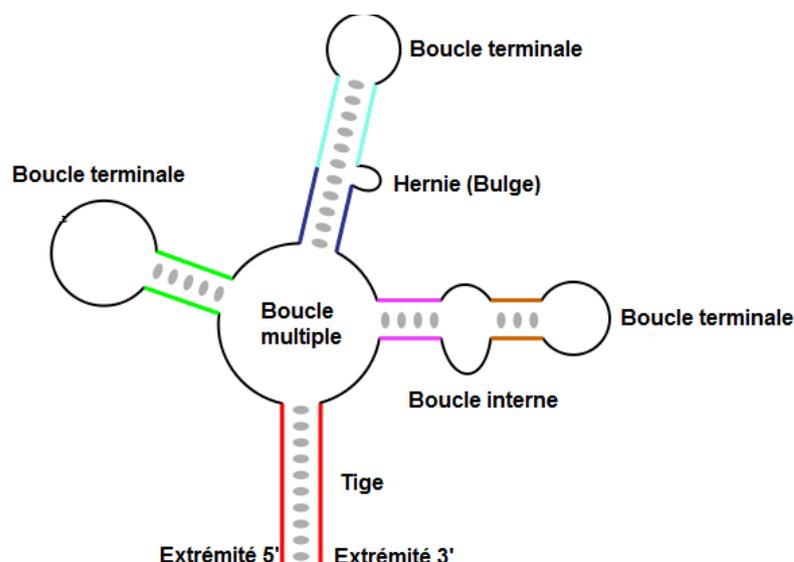


Figure 5 : Représentation des principaux motifs d'une structure 2d de l'ARN

4-2- Prédiction des structures 2d des ARN

Un bon nombre de méthodes permettent de prédire la structure 2d d'une molécule d'ARN. Mais deux principales méthodes de prédiction de structure sont les plus employées pour des fins d'annotation :

- a. Analyse comparative des séquences / Analyse de covariation : prend en compte les modèles de paires de bases conservés durant l'évolution moléculaire des séquences, car celles-ci varient en même temps au cours de l'évolution tout en maintenant l'intégrité structurale.
- b. Méthode du minimum d'énergie libre (ΔG°) : détermine la structure des régions complémentaires qui sont énergétiquement stables. La structure 2d selon cette procédure est la celle avec le minimum d'énergie libre. L'hypothèse fondamentale de cette théorie est que l'énergie de chaque paire de bases est indépendante de toutes les autres paires ; en conséquence de quoi, l'énergie libre totale de toute la molécule est la somme de toutes les énergies des paires de bases (figure 6). Les programmes (logiciels) les plus performants sont :
 - RNA Fold Web Server (<http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi>)
 - MFold (<http://unafold.rna.albany.edu/?q=mfold>).

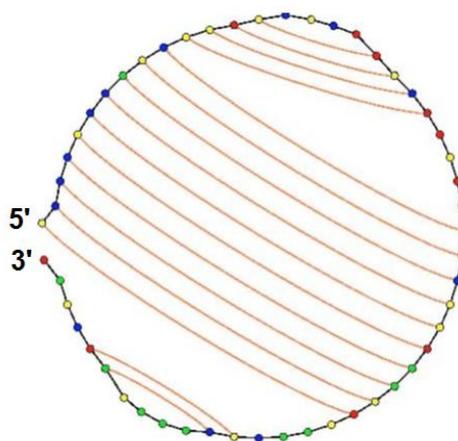


Figure 6 : Naissance des interactions entre les nucléotides pour former la structure 2d ayant le minimum d'énergie libre.

L'estimation de l'énergie pour une structure 2d donnée est calculée à la base des valeurs d'énergies standards (Tableaux 1 et 2) calculées pour les différentes paires de bases possibles selon Turner *et al.*, (1988).

Tableau 1 : Énergie d'empilement de paires de base (kcal/mole à 37°C) (Turner *et al.*, 1980)

	A/U	C/G	G/C	U/A	G/U	U/G
A/U	-0.9	-1.8	-2.3	-1.1	-1.1	-0.8
C/G	-1.7	-2.9	-3.4	-2.3	-2.1	-1.4
G/C	-2.1	-2.0	-2.9	-1.8	-1.9	-1.2
U/A	-0.9	-1.7	-2.1	-0.9	-1.0	-0.5
G/U	-0.5	-1.2	-1.4	-0.8	-0.4	-0.2
U/G	-1.0	-1.9	-2.1	-1.1	-1.5	-0.4

Tableau 2 : Énergie de déstabilisation pour les autres boucles (kcal/mole à 37°C) (Serra et Turner, 1995)

Nombre de bases	1	5	10	20	30
Boucle interne	-	5.3	6.6	7.0	7.4
Bulge	3.9	4.8	5.5	6.3	6.7
Épingle à cheveux	-	4.4	5.3	6.1	6.5

4-3 L'ARN 16S et ses fonctions :

L'ARNr 16S possède de nombreuses fonctions parmi les plus rencontrés sont :

- Rôle structural : définissant les positions des protéines ribosomales,
- L'extrémité 3' contient une séquence "anti-Shine-Dalgarno" (AGGAGG) qui se lie en amont du codon AUG (Start codon) de l'ARNm. L'extrémité 3' de l'ARNr 16S se lie aux protéines S1 et S21 connues pour être impliquées dans l'initiation de la synthèse des protéines,
- Interagit avec le 23S, contribuant ainsi à la liaison des deux sous-unités ribosomiques (30S et 50S),
- stabilise l'appariement codon-anticodon sur le site A du ribosome par l'intermédiaire de la formation d'une liaison hydrogène entre l'atome N1 des résidus Adénine 1492 et 1493 et le groupe 2'OH de l'épine de l'ARNm.

MATERIEL ET METHODES

A- MATERIEL

1- Description des sites de prélèvement :

Un premier site de prélèvement de sol est situé dans une ferme (Ferme 1) entre la commune de Ain S'mara (de coordonnées 36° 16' 03'' Au Nord et 6° 30' 05'' au Sud et à une altitude de 629 m) et la nouvelle ville Ali Mendjeli (à 36° 14' 45'' Nord et 6° 34' 02'' Sud). C'est un sol qui n'a pas été enrichi d'engrais pour l'année en cours et dans le quel a été cultivé du blé dur (*Triticum durum*). Le deuxième site est situé dans une ferme (Ferme 2) entre Ain S'mara et la localité de Arama en allant vers la commune de Oued Seguin (Daira de Téleghma, Wilaya de Mila). Ce deuxième site a été enrichi d'engrais en mois de février de cette même année.

2- Echantillonnage : Les prélèvements se sont effectués en mois de Février 2016 au niveau des deux fermes. Les débris de surface ont été écartés, puis les prélèvements se sont effectués, sur une profondeur de 10 cm, à l'aide d'une spatule préalablement stérilisée au four dans du papier aluminium. Ensuite, les échantillons ont été transportés au laboratoire de microbiologie où un tamisage à 2mm a été effectué pour homogénéiser les échantillons.

3- Matériel bioinformatique : Le séquençage rRNA 16S des différents isolats a été effectué par le Dr M.Kerkoud, responsable scientifique DIAG-GENE à Angers Technopôle (France). Les séquences ont fait l'objet de matériel d'études biomoléculaires et bioinformatiques. Une gamme de programmes bioinformatiques a été utilisée pour traiter les séquences 16S Rrna :

- RNAFold
- EMBOSS
- BLAST
- MEGA 5.2

B- METHODES

1- Analyse microbiologique du sol

1-1 Isolement et purification : L'échantillon a été transporté au laboratoire de biotechnologie (Faculté SNV, Université Frères Mentouri Constantine) où il a été séché à température ambiante pendant 24h (photo 1), puis tamisé à 2 mm (photos 2 et 3).



Figure 7 : Séchage du sol à température ambiante



Figure 8 : Tamisage à 2 mm du sol sans engrais (Ferme 1)



Figure 9 : Tamisage du sol à 2 mm avec engrais (Ferme 2)

La réalisation des dilutions, à partir de l'échantillon, a été faite géométriquement à raison de dix (Bastide *et al.*, 1986) :

- Peser 1g de sol tamisé et frais dans un tube stérile contenant 9mL d'eau physiologique stérile (C'est la dilution 10^{-1}).
- Mélanger pendant 5 min à l'aide d'un vortex pour obtenir une solution homogénéisée et détacher les cellules bactériennes fixées sur les particules du sol.
- Effectuer des dilutions géométriques successives en cascades jusqu'à obtention de la dilution 10^{-6} .

L'inoculation (en triplicata) s'est effectuée par étalement, à l'aide d'un râtelier en verre fabriqué à partir d'une pipette Pasteur stérile, de 0,1ml prélevé des trois dilutions (10^{-4} , 10^{-5} et 10^{-6}), sur boîtes de Pétri contenant de la GN.

Après incubation de 24h à 37°C, les colonies développées ont été repiquées sur GN dans un but de purification.

1-2 Caractérisation présomptive des isolats bactériens

Des tests morphologiques, physiologiques et biochimiques ont été réalisés pour une caractérisation des isolats.

1-2-1 Etude des caractères morphologiques

Cette étude a été basée sur des observations macroscopiques et microscopiques optiques (Gx100) pour différencier l'aspect des colonies sur boîtes de Pétri, le type de Gram, la forme cellulaire ainsi que la disposition (isolées ou regroupées) et la mobilité des cellules bactériennes.

- Examen macroscopique : L'observation de l'aspect macroscopique des colonies sur GN a concerné les éléments suivants (Thomas *et al.*, 2002) :

- La forme des colonies : circulaires, avec contours régulier ou irrégulier,
- La couleur de la colonie,
- L'élévation : convexes, concaves, plates, à centre élevé,
- L'opacité : opaque, translucide ou transparente,
- La surface : lisse, rugueuse, sèche, dentelée.
- Aspect microscopique : Observation des bactéries après coloration de Gram selon le protocole de Nicolle (1895):
- Préparer et fixer un frottis bactérien à la chaleur du bec Bunsen,
- Recouvrir au violet de Gentiane pendant 1 à 5 min. Eliminer l'excès par l'eau courante,
- Ajouter du Lugol : deux bains de 4 à 6 secondes, Laver l'excès par l'eau courante,
- Traiter au mélange alcool/acétone (3/1), puis rinçage à l'eau,
- Recolorer à la Fuch sine pendant 20 secondes, puis rinçage à l'eau, et en fin, séchage.
- L'observation à l'huile à immersion.

1-2-2 Métabolisme glucidique

Les tests réalisés sont les plus classiques de la microbiologie fondamentale, à savoir :

- Test de Mannitol Mobilité
- Gaz en glucose, lactose, production d'H₂S sur milieu KIA

1-2-3 Etude des enzymes respiratoires terminales (tests physiologiques)

- Recherche de l'oxydase, Recherche de la catalase, Test de la nitrate réductase, Test ONPG

1-2-4 Étude des voies fermentatives intermédiaires des acides complexes : RM et VP

1-2-5 Métabolisme protéique :

- Recherche de l'indole et mise en évidence de l'uréase et de la TDA
- Recherche des décarboxylases : LDC, ODC et ADH

2- Analyse protéomique des isolats

Afin d'apporter une meilleure corroboration aux résultats phénotypiques, une électrophorèse SDS-PAGE des protéines totales a été réalisée au niveau du laboratoire de recherche à l'INAATA, Université Frères Mentouri Constantine où nous avons passé un stage pratique d'une durée d'un mois sous l'honorable encadrement du Dr. Gagaoua M. qui nous a aimablement assuré les meilleures conditions d'apprentissage théorique et surtout pratique et la documentation pertinente en relation avec la nature de notre stage.

Vue les contraintes expérimentales et surtout le manque ressenti des produits chimiques indispensables à une telle étude, nous avons choisi neuf souches parmi les 24 isolées des deux fermes. Le choix des souches a été effectué sur la base de leur aspect et des résultats de l'analyse microbiologique.

La SDS-PAGE s'est réalisée selon deux étapes principales :

- Extraction et précipitation des protéines bactériennes (Figure 6), (Gagaoua *et al.*, 2015)
- Electrophorèse SDS-PAGE (Laemmli, 1970)

Une fois les échantillons des neuf souches concentrées, une électrophorèse est appliquée selon le procédé précédent.

La migration électrophorétique en conditions dénaturantes SDS-PAGE a été effectuée selon les étapes décrite dans le procédé de Gagaoua (2015) :

a- Préparation de la cuve d'électrophorèse et des gels (Gagaoua, 2015) :

- Monter les plaques de la cuve avec toutes les précautions d'usage en utilisant les espaceurs adéquats. Vérifier la propreté des plaques.
- Préparer le gel de séparation ou Résolving gel (12 %) dans un bêcher de 25mL avec un barreau magnétique selon le tableau suivant :

Tableau 3: Volumes des composants ajoutés pour préparer le gel de séparation.

	Volumes pour 2 gels (mL)
Acrylamide / Bisacrylamide	6,0
Tris 1,5 M pH 8,8	3,75
Eau bi-distillée	4,95
SDS 10%	0,1
Agiter avant d'ajouter le TEMED et l'APS en μ L	
TEMED	12
APS 30 mg/mL	100
Volume final (mL)	~15

- Couler 5000 μ l de cette solution pour un gel
- En fin de coulage recouvrir le gel avec 400 μ L d'éthanol à 70 %.
- Laisser polymériser environ 25-30 minutes.
- Préparer le gel de concentration ou Stacking gel (3,75 %) dans un bêcher de 25mL avec un barreau magnétique selon le tableau suivant :

Tableau 4 : Volumes des composants ajoutés pour préparer le gel de concentration.

	Volumes pour 2 gels (mL)
Acrylamide / Bisacrylamide	0,75
Tris 0.5 M pH 6.8	1,5
Eau bi-distillée	3,6
SDS 10%	0,6
Agiter avant d'ajouter le TEMED et l'APS en μL	
TEMED	10
APS 30 mg/mL	80
Volume final (mL)	~6,55

- Couler le stacking.
- Insérer immédiatement les peignes appropriés et laisser polymériser au moins 15 à 20 minutes.
- Déposer avec précaution 100 μL de chaque culot de nos échantillons dans son puits approprié.
- Déposer la solution de migration dans la chambre extérieure et intérieure de la cuve.
- Ajouter 200 μL de solution β mercaptoéthanol dans la solution de migration.
- Brancher et allumer le générateur à 80V pendant 30min, puis 145V pendant 3 heures.

b- Fixation des gels :Après trois heures de migration électrophorétique, les gels sont alors récupérés de la cuve et placés dans un bac contenant une solution de fixation (40% alcool éthylique, 10% acide acétique) pendant 10 min.

c- Préparation des solutions pour une coloration au nitrate d'argent (AgNO_3) : Après fixation, les gels seront colorés au nitrate d'argent. Pour cela, une préparation des solutions a été faite selon le protocole suivant:

a- Solution de fixation : Acide acétique 10% 10 mL Ethanol 40% 40 mL H ₂ O 50 mL	c- Solution de coloration au nitrate d'argent : Nitrate d'argent 0,25% 0,125 g + 5 mL H ₂ O H ₂ O 45 mL
b- Solution de sensibilisation pour 50 mL : Ethanol 30% 15 mL Sodium thiosulfate 0.2 % 0,1 g + 2mL H ₂ O Sodium acétate 6,8 % 3,4 g H ₂ O 33mL	d- Solution de développement : Sodium carbonate 2,5 % 1,25 g H ₂ O 50 mL Formaldéhyde 0,04 % 20µL
e- Solution stop: EDTA 1,46 % 0,73 g H ₂ O 50 mL	

d- Coloration au nitrate d'argent : Adapté aux petits gels (50mL). Notre gel a été mis successivement dans plusieurs solutions selon le protocole de Yan *et al.*, (2000).

