Avant de commencer, quelques notions clés :

* **FASTA** est un format de fichier texte simple qui représente une séquence d’ADN (ou de protéines). La première ligne commence par « > » suivi d’un identifiant, puis viennent, sur les lignes suivantes, les lettres A, T, C et G qui forment la séquence nucléotidique.
* **BLAST** (Basic Local Alignment Search Tool) est un outil en ligne qui compare votre séquence à celles de bases de données publiques. Il retrouve les meilleures correspondances en évaluant le pourcentage d’identité et la valeur e (un indicateur statistique de la fiabilité de la recherche).

**Instructions**

1. **Connectez-vous** sur BLAST (blastn)
2. **Copier-collez** votre fichier FASTA *mystery\_gene.fasta*
3. **Lancez** la recherche pour identifier le gène et l’espèce.

Pensez bien à noter le nom du gène, l’organisme, le pourcentage d’identité et la valeur e pour en discuter ensuite.

1. Quel gène mystère avez-vous identifié ?Indiquez son nom et l’espèce correspondante.
2. Quel est le pourcentage d’identité et la valeur E associés au meilleur hit ? Expliquez ce que ces deux chiffres signifient pour la fiabilité de la recherche.