



Savanturiers

● Défis de la rentrée en sciences

La génétique du surimi



NIVEAU
PRIMAIRE

NIVEAU
COLLEGE

NIVEAU
LYCÉE

AFFPER
créer et transmettre

CONTRIBUTIONS:

Quentin GOUIL, Stéphane DEBOVE et l'équipe AFPER.

Table des matières

Pour les enseignants

Présentation du défi

Déroulé de l'activité

Liste du matériel

Conseils pour un bon déroulement de l'activité

Table des matières

Pour les élèves

Niveau primaire

Niveau collège

Niveau lycée

Présentation et objectifs du défi

Documents à votre disposition

Instructions

Présentation du défi

Ce défi a pour but d'aider le généticien Quentin Gouil à identifier les espèces de poissons présentes dans un bâton de surimi, à partir des séquences génétiques trouvées dans ce bâton. Ce défi est conçu pour être réalisé en une après-midi par des élèves d'école élémentaire, de collège ou de lycée, en adaptant la version choisie :

Version simple

NIVEAU
PRIMAIRE

NIVEAU
COLLEGE

Les élèves recevront des séquences génétiques simplifiées (des suites de 30 à 70 caractères A,T,C,G) qu'ils devront comparer à des séquences de référence fournies. Le travail consistera à aligner et faire correspondre les séquences pour trouver les séquences avec le plus haut pourcentage de similarité.

Version complexe

NIVEAU
COLLEGE

NIVEAU
LYCÉE

L'objectif sera le même mais les élèves devront utiliser le service en ligne BLAST <https://blast.ncbi.nlm.nih.gov/> (« Basic Local Alignment Search Tool ») pour identifier les espèces présentes dans le surimi. Il leur sera fourni un fichier contenant l'ensemble des séquences d'ADN détectées par Quentin dans le surimi et ils devront les copier/coller dans l'outil en ligne, puis sélectionner la meilleure correspondance trouvée sur la base de différents indicateurs de similarité.



Déroulé de l'activité

- Regardez avec vos élèves la vidéo de Quentin Gouil.



- Facultatif : expliquez à vos élèves les mots ou concepts utilisés par le chercheur que vos élèves pourraient ignorer.

- Fournissez-leur les pages de cette brochure détaillant les objectifs, le matériel et les instructions pour les élèves (version simple ou complexe).

- Laissez-leur du temps pour réaliser le défi.

- Débriefez l'activité en leur demandant de discuter et d'interpréter leurs observations.



Liste du matériel

Version simple

NIVEAU
PRIMAIRE

NIVEAU
COLLEGE

1. Impression des séquences de nucléotides :

- Pour chaque groupe d'élèves, imprimez les séquences d'ADN à identifier fournies ci-dessous.
- Fournissez également les séquences de référence avec le nom des espèces déjà connues.

2. Matériel pour comparer :

- Feuilles de papier, stylos et règles pour faciliter l'alignement et la comparaison des séquences

Liste du matériel

Version complexe

NIVEAU
COLLEGE

NIVEAU
LYCÉE

1. Fichier séquences_adn_surimi.txt :

- Téléchargez le fichier séquences_adn_surimi.txt et assurez-vous qu'il soit disponible pour les élèves.

2. Ordinateurs ou tablettes :

- avec accès à Internet pour utiliser le service BLAST https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome.
- avec un logiciel permettant d'ouvrir le fichier .txt (un logiciel de traitement de texte comme Word/Libre office ou un éditeur de texte comme Notepad++ suffit).

Conseils pour un bon déroulement de l'activité



- **Choix de la version du défi** : si vous enseignez au collège, les deux versions du défi (simple et complexe) peuvent convenir. Choisissez celle que vous préférez en fonction de l'âge exact et de la débrouillardise de vos élèves. La version simple sera plus facile à mettre en place mais plus éloignée du travail des chercheurs.
- **Vos élèves n'ont pas besoin de connaissances particulières en génétique pour réaliser ce défi.** Il leur suffit de savoir que tout être vivant est caractérisé par une séquence d'ADN dont une partie est commune à d'autres espèces. En comparant les séquences des espèces (représentées par des suites de lettres A,T,C,G), on peut donc arriver à les différencier.
- Dans la version simple du défi, l'identification des espèces ne présente pas de difficulté particulière : elle se fait par simple alignement des séquences (les découper et les faire glisser parallèlement deux à deux sur le bureau jusqu'à ce qu'elles s'alignent) et calcul du nombre de différences. La seule chose sur laquelle vous devrez peut-être insister est qu'il ne faut pas nécessairement chercher une séquence référence identique à la séquence à identifier (la séquence à identifier pouvant contenir des erreurs de séquençage) mais simplement une séquence référence avec le plus grand pourcentage de similarité.