<u>a- Fixation :</u> - 15 min dans la solution de fixation, jeter la solution. - 15 min dans une nouvelle solution de fixation.	<u>e- Lavage :</u> - 1 min dans l'eau distillée, jeter l'eau. - 1 min dans l'eau distillée.
<u>b- Sensibilisation :</u> - 30 min dans la solution de sensibilisation.	<u>f- Développement:</u> - Mettre le gel dans la solution de développement jusqu'à l'apparition des bandes protéiques.
<u>c- Lavage:</u> - 5 min dans l'eau distillée, jeter l'eau. - 5 min dans l'eau distillée, jeter l'eau. - 5 min dans l'eau distillé.	<u>g- Stop:</u> - Mettre le gel dans la solution stop pendant 10 min.
<u>d- Coloration au nitrate d'argent :</u> - 20 min dans la solution nitrate d'argent.	<u>h- Lavage :</u> - 5 min dans l'eau distillée, jeter l'eau. - 5 min dans l'eau distillée, jeter l'eau. - 5 min dans l'eau distillé.

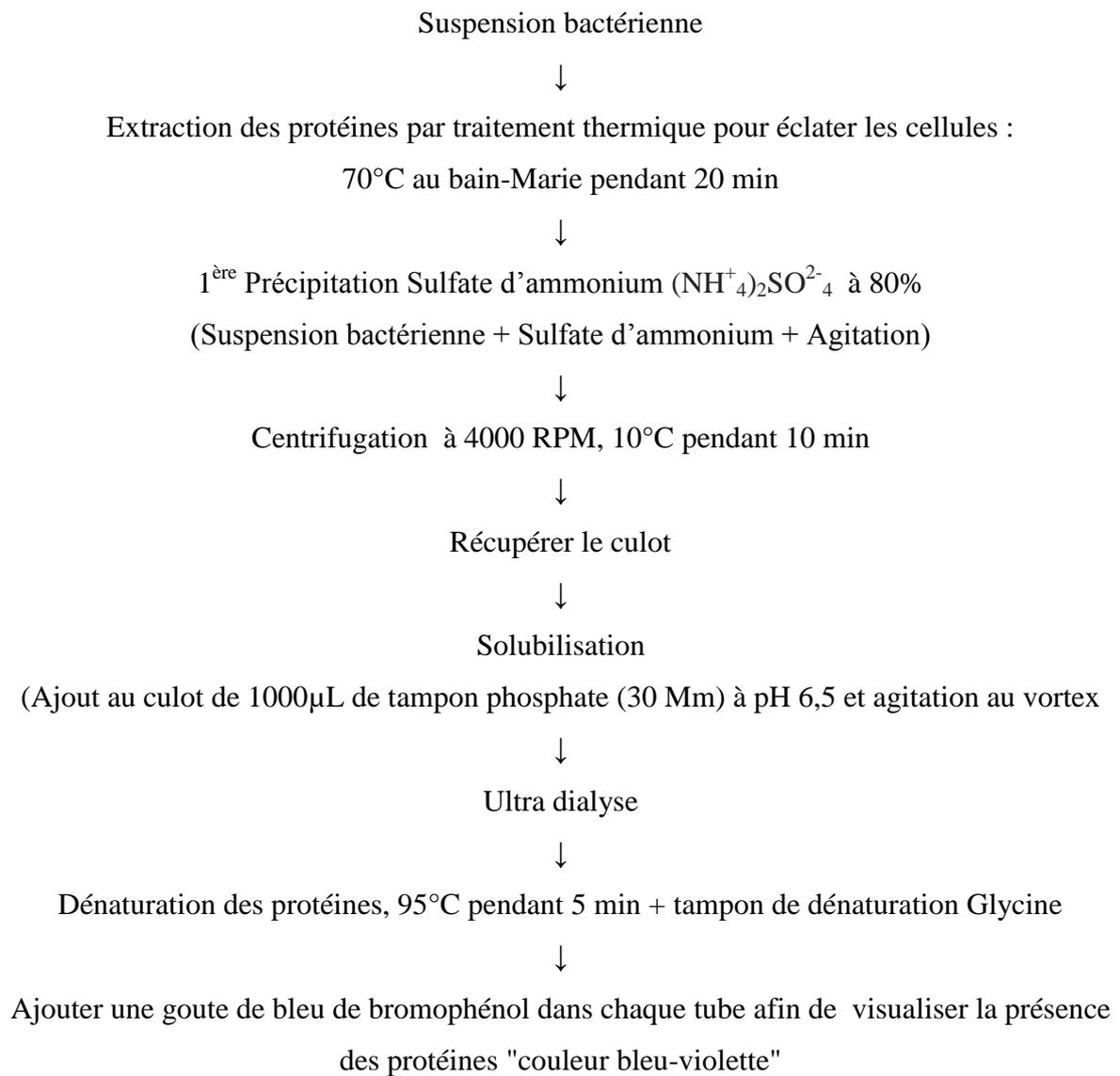


Figure 10: Etapes d'extraction et de précipitation des protéines totales des neuf isolats (Gagaoua, 2016)

3- Analyse biomoléculaire par séquençage rRNA 16S

3-1 Préparation des échantillons pour PCR et séquençage

Tous les isolats ont été traités (figure 7. Communication personnelle, Dr Kerkoud M. Responsable Scientifique DIAG-GENE, Angers Technopole), avant l'amplification et le séquençage, dans le but de détruire les cellules bactériennes et d'en extraire le génome total :

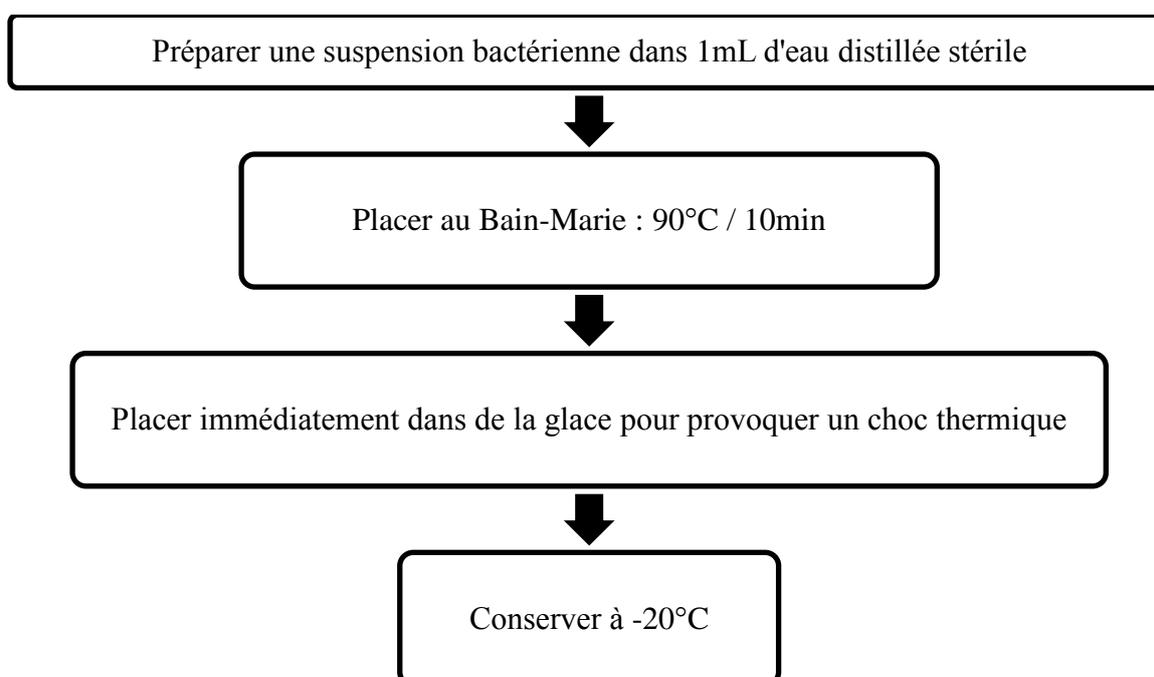


Figure 11 : Etapes préliminaires pour extraction de l'ADN bactérien

Les 24 isolats ont été envoyés au laboratoire d'Angers Technopole pour séquençage du gène rRNA 16S. L'amorce universelle employée pour tous les isolats est la 16S-27f : AGA GTT TGA TCM TGG CTC AG (M étant souvent une Cytosine). C'est une amorce spécifique aux séquences moyennement dépassant les 1466 pb. Son amorce inverse est la 1492R : TAC GGY TAC CTT GTT ACG ACT T (Y étant souvent une Cytosine).

Les séquences 16 rRNA de 17 isolats ont été alignées sur NCBI pour vérifier leur homologie avec les séquences de GenBank. L'approche employée est un BLAST contre la base de données 16S ribosomal RNA séquences (Bacteria and Archaea). Nous notons à ce stade que six isolats non pas donné de résultats PCR qui permettraient un séquençage, éventuellement à cause de la dégradation de leurs ADN.

3-2 Prédiction des structures 2d des rRNA 16s

Chacune des 17 séquences 16S rRNA a fait l'objet d'une prédiction de structure 2d par la méthode du minimum d'énergie libre (ΔG°). L'outil de prédiction structurale utilisé est le RNA Fold du package Vienna RNA Web Server (<http://rna.tbi.univie.ac.at/>). Le choix est porté sur cet outil car il est parmi les rares logiciels à prédire la structure 2d sur la totalité de la séquence. En effet, beaucoup de programmes bioinformatiques font leurs calculs prédictionnels sur une portion plus courte de nucléotides. Par exemple, le logiciel MFold ; pourtant un des plus anciens logiciels de bioinformatique et de biologie moléculaire (Introduit en 1995 à l'école de médecine à Washington) permet de prédire des structures 2d à partir de séquences de longueur pouvant atteindre 800 nucléotides ; d'autres n'en permettent que 400.

4- Traitement bioinformatique des séquences 16S rRNA

Le traitement bioinformatique des séquences 16 rRNA a concerné le Blast et la prédiction des structures 2D de ces séquences.

Pour le traitement par Blast, l'analyse bioinformatique des séquences a été faite par confrontation aux séquences de la banque de données Genbank de NCBI par le programme informatique Blast. Celui-ci permet d'aligner notre séquence requête rRNA 16S avec les séquences rRNA 16S qui lui sont homologues. En fin d'alignement, Blast propose un ensemble de séquences rRNA qui ont un taux d'identité (% de ressemblance) total ou presque. Nous attribuons alors un nom systématique à notre isolat en fonction du taux le plus élevé proposé par Genbank. Les paramètres à prendre en considération pour une telle décision sont la valeur de E qui représente la probabilité de proposer une autre séquence de la base de données à la place de la séquence homologue à la nôtre. Plus cette valeur tend vers zéro, plus la séquence proposée par Blast est la bonne. Un autre paramètre aide à la décision. Il s'agit du taux d'identité (%identité) qui représente la correspondance des nucléotides des deux séquences comparées.

La prédiction des structures 2D des séquences 16S rRNA a été réalisée par le programme RNA Folding (version 2.3 énergies) du serveur Web MFold hébergé à l'Université de New York, au RNA Institute. Ce programme estime plusieurs possibilités de structures 2D pour une séquence rRNA 16S donnée et propose la meilleure prédiction en fonction de la meilleure énergie libre ΔG° synonyme d'une meilleure stabilité conformationnelle de cette structure 2D.

RESULTATS ET DISCUSSION

A- RESULTATS : Trois types de résultats ont été obtenus durant ce travail :

- Caractérisation phénotypique basée sur quelques caractères morphologiques, biochimiques et physiologiques,
- Analyse du profil électrophorétique des protéines totales pour neuf isolats,
- Prédiction des structures 2D des séquences rRNA 16s et Blasting sur NCBI.

1- Caractérisation phénotypique des isolats : Les résultats obtenus concernent deux types de sols des fermes 1 (sans engrais) et 2 (avec engrais). Au point de vue morphologique, les résultats obtenus pour les isolats des deux fermes sont rapportés sur les tableaux 5 et 6.

Tableau 5 : Résultats de la caractérisation morphologique des colonies des isolats du sol de la ferme 1 (sans engrais)

	Forme	Couleur	Surface
AV-4S1	Circulaire, Contour irrégulier, Plate	Beige, Translucide	Sèche
AV-4S2	Circulaire, Contour régulier, Convexe	Beige, Translucide	Lisse, Brillante
AV-4S3	Circulaire, Contour régulier, Convexe	Beige, Translucide	Lisse, Brillante
AV-4S4	Circulaire, Contour régulier, Convexe	Beige, Translucide	Sèche
AV-4S5	Circulaire, Contour régulier, Convexe	Beige, Translucide	Lisse, Brillante
AV-5S1	Circulaire, Contour irrégulier, Plate	Beige, Translucide	Sèche
AV-5S2	Circulaire, Contour régulier, Convexe	Beige, Translucide	Lisse, Brillante
AV-5S3	Circulaire, Contour régulier, Convexe	Beige, Translucide	Lisse, Brillante
AV-5S4	Circulaire, Contour régulier, Convexe	Beige, Translucide	Lisse, Brillante
AV-6S1	Fusifforme Contour régulier Convexe	Beige, Translucide	Lisse, Brillante
AV-6S2	Circulaire, Contour irrégulier, Plate	Beige, Translucide	Sèche
AV-6S3	Circulaire, Contour régulier, Convexe	Beige, Translucide	Lisse, Brillante

Tableau 6 : Résultats de la caractérisation morphologique des colonies des isolats du sol de la ferme 2 (avec engrais)

	FORME	COULEUR	SURFACE
AP-4S1	Circulaire, Contour régulier, Convexe	Blanchâtre, Opaque	Lisse, Brillante
AP-4S2	Circulaire, Contour régulier, Convexe	Beige, Translucide	Lisse, Brillante
AP-4S3	Circulaire, Contour régulier, Convexe	Blanchâtre, Opaque	Lisse, Brillante
AP-4S4	Circulaire, Contour régulier, Convexe	Beige, Translucide	Lisse, Brillante
AP-5S1	Circulaire, Contour régulier, Convexe	Beige, Opaque	Lisse, Brillante
AP-5S2	Circulaire, Contour régulier, Convexe	Beige, Translucide	Lisse, Brillante
AP-5S3	Circulaire, Contour régulier, Convexe	Blanchâtre, Translucide	Lisse, Brillante
AP-5S4	Circulaire, Contour régulier, Convexe	Beige, Translucide	Lisse, Brillante
AP-6S1	Circulaire, Contour régulier, Convexe	Beige, Translucide	Lisse, Brillante
AP-6S2	Circulaire, Contour régulier, Convexe	Beige, Translucide	Lisse, Brillante
AP-6S3	Circulaire, Contour régulier, Convexe	Blanchâtre, Translucide	Lisse, Brillante
AP-6S3	Circulaire, Contour régulier, Convexe	Beige, Translucide	Lisse, Brillante

2- Caractérisation métabolique des isolats : Les résultats obtenus sur les tests biochimiques pour les isolats des deux fermes 1 et 2 sont portés sur les tableaux 7 et 8.

