

AMARA KORBA RAOUF/toxicologie/Semestre 1/Bio-informatique/section 01					
Matricule	Nom	Prénom	EMD	CC	MG
181833051101	AKMOUM/اقموم	AMANI SABRINE/اماني صابرين	3,75	7,50	5,25
171833060480	ALI/علي	DALILA/دليلة	5,00	10,00	7,00
191933050296	AMARA/عمارة	ACHREF/اشرف	6,50	14,50	9,70
171733062368	AMARA/عمارة	OUSSAMA/أسامة	3,25	12,00	6,75
191933049064	BAABOUCHE/بعبوش	RAZIKA/رزيقة	3,75	9,00	5,85
191933051556	BELAIBA/بلعاببة	RAYENE/ريان	10,25	14,00	11,75
191933050674	BELFERROUM /بلفروم	TAOUES/الطاوس	6,75	12,50	9,05
181833053151	BELHADAD/بلحداد	ANISSA/أنيسة	3,50	9,00	5,70
191933051795	BENFEDILA/بن فضيلة	AMAL/أمال	0,00	0,00	0,00
161633064620	BENKERRI/بن قري	RAHMA/رحمة	1,50	9,75	4,80
191933051510	BENZIANE/بن زيان	ICHERAK/إشراق	5,75	12,50	8,45
191933047448	BOUAFIA/بوعافية	CHAIMA/شيماء	6,50	10,00	7,90
191933052328	BOUBAITI/بوبيعيثي	AHLEM/أحلام	5,75	11,50	8,05
191933045386	BOUDEROUAZ/بودرواز	ASSIA/آسيا	4,00	10,00	6,40
191933051800	BOUKHARI/بوخاري	AYA/آية	0,00	0,00	0,00
171733055548	BOUMEZBEUR/بومزير	ASMA/اسماء	2,50	7,50	4,50
191933046688	BOUNAB/بوناب	HIBA/هبة	4,50	14,00	8,30
181833053790	BOUSSOUAR/بوصوار	OUNISSA/اونيسة	5,75	12,50	8,45
191933046491	DADOUCHE/دعدوش	CHOUROUK/شروق	0,00	0,00	0,00
191933045506	DEFFAF/دفاف	KHAOULA/خولة	9,00	11,00	9,80
181833052242	DERARDJA/درارجة	AICHA/عائشة	7,00	13,50	9,60
181833053103	DERRECHE/درش	ASSIA/آسية	6,50	16,00	10,30
191933051581	FARSI/فارسي	ABD ELHAMID/عبد الحميد	3,00	9,50	5,60
191933050343	FORTAS/فرطاس	ZAHRA ERAYANE/زهرة الريان	9,50	10,00	9,70
191933055574	HAMMOUDI/حمودي	KHOULOUD/خلود	3,00	9,50	5,60
191933047300	LAKHAL/لكحل	AYA/آية	4,50	8,00	5,90
191933046430	LOUALA/الوالع	DOUA/دعاء	6,50	9,00	7,50
191933046484	LOUNICI/لونيسي	Souhila/سهيلة	4,50	7,50	5,70
181833055018	MEBARKI/مباركي	KAOUTHAR/كوثر	3,00	8,00	5,00
22088014051	MENAD/مناد	Wahiba/وهيبة	0,00	3,00	1,20
191933046533	MENDOUD/مندود	AICHA/عائشة	7,75	12,50	9,65
161633062831	OUKEMOUM/أوقموم	IMENE/إيمان	10,00	13,00	11,20
181833055945	TABI/طي	LAMIS/لميس	4,75	8,00	6,05
191933050763	YAKOUBI/يعقوبي	MAISSA/ميساء	5,00	9,50	6,80

Corrigé Type de l'EMD : Semestre (1) – 2022/2023
Niveau 1^{ère} Année Master
Spécialités : Biochimie / Toxicologie
Matière : Bioinformatique
Date : Le 09/01/2023 / Durée : 1h30

Partie I : Question de synthèse (06 pts)

Dans un texte de plus de 350 mots, décrivez l'importance des réseaux en bioinformatique.

Le rôle du réseau en bioinformatique est avant tout de servir de **plateforme pour partager les données (0,5pt)**, mais aussi de fournir **les ressources logicielles (GUI et Web Tools) (0,5pt)** à l'ensemble des membres d'un laboratoire de recherche et de la communauté scientifique d'une manière générale.