Conseils pour un bon déroulement de l'activité



Pour la version complexe, nous vous recommandons de vous familiariser avec l'outil BLAST et l'interprétation de ses résultats avant de réaliser l'activité en classe. Commencez par regarder cette vidéo <https://www.youtube.com/watch?v=yQvG2ojWZjo> puis suivez ces consignes :

- Accédez à BLAST : rendez-vous sur le site BLAST NCBI . Il existe différents types d'analyse BLAST, celle que vous devez utiliser est Blastn (celle sélectionnée par défaut si vous cliquez sur le lien ci-dessus).
- Copier-coller une séquence d'ADN : ouvrez le fichier séquences_adn_surimi.txt et copiez-collez une séquence d'ADN (une des lignes composées de lettres A,T,C,G) dans le champ « Enter accession number(s), gi(s), or FASTA sequence(s) » (voir capture d'écran 1). Cliquez ensuite sur « BLAST ».
- Au bout de quelques secondes, les résultats apparaissent sous forme de tableau (voir capture d'écran 2). L'outil fait ressortir toutes les séquences de sa base de données qui ressemblent à la séquence que vous avez entrée. Il est normal que des séquences d'espèces différentes apparaissent, car les espèces partagent de nombreux gènes. Le plus important est de déterminer l'espèce qui ressemble le plus. Pour vous aider dans cette tâche, voilà les colonnes les plus intéressantes à analyser :

Conseils pour un bon déroulement de l'activité



- Score : indique la qualité de l'alignement entre la séquence à identifier et une séquence de référence de la base de données. Plus le score est élevé, meilleure est la correspondance. Vous pouvez ne vous focaliser que sur cet indicateur et ignorer les autres colonnes pour simplifier l'activité.
- Query cover : représente, en pourcentage, la longueur de séquence à identifier qui a réussi à être alignée avec la séquence de référence.
- Percentage identity : indique le pourcentage de nucléotides identiques entre la séquence à identifier et la séquence de référence, à l'intérieur de l'alignement trouvé. Un pourcentage élevé signifie une forte similarité.

Ne vous inquiétez pas si vous trouvez de l'ADN de plancton ou de bactéries dans les résultats ! Cela ne signifie pas forcément que le bâton de surimi était impropre à la consommation, car une correspondance ne veut pas dire qu'une espèce est présente en grande quantité. Des bactéries, virus et planctons peuvent se retrouver à l'état de traces.

NIVEAU
PRIMAIRE

NIVEAU
COLLEGE

Présentation du défi

Aidez Quentin Gouil à identifier les espèces de poissons présentes dans un bâton de surimi !

La génétique est une science passionnante qui nous permet de comprendre la diversité de la vie sur Terre. Chaque espèce a une séquence génétique unique, un peu comme une empreinte digitale. En étudiant ces séquences, les scientifiques peuvent identifier les espèces, comprendre leurs relations et même découvrir de nouvelles espèces !



NIVEAU
PRIMAIRE

NIVEAU
COLLEGE

En quoi consiste le défi

On va vous donner des séquences génétiques que Quentin a trouvé dans le surimi. Les séquences génétiques sont comme des codes secrets qui caractérisent chaque poisson. Chaque séquence est composée de quatre lettres : A, T, C, et G. Par exemple :

ATGGCAATTAGGC

Vous allez aussi avoir des « séquences de référence », qui sont des séquences d'ADN pour lesquelles on connaît déjà l'espèce.

Votre mission sera de trouver quelles séquences de référence sont les plus proches des séquences à identifier, pour pouvoir déterminer quels poissons ont été mis dans le surimi !