Tableau 7 : Caractères biochimiques des isolats de la ferme 1 (sans engrais)

Isolats	MTL	GAZ	LCT	RM	VP	URS	NDL	CTR	ONPG	LDC	ODC	ADH	H ₂ S	TDA
AV-4 S ₁	+	-	-	+	-	-	-	-	-	-	+	+	-	-
AV-4 S ₂	+	-	-	+	-	-	-	+	+	-	+	+	-	-
AV-4 S ₃	+	+	-	+	-	-	-	-	-	-	+	+	-	-
AV-4 S ₄	+	+	+	+	-	-	-	+	+	+	-	-	-	-
AV-4 S ₅	+	-	-	+	-	-	-	-	-	-	+	+	-	-
AV-5 S ₁	+	-	-	+	-	-	-	-	+	-	+	+	-	-
AV-5 S ₂	+	-	+	+	-	-	-	-	-	-	+	+	-	-
AV-5 S ₃	+	-	-	+	-	-	-	-	+	-	+	+	-	-
AV-5 S ₄	-	-	-	+	-	-	-	-	+	+	+	+	-	-
AV-6 S ₁	-	-	-	+	-	-	-	-	+	-	+	+	-	-
AV-6 S ₂	+	-	-	+	-	-	-	-	+	+	+	+	-	-
AV-6 S ₃	+	-	-	+	-	-	-	-	-	-	+	+	-	-

(+) : présence du caractère étudié ; (-) : absence du caractère étudié

Tableau 8 : Caractères biochimiques des isolats de la ferme 2 (avec engrais)

Isolats	MTL	GAZ	LCT	RM	VP	URS	NDL	CTR	ONPG	LDC	ODC	ADH	H ₂ S	TDA
AP-4 S ₁	+	-	-	+	-	-	-	-	-	+	+	+	-	-
AP-4 S ₂	+	-	-	+	-	+	-	-	+	-	+	+	-	-
AP-4 S ₃	+	+	+	+	-	-	-	+	+	+	+	+	-	-
AP-4 S ₄	+	+	+	+	-	-	-	+	+	+	-	+	-	-
AP-4 S ₅	+	-	-	+	-	-	-	-	-	-	+	+	-	-
AP-5 S ₁	+	-	+	+	-	-	-	-	-	+	+	+	-	-
AP-5 S ₂	+	-	-	+	-	+	-	-	+	+	+	+	-	-
AP-5 S ₃	-	-	-	+	-	-	-	-	+	+	+	+	-	-
AP-5 S ₄	+	+	+	+	-	-	-	+	+	-	+	+	-	-
AP-6 S ₁	+	-	-	+	-	+	-	-	-	+	+	+	-	-
AP-6 S ₂	+	-	-	+	-	-	-	-	+	+	+	+	-	-
AP-6 S ₃	+	-	-	+	-	-	-	-	-	-	+	+	-	-

(+) : présence du caractère étudié ; (-) : absence du caractère étudié

3- Caractérisation physiologique des isolats : Les résultats obtenus sur les tests physiologiques pour les isolats des deux fermes 1 et 2 sont portés sur les tableaux 9 et 10.

Tableau 9 : Caractères physiologiques des isolats de la ferme 1 (sans engrais)

Isolats	CTL	XDS	MEVAG H	MEVAG SH	VF	GAZ VF	T45 ⁰ C	NR
AV-4 S ₁	+	-	+	+	AAF	-	+	+
AV-4 S ₂	+	-	NF	NF	AAF	+	-	+
AV-4 S ₃	+	+	-	-	AAF	+	-	+
AV-4 S ₄	+	-	-	-	AAF	-	+	+
AV-4 S ₅	+	-	NF	NF	AAF	++	-	+
AV-5 S ₁	+	+	NF	NF	AAF	++	-	+
AV-5 S ₂	+	+	NF	NF	AAF	+	-	+
AV-5 S ₃	+	+	NF	NF	AAF	++	-	+
AV-5 S ₄	+	-	NF	NF	AAF	++	-	+
AV-6 S ₁	+	-	NF	NF	AAF	++	-	+
AV-6 S ₂	+	-	NF	NF	AAF	++	+	+
AV-6 S ₃	+	+	NF	NF	AAF	++	-	+

(+): présence du caractère étudié; (-): absence du caractère étudié; (++) : Caractère étudié fortement présent (NF) : test non fait

Tableau 10 : Caractères physiologiques des isolats de la ferme 2 (avec engrais)

Isolats	CTL	XDS	MEVAG H	MEVAG SH	VF	GAZ VF	T45 ⁰ C	NR
AP-4 S ₁	+	+	+	+	AAF	-	+	+
AP-4 S ₂	+	+	NF	NF	AAF	++	+	+
AP-4 S ₃	+	-	+	+	AAF	++	+	+
AP-4 S ₄	-	+	+	+	AAF	++	-	+
AP-4 S ₅	+	-	NF	NF	AAF	+	-	+
AP-5 S ₁	+	-	NF	NF	AAF	++	+	+
AP-5 S ₂	+	+	NF	NF	AAF	++	-	+
AP-5 S ₃	+	-	NF	NF	AAF	++	-	+
AP-5 S ₄	+	-	NF	NF	AAF	+	-	+
AP-6 S ₁	+	-	NF	NF	AAF	++	-	+
AP-6 S ₂	+	+	NF	NF	AAF	++	-	+
AP-6 S ₃	+	-	NF	NF	AAF	++	-	+

(+): présence du caractère étudié; (-): absence du caractère étudié; (++) : Caractère étudié fortement présent (NF) : test non fait

4-Analyse du profil électrophorétique des protéines totales pour neuf isolats

La migration électrophorétique des protéines totales est représentée sur la figure 12. Certains isolats ont montré un meilleur profil ; mais insuffisant pour réaliser une taxinomie numérique.

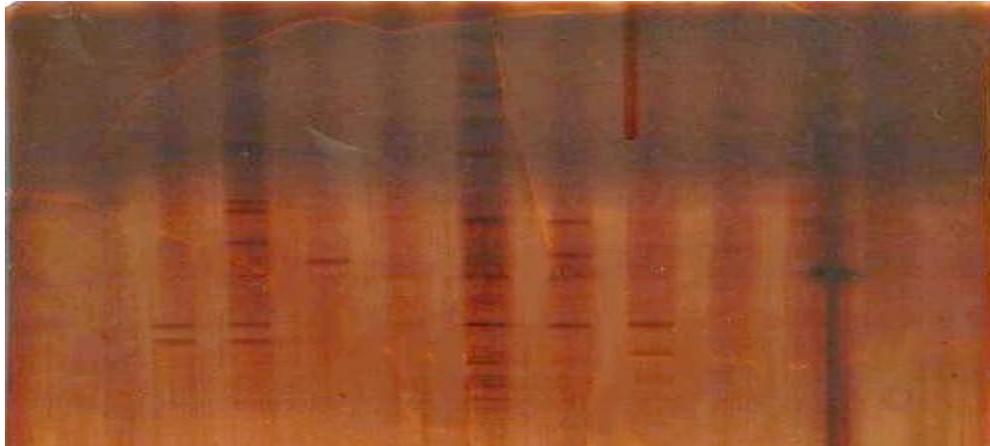


Figure 12 : Gel d'électrophorèse de type SDS-page

Tableau 11 : Fantôme du gel d'échantillons bactériens représente l'ordre de dépôt des isolats dans les puits du gel.

Puits	1	2	3	4	5	6	7	8	9	10	11	12
Echantillons	Vide	AV-4 S ₅	AP-5 S ₁	AV-6 S ₃	AP-4 S ₂	AP-5 S ₂	AV-6 S ₁	AV-4 S ₃	AV-5 S ₃	AP-4 S ₁	Vide	Vide
Quantité		40µL										

5- Traitement bioinformatique des séquences rRNA 16S

5-1 Exploitation bioinformatique des séquences rRNA 16S : Le Blasting réalisé sur le portail NCBI a donné les résultats montrés sur le tableau suivant :

Tableau 12 : Résultats du Blast pour les isolats des deux fermes

	Isolats	Définition	Accession	Identité %	E-value
1	AV-4S1	<i>Staphylococcus haemolyticu sstrain B-16</i>	KC139451.1	99	0.0
2	AV-4S2	<i>Bacterium N10 16S ribosomal</i>	GQ389785.1	92	0.0
3	AV-4S3	<i>Acinetobacter sp. VZ4C</i>	JQ618100.1	77	5e ⁻⁹²
4	AV-4S5	<i>Staphylococcus haemolyticus strain B-16</i>	KC139451.1	79	0.0
5	AV-5S4	<i>Staphylococcus aureus strain RCB1010</i>	KT261222.1	96	0.0
6	AV-6S1	<i>Acinetobacter lwoffii strain MR-18-2</i>	KF848508.1	90	0.0
7	AV-6S2	<i>Acinetobacter lwoffii strain 272XG8</i>	KF818627.1	94	0.0
8	AV-6S3	<i>Arthrobacter sp. 48 TDS</i>	EU487613.1	83	1e ⁻¹⁴⁸
9	AP-4S3	<i>Staphylococcus hominis strain +Y112</i>	JX077103.1	99	0.0
10	AP-4S4	<i>Bacillus drentensis strain BF11</i>	KJ524509.1	98	0.0
11	AP-5S1	<i>Bacterium L35 (non identifiée)</i>	KU221401.1	99	0.0
12	AP-5S2	<i>Acinetobacter lwoffii strain 266XY3</i>	KF818657.1	95	0.0
13	AP-5S3	<i>Sanguibacter soli. CW-44</i>	JX997904.1	90	0.0
14	AP-5S4	<i>Acinetobacter lwoffii strain 265ZG8</i>	KF831409.1	94	0.0
15	AP-6S1	<i>Bacillus sp. QW52</i>	KF737385.1	97	0.0
16	AP-6S2	<i>Mobilitalea sibiricaP3M-3</i>	NR134091.1	100	4.8
17	AP-6S4	<i>Staphylococcus sp. C9I1</i>	HQ717320.1	73	2e ⁻⁵⁶

Les résultats de Blast (voir Annexe) ont montré une variabilité systématique de nos isolats. Ainsi, dans le sol de la ferme 1 (sans ajout d'engrais), sur les huit isolats, nous avons identifié trois genres différents :

- *Staphylococcus* : Trois isolats
- *Acinetobacter* : Trois isolats
- *Arthrobacter* : Un isolat

Le sol de la ferme 2 (avec engrais) présente une plus grande variabilité bactérienne. En effet, sur un total de neuf isolats, nous avons isolés cinq genres :

- *Staphylococcus* : Deux isolats
- *Bacillus* : Deux isolats
- *Acinetobacter* : Deux isolats
- *Sanguibacter* : Un isolat
- *Mobilitalea* : Un isolat

Ce type de résultats a été observé par Svirskene (2003), où il a constaté un changement de la population bactérienne (y compris le genre *Azotobacter*) qui s'est manifesté par un changement des statuts pédologique et microbiologique suite à l'ajout de fertilisants. De plus, Gangwar (1992), en étudiant l'effet de fertilisants sur des mûraies, a constaté une activité microbiologique plus importante se manifestant par un rendement remarquable et une bonne qualité des feuilles.

L'isolat AP-6S2 (ferme 2, engrais), vue sa très courte séquence de 236 pb, est difficile à identifier par Blast. L'organisme renvoyé par Genbank (*Mobilitalea sibirica* P3M-3) peut être totalement différent de notre souche, car E-value est trop élevée (E= 4,8). La comparaison entre notre isolat (236pb) et *Mobilitalea sibirica* P3M-3 (1444 pb) n'a concerné qu'une taille de 16 nucléotides ; ce qui est insuffisant pour apporter une bonne appréciation du résultat de ce Blast. La séquence de notre isolat est, soit à refaire pour confirmation, soit il s'agit d'une souche non encore répertoriée dans les bases de données. Si la répétition du séquençage confirme cette dernière hypothèse, notre isolat aura à subir les techniques des systématiques polyphasiques pour l'identifier définitivement et l'ajouter au catalogue bactérien.

Dans le cas de l'isolat AP-5S1, Blast a rendu une séquence similaire à la nôtre (E=0,0), mais les auteurs de cette séquence qui est homologue à la nôtre, n'ont pas encore identifié leur souche. Ceci rend compte de la possibilité que l'isolat AP-5S1 soit nouveau et sur lequel une polyphasique doit être appliquée.

L'isolat AV-4S2 (ferme 1 sans engrais) est en situation identique à celle de l'isolat AP-5S1. Il faut pousser l'identification par une approche polyphasique.

Nous pouvons regrouper ces données phylogénétiques (Tableau 13) en fonction de l'emplacement systématique de chacun des isolats.

Tableau 13 : Synthèse de la place systématique des différents isolats

	Domaine	Phyla	Classe	Ordre	Famille	Genre	Espèce
1	Bacteria	Firmicutes	Bacilli	Bacillales	Staphylococcaceae	<i>Staphylococcus</i>	<i>S. haemolyticus</i> <i>S. aureus</i> <i>S. hominis</i>
2	Bacteria	Firmicutes	Bacilli	Bacillales	Bacillaceae	<i>Bacillus</i>	<i>B. soli</i>
3	Bacteria	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae	<i>Mobilitalea</i>	<i>M. sibirica</i>
4	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae	<i>Acinetobacter</i>	<i>A. Lwoffii</i>
5	Bacteria	Actinobacteria	Actinobacteria	Micrococcales	Sanguibacteraceae	<i>Sanguibacter</i>	<i>S. soli</i>
6	Bacteria	Actinobacteria	Actinobacteria	Micrococcales	Micrococcaceae	<i>Arthrobacter</i>	<i>Sp.</i>

Il ressort que sur le total des 17 souches identifiées par BLAST, le phylum des Firmicutes est plus présent avec 50% de distribution, suivi des Actinobacteria avec une fréquence de 33,33% et le phylum des Protéobacteria est représenté dans 16,7%.

D'après le phylogramme obtenu (Figure N°13), nous constatons que plusieurs clades sont homogènes et regroupent des espèces bien apparentées. La souche de *Mobilitalea sibirica* fait office du groupe externe (Out Group) car elle occupe la position d'une racine et elle ne forme aucun clade. Ceci est due au fait que la taille de la séquence alignée est nettement inférieure.