En génomique exploratoire l'acquisition de l'information biologique sous forme numérique passe par les plateformes de séquençages. Ces données ont deux chemins possibles :

- 1- Les données peuvent être directement déposées dans les **bases de données publiques ou sites portails** (NCBI, EMBL, DDBJ ... etc.) **(1pt)**
 - 2- Stockage dans un réseau local → Traitement et Analyses des données → Stockage dans les réseaux public
- En premier lieu, les données sont stockées dans un **réseau local** (Serveurs des laboratoires de recherche (CRBT, CNE, IPA ... etc.) **(1pt)**
 - Ensuite les données sont annotées à l'aide de divers outils bioinformatique (Web, GUI, ou CLI). Ex : Expasy Translate, ORF Finder, BLAST, GeneScan ... etc. **(1pt)**
 - Les données peuvent être soumis (**soumission**) aux bases de données publiques. Après **vérifications**, ces données seront validées et peuvent enfin être mis à la disposition de la communauté scientifique **(0,5pt)**

Si des modifications sont nécessaire après dépôt le chercheur peut éventuellement demander à effectuée un **mis à jour** des données qu'il a soumis (**Rapatriement**). **(0,5pt)**

L'utilisation des réseaux en bioinformatique est indispensable pour la **circulation**, le **partage**, **l'analyse** et le **traitement** de l'information biologique. **(1pt)**

Partie II : Répondez par Vrai ou Faux, justifier votre réponse (08pts)

Q1) RStudio est un outil bioinformatique de type Web (Point and Click)

FAUX (0,5 pt)

Justification : RStudio est un outil bioinformatique de type GUI, d'un niveau d'utilisation intermédiaire. Il sert d'environnement de développement pour le langage de programmation R et également utilisé pour le traitement de données et l'analyse statistique.

Q2) La recherche gènes chez les procaryotes se base uniquement sur la détection des signaux ORF et des CDS

FAUX (0,5 pt)

Justification : ORF, CDS, RBS, initiateur de la transcription, terminateurs de la transcription (0,5 pt)

Q3) Le principe de séquençage de la technique Ion Torrent c'est : une hybridation avec ligature d'amorces et suivi de la fluorescence des oligonucléotides hybridés

FAUX (0,5 pt)

Justification : Synthèse enzymatique, Détection d'ions hydrogènes libérés durant la polymérisation de l'ADN (0,5 pt)

Q4) Il existe deux approches complètement différentes de la bioinformatique

FAUX (0,5 pt)

Justification : il existe deux approches complémentaires de la bioinformatique : les outils Web, GUI et les outils lignes de commande (CLI) (0,5 pt)

Q5) Les site portails offrent seulement l'accès a plusieurs bases de données biologiques

FAUX (0,5 pt)

Justification : Les site portails offrent non seulement l'accès a plusieurs bases de données biologiques mais offres aussi plusieurs outils bioinformatiques (Expasy Translate, GeneScan, ORF Finder ... etc.)

Q6) Dans la technique de Sanger les ddNTP sont utilisés pour le suivi de la fluorescence du nucléotide incorporé

FAUX (0,5 pt)

Justification : Dans la technique de Sanger les ddNTP sont utilisés (0,5 pt)

Q7) Les programmes de recherche de la fonction des objets génétiques prédit par l'annotation syntaxique sont basés sur des algorithmes d'alignement multiple (0,5 pt)

FAUX..... (0,5 pt)

Justification : Les programmes de recherche de la fonction des objets génétiques sont basés sur des algorithmes d'alignement simple (0,5 pt)

Q8) Les NGS ont une longueur de lecture supérieure aux techniques de séquençage de première génération.

FAUX

Justification : Les NGS ont une longueur de lecture supérieure aux techniques de séquençage de première génération.

Partie III : Exercice (06 Pts)

En utilisant le langage de votre choix, écrivez un algorithme de détection d'ORFs dans une séquence génomique d'un organisme procaryote

Déclaration des variables

```
SequenceADN = (Mettre votre séquence ICI)
StopCodon = (TAA, TAG, TGA)
StartCodons = (ATG, GTG, TTG)
ORFsequence
CDS
```

initialisation et affectation des valeurs

```
StopCodon = 0
StartCodons = 0
ORFsequence = 0
CDS = 0
```

Instructions de contrôle

Divise la séquence d'ADN en 6 phases de lecture différents (+1, +2, +3, -1, -2 et -3).

Dans chaque phase de lecture, recherche les triplets stop

Si la distance entre 2 triplets stop consécutifs est supérieure au seuil (300 Nucléotides), enregistrer cette ORF

Dans cette ORF, recherche le premier codon start, de façon à ce que la protéine codée soit de taille maximale.

Instructions de d'affichage

Affiche le nombre et la liste des ORFs

Le chargé de la matière

M. Amara Korba Raouf

