

Nom, Prénom : Boukharouba Boumediene

Matricule : 20032279

Nom, prénom: Kassis Charbel

Matricule: 0976458

---

**Question 1**

Pour générer la réponse, il faut lancer le programme question1.py. Ce programme va générer quatre fichiers dans [repertoire\_courant]/question1/outputX.txt où  $X \in \{0,1,2,3\}$ . Chaque fichier contient l'output de chacun des inputs extraits du fichier unknown.fasta

*python question1.py***Question 2**

Sequence 0 : M|cat|

Sequence 1 : R|tct|

Sequence 2 : pas possible, pas assez de similarite

Sequence 3 : R|tct

**Question 3****Question 4**

La taille de la graine va influencer l'algorithme de recherche de kmer dans les séquences de la base de donne, plus le kmer est grand moins de comparaison on va faire, car on va arrêter les comparaisons dès qu'il y a un mismatch si la graine à cette position est égale à 1. Plus il y a de 1 dans la graine, plus l'algorithme sera rapide. De plus, plus on a de 1 et plus la taille du kmer est grand, plus la précision augmente, mais l'algorithme risque de rater des résultats pertinents : Plus on a de 1, plus la sélection des HSP sera stricte.

## **Question 5**

Tout comme la question 1, il faut lancer le programme question5.py pour générer les fichiers output dans le répertoire question5. On obtient beaucoup plus de résultats pertinents avec cette version. Ceci est normal, car on a introduit des 0 dans la graine, ce qui rend l'algorithme moins strict sur les positionnements des HSP que celui par défaut, qui a onze 1.