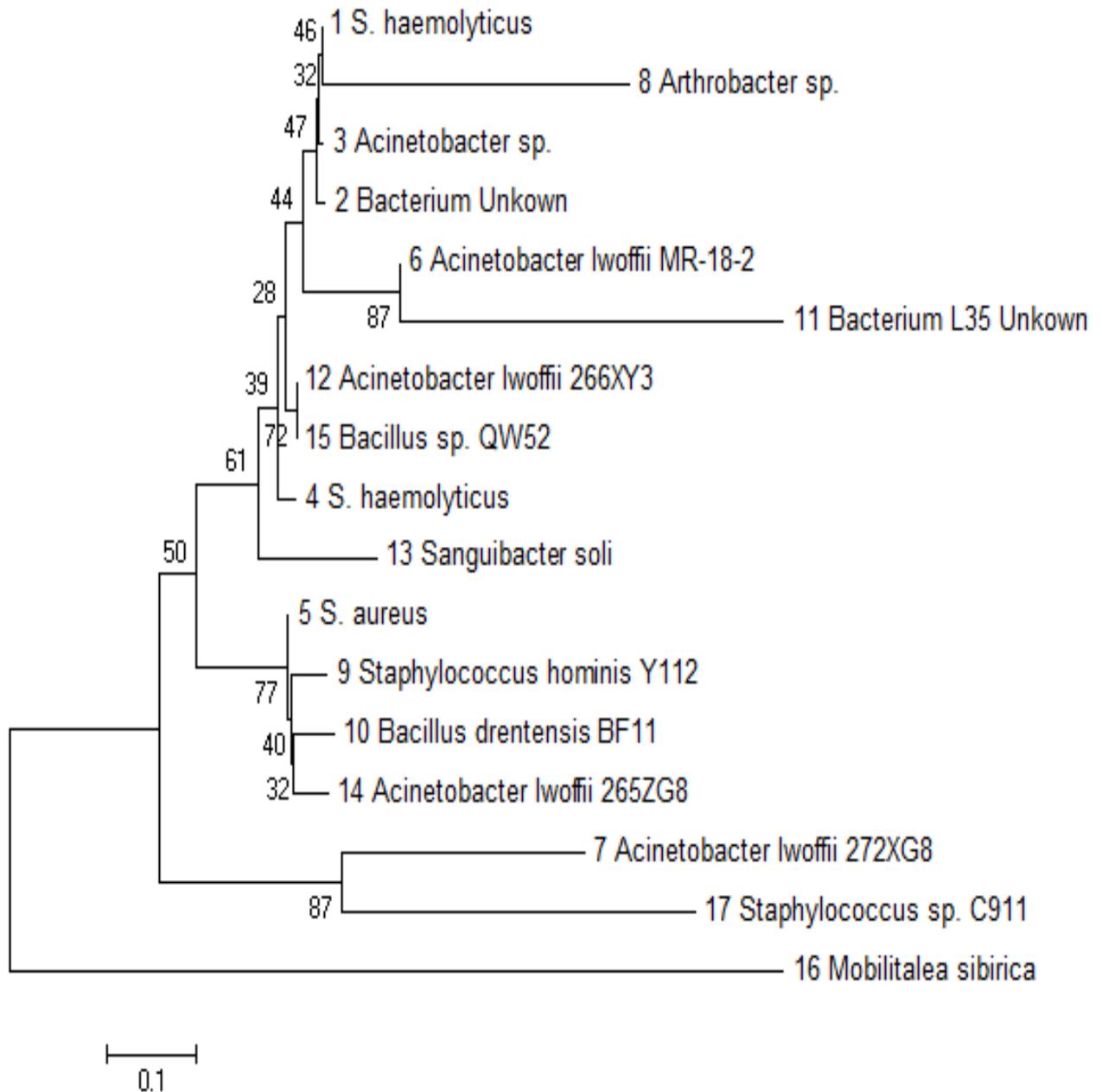
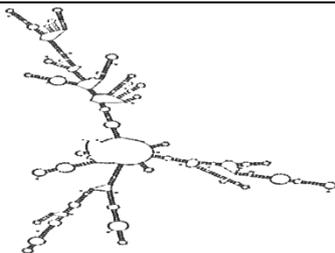
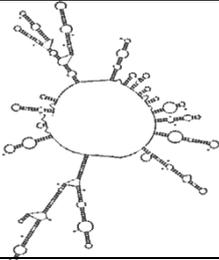
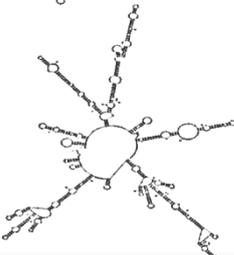
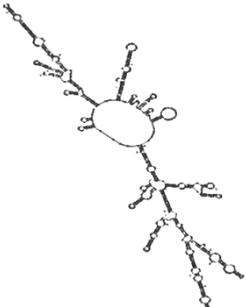


Figure 13 :Phylogramme rRNA 16S des 17 souches.

5-2 Prédiction des structures 2D

Les résultats des prédictions 2D des séquences rRNA 16S sont représentés sur le tableau N°11 résumant les énergies libres ΔG° des différentes structures 2D estimées par les calculs thermodynamiques.

Tableau 14 : Energies libres ΔG° des différentes structures 2D des rRNA 16S pour les isolats de la ferme 1 (sans engrais)

	Structure 2D de l'ARN 16S	Energies
AV-4S1 <i>Staphylococcus haemolyticus strain B-16</i>		$\Delta G = -288,9$ Kcal/mol $\Delta H = -2361,40$ Kcal/mol $\Delta S = -6846,3$ cal/(K.mol) $T_m = 71,7^\circ\text{C}$
AV-4S2 <i>Bacterium N10</i>		$\Delta G = -240,80$ Kcal/mol $\Delta H = -2233,40$ Kcal/mol $\Delta S = -6424,6$ cal/(K.mol) $T_m = 74,4^\circ\text{C}$
AV-4S3 <i>Acinetobacter sp. VZ4C</i>		$\Delta G = -222,90$ Kcal/mol $\Delta H = -2334,00$ Kcal/mol $\Delta S = -6806,7$ cal/(K.mol) $T_m = 69,7^\circ\text{C}$
AV-4S5 <i>Staphylococcus haemolyticus strain B-16</i>		$\Delta G = -212,6$ Kcal/mol $\Delta H = -2194,00$ Kcal/mol $\Delta S = -6388,5$ cal/(K.mol) $T_m = 70,2^\circ\text{C}$

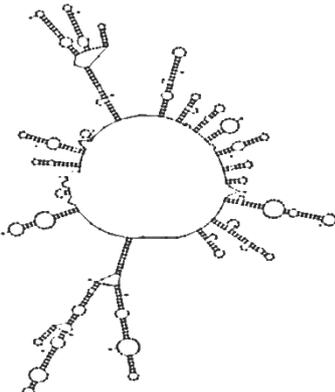
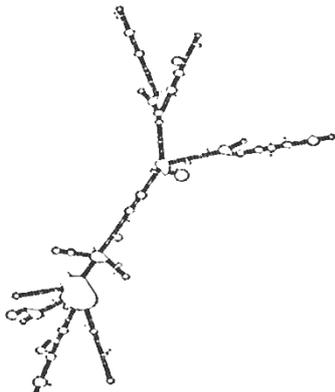
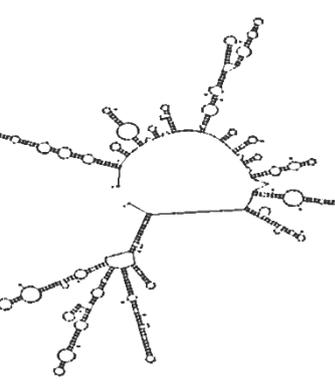
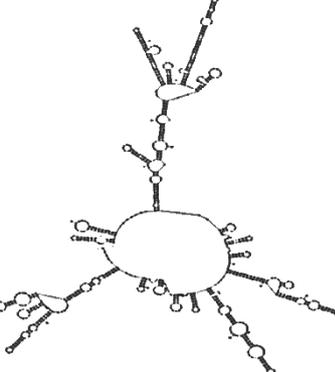
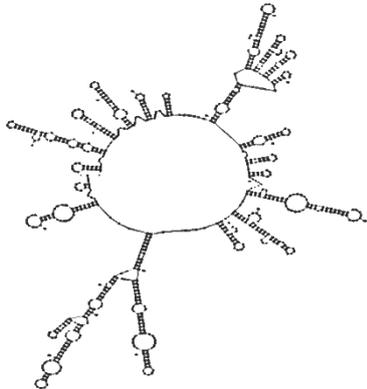
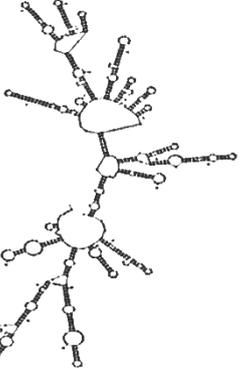
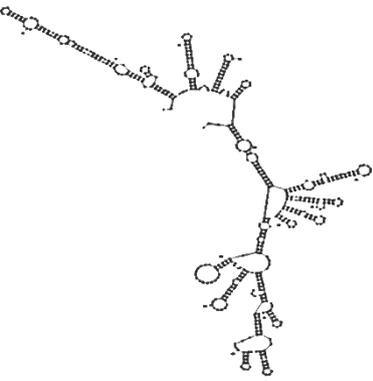
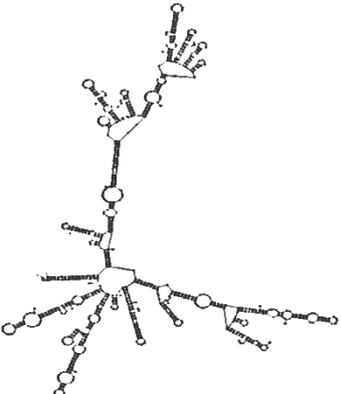
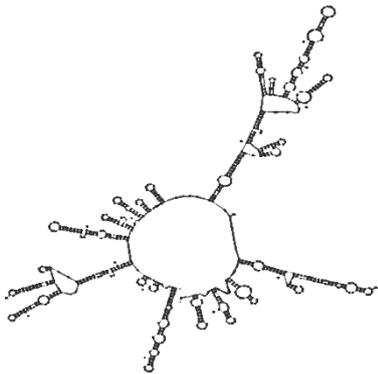
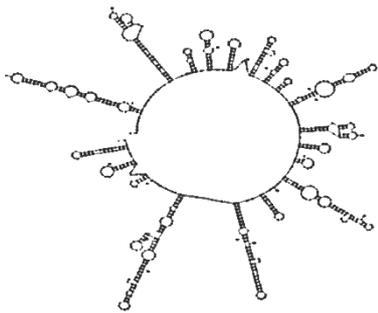
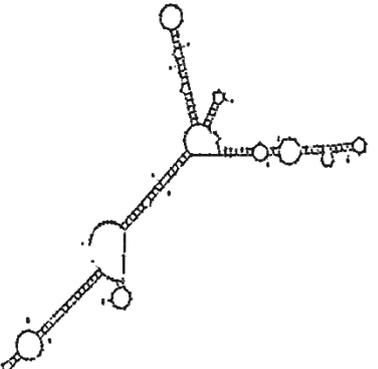
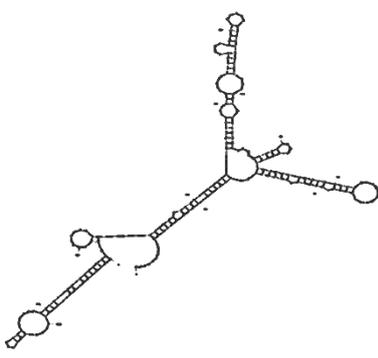
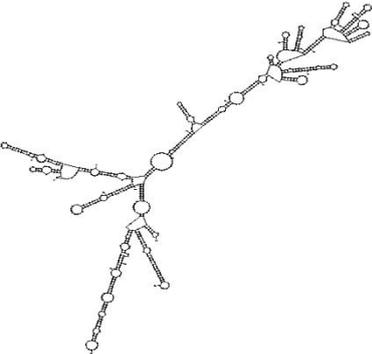
AV-5S4 <i>Staphylococcus aureus</i> strain RCB1010	 <p>The diagram shows a central circular core with numerous short, branched chains extending outwards. Each chain consists of small spheres (atoms) connected by lines (bonds), representing a complex, multi-branching molecular structure.</p>	$\Delta G = -241,40 \text{ Kcal/mol}$ $\Delta H = -2260,90 \text{ Kcal/mol}$ $\Delta S = -6511,3 \text{ cal/(K.mol)}$ $T_m = 74^\circ\text{C}$
AV-6S1 <i>Acinetobacter lwoffii</i> strain MR-18-2	 <p>The diagram shows a central circular core with several long, thin, branched chains extending outwards. The chains are composed of small spheres and connecting lines, forming a sparse, tree-like structure.</p>	$\Delta G = -228,10 \text{ Kcal/mol}$ $\Delta H = -2377,90 \text{ Kcal/mol}$ $\Delta S = -6931,4 \text{ cal/(K.mol)}$ $T_m = 69,9^\circ\text{C}$
AV-6S2 <i>Acinetobacter lwoffii</i> strain 272XG8	 <p>The diagram shows a central circular core with many short, branched chains extending outwards. The chains are composed of small spheres and connecting lines, forming a dense, multi-branching molecular structure.</p>	$\Delta G = -242,30 \text{ Kcal/mol}$ $\Delta H = -2291,30 \text{ Kcal/mol}$ $\Delta S = -6606,4 \text{ cal/(K.mol)}$ $T_m = 73,6^\circ\text{C}$
AV-6S3 <i>Arthrobacter sp.</i> 48 TDS	 <p>The diagram shows a central circular core with several long, thin, branched chains extending outwards. The chains are composed of small spheres and connecting lines, forming a sparse, tree-like structure.</p>	$\Delta G = -269,00 \text{ Kcal/mol}$ $\Delta H = -2379,60 \text{ Kcal/mol}$ $\Delta S = -6805 \text{ cal/(K.mol)}$ $T_m = 76,5^\circ\text{C}$

Tableau 15 : Energies libres ΔG° des différentes structures 2D des rRNA 16S pour les isolats de la ferme 2 (avec engrais)

	Structure 2D de l'ARN 16S	Energie
AP-4S3 <i>Staphylococcus hominis</i> strain +Y112		$\Delta G = -241,50 \text{ Kcal/mol}$ $\Delta H = -2278,30 \text{ Kcal/mol}$ $\Delta S = -6567,1 \text{ cal/(K.mol)}$ $T_m = 73,7^\circ\text{C}$
AP-4S4 <i>Bacillus drentensis</i> strain BF11		$\Delta G = -286,00 \text{ Kcal/mol}$ $\Delta H = -2641,90 \text{ Kcal/mol}$ $\Delta S = -7596 \text{ cal/(K.mol)}$ $T_m = 74,6^\circ\text{C}$
AP-5S1 <i>Bacterium</i> L35		$\Delta G = -191,10 \text{ Kcal/mol}$ $\Delta H = -1686,60 \text{ Kcal/mol}$ $\Delta S = -4821,8 \text{ cal/(K.mol)}$ $T_m = 76,6^\circ\text{C}$
AP-5S2 <i>Acinetobacter lwoffii</i> strain 266XY3		$\Delta G = -240,70 \text{ Kcal/mol}$ $\Delta H = -2409,70 \text{ Kcal/mol}$ $\Delta S = -6993,3 \text{ cal/(K.mol)}$ $T_m = 71,4^\circ\text{C}$

<p>AP-5S3</p> <p><i>Sanguibacter sp. CW-44</i></p>	 <p>A complex, branched molecular structure with a central circular core and multiple radiating arms, each ending in a small circular head. The structure is highly detailed with numerous atoms and bonds visible.</p>	<p>$\Delta G = -275,40 \text{ Kcal/mol}$</p> <p>$\Delta H = -2408,00 \text{ Kcal/mol}$</p> <p>$\Delta S = -6876 \text{ cal/(K.mol)}$</p> <p>$T_m = 77^\circ\text{C}$</p>
<p>AP-5S4</p> <p><i>Acinetobacter lwoffii strain 265ZG8</i></p>	 <p>A complex, branched molecular structure with a central circular core and multiple radiating arms, each ending in a small circular head. The structure is highly detailed with numerous atoms and bonds visible.</p>	<p>$\Delta G = -247,90 \text{ Kcal/mol}$</p> <p>$\Delta H = -2228,60 \text{ Kcal/mol}$</p> <p>$\Delta S = -6386,2 \text{ cal/(K.mol)}$</p> <p>$T_m = 75,8^\circ\text{C}$</p>
<p>AP-6S1</p> <p><i>Bacillus sp. QW52</i></p>	 <p>A complex, branched molecular structure with a central circular core and multiple radiating arms, each ending in a small circular head. The structure is highly detailed with numerous atoms and bonds visible.</p>	<p>$\Delta G = -61,30 \text{ Kcal/mol.}$</p> <p>$\Delta H = -635,90 \text{ Kcal/mol.}$</p> <p>$\Delta S = -1852,6 \text{ Cal/(K.mol.)}$</p> <p>$T_m = 70^\circ\text{C}$</p>
<p>AP-6S2</p> <p><i>Mobilitalea sibirica P3M-3</i></p>	 <p>A complex, branched molecular structure with a central circular core and multiple radiating arms, each ending in a small circular head. The structure is highly detailed with numerous atoms and bonds visible.</p>	<p>$\Delta G = -61,30 \text{ Kcal/mol.}$</p> <p>$\Delta H = -635,90 \text{ Kcal/mol}$</p> <p>$\Delta S = -1852,6 \text{ cal/(K.mol)}$</p> <p>$T_m = 70^\circ\text{C}$</p>

<p>AP-6S4</p> <p><i>Staphylococcus sp. C9II</i></p>		<p>$\Delta G = -244,10$ Kcal/mol</p> <p>$\Delta H = -2361,90$ Kcal/mol</p> <p>$\Delta S = -6828,3$ cal/(K.mol)</p> <p>$T_m = 72,7^\circ\text{C}$</p>
---	---	--

La prédiction des structures 2D révèle que les structures ont toutes des énergies libres favorables à des structures secondaires stables du fait que toutes les énergies estimées sont considérablement négatives ; synonyme d'une structure secondaire très stable. La plus faible énergie libre ΔG° est observée chez l'isolat AV-4S1 avec un $\Delta G^\circ = -288,9$ Kcal/mol ; alors que les individus AP-6S1 et AP-6S2 ont exprimé les énergies libres les plus élevées avec $\Delta G^\circ = -61,3$ Kcal/mol.

