

**M1 MABS**  
**Examen terminal d'Evolution Moléculaire (durée 2h)**  
**(EM7BMAAME2) – Janvier 2013**

**Questions de cours**

Pour les questions suivantes, choisissez la bonne réponse, et expliquez pourquoi (vous pouvez illustrer avec des schémas):

1. La sélection positive agit sur une mutation (dans un gène) qui est favorable à l'individu, et qui va se fixer à terme dans la population: vrai ou faux ?
2. La sélection négative sur un gène génère un excès d'allèles rares aux sites polymorphes de ce gène : vrai ou faux ?
3. La sélection négative n'opère jamais sur un domaine protéique important dont la conservation est requise pour la fonction de la protéine: vrai ou faux ?
4. La sélection balancée maintient le polymorphisme dans un gène, mais ~~pas~~ par un mécanisme d'avantage aux hétérozygotes.
5. Quand on fait un test de neutralité de McDonald & Kreitman sur un alignement de séquences codantes, on effectue en fait un test de Student pour tester la relation entre type de substitution (synonyme / non synonyme) et type de polymorphisme (intra-espèce, inter-espèce): vrai ou faux ?
6. Dans un alignement d'une séquence codante, le ratio dN/dS (proportion de substitutions non synonymes sur la proportion de substitutions synonymes) est égal à 1 en cas d'évolution neutre, >1 en cas de sélection positive (adaptative), et < 1 en cas de sélection négative (purifiante) : vrai ou faux ?

**Problème 1**

Une partie de l'alignement de la séquence nucléotidique de l'exon 2 d'un gène de 8 individus de *Phytophthora infestans* (pathogène de la pomme de terre) est présenté en Figure 1. Ces individus sont capables d'infecter 2 génotypes de pomme de terre (A et B). Les sites variables sont indiqués en gras et soulignés.

Figure 1: séquence de l'exon 2, chez 8 individus de *Phytophthora infestans*:

```
1 - ACT GCG CTG TCC CCG CGC TCA GGC TCT
2 - ACT GCG CTG TCC CCG CGC TCA GGC TCT
3 - ACT GCG CTG TCC CCG CGC TCA GGC TCT
4 - ACT GCG CTG TCC CCG CGC TCA GGC TCT
5 - TCT GCT CTG TCC CCA CGC TCA GGC ACT
6 - TCT GCT CTG TCC CCA CGC TCA GGC ACT
7 - TCT GCT CTG TCC CCA CGC TCA GGC ACT
8 - TCT GCT CTG TCC CCA CGC TCA GGC ACT
```

Code génétique:

ACT = Thr, TCT = Ser,  
GCG = Ala, GCT = Ala,  
CCA = Pro, CCG = Pro

Un test de neutralité de Tajima a été effectué, basé sur cet alignement. La formule de la statistique  $D$  du test de Tajima est :

$$D = \frac{\hat{\theta}_{\pi} - \hat{\theta}_{\varsigma}}{SE(\hat{\theta}_{\pi} - \hat{\theta}_{\varsigma})}$$

où  $\theta_{\pi}$  et  $\theta_{\varsigma}$  sont deux estimateurs de la diversité au niveau nucléotidique sur une séquence (SE = écart-type).



---

Justifiez à chaque fois vos réponses:

7. Lequel de ces deux estimateurs évalue la probabilité qu'un nucléotide soit polymorphe ? Expliquez comment vous feriez pour le calculer.
8. Lequel de ces deux estimateurs évalue la probabilité d'hétérozygotie par nucléotide? Expliquez comment vous feriez pour le calculer.
9. Le test de Tajima donne une valeur positive de 5 et le test est significatif ( $p\text{-valeur}=10^{-6}$ ) : que concluez-vous ?
10. Quel(s) d'acide(s) aminé(s) est (sont) sous sélection (et quel type sélection) ?
11. Combien d'individus sont capables d'infecter chaque génotype de pomme de terre ? Sont-ils les mêmes pour le génotype A et le génotype B?
12. Concluez sur la stratégie évolutive employée par le pathogène pour infecter son hôte.

## **Problème 2**

Lin et collaborateurs (Nucleic Acids Res., 2007, 1-13) ont conduit une analyse phylogénétique des composants clés du système de réparation des mésappariements au niveau de l'ADN. Ces mésappariements sont introduits par l'ADN polymérase lors de la division cellulaire et leur non correction entraînerait des mutations. Le gène codant pour la protéine MutS est un acteur de ce système de réparation. Chez *Escherichia coli* il a été démontré qu'un homodimère de MutS se lie sur les nucléotides mal appariés et forme un complexe MutS-ADN. Cependant plusieurs gènes homologues à *mutS* existent dans *E. coli* et une recherche de similarité dans 461 génomes bactériens a montré qu'on pouvait distinguer 4 sous-familles de protéines MutS, MutS1 (correspondant à la protéine MutS dont la fonction est décrite ci-dessus) à MutS4. Chez les eucaryotes, 7 gènes homologues aux gènes MutS bactériens ont été identifiés. Pour étudier l'histoire évolutive de ces gènes, une reconstruction phylogénétique sur les produits des gènes de la famille MutS a été réalisée à partir d'un ensemble d'espèces représentatives des trois domaines du vivant et a conduit à l'arbre présenté Figure 1.

La reconstruction phylogénétique a été réalisée à l'aide de deux méthodes : la Neighbor Joining (NJ) et maximum de vraisemblance (PhyML). Les deux topologies étant congruentes, *i.e.* les mêmes, seul l'arbre obtenu par la NJ est présenté mais les valeurs de bootstrap obtenues avec chacune des méthodes sont reportées (NJ et PhyML).

- Par l'analyse de l'arbre de la figure 1, répondez aux questions suivantes
  1. Indiquer sur l'arbre les différents clusters correspondant aux différentes sous-familles MutS bactériennes en incluant leurs homologues présents chez les eucaryotes et/ou archaea.
  2. Quel(s) est (sont) les sous-familles bactériennes pouvant avoir été à l'origine des sous-familles MSH eucaryotes ? Justifier votre réponse.
  3. Les différentes sous-familles MSH ont-elles la même distribution taxonomique? Pouvez-vous identifier des sous-familles présentes uniquement dans certains groupes taxonomiques ? Expliciter vos réponses.
  4. Quelles sont les sous-familles MutS présentes chez les archaea ? Une de ces sous-familles pourrait-elle être ancestrale ? Si oui, laquelle et pourquoi.
  5. Pouvez-vous identifier des événements de duplication ou de transfert horizontaux qui auraient pu toucher les gènes codant les protéines MutS d'archaea ? Développez votre réponse.
  6. Discuter l'évolution des différentes sous-familles MutS chez les bactéries.