

# Université PARIS XI - Enseignement de M2 - UE ABA

## EXAMEN

### ***Partie D. Gautheret - Durée de l'épreuve : 1h30 - Documents autorisés***

Un nouveau génome bactérien vient d'être séquencé. Les gènes protéiques (ORF) ont été annotés au moyen d'une recherche de tous les cadres de lecture de plus de 150 aminoacides, puis d'un Blastp contre une banque de protéines. Aucun autre type de gène n'a été recherché. Vous êtes en charge de l'annotation des ARN non codants (ARNnc) dans ce génome.

(la plupart des questions admettent plusieurs bonnes réponses. Tout protocole cohérent et efficace sera accepté).

#### **Question 1. 2pt**

Vous devez commencer par une approche de génomique comparative. Pour identifier le plus grand nombre de séquences non codantes, vaut-il mieux comparer le génome étudié à :

1. une souche différente de la même espèce
2. un génome d'espèce différente mais proche
3. un génome d'espèce différente mais éloignée

Justifiez votre réponse.

#### **Question 2. 1,5pt**

Proposez un protocole rapide permettant de récupérer les séquences non-codantes conservées entre les deux génomes.

#### **Question 3. 1,5pt**

Vous obtenez dans l'étape précédente 350 fragments conservés de 30 à 1500 nt. Proposez un protocole pour identifier les ARNnc connus dans ce jeu de séquence.

#### **Question 4. 1pt.**

Quels types d'ARNnc devez-vous retrouver à coup sûr ?

#### **Question 5. 1pt.**

Au cours des opérations réalisées aux questions 2 et 3, avez-vous pu rater des ARNnc réels ? Pourquoi ?

#### **Question 6. 3pt**

Il vous reste 280 fragments conservés non identifiés. De quel type d'élément peut-il s'agir ? Proposez des critères pour différencier ces différents types d'éléments. Proposez ensuite un protocole permettant de les classer le mieux possible.