Les structures 2D des séquences 16S des isolats révèlent des moyennes d'énergies spécifiques à chaque catégorie de sol. En effet, les séquences des germes isolés de la ferme 1 (sans engrais) expriment une énergie libre moyenne ΔG° de $-243,25$ Kcal/mol., alors que le groupe d'individus isolés du sol de la ferme 2 ont une énergie libre supérieure ($\Delta G^\circ = -206,37$ Kcal/mol.).

Les différences enregistrées dans les topologies 2D entre les différentes séquences serait due au programme informatique MFold qui est limité aux prédictions 2D pour les tailles de séquences ne dépassant pas les 800 pb, d'une part, et d'autre part, les tailles des séquences rRNA 16S sont partielles et de longueurs différentes. La prédiction ne se faisant donc pas sur les mêmes portions rRNA.

Selon Zuker (1991), dans la plupart des cas prédits, une structure estimée correcte à 80% est retrouvée avec une perte d'énergie libre ΔG° de 2% de l'énergie libre ΔG° de la plus faible structure prédite. Ceci révèle que la prédiction 2D reste une supposition théorique basée sur des calculs thermodynamiques qui proposent une structure 2D de la séquence qui pourrait exister à l'intérieur de la cellule bactérienne.

CONCLUSION

Ce travail représente une contribution dans un travail de recherche destiné à un Doctorat en écologie microbienne.

Nous avons tenté d'identifier le microbiote du sol par une approche phénotypique supportée par le séquençage du gène 16S rRNA et de prédire les structures 2D de ces séquences.

L'analyse phénotypique reste insuffisante pour pouvoir se prononcer au niveau de l'espèce, cependant certains germes ont pu être identifiés par séquençage 16S.

En tenant compte des résultats obtenus, il ressort que les deux sols des deux fermes 1 et 2 (sans engrais et avec engrais respectivement) ont des genres bactériens en commun tels que *Staphylococcus* et *Acinetobacter*. Cependant le sol de la ferme 2 contient des genres absents dans le sol de la ferme 1 tels que *Sanguibacter* et *Bacillus*.

Il convient également de dire que les Firmicutes occupent une place dominante (50%) dans les sols des deux fermes, ainsi que les Actinobacteria. Les Protéobacteria ne sont pas largement distribués dans les deux sols comparativement aux Firmicutes et aux Actinobacteria.

L'analyse des protéines totales par SDS-PAGE doit être améliorée pour le cas des bactéries afin de mieux mettre en exergue les différences interspécifiques bactériennes.

Les structures 2D sont un outil important à ajouter dans l'identification des bactéries, car les valeurs thermodynamiques sont spécifiques des groupes bactériens.

En perspective, nous préconisons une analyse plus poussée de l'ADN des isolats AV-4S2 (ferme 1 sans engrais), AP-6S2 et AP-5S1 (ferme 2 avec engrais) afin de confirmer ces espèces car elles sont inconnues sur toutes les bases de données génomiques (GenBank, EMBL et DDBJ). Il peut s'agir donc de nouvelles espèces qu'il faut identifier et les insérer dans le catalogue bactérien.

Enfin, ce travail ne constitue pas une confirmation absolue de l'état bactérien des sols des deux fermes, mais apporte une appréciation et une idée générale sur les interactions éventuelles entre les différentes communautés bactériennes.

Références Bibliographiques

1. Clarridge, J. E., (2004). Impact of 16S rRNA Gene Sequence Analysis for Identification of Bacteria on Clinical Microbiology and Infectious Diseases. *Clin Microbiol. Rev.* **17** : 840-862
2. Chun J., Good fellow M., (1995). A phylogenetic analysis of the genus *Nocardia* with 16S rRNA gene sequences. *Int. J. Syst. Bacteriol.* **45** : 240-245
3. Embley T. M., and Stackbrandt E., (1994). The molecular phylogeny and systematics of the actinomycetes. *Annu. Rev. Microbiol.* **48** : 257-289
4. Gagaoua, M., Hoggas, N., Hafid, K. (2015). Three phase partitioning of zingibain, a milk-clotting enzyme from *Zingiber officinale* Roscoe rhizomes. *Int. J. Biol. Macromolecules.* **73**: 245-252
5. Janda, J. M., Abbott, S. L. (2007). 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls. *J. Clin. Microbiol.* **45** : 2761-2764
6. Kataoka M., Ueda K., Kudo T., Seki T., Yoshida T., (1997). Application of the variable region in 16S rDNA to create an index for rapid species identification in the genus *Streptomyces*. *FEMS Microbiol. Lett.* **151** : 249-255
7. Khelili K. (2015). *Application des outils biomoléculaires 16S, MALDI-TOF et métagénomique pour la détermination des variantes taxonomiques bactériennes telluriques*. Thèse Doctorat. Université Frères Mentouri Constantine. 108p.
8. Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227** : 680-685
9. Laurent F.J., Provost F., Boiron P., (1999) Rapid identification of clinically relevant *Nocardia* species to genus level by 16S rRNA gene PCR. *J. Clin. Microbiol.* **37** : 99 -102
10. Olsen, G. J. & Woese, C. R. (1993). Ribosomal RNA: a key to phylogeny. *FASEB J* **7** : 113-123

11. Rossello Mora R. Amann R., (2001). The species concept for prokaryotes. *FEMS Microbiol. Rev.* **25** : 39-67
12. Serra M. J., Turner D. H. (1995). Predicting thermo dynamic properties of RNA. *Methodes Enzymol.***259** : 242-61
13. Stackebrandt E., Goebel B. M., (1994). Taxonomic note : a place for DNA-DNA reassociation and 16S rRNA sequences analysis in the present species definition in bacteriology. *Int. J. Syst. Bacteriol.***44** : 846-849
14. Svirskene A. (2003). Microbiological and biochemical indicators of anthropogenic impacts on soils. *Eurasian Soil Science.* **36** : 192-200
15. Turner D. H., Sugimoto N., Freier S. M. (1980). RNA Structure Prediction *Annual Review of Biophysics and Biophysical Chemistry.***17**: 167-192
16. Weisburg W.G., Barns S.M., Pelletier D.A., Lane D.J. (1991). 16S ribosomal DNA amplification for phylogenetic study. *J. Bacteriol.* **173** : 697-703
17. Woese, C. R. (1987). Bacterial evolution. *Microbiol. Rev.* **51** : 222-270
18. Yan J. X., Harry R. A., Spibey C., Dunn M. J. (2000). Post electrophoretic staining of proteins separated by two-dimensional gel electrophoresis using SYPRO dyes. *Electrophoresis***21** : 3666-3672
19. Zhang Lu., Wang L, Shi Y. Liu Z. Quintana E. T. Goodfellow M. (2003). *Actinomadura catellatispora* sp. nov. and *Actinomador aglanciflava* sp. nov., from a sewage ditch and soil in southern China. *Int. J. Syst. Evol. Microbiol.* **53** : 137-142
20. Zuker M., Jaeger J. A., Turner D. H. (1991). A comparison of optimal and suboptimal RNA secondary structures predicted by free energy minimization with structures determined by phylogenetic comparison. *Nucleic Acids Research.* **19** : 2707 – 2714

Annexe 1 : Séquences des 16S rRNA au format EMBL

1 SQ	Sequence 1396 BP; 368 A; 320 C; 398 G; 310 T; 0 other;						
	GGGATGCGGG	TGGCTATACA	TGCAGTCGAG	CGGACAGATA	AGGAGCTTGC	TCCTTTGACG	60
	TTAGCGGCGG	ACGGGTGAGT	AACACGTGGG	TAACCTACCT	ATAAGACTGG	GATAACTTCG	120
	GGAAACCGGA	GCTAATACCG	GATAATATTT	GGAACCGCAT	GGTTCGATAG	TGAAAGATGG	180
	TTTTGCTATC	ACTTATAGAT	GGACCCGCGC	CGTATTAGCT	AGTTGGTAAG	GTAACGGCTT	240
	ACCAAGGCGA	CGATACGTAG	CCGACCTGAG	AGGGTGATCG	GCCACACTGG	AACTGAGACA	300
	CGGTCCAGAC	TCCTACGGGA	GGCAGCAGTA	GGGAATCTTC	CGCAATGGGC	GAAAACCTGA	360
	CGGAGCAACG	CCGCGTGAGT	GATGAAGGTC	TTCGGATCGT	AAAACCTCTGT	TATTAGGGAA	420
	GAACATACGT	GTAAGTAACT	GTGCACGTCT	TGACGGTACC	TAATCAGAAA	GCCACGGCTA	480
	ACTACGTGCC	AGCAGCCGCG	GTAATACGTA	GGTGGCAAGC	GTTATCCGGA	ATTATTGGGC	540
	GTAAAGCGCG	CGTAGGCGGT	TTTTTAAGTC	TGATGTGAAA	GCCCACGGCT	CAACCGTGGA	600
	GGGTCAATTGG	AAACTGGAAA	ACTTGAGTGC	AGAAGAGGAA	AGTGGAATTC	CATGTGTAGC	660
	GGTGAAATGC	GCAGAGATAT	GGAGGAACAC	CAGTGGCGAA	GGCGACTTTC	TGGTCTGTAA	720
	CTGACGCTGA	TGTGCGAAAG	CGTGGGGATC	AAACAGGATT	AGATACCCTG	GTAGTCCACG	780
	CCGTAAACGA	TGAGTGCTAA	GTGTTAGGGG	GTTTCCGCCC	CTTAGTGCTG	CAGCTAACGC	840
	ATTAAGCACT	CCGCCTGGGG	AGTACGACCG	CAAGGTTGAA	ACTCAAAGGA	ATTGACGGGG	900
	ACCCGCACAA	GCGGTGGAGC	ATGTGGTTTA	ATTCAAGCA	ACGCGAAGAA	CCTTACCAAA	960
	TCTTGACATC	CTTTGACAAC	TCTAGAGATA	GAGCTTTCCC	CTTCGGGGGA	CAAAGTGACA	1020
	GGTGGTGCAT	GGTTGTTCGT	AGCTCGTGTG	GTGAGATGTT	GGGTAAAGTC	CCGCAACGAG	1080
	CGAACCCCTT	AAGCTTAGTT	GCCATCATT	AGTTGGGCAC	TCTAAGTTGA	CTGCCGGTGA	1140
	CAAACCGGAA	GAAAGGTGGG	GATGACGTC	AATCATCATG	CCCCTTATGA	TTTGGGCTAC	1200
	CCCCGTGCTA	CAATGGACAA	TACAAAGGGC	ACCTAAACCG	CGAGGTCAAG	CAAAATCCAT	1260
	AAAGTTGTTT	CCATTTCCGG	TTTGGATTCC	GCAACCCTCC	AACTTGGATC	TTGGAACCCC	1320
	TGTTAACCTT	GAAACCCCTG	GCTCTCTGGA	GAACGTTTCC	AGGACCTTGT	ACACCCCGAC	1380
	TATCCTCCAC	CCTAAG					1396
2 SQ	Sequence 1193 BP; 323 A; 278 C; 333 G; 259 T; 0 other;						
	TTGCAGCGGC	AGCTATACAT	GCAGTCGAGC	GAACAGATTA	GGAGCTTGCT	CCTTTGACGT	60
	TAGCGGCGGA	CGGGTGAGTA	ACACGTGGAT	AACCTACCTG	TGACACTGGT	ATATCTCCGG	120
	GAAACCGGAG	CTAATGCCGG	ATATCTTTTC	TAACCGCGTG	GTTCAATAGT	GAAAGACGGT	180
	TTTGCTTTCA	CTTATAAATG	GACCCGCGCC	GTATTAGCTA	GTTGGGAGGG	TAACGGCTTA	240
	CCAAGGCAAC	GATCCGTACC	CAACCTGAAA	GGGGGATCGG	CCCCCTGGA	ACTGAAACCC	300
	GGTCCAAACT	CCTACGGGAG	GCAGCAGTAG	GGAATCTTCC	GCAAGGGGCA	AAAGCCTGAC	360
	GGACCACCGC	CCCGGGAGTG	ATGAAGGTCT	TGCAATCGTA	AAACTCTGTT	ATTAGGGAAA	420
	AACAAATGTG	TAAGTAACTA	TGCACGTCTT	GACGGTACCT	AATCAAAAAG	CCCCGGCTAA	480
	CTACTTGCCA	CCACCCGCGG	TAATACGTAG	GGGGCAAGCG	TTATCCGGA	TTATTGGGCG	540
	TAAAGCGCGC	GTAGGCGGTT	TTTTAAGTCT	GATGTGAAAG	CCCACGGCTC	AACCGTGGAG	600
	GGTCATTGGA	AACGTGAAAA	CTTGAGTGCA	GAAGAGAAAA	GTGGAATTCC	ATGTGTAGCG	660
	GTGAAATGCG	CAGAGATATG	GAGGACCACC	AGTGGCAAAG	CCGACTTTCT	GGTCTGTAAC	720
	TGACGCTGAT	GTGCGAAAGC	GTGGGGATCA	AACAGGATTA	GATACCCTGG	TAGTCCACGC	780
	CGTAAACGAT	GAGTGCTAAG	TGTTAAGAGG	GGTTTCCGCC	CCTTAGTGCT	GCAGCTAACG	840
	CAATAAGCAC	TCCGCCTGGG	GAGTACGACC	GCAAGGTTGA	ACTCAAAGGA	ATTGACGGGG	900
	ACCCGCACAA	GCGGTGAGCA	TGTGTTTAAT	CGAAGCAACG	CGAAGAACTT	ACAGATCTGA	960
	CATCTCTGAC	ACTCTAGAAA	ATAGAGCCTT	CCCCTCGAGG	ACAAGTGACA	GTGTGCATGC	1020
	TGTCGTGACG	TCGTGTCTGA	AATGTGGGTA	AGTCCGCACG	AGCCACCTGA	CTAGTGGCAG	1080
	TCATAATTGC	ATCTAGTGAC	TGCGTGACAA	CGAGAGTGGG	ATGACTCATC	TCATGCCTTA	1140
	GACTGCTACG	TTCGATGCAT	CAGACGTACT	GAAGTTCAGC	AATTCCATAG	GCT	1193
3 SQ	Sequence 1039 BP; 287 A; 245 C; 273 G; 234 T; 0 other;						
	GTGCATGGCG	CGGCTACACA	TGCAGTCGAG	CGGGGATAGG	GTGCTTGAC	CTGAGTCCTA	60
	GCGGCGGAAG	GGTGAATTAT	GCTTAGGAAT	CTGCGTATTA	GTGGGGGACA	ACGTTTCGAA	120
	AAGGACGCTA	ATACAGCGGA	CATCCTATTG	TATAAAGCTG	GGGATCATCG	AACCTTGTTT	180
	TAATCTGTGA	CTCTAAGACG	GATTCGCTAC	GTGGTGGCTA	AATGGCCTAC	CAATGCCACA	240
	ACCTGGAGAC	GATCCGAAAC	CAAGATCCGA	CGGTGATCGA	CCGAAATGCG	GCCCAGACTC	300
	CTACCGGACG	CCTACGGGAG	GAATATTGAA	GGAAGCTTCC	AACCCGGATC	AAACCCCTGAC	360
	GCATGAGCGA	CGAATGACTT	TTGAATGTCT	TCCACTTTAA	ACGATCAGGA	GGCTTACCAG	420

