

Exercices d'algorithmique

Application à la biologie

Langage : Python

1. Énoncé :

Quelle est la taille en acides aminés (aa) d'une protéine, sachant que la taille en nucléotides (nt) du gène associé est de 2778 nt ? On rappelle que le gène commence par ATG et se termine par un codon stop TAA, TAG ou TGA.

2. Étape de réflexion :

À partir de l'énoncé on doit définir les flux entrants (données), les flux sortants (résultats) et le moyen de passer des uns aux autres.

Données : taille du gène **Tg**, en nt

Sortie : taille de la protéine **Tp**, en aa

Traitement : pour obtenir la taille de la protéine, il faut retrancher la taille du codon stop à la taille du gène puis diviser cette taille par trois (1 aa = 3 nt), donc $Tp = (Tg - 3) \div 3$

3. Conception de l'algorithme

définir Tg

$Tp = (Tg - 3) / 3$

Ecrire Tp

4. Ecriture du programme

```
Tg = 2778
```

```
Tp = (Tg - 3) / 3
```

```
print Tp
```

5. Validation par la phase de tests

Pour valider le programme il faut effectuer un certain nombre de tests en modifiant les données initiales, ici la taille du gène (Tg).

Que constatez-vous si l'on prend $T_g=15$ et $T_g=16$?

Que se passe-t-il si la taille du gène n'est pas multiple de 3 ?

On peut constater qu'il n'y a pas de message d'erreur alors que ce gène a une taille non « traduisible en acides aminés ».

Dans python, une division réalisée avec 2 entiers $x \div y$ ($x, y \in \mathbb{N}$) est une division entière, ($T_g - 3$) et 3 étant tous deux des entiers, on obtient 4 au lieu de 4.333333. Il faut donc rajouter un test afin de vérifier si T_g est un multiple de 3.

Que se passe-t-il quand on réalise la division suivante : « $(T_g - 3) / 3.0$ »

On obtient bien 4.33333. Ainsi, si au moins l'un des 2 nombres x ou y (numérateur ou dénominateur) est un « floatant » (x ou $y \in \mathbb{R}$), python réalise une division « réelle ».

Pour vérifier qu'un nombre est multiple de 3, on peut donc comparer le résultat de la division par un entier (3) avec le résultat de la division par un réel (3.0).

Qu'affiche python quand on tape « $13/3 == 13/3.0$ » et « $12/3 == 12/3.0$ » dans l'interpréteur de commandes ?

Rajoutez un test à votre programme qui vérifie que la taille du gène est bien multiple de 3 et dans ce cas réalise le calcul de T_p , sinon le programme affiche « taille du gene non multiple de 3 ».

Exercices de programmation

Pour les exercices suivants vous réaliserez la conception (définition des données, des résultats et du traitement) par écrit avant de programmer.

1. L'ADN d'*Escherichia coli* fait 4 639 675 paires de bases (pb). En phase de croissance, la bactérie se divise toute les heures environ. Combien de paires de bases d'ADN peut-on trouver dans le milieu après 8 heures de culture ? On affichera pour chaque heure, le nombre de cellules et le nombre de pb total.
2. Les polymérases, malgré leur précision, peuvent faire des erreurs de réplication. On estime qu'une pb sur 10^6 est mutée sur les brins d'ADN nouvellement synthétisés. Combien de pb sont mutées à chaque temps ? Du coup, combien de mutations sont cumulées au cours du temps ? On supposera insignifiant le nombre de pb mutées 2 à plusieurs fois. On rajoutera à l'affichage existant le nombre de pb mutées total.
3. Ces mutations ont une probabilité de 0.001 d'être délétères. On réalisera cette fois-ci

un tirage aléatoire pour déterminer le nombre de mutations délétères apparaissant à chaque cycle.

Pour cela, il faut, pour chaque pb mutée, tirer un nombre aléatoire entre 0 et 1. Si ce nombre aléatoire est inférieur à 0.001 alors on comptera une mutation délétère de plus.

Pour tirer un nombre aléatoire entre 0 et 1, il faut (i) importer le module random (`import random`) en début de programme, (ii) utiliser la fonction `random()` du module random. Tapez plusieurs fois dans l'interpréteur `random.random()`, que constatez-vous ?

Rajoutez à chaque temps le nombre de mutations délétères cumulées au cours du temps juste après le nombre de mutations.

4. À partir du moment où une mutation délétère apparaît, une bactérie meure. Effectuer les corrections sur toutes les variables adéquates à chaque temps selon le nombre de mutations délétères apparues. On estimera qu'il y a environ le même nombre de pb mutées par bactérie.