Présentation du défi

NIVEAU
COLLEGE

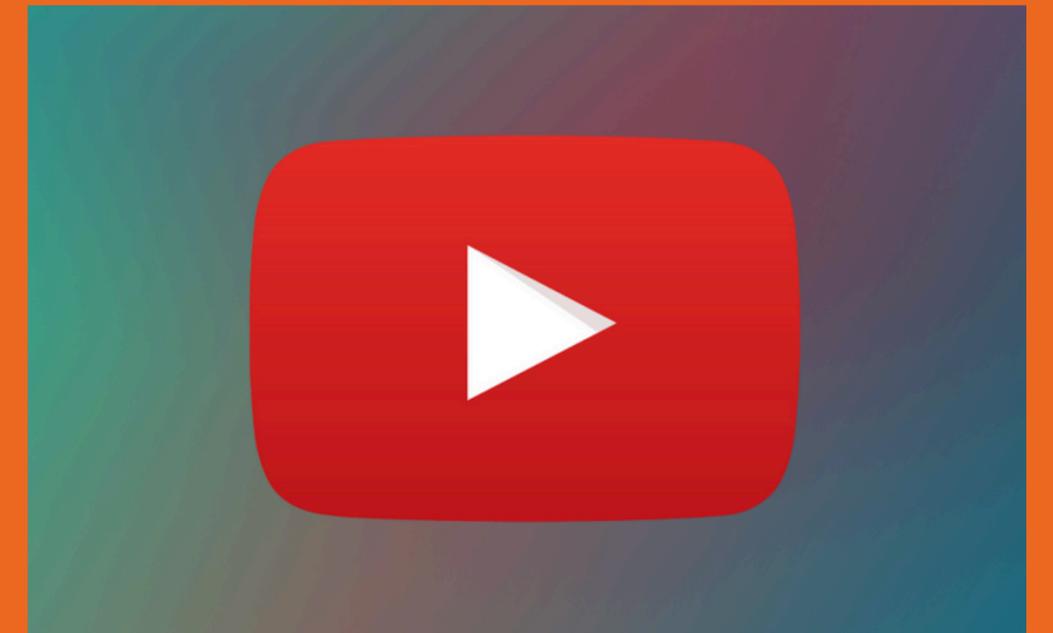
NIVEAU
LYCÉE

Aidez Quentin Gouil à identifier les espèces de poissons présentes dans un bâton de surimi !

La génétique est une science passionnante qui nous permet de comprendre la diversité de la vie sur Terre. Chaque espèce a une séquence génétique unique, un peu comme une empreinte digitale. En étudiant ces séquences, les scientifiques peuvent identifier les espèces, comprendre leurs relations et même découvrir de nouvelles espèces !

En quoi consiste le défi

Vous allez recevoir un fichier .txt contenant toutes les séquences génétiques que Quentin a trouvé en analysant un échantillon de surimi. Votre mission sera d'utiliser le service en ligne BLAST qu'utilisent les chercheurs eux-mêmes pour identifier à quelles espèces appartiennent ces séquences.



Documents à votre disposition

NIVEAU
PRIMAIRE

NIVEAU
COLLEGE

- **Des séquences génétiques à identifier** : des séquences de lettres (A,T,C,G) qui représentent des morceaux d'ADN des espèces présentes dans le surimi.
- **Des séquences de référence** pour lesquelles on connaît le nom de l'espèce associée
- **Feuilles et stylos** : pour recopier les séquences et pouvoir facilement les comparer.

Documents à votre disposition

NIVEAU
COLLEGE

NIVEAU
LYCÉE

- **Ordinateur avec accès à Internet** : pour accéder au service BLAST à l'adresse https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome.
- **Fichier sequences_adn_surimi.txt** : ce fichier contient toutes les séquences génétiques que Quentin a trouvées dans le surimi.
- **Bloc-notes ou logiciel de traitement de texte** : nécessaire pour ouvrir le fichier .txt
- Une vidéo pour vous familiariser avec le service BLAST : <https://www.youtube.com/watch?v=yQvG2ojWZjo>

Instructions

NIVEAU
PRIMAIRE

NIVEAU
COLLEGE

Étape 1 : aligner les séquences fournies

Découpez les séquences d'ADN à identifier. Observez-les bien, puis faites-les « glisser » sur la table en face de chacune des séquences références, jusqu'à trouver une position où beaucoup de lettres sont les mêmes. Bravo, vous avez réussi à « aligner » les séquences !