AACAATACCC	GGAATAACTG	GACGTTACTT	GACGAAACCT	CACCGGAAAA	CTCTGTGCCA	480
CTACGTGCCA	GCATACGCAG	GGTGAAAGCG	TTAATCGGAT	TTACTGGGAA	TATTGGCGCG	540
TAAGTGGCTA	ATTAAGGTTC	CTGTGAAATC	CCCGAGCTTA	ACTTGGGAAT	TGCATTC AAT	600
GCTGGTTAAA	ACTGGGAAAC	TTGAGGAAGG	TAGAATTCCA	GGGAATTCCA	AGTGTATGCG	660
TGAAAATCTG	GAAGAATTCC	AATGAACACC	AGTGGCCATC	GGGACTTCT	GTGCCACTGA	720
TGAGCCTGAG	GCTGGAGAGC	ATACAGGATT	AAATACACTA	GATACCCTGG	CCGTACACGA	780
CGTCTACTAT	GAGTGGTAAC	CGTTGAAGGG	TTACCGGCC	CTTCTAACTC	CATCAAATAG	840
CATCTCCTGC	TGAGTACTGG	GCGAGTACAG	TACAAC TGAA	TGAAACTTCA	AAGGGAGTCC	900
CCACGAGCCG	CCGCACCCGT	GGTATCATGT	CGATTTAATT	CCAAGCAACC	CTAACGAACC	960
TTTCTCAGTC	TTAGACATCT	TCTAGACAAT	CGCATGGAGC	GTGTCTGGAC	CCTTCGGCGC	1020
AGTAGATGGC	ATAGGGTGT					1039

4 SQ Sequence 1294 BP; 334 A; 315 C; 324 G; 321 T; 0 other;

CAGCAGTACG	CAGCTACACT	GCAGTCGAGC	GAACAGATAG	TAGATCTGTG	CTCTTTTGTC	60
GTTGCCGACG	GATGGGTGAG	GATCACGTGT	CTGACCTTCC	TTGGGGACTG	GCATCTCTTC	120
GGGAATCCGG	AGCTAATACC	GGTCCTTATT	TCAACCGCA	GGGATCTTCG	GGGCTTGATG	180
GTTTTGCTGA	CCCTTATTCA	TGGACCCTCG	CCGGAGTAGC	TAATTGCTAC	CGTAGCGACT	240
TACCGTGGCG	ACTATGAGAA	GACGACCTGA	CAGGGTGATA	CTGAAAAC TG	GAAC TGACTC	300
CTAGTCCAGA	CTCCTACGGG	AAATATTGGA	CAATGAATCT	TAACCCTGGT	CCGAACGCCT	360
GACGGATGTG	CGCCACGCCT	TTGATGAAGG	TCTTCTTAAA	GCGAACTCTG	TTATTACGAA	420
ATAACATAAC	TGGAATTAAG	TGTGTTCCCT	TTGAAAGTAC	CTAATCATAA	CTCCGCGGCT	480
AACTACGTGC	CAACACCCAC	GGTGATAACCT	ATGTGGCAAT	TTTTAGCCGG	AATTATTGGC	540
CGTAATGCGC	CGGTAAGCGG	TTTGTTAAAT	CTGATATCTT	AGCCCGCGAA	TCGCCCTCGA	600
AACTGCATTG	CTAACTGATG	GAATAGAATG	CTAAAAATGA	AAGTGTAATT	CCGTGTGGAG	660
CGGAGATCTG	CACGAATATC	TGGAGGAACA	CCAGTGGTCA	GGGCTACTTT	CTGGTCTGTA	720
ACTGACGCTG	ATGTGCGAAA	GCGTGGATAT	CATACCCGAT	TAGATACCCT	GGTAGACCAT	780
GCCGTATACG	ATGAGTGCTA	AGTGTTAGCT	GGGTTTCCGC	CCCTTAGTGC	TGCAGCTAAC	840
GCATTAAGCA	ATCCGCCTGG	CGAGAACGAC	CGCAGGTTGA	TACTCAAAGG	AATTGACAGG	900
GACCCGCACA	AGCGGTGGAG	CATGTGGTAT	AATTCGAACC	ATACCTAGAA	CCGTACCAAA	960
ACTAGACATC	CTTGACAAC T	CTATAGCTTC	AGCTATCCCT	TCGGGGGACT	ACGTGACTGG	1020
TCGTGCGATG	GGTGTGCTCA	GACTTCTGTC	TAGATATCTC	GAGCTAACGT	CACCGCATCT	1080
ACGTCAATCG	CCTAACGCGT	TATAATGCAG	TCATTATGAT	CGGATACTGC	CAAGGTGAAC	1140
ATGCCGGATG	AAAGACGGGA	GAACGTTCTGA	GACTGACGTG	CCACACTTAC	AATGCACGCT	1200
TAACATCGGG	CTTACCATCG	TGCTACATGT	CATGTCAACT	GCGATGCGAG	AATCGCTAGA	1260
TAGCTCAAGA	CGTAACTGCC	CACTGTAAAG	CTTG			1294

5 SQ Sequence 1225 BP; 331 A; 267 C; 362 G; 265 T; 0 other;

GGGCTGCGGT	GCTATACATG	CAGTCGAGCG	AACGGACGAG	AAGCTTGCTT	CTCTGATGTT	60
AGCGGCGGAC	GGGTGAGTAA	CACGTGGATA	ACCTACCTAT	AAGACTGGGA	TAACTTCGGG	120
AAACCGGAGC	TAATACCGGA	TAATATTTTG	AACCGCATGG	TTCAAAGTG	AAAGACGGTC	180
TTGCTGTCAC	TTATAGATGG	ATCCGCGCTG	CATTAGCTAG	TTGGTAAGGT	AACGGCTTAC	240
CAAGGCAACG	ATGCATAGCC	GACCTGAGAG	GGTGATCGGC	CACACTGGAA	CTGAGACACG	300
GTCCAGACTC	CTACGGGAGG	CAGCAGTAGG	GAATCTTCCG	CAATGGGCGA	AAGCCTGACG	360
GAGCAACGCC	GCGTGAGTGA	TGAAGGTCTT	CGGATCGTAA	AACTCTGTTA	TTAGGGAAGA	420
ACATATGTGT	AAGTAACTGT	GCACATCTTG	ACGGTACCTA	ATCAGAAAGC	CACGGCTAAC	480
TACGTGCCAG	CAGCCGCGGT	AATACGTAGG	TGGCAAGCGT	TATCCGGAAT	TATTGGGCGT	540
AAAGCGCGCG	TAGGCGGTTT	TTTAAGTCTG	ATGTGAAAGC	CCACGGCTCA	ACCGTGGAGG	600
GTCATTGGAA	ACTGGAAAAC	TTGAGTGCAG	AAGAGGAAAG	TGGAATTCCA	TGTGTAGCGG	660
TGAAATGCGC	AGAGATATGG	AGGAACACCA	GTGGCGAAGG	CGACTTTCTG	GTCTGTAACT	720
GACGCTGATG	TGCGAAAGCG	TGGGGATCAA	ACAGGATTAG	ATACCCTGGT	AGTCCACGCC	780
GTAAACGATG	AGTGCTAAGT	GTTAGGGGGT	TTCCGCCCTT	TAGTGCTGCA	GCTAACGCAT	840
TAAGCACTCC	GCCTGGGGAG	TACGACCGCA	AGGTTGAACT	CTAAGGAATT	GACGGGGACC	900
CGCACAAGCG	GTGGAGCATG	TGGTTAATTC	GAAGCAACGC	GAAGAACTTA	CCAAATCTTG	960
ACATCCTTTG	AAACTCTAGA	GATAGAGCTT	TCCCTTCGGG	GGACAAGTGA	CAGTGGTGCA	1020
TGGTTGGCCG	CACTCGTGTG	GTAGATGTG	GGTAAGTCCC	GCACGAGCGC	ACCCTTAGCT	1080
TAGTTGCATC	ATAATTTGGG	ACTCTAGTGA	CTGCCGTGAC	AACCGGAGGA	AGTGGGGATA	1140
CGTCAATCAC	CTGCCCTTAA	GATTGGGCTA	CCCCCTGCTA	AATGGAAATT	AAAGGGACGG	1200
AACCGAAAGC	CAGGAAACCT	TAATT				1225

6 SQ Sequence 1284 BP; 329 A; 300 C; 378 G; 277 T; 0 other;

CGCGTGCGCG	GCTATCATGC	AGTCGAGCGG	GGGATAGTTT	TTGCTACTTT	ACCTAGGGGC	60
GGACGGGTGA	GTGTGCTTAG	GAAGCTGGCT	AGTAGGGGGG	GACGACATCT	CTAAGGGGAT	120
GCTAATACCG	CATAGGTCCT	ATGGGATAAA	GCGAGGGACC	TTCGGGCCTT	GCGCTAATAG	180
ATGAGCCTAA	GTCGGATTAG	CTAGTTGGTG	GTGTAAAGGT	CTACCAAGGC	CACGATCTGT	240
AGCGGGTCTG	AAAGGATGAT	CCGCCACGCT	GGGACTGAGA	CACGGCCCCG	ACTCCTACGG	300
CAGGCAGCAG	TGGGGAATAT	TGAACAATGG	GGGGAACCCT	GATCCAGCCA	TGCCGCGTGA	360
GTGAAGAAGG	CCTTTTGGTT	GTGAAGCTCT	TTAAGCGAGG	AGGAGGCTAC	CGAGATTAAT	420
ACTCTTGAAT	ACTGCACGTT	ACTCAAAGAA	TAAACACCGG	CTAACTCTGT	GCCATCACCC	480
GCGGTAATAC	AGAGGGTGCA	AGCGTTAATC	GGATTTACTG	GGCGTAAAGC	GCGCGTAGGT	540
GGCCAATTAA	GTCAAATGTG	AAATCCCCGA	GCTTAACTTG	GGAATTGCAT	TCGATACTGG	600
TTGGCTAGAG	TATGGGAGAG	GATGGTAGAA	TTCCAGGTGT	AGCGGTGAAA	TGCGTAGAGA	660
TCTGGAGGAT	ACCGATGGCG	AAGGCAGCCA	TCTGGCCTAA	TACTGACACT	GATGTGCCAC	720
AGCATGCGGG	AGCAAACAGG	ATTAGATAAC	CTGGTAGTAC	ATGCCGTAAA	CGATGTCTAC	780
TAGCCGTTGG	CGCCGTTGAC	GCTTTAGTGG	CGCACCTAAC	GCGATAATAA	ACCGCCTGGC	840
GAGCACGGTC	GCAAGAATAA	TACTCAAATG	AATTGACTGA	GGGACCGCAC	AAGCGGCGCA	900
GCATGTGGTT	GAATTTTCGAT	GCAACGCTCA	AAAACCATCA	CCTAGGACCT	GACCTACGTA	960
CGTACTTTCC	ATCATGAGGG	ATTGGTGCCA	TTCGTGCAGC	TCACATACAG	CTGCTGCATG	1020
GACTGGTCGT	CAGCTCGTCC	CTGAGAATGT	TCGGTTAATT	CCGCATCAAG	CGCCACCACT	1080
ATCGCTTATC	TGCCAGCTGG	TTAACCCGGG	TAACCTCAAG	GATTACTGAC	CAGTGACAGC	1140
ACCTGGGAGG	AGGCGAGGGA	TGACGGCGAG	TCGTCATGGC	CTTACGGACC	AGGTGCCTAC	1200
TCAGCTGTAC	TACACGTGTC	GGATACGACG	GTAGACATAC	TCAGCGGAAG	CATGCCCATA	1260
ACGTTTACAA	CAAAAAGCCG	TAGC				1284

7 SQ Sequence 1202 BP; 305 A; 273 C; 355 G; 269 T; 0 other;

GGGCATTGCG	CGGCTTACAA	TGCAGTCGAG	CGGGGAAAGT	AGCTTGCTAC	TTGACCTAGC	60
GGCGGAGGGG	TGAGTAATGC	TTAGGAATCT	GCCTATTAGT	GGGGGACAAC	ATCTCGAAAAG	120
GGATGCTAAT	ACCGCATACG	TCCTACGGGA	GAAAGCAGGG	GACCTTCGGG	CCTTGCGCTA	180
ATACATGAGC	CTAATTCGGA	TTACCTAGTT	GGTGGGGTAA	AGGCCTACCA	AGGCGACGAT	240
CTGTAGCGGG	ACTGAGAGGA	TGATCCGCCA	CACTGGGACT	GACACACGGG	CCACTCCTCT	300
ACGGGAGGCA	CCACTGGGGA	ATATTGGACA	ATGGGGGCAA	CCCTGATCCA	GCCATGCCGC	360
GTGTGTGAAC	AAGGCCTTTT	GGTTGTAAAT	CACTTTAAGC	GAGGAGGATG	CTACCGAGAT	420
TAATACTCTT	GGATAGTGGA	CGTTACTCGC	ATAATAAGCA	CCGGCTAACT	CTGTGCCAGC	480
AACCGCGGTA	ATACAGAGGG	TGCAAGCGTT	AATCGGATTT	ACTGGGCGTA	AAGCGCGCGT	540
AGGTGGCCAA	TTAAGTCAAA	TGTGAAATCC	CCGAGCTTAA	CTTGGGAATT	GCATTCGATA	600
CTGGTTGGCT	AGAGTATGGG	AGAGGATGGT	AAAATTCCAG	GTGTAGCGGT	GAAATGCGTA	660
GAGATCTGGA	GGAATACCGA	TGGCGAACGC	AGCCATCTGG	CCTAATACTG	ACACTGAGGT	720
GCGAAAGCAT	GGGGGAGCAA	ACAGGATTAG	ATACCCTGGT	AGTCCATGCC	GTAAACGATG	780
TCTACTAACC	GATTGGTGCC	TTTGAGGCTT	TAGTGGCGCA	TCTAACGCGA	TAAGTATACC	840
GCTATGGAGAG	TACCGTCGGA	AGACTAAAAC	TCAAATGAAT	TGACTGGGGC	CCGCACAAGC	900
GGTGGAGCAT	TGGGGTTACA	TTTCGATGCAA	CTTCGTAGAA	CCTTACCTGA	TCATGACATA	960
GTATGACTTT	CCCAGAGATG	GACTGGTGCC	TTCGGAACCT	ACCTACAGGT	GCTGCATGGC	1020
TGTCGTCAGC	TCGTGTCGTG	ACATGTTGGG	TAAGTCCC GC	ACGAACCCCA	ACCCTTTTCC	1080
GTACTCTGCC	ATCGTGTTAG	CCGGTACTTT	AAGGATACCT	GTCCAGTGAC	GGACCTGGCA	1140
GGAGCGCGGG	GACGACGGTC	GAGTCATCAA	GTGCTCATT	CCGATCGCGG	GTCTTAACAC	1200
TG						1202

