

**Génomique et Bioinformatique**  
**ESIL Biotechnologie 2<sup>ème</sup> année**  
**Février 2005**

*Durée: 2 heures*

*Documents non autorisés*

D. Gautheret, JC. Guillemot

**1. (5 pt)**

Préciser les principales techniques d'étude du

- Transcriptome
- Protéome,
- Interactome.

Choisir une technique, la détailler, en donnant les avantages et limites.

**2. (5 pt)**

Détailler la méthode d'analyse quantitative de complexes de protéines par ICAT (Isotope-Coded Affinity Tags).

**3. (5 pt)**

Quelles sont les deux types de séquences répétées les plus répandues dans le génome humain ?  
Donnez-en quelques caractéristiques.

**4. (5 pt)**

Donnez en une seule ligne la définition de chaque banque ou outil suivant :

- SRS
- Genbank
- Swissprot
- Ensembl
- Entrez

**5. (5 pt)**

Donnez trois façons de représenter un ensemble de séquences protéiques. Pour chacune des façons donnez au moins un avantage ou un inconvénient.

**6. (5 pt)**

Un laboratoire identifie chez l'homme un gène dont les mutations semblent associées à un certain type de cancer. La fonction de ce gène est inconnue. Le laboratoire vous recrute comme consultant en bioinformatique-génomique pour élucider cette fonction. Quelle démarche proposez-vous dans les deux cas de figure suivants ?

1 – Une recherche par Blast contre les banques de protéine identifiées (avec une bonne E-value) des protéines de fonction bien caractérisée.

2 - Une recherche par Blast contre les banques de protéine non identifiées (avec une bonne E-value) que des protéines de fonction mal caractérisée.

(les démarches peuvent être expérimentales ou informatiques. L'expérimentation chez l'humain est encore interdite! Maximum une page)

### 7. (5 pt)

On réalise un Blast d'une protéine inconnue de souris contre la banque nr protéique. On obtient les résultats suivants. Répondez aux questions brièvement SANS justifier vos réponses.

Type de protéine	Espèce	E-value
P1	Ratus norvegicus	1e-100
P1	Homo sapiens	1e-90
P2	Mus musculus	1e-10
P1	Xenopus laevis	1e-8
P2	Homo sapiens	1e-7
P1	Zebrafish	1e-5
P2	Arabidopsis thaliana	0.1
P3	Drosophila melanogaster	0.1

- Quel est le type de la protéine inconnue ?
- Quelle est la relation probable entre P1 et P2 ?
- Quelle est la relation probable entre P1 et P3 ?
- Quelle est la relation probable entre P2 et P3 ?
- Que signifie une E-value de 0.1 ? Est-ce significatif ?

### 8. (5 pt)

Dans l'arbre suivant, Pa et Pb sont des familles de protéines distinctes. Répondez par vrai ou faux à chacune des questions suivantes :

- La protéine X est une Pa
- Y est paralogue de Pb
- Y est orthologue de Pb
- Y n'est pas une Pb
- Trois événements de duplication apparaissent dans cet arbre