Étape 2 : calculer le nombre de différences

Même si deux séquences sont alignées et partagent beaucoup de lettres, quelques-unes ne correspondront pas. Calculez combien il y a de telles positions où deux lettres sont différentes. Faites ensuite la même chose avec la séquence référence suivante : alignez la séquence à identifier, puis calculez le nombre de différences. La séquence de référence qui ressemble le plus à la séquence à identifier (qui a le moins de lettres différentes) sera celle du poisson que l'on cherche !

Instructions

NIVEAU
PRIMAIRE

NIVEAU
COLLEGE

Étape 3 : identifier chaque poisson

Faites la même chose pour chaque séquence à identifier : alignez-la avec chaque séquence de référence pour trouver le poisson dont il s'agit. Notez le nom de l'espèce de poisson que vous avez identifiée dans un tableau comme ça :

Séquence à identifier	Espèce identifiée
ATCGTACGATCGTACG	Poisson A
CGTACGATCGTACGAT	Poisson B

Instructions

NIVEAU
COLLEGE

NIVEAU
LYCÉE

Le service BLAST (Basic Local Alignment Search Tool) va vous permettre de comparer les séquences d'ADN trouvées dans le surimi avec celles d'une base de séquences connues. Commencez par regarder cette vidéo pour vous familiariser avec l'outil : <https://www.youtube.com/watch?v=yQvG2ojWZjo>

Puis, voici comment procéder en détail :

- **Accédez à BLAST** : rendez-vous sur le site [BLAST NCBI](#). Il existe différents types d'analyse BLAST, celle que vous devez utiliser est Blastn (celle sélectionnée par défaut si vous cliquez sur le lien ci-dessus).
- **Copiez-collez une séquence d'ADN** : ouvrez le fichier `sequences_adn_surimi.txt` et copiez une séquence d'ADN (une des lignes composées de lettres A,T,C,G) dans le champ « Enter accession number(s), gi(s), or FASTA sequence(s) » (voir capture d'écran 1). Cliquez ensuite sur « BLAST ».

Instructions

NIVEAU
LYCÉE

- Au bout de quelques secondes, les résultats apparaissent sous forme de tableau (voir capture d'écran 2). L'analyse fait ressortir toutes les séquences de la base de données qui ressemblent à celle que vous avez entrée. Il est normal que des séquences de plusieurs espèces puissent ressortir, car différentes espèces partagent de nombreux gènes. Le plus important est de déterminer la séquence avec la plus forte similarité. Pour cela, voici les colonnes les plus intéressantes à analyser :
 - **Score** : indique la qualité de l'alignement entre la séquence à identifier et la séquence de référence dans la base de données. Plus le score est élevé, meilleur est l'alignement. Vous pouvez ne vous focaliser que sur cet indicateur et ignorer les autres colonnes pour simplifier l'activité. La correspondance avec le score le plus élevé est probablement l'espèce de poisson que vous recherchez. Vous n'avez plus qu'à chercher sur internet le nom de l'espèce en latin pour savoir de quel poisson il s'agit.
 - **Query cover** : représente, en pourcentage, la longueur de séquence à identifier qui a réussi à être « alignée » avec la séquence de référence.
 - **Percentage identity** : indique le pourcentage de nucléotides (les lettres A,T,C,G) identiques entre la séquence à identifier et la séquence de référence, à l'intérieur de l'alignement. Un pourcentage élevé signifie une forte similarité.
 -

Instructions

NIVEAU
LYCÉE

- **Notez vos résultats**

Pour chaque séquence analysée, notez vos résultats dans un tableau comme celui-ci :

Numéro de la ligne du fichier	Espèce identifiée	Score
4813	Poisson A	98.5
4818	Poisson B	97.2

- **Après analyse d'un certain nombre de séquences, partagez vos résultats avec vos camarades. Avez-vous trouvé les mêmes espèces qu'eux ? Y en a-t-il qui vous ont étonné ?**