8 SQ Sequence 1190 BP; 289 A; 312 C; 353 G; 236 T; 0 other;

GGCAGGGCGG	TGCTATACTG	CAGTCGAGCG	GACTTGTTGA	GGTTGGTCCG	TTTAGTTAGC	60
GGCGGGAGGG	TGGATTACCC	GTGGGGCACC	TGGCCCCGGC	AATGGGAATA	CTTCCGGAAA	120
ACGGGGGTTA	TACCGAAAAA	CCTTGGAAC	CTCCTGGGTT	TCTTTTTAAA	AGGGGCTTTC	180
GGTGGCCCTG	GCAGAAGGGC	CCCCCGCCCC	ATTACCTATT	GGGGGAGGAA	CCGGTCCCCC	240
AGGGCACCAA	TCCTAACCCA	ACTGGAAAGG	TGGACCGGCC	CCCTGGGAAT	GAAAACCCGG	300
CCCAAACCCC	ACCGGAAGGC	ACCATAAGGA	ATCCTCCCCC	ATGGAACAAA	GTCTGAACGA	360
ACCACCCCCC	CGGGATGGAA	AAAGGTTTCC	GAACCTAAAA	CTTTGGTGCC	CAGAAAAAAC	420
CAATTCCGGA	ATTACTGGAC	GGACCTTGGC	CGGACCTGGG	CCGAAAACCC	CCGTAACCTT	480
CCTGGCCACC	CCCCCGGTAA	TTCTAGGTG	GGCAGCGTTG	GCCGGAATTT	TTGGGCGTTA	540
AGCGCGCCCC	AGCCGGCTTT	TAAGTCTGAA	GGGGAAACCC	CCGGGTCAAC	CCGGGAAGGT	600
CAATTGAAAC	TTGAAAGCTT	GAATGGAAAA	GAAAAAGAAG	GGAATTTTAC	GTGGTACCGT	660

GGAATGGCGT	AAGATGTTGA	AGAAAACAGT	GGCCGAGGCC	GTCCTTGGT	TTTAACTGA	720
CGGCTAAGCC	CGAAGCCGTG	GGGAACGAAC	AGGATAGATA	CCCTGGGTAG	TCCACGCCGT	780
AAACGATGAG	TGCTAATGTT	GGGGGGGTTT	TCGCCCTTC	ATTGCTGCAG	CTTACGCATA	840
AGCACCTCCG	CCTGGGGAGT	ACGGCGCAAG	CCTGAACTCA	AAGAATTGAC	GGGGCCCGCC	900
ACAAGCCGTG	AGCATGTGTT	TAATCGAAGC	ACGCGAGACT	ACAGTCTTGA	CATCCGCTGA	960
TCGGTCTGGA	GAACGATCTC	CCTCGGGCAG	CGTGACCAGG	TGTGCATGAT	GTCGTCAGCT	1020
CGTGTCTGTA	AATGTTGGTA	AGTCCGCAAC	GAGCGCAACC	TGATCTAGTT	GCCAGCATTC	1080
AGTGGCACTC	TAGTGACTGG	CGTGACACGA	GATGTGGTAT	GACTCATATC	ATGCCCTTAG	1140
ACTGCCTAAC	GTCCAATGAT	GTCAGTCTGC	GAGACGCAGG	CTAAGGCATC		1190

9 SQ Sequence 1343 BP; 365 A; 297 C; 390 G; 291 T; 0 other;

GGTGCATGCG	GGTGCTATAA	CTGCAAGTCG	AGCGGACAGA	CGAGGAGCTT	GTCCTTTGA	60
CGTTAGCGGC	GGACGGGTGA	GTAACACGTG	GGTAACCTAC	CTATAAGACT	GGGATAACTT	120
CGGGAAACCG	GAGCTAATAC	CGGATAATAT	TTCGAACCGC	ATGGTTCGAT	AGTGAAAGAT	180
GGCTCTGCTA	TCACTTATAG	ATGGACCTGC	GCCGTATTAG	CTAGTTGGTA	AGGTAACGGC	240
TTACCAAGGC	AACGATACGT	AGCCGACCTG	AGAGGGTGAT	CGGCCACACT	GGAACTGAGA	300
CACGGTCCAG	ACTCCTACGG	GAGGCAGCAG	TAGGGAATCT	TCCGCAATGG	GCGAAAACCT	360
GACGGAGCAA	CGCCGCGTGA	GTGATGAAGG	TCTTCGGATC	GTAAAACTCT	GTTATTAGGG	420
AAGAACAAAC	GTGTAAGTAA	CTGTGCACGT	CTTGACGGTA	CCTAATCAGA	AAGCCACGGC	480
TAACTACGTG	CCAGCAGCCG	CGGTAATACG	TAGGTGGCAA	GCGTTATCCG	GAATTATTGG	540
GCGTAAAGCG	GCGGTAGGCG	GTTTTTTAAG	TCTGATGTGA	AAGCCACGG	CTCAACCGTG	600
GAGGGTCATT	GGAAACTGGA	AAACTTGAGT	GCAGAAGAGG	AAAGTGGAAT	TCCATGTGTA	660
GCGGTGAAAT	GCGCAGAGAT	ATGGAGGAAC	ACCAGTGGCG	AAGCGACTT	TCTGGTCTGT	720
AACTGACGCT	GATGTGCGAA	AGCGTGGGGA	TCAAACAGGA	TTAGATAACC	TGGTAGTCCA	780
CGCCGTAAAC	GATGAGTGCT	AAGTGTTAGG	GGGTTTCCGC	CCCTTAGTGC	TGCAGCTAAC	840
GCATTAAGCA	CTCCGCCTGG	GGAGTACGAC	CGCAAGGTTG	AAACTCAAAG	GAATTGACGG	900
GGACCCGCAC	AAGCGGTGGA	GCATGTGGTT	TAATTGCAAG	CAACGCGAAG	AACCTTACCA	960
AATCTTGACA	TCCTTTGACC	CTTCTAGAGA	TAGAAGTTTC	CCCTTCGGGG	GACAAAAGTGA	1020
CAGGTGGTGC	ATGGTTGTGC	TCAGCTCGTG	TCGTGAGATG	TTGGGTAAAG	TCCCACAACG	1080
AGCGCAACCC	TTAAGCTTAG	TTGCCATCAT	TAAGTTGGGC	ACTCTAAGTT	GACTGCCGGT	1140
GACAAACCGG	AGGAAGGTTG	GGGATGACGT	CAAATCATCA	TGCCCTTAT	GATTTGGGCT	1200
ACACACGTGC	TACAATGGAC	AATACAAAGG	GCAGCGAAAC	CGCGAGGTCA	AGCAAATCCC	1260
ATAAAGTTGT	TCCCAGTTCG	GATTGGAGTC	TGCACTCGAC	TAAGTGAATC	TGGAACCCCC	1320
TAGAACCCTA	AAATACATAT	CAG				1343

10 SQ Sequence 1320 BP; 327 A; 306 C; 420 G; 267 T; 0 other;

GGCGCTGGGG	GGTGGCTATA	CTGCAAGTCG	AGCGATCAAT	AGGAGCTTGC	TCCTGTTGGT	60
TAGCGGCGGA	CGGGTGAGTA	ACACGTGGGC	AACCTGCCTG	TAAGACTGGG	ATAACACCGG	120
GAAACCGGTG	CTAATACCGG	ATAATCCTTT	TCCTCTCATG	AGGAAAAGCT	GAAAGTCGGT	180
TTCGGCTGAC	ACTTACAGAT	GGGCCGCGG	GCATTAGCT	AGTTGGTGAG	GTAACGGCTC	240
ACCAAGGCGA	CGATGCGTAG	CCGACCTGAG	AGGGTGATCG	GCCACACTGG	GACTGAGACA	300
CGGCCAGAC	TCCTACGGGA	GGCAGCAGTA	GGGAATCTTC	CACAATGGAC	GAAAAGTCTGA	360
TGGAGCAACG	CCGCGTGAGC	GATGAAGGCC	TTCGGGTCTG	AAAGCTCTGT	TGTTAGGGAA	420
GAACAAGTAC	CGGAGTAACT	GCCGGTACCT	TGACGGTACC	TAACCAGAAA	GCCACGGCTA	480
ACTACGTGCC	AGCAGCCGCG	GTAATACGTA	GGTGGCAAGC	GTTGTCCGGA	ATTATTGGGC	540
GTAAAGCGCG	GCGAGGCGGT	CCTTTAAGTC	TGATGTGAAA	GCCCACGGCT	CAACCGTGGA	600
GGGTCAATTG	AAACTGGGGG	ACTTGAGTAC	AGAAGAGGAA	AGCGGAATTC	CACGTGTAGC	660
GGTGAAATGC	GTAGAGATGT	GGAGGAACAC	CAGTGGCGAA	GGCGGCTTTC	TGGTCTGTAA	720
CTGACGCTGA	GGCGCGAAAG	CGTGGGGAGC	AAACAGGATT	AGATAACCTG	GTAGTCCACG	780
CCGTAAACGA	TGAGTGCTAA	GTGTTAGGGG	GTTTCCGCCC	CTTAGTGCTG	CAGCTAACCG	840
ATTAAGCACT	CCGCCTGGGG	AGTACGGCCG	CAAGGCTGAG	ACTCAAAGGA	ATTGACGGGG	900
GCCCGCACAA	GCGGTGGAGC	ATGTGGTTTA	ATTGCAAGCA	ACGCGAAGAA	CCTTACCAGG	960
TCTTGACATC	CTCTGACACT	CCTAGAGATA	GGATTTTCCC	CTTCGGGGGA	CAGAGTGACA	1020
GGTGGTGCAT	GGTTGTCTGC	AGCTCGTGTG	GTGAGATGTT	GGGTAAAGTC	CCGCAACGAG	1080
CGCAACCCTT	GATCTTAGTT	GCCAGCATTC	AGTTGGGCAT	TCTAAGGTGA	CTGCCGGTGA	1140
CAAACCGGAA	GAAAGGTGGG	GATGACGTCC	AATCATCATG	GCCTCTTATG	ACCTGGGCTT	1200
AACACTCGGC	TACAAAGGTA	TGGTACAAAG	GGCTGCTAAA	ACCGCCAGGT	TTAGCCAAGC	1260
CCAGAAAAGC	ATTTCCCAT	TGGGAGTGCA	GGGTGAGAAT	CCGCCGGTAT	TTAAGCGCGA	1320

11 SQ Sequence 518 BP; 131 A; 121 C; 169 G; 97 T; 0 other;

GTGCGGGCGG	CGTGCTATAC	ATGCAAGTCG	AGCGGACAGA	AGGGAGCTTG	CTCCCGGATG	60
TTAGCGGCGG	ACGGGTGAGT	AACACGTGGG	TAACCTGCCT	GTAAGACTGG	GATAACTCCG	120
GGAAACCGGA	GCTAATACCG	GATAGTTCCT	TGAACCGCAT	GGTTCAAGGA	TGAAAGACGG	180
TTTTCGGCTGT	CACTTACAGA	TGGACCCGCG	GCGCATTAGC	TAGTTGGTGA	GGTAACGGCT	240
CACCAAGGCG	ACGATGCGTA	GCCGACCTGA	GAGGGTGATC	GGCCACACTG	GGCCACACTG	300
ACGGCCCAGA	CTCCTACGGG	AGGCAGCAGT	AGGGAATCTT	CCGCAATGGA	CGAAAGTCTG	360
ACGGAGCAAC	GCCGCGTGAG	TGATGAAGGT	TTTCGGATCG	TAAAGCTCTG	TTGTTAGGGA	420
AGAACAAGTG	CAAGAGTAAC	TGCTTGACCC	TTGACGGTAC	CTAACCAGAA	AGCCACGGGT	480
AACTACATGC	CAGCAGCCGC	GGGTAATACT	TAGGTGGC			518

12 SQ Sequence 1227 BP; 313 A; 270 C; 374 G; 270 T; 0 other;

GGGCGCCGGC	GGCTACACAT	GCAGTCGAGC	GGGGAAGGTA	GCTTGCTACT	TTACCTAGCG	60
GCGGACGGGT	GAGTAATGCT	TAGGAATCTG	CCTATTAGTG	GGGACAACA	TCTCGAAAGG	120
AATGCTAATA	CCGCATACGT	CCTACGGGAG	AAAGCAGGGG	ATCTTCGGAC	CTTGCCTAA	180
TAGATGAGCC	TAAGTCGGAT	TAGCTAGTTG	GTGGGGTAAA	GGCCTACCAA	GGCGACGATC	240
TGTAGCGGGT	CTGAGAGGAT	GATCCGCCAC	ACTGGGACTG	AGACACGGCC	CAGACTCCTA	300
CGGGAGGCAG	CAGTGGGGAA	TATTGGACAA	TGGGCGGAAC	CCTGATCCAG	CCATGCCGCG	360
TGTGTGAAGA	AGGCCTTTTG	GTTGTAAAGC	ACTTTAAGCG	AGGAGGAGGC	TACTTAAAT	420
AATACTCTTG	GATAGTGGAC	GTTACTCGCA	GAATAAGCAC	CGGCTAACTC	TGTGCCAGCA	480
GCCGCGGTAA	TACAGAGGGT	GCAAGCGTTA	ATCGGATTTA	CTGGGCGTAA	AGCGCGCGTA	540
GGTGGCCTTT	TAAGTCGAAT	GTGAAATCCC	TGAGCTTAAC	TTAGGAATTG	CATTGATAC	600
TGGATGGCTA	GAGTATGGGA	GAGGATGGTA	GAATTCCAGG	TGTAGCGGTG	AAATGCGTAG	660
AGATCTGGAG	GAATACCGAT	GGCGAAGGCA	GCCATCTGGC	CTAATACTGA	CACTGAGGTG	720
CGAAAGCATG	GGGAGCAAAC	AGGATTAGAT	ACCCTGGTAG	TCCATGCCGT	AAACGATGTC	780
TACTAGCCGT	TGGGGCCTTT	GAGGCTTTAG	TGGCGCAGCT	AACGCGATAA	GTAGACCGCC	840
TGGGGAGTAC	GGTGCACAAG	CTAACACTCA	AATGAATTGA	CGGGGGCCCG	CACAAGCGGT	900
GGAACATGTG	GTTTAATTTC	ATGCAACGCG	AAGAACCCTA	CCTGGTCTTG	ACATACTAAG	960
AACTTTCCAG	AGATGGATTG	GTGCCTTCGG	GAACCTCATA	CAGGTGCTGC	ATGGCTGTCT	1020
CAGCTCTTGT	CGTGAGATGT	GGGTAAATTC	GCAACGAGCG	CACCCTTTTC	TTATTTGCCA	1080
CGGTAAAGCC	GGAACCTAAG	GATCTGCCGT	GACAACCTGA	GGAAGCGGGA	GACGTCAATC	1140
ATCATGCCCT	TACGGCACGG	CTACCCCGTG	CTACATGGTC	TGGCAAAGGG	TTCTACCTCG	1200
CAAAAGGTCC	GAACCTAAAG	CCGATCG				1227

