

Génomique. ESIL 2^{ème} année

Décembre 2002

Durée: 2 heures

Documents non autorisés

D. Gautheret, D. Thieffry

1. (1 pt)

Pourquoi l'obtention de la séquence d'un génome complet implique-t-elle de séquencer plusieurs fois la longueur de ce génome ?

1. (2 pt)

En quoi le séquençage du génome humain par Celera diffère-t-il de celui effectué par le Consortium International ? Expliquer.

2. (3 pt)

Pourquoi les banques d'EST comprennent-elles plusieurs fois les mêmes séquences ? Comment tente-t-on de réduire cette redondance ? Dans quel cas cette redondance présente-t-elle un intérêt ? Dans quel cas veut-on l'éviter ?

3. (2 pt)

Quelles régions d'un gène peut-on espérer retrouver sous forme d'EST ? (codant, UTR, introns, exons, promoteur) ?

4. (2 pt)

Afin de caractériser fonctionnellement une protéine inconnue, on recherche dans les banques Swissprot et NR protéique des séquences semblables. Aucune protéine semblable n'est trouvée. Dans quelles autres banques peut-on chercher ? Pourquoi ce qui n'est pas trouvé dans les premières banques pourrait il être trouvé dans les secondes ?

5. (2 pt)

Les puces à ADN peuvent être utilisées pour caractériser les régions impliquées dans la fixation de facteurs de transcription. Comment? Que doit-on alors mettre sur les puces?

6. (4 pt)

Comparer (par exemple à l'aide d'un tableau spécifiant les avantages, limites, etc.) les techniques de caractérisation de l'expression des gènes à grande échelle, au niveau transcriptionnel et au niveau protéique.

7. (4 pt)

Décrire brièvement deux approches différentes à haut débit pour caractériser un maximum de protéines susceptibles d'interagir avec une protéine donnée.