13 SQ Sequence 1277 BP; 302 A; 322 C; 395 G; 258 T; 0 other;

CGGCCAGTGG	CGCGTGCTAC	ACATGCAAGT	CGAGCGGTGA	CGGCGGAGCT	TGCTCTGGCT	60
GATCAGTGGC	GAACGGGTGA	GTAACACGTG	AGTAACCTGC	CCCTCACTTC	GGGATAACCT	120
CGGGAAATCG	GGGCTAATAC	CGGACTACTAC	ATCTGACTGC	ATGGTCGGGT	CTGAAAAGAT	180
TTATCGGGGG	GGGATGGACT	CGCGGCTAT	CAGCTTGGTG	GTGAGGTAAC	GGTCACCAA	240
GGCGACGACG	GGTAGCCGGC	CTGAGAGGGC	GACCGGCCAC	ACTGGGACTG	AAATAGCGCC	300
CAGACTCCTA	CGGGAGGCAG	CAATGGGGAA	TATTGCACAA	TGGGCGAAAAG	CCTGATGCAG	360
CGACGCCGCG	TGAGGGATGA	AGGCCTTCGG	GTTGTAAACC	TCTTTCCTAT	GGAAAGAAGCG	420
AAAATGACGG	TACCTGCGGA	AAAATCGCCG	GCTAACTACA	TGCCAGCAGC	CGCGGTAATA	480
CATATGGCGC	ACGCGTTGTC	CGGAATTATT	GGGCGTAAAA	AGCTCGTACG	CGGTTTGTTCG	540
CGTCTGGGGT	GAAAACCCCA	GGCTCTACCT	TGGGCTTGCA	TCGGGTACGG	GTGACTAGA	600
GAGCGGTATG	GGTACTGGA	ATTCTTGGTG	TAGCGGTGGA	ATGCGCAAAT	ATCACGAGGA	660
ACACCGATGT	GGAAAGCGAG	TCACTGGGCC	GCAACTGACG	CTGAGGAGCG	AAAGCATGGG	720
TAGCGAACAT	GATTAGATAC	ACTGTGAGTA	CATGCCGTGA	ACATTGGGCA	CGATGTGTGG	780
GGCTCATTCC	ACGAATTCCG	TGCCCGAGCA	AACGCATTAA	CTGTCCC GCC	TGGGGAGTAC	840
GCCCGCAACG	CTCAAACCTCA	AAAGAATTGT	CGGGGGCCCG	CACCACCGGC	GGAGCATGCG	900
GATTACTTCT	ATGCGACGCG	AAAAGCCTTA	TAATGCTTGA	CATACACAGC	ACTCTTGCAG	960
AGATGTGTGT	GTCCTCACAC	TGGTGTACAT	GTGGCTGCTT	GCTTGTGCTC	AACTCGGCTC	1020
GTGGAATGAT	GGGTTAAGTC	CCGCAACGAC	GCAGCCCTCG	TCCTAATGTT	GCCCAGCACA	1080
TCATGCCGAT	GGACTCATAG	TAGACTGCCA	TGCGTGACTC	GCACGTAAGG	TGGATGACTG	1140
ACGGTCAATC	ATCATGCACC	CTGATGATCT	ATGGCTTGCC	GCCATGCTAC	ATGGCACGTG	1200
ACGGAAGCAT	GCGATCGATA	GAGTGGACCG	ATGCCTAAAA	GAACGACTTC	AGTTCGGGCC	1260
AATTGCGCGG	TCCTGAC					1277

14 SQ Sequence 1238 BP; 311 A; 288 C; 360 G; 279 T; 0 other;

CGGCGTGGCG	GCGGCTTACC	ATGCAGTCGA	GCGGGGAAAA	GTACCTTGCT	ACTTGACCTA	60
GCGGCGGACG	GGTGAGTAAT	GCTTATGAAT	CGGCCTATTA	TCGGGGGACA	ACATCTCGAA	120
AGGGATGCTA	ATACCGCATA	CGTCCTACGG	GAGAAGGCAG	GGGACCTTCG	GGCCTTGCGC	180
TAATAGATGA	GCCTAAGTCG	GATTAGCTAG	TTGGTGGGGT	AAAGGCCTAC	CAAGGCGACA	240
ATCTGTAGCG	GGTCTGAGAG	GATGATCCGC	CACACTGGGA	CTGAGACACG	GCCCACACTC	300
CTACGGGAGG	CAGCAGTGGG	GAATATTGGA	CAATGGGGGG	AACCCTGATC	CACCCATGCC	360
CCGTGTGTGA	AAAAAGCCTT	TTGGTTGTGA	AGCACTTTTT	CCGAGGAGGA	GGCTACCGAG	420
ATTAATACTC	TTGGATAGAG	GACGTTACTC	GCAGAATAAC	CACCGGCTAA	CTCTGTGCCA	480
GCAGCCGCGG	TAATACAGAG	GGTGCAAGCG	TTAATCGGAT	TTACTGGGCG	TAAAGCGCGC	540
GTAGGTGGCC	AATTAAGTCA	AATGTGAAAT	CCCCGAGCTT	AACTTGGGAA	TTGCATTCGA	600
TACTGGTTGG	CTAGAGTATG	GGAGAGGATG	GTAGAATTCC	AGGTGTAGCG	GTGAAATGCG	660
TATAGATCTG	GAGGAATACC	GATGGCGAAC	GCAGCCATCT	GGCCTAATAC	TGACACTGAG	720
GTGCGAAAGC	ATGGGGAGCA	AACAGGATTA	CATACCCTGG	TAGTCCATGC	CGTAAACGAT	780
GTCTACTATC	CGTTGGGGCC	TTTGAGGCTT	TATTGGCGCA	GCTAACGCGA	TAAGTAGACC	840
GCCTGGGGAG	TACGGTCGCA	AGACTAAAAC	TCACATGAAT	TGACGGGGGC	CCGCACATTC	900
GGTCGGAGCA	TGTGGTTTAA	TTTCGATGCAA	CGCGAAGCAT	CCTTACCTGG	TCTTGACATA	960
ATAGAACTTT	CCATAGATGC	ATTGTTGCCT	TCGGGAACGT	ACATACACGT	GCTGCATGGC	1020
TGTCGTCAGC	TCGTGTCTCTG	AGATGTTGCG	CTTACCTCCC	GTCAACTGAG	TCGCACCTTT	1080
TTTCCTTTAT	TTGACTGCGA	GTAAAGCGGG	ACCTTATGGA	TACTGACAGT	GACAACCTGG	1140
AGATGGCGGG	ACCGACTTCA	GTCATCATGC	CTACCATCAG	CTACCCTGAG	CTACATGATC	1200
GGTACAAGAG	GTATGCCTAA	CCCTCCGCGA	AAGATAGC			1238

15 SQ Sequence 1415 BP; 351 A; 341 C; 428 G; 295 T; 0 other;

GGGCGATGCG	GTGGCTATAC	ATGCAGTCGA	GCGGACGAAG	GGAGCTTGCT	CCCGGATGTT	60
AGCGGCGGAC	GGGTGAGTAA	CACGTGGGTA	ACCTGCCTGT	AAGACTGGGA	TAACCTCGGG	120
AAACCGGAGC	TAATACCGGA	TAGTTCCTTG	AACCGCATGG	TTCAAGGATG	AAAGACGGTT	180
TCGGCTGTCA	CTTACAGATG	GACCCGCGGC	GCATTAGCTA	GTTGGTGAGG	TAACGGCTCA	240
CCAAGGCGAC	GATGCGTAGC	CGACCTGAGA	GGGTGATCGG	CCACACTGGG	ACTGAGACAC	300
GGCCCAGACT	CCTACGGGAG	GCAGCAGTAG	GGAACTTTCC	GCAATGGACG	AAAGTCTGAC	360
GGAGCAACGC	CGCGTGAGTG	ATGAAGGTTT	TCGGATCGTA	AAGCTCTGTT	GTTAGGGAAG	420
AACAAGTGCA	AGAGTAACTG	CTTGCACCTT	GACGGTACCT	AACCAGAAAG	CCACGGCTAA	480
CTACGTGCCA	GCAGCCGCGG	TAATACGTAG	GTGGCAAGCG	TTGTCCGGAA	TTATTGGGCG	540
TAAAGGGCTC	GCAGGCGGTT	TCTTAAGTCT	GATGTGAAAG	CCCCCGGCTC	AACCGGGGAG	600
GGTCATTGGA	AACTGGGAAA	CTTGAGTGCA	GAAGAGGAGA	GTGGAATTCC	ACGTGTAGCG	660
GTGAAATGCG	TAGAGATGTG	GAGGAACACC	AGTGGCGAAG	GCGACTCTCT	GGTCTGTAAC	720
TGACGCTGAG	GAGCGAAAGC	GTGGGGAGCG	AACAGGATTA	GATACCCTGG	TAGTCCACGC	780
CGTAAACGAT	GAGTGCTAAG	TGTTAGGGGG	TTTCCGCCCC	TTAGTGCTGC	AGCTAACGCA	840
TTAAGCACTC	CGCCTGGGGA	GTACGGTCGC	AAGACTGAAA	CTCAAAGGAA	TTGACGGGGG	900
CCCGCACAAAG	CGGTGGAGCA	TGTGGTTTAA	TTCGAAGCAA	CGCGAAGAAC	CTTACCAGGT	960
CTTGACATCC	TCTGACAACC	CTAGAGATAG	GGCTTTCCCT	TCGGGGACAG	AGTGACAGGT	1020
GGTGCATGGT	TGTCCTCAGC	TCGTGTCTGTG	AGATGTTGGG	TTAAATCCCC	CAACGAACCG	1080
CAACCTTTGA	TCTTAGTTGC	CAGCATTCAA	GTTGGGCACT	CTAAAGTTGA	CTGCCCGGTG	1140
ACAAACCGGA	AGGAAGGTTG	GGGATAACGT	CCAATCATCA	TGCCCTTAT	GACCTGGGGC	1200
TACCACGCTT	CTACAATGGA	CGAAAATAAG	GGGCTGCGAA	ACCCCGAAGG	TTTAGGCCAA	1260
TCCCCAGAAA	TCCGTTTCCA	GATTTCGAAT	CGCATGCTCG	CAACCCTCCG	TGCTTGAAAC	1320
TTGTAATCCG	TGGGAAATCC	CGGGAACCTC	TATGCCCCCT	GTAGAAATCT	TTTCCCTGGG	1380
CCGTGTGACA	ACCTCTCCTT	CTAATCCGTC	TACGG			1415

16 SQ Sequence 236 BP; 56 A; 52 C; 64 G; 64 T; 0 other;

TGATATTTTT	GGCTGGGTGT	AAAAGTAACA	TGTCAGACGA	GCGAAGTTAG	TTGTTCCCTTT	60
CTCGTTGGCG	TTGGGGGGGC	CCCTGTGAAT	GCGTATGTGC	CTCTTTCCTT	GAGATCTGCT	120
CGGTACCGGA	AAACATTGCA	AAAATACGTA	AATATTTTCA	CACGCCGGCT	TGCAAGTCGG	180
AGCTGCCGCC	CCCCCAATC	AGGATCCACA	AATTGTAAAA	TGATCAGGAG	GGGTGG	236

Résumé

Cette étude repose en premier lieu sur la caractérisation phénotypique classique (morphologie, biochimie et physiologie des isolats) suivie d'une analyse protéomique par la technique de SDS-page et enfin, corroborée par l'analyse bioinformatique des séquences 16S rRNA par la prédiction des structures 2D et la détermination des liens phylogénétiques. Ces techniques, apportent une valeur ajoutée dans l'identification bactérienne. La comparaison d'une séquence contre ces banques (BLAST) peut biaiser le résultat. Le résultat qui ressort de cette étude est la possibilité d'être face à trois germes nouveaux encore non identifiés sur toutes les bases de données.

Mots clés : *Sol agricole, ARNr 16S, Structure 2D, SDS-page, BLAST.*

Abstract

This study is primarily based on classical phenotypic characterization (morphology, physiology and biochemistry of isolates) followed by proteomic analysis by SDS-page technique and finally supported by bioinformatics analysis of 16S rRNA sequences by predicting 2D structures and determination of phylogenetic relationships. These techniques add value in bacterial identification. Comparing a sequence against these banks (BLAST) can skew the result. The results from this study is the possibility of being faced with three new germs as yet unidentified on all databases.

Keywords : *Agricultural soil, 16S rRNA, 2D structure, SDS-PAGE, BLAST*

المخلص

تستند هذه الدراسة أساساً على الخصائص المظهرية (التشكل، علم وظائف الأعضاء والكيمياء الحيوية من العزلات)، يليه تحليل البروتين بواسطة تقنية SDS-PAGE وأخيراً بدعم من تحليل المعلومات البيولوجية تسلسل 16S الريبوسومي من خلال التنبؤ بالهيكل 2D وتحديد علاقات النشوء والتطور. هذه التقنيات تضيف قيمة في تحديد البكتيريا. نستنتج من هذه الدراسة إمكانية تواجد ثلاثة أفراد بكتيرية جديدة ومجهولة الهوية حتى الآن على كافة قواعد البيانات.

الكلمات المفتاحية : *التربة الفلاحية - 16S rRNA - البنية - الثانوية - SDS-PAGE - BLAST*

Présenté par : **SAYAH Djaber**

Analyses phénétique et bioinformatique des isolats bactériens de deux sols cultivés de blé dur : rôle de la structure 2D des ARNr 16S dans l'identification des isolats

Mémoire de fin de cycle pour l'obtention du diplôme de Master en **Ecologie microbienne**

Cette étude repose en premier lieu sur la caractérisation phénotypique classique (morphologie, biochimie et physiologie des isolats) suivie d'une analyse protéomique par la technique de SDS-page et enfin, corroborée par l'analyse bioinformatique des séquences 16S rRNA par la prédiction des structures 2D et la détermination des liens phylogénétiques. Ces techniques, apportent une valeur ajoutée dans l'identification bactérienne. La comparaison d'une séquence contre ces banques(BLAST) peut biaiser le résultat. Le résultat qui ressort de cette étude est la possibilité d'être face à trois germes nouveaux encore non identifiés sur toutes les bases de données.

Mots clés : *Sol agricole, ARNr 16S, Structure 2D, SDS-page, BLAST.*

Laboratoire de recherche : Biologie Moléculaire et Cellulaire

Jury d'évaluation :

Président du jury : R. Alatou (Maître de Conférences A - UFM Constantine)

Rapporteur : M. A. Hamidechi (Professeur - UFM Constantine)

Examineur : M. Gagaoua (Maître de Conférences A -INATAA – UFM Constantine)

Examinatrice : I. Guergouri (Maître-Assistante A - UFM Constantine)

Date de soutenance : 16/06/2016