

Compte rendu

Initiation à la Bioinformatique TP2. Applications de la bioinformatique

- Tous les champs du formulaire doivent être remplis ;
- Une fois que votre compte rendu est complété, sauvegardez le en appuyant sur "enregistrer" en haut à gauche de votre lecteur pdf (Veuillez changer le nom du fichier).

Date

Adresse électronique

Nom et Prénom

1. Application en phylogénie: construction d'arbres phylogénétiques (Dendrogrammes) avec MEGA 6.06

1-1-Choix des données: les séquences utilisées en phylogénie sont des séquences nucléiques ou protéiques appartenant au même type de marqueurs moléculaires (molécules ayant un faible taux de mutation).

Exemple: ARN 16 S L'analyse des séquences du gène d'ARNr 16S est une des méthodes les plus utilisées pour identifier et caractériser des espèces ou des communautés procaryotes dans des domaines aussi divers que l'écologie, la médecine ou la microbiologie alimentaire. Dans cet exemple, nous souhaitons établir les liens de parenté entre 23 taxons (UTO) représentant 06 genres de la famille des *Bacillaceae* par l'analyse phylogénétique de leurs séquences ARNr 16S respectives (voir Annexe 2).

1-2-Installez le logiciel [MEGA 6.06](#) à partir du fichier MEGA6.06_setup.exe

1-3-Alignement des séquences: réalisez un alignement de toutes les séquences avec l'outil Clustal W de MEGA, une fois fait, enregistrez le fichier de l'alignement sous le format .MEG

1-4-Utilisez le fichier de l'alignement pour construire un arbre phylogénétique selon la méthode du Neighbor-Joining (NJ)

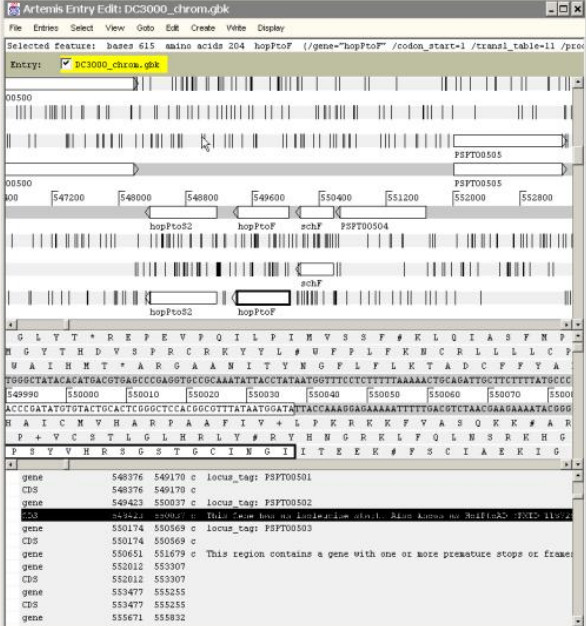
1-5-Commentez les liens de parenté entre les genres de bacilles de l'arbre.

Quel est le taxon le plus proche de la souche *Bacillus amyloliquefaciens* strain XS-1 . Justifiez?

2. Application dans l'annotation des séquences: annotation d'une séquence ADN avec le logiciel [ARTEMIS](#)

ARTEMIS est un logiciel de visualisation de séquences ADN et un outil d'annotation qui permet la visualisation des caractéristiques structurales et fonctionnelles de séquences génomiques partielles ou complètes.

2.1. Lancez Artemis et ouvrez le fichier "Artemis exemple.txt" contenant une séquence génomique en allant à File---> Open ---> Choisir d'afficher tous les fichiers puis cliquer dessus. La fenêtre d'ARTEMIS s'affiche sur l'écran :



The screenshot displays the Artemis software interface. The main window shows a genomic sequence with various features and annotations. The top panel, labeled "Overview window", shows a summary of the sequence with features like "hopPtoF" and "schF" overlaid on the DNA strands. The middle panel, labeled "DNA view window", shows the sequence in detail, including the DNA strands and the translation of the sequence into amino acids. The bottom panel, labeled "Feature list", shows a list of features with their coordinates and descriptions. The "Selected feature" panel on the right shows details for the selected feature, "hopPtoF".

Selected feature (hopPtoF)

Active entry
annotated sequence of DC3000
chromosome downloaded from
NCBI

Overview window
shows features specified in the
active entry (in this case, genes
and CoDing Sequences or CDS)
overlaid on the two DNA strands
and six translation frames

DNA view window
genes and CDSes as above,
shown on the sequence level

Feature list
shows annotation record for
all features in the active entry

La partie du haut, "Overview", montre une vue générale du fragment génomique présentant les deux brins d'ADN sous forme de ligne grise. Les 3 lignes gris clair du haut représentent les 3 cadres de lecture sur le brin direct (+) et les 3 lignes du bas montrent les 3 cadres sur le brin complémentaire (-). Chaque ligne noire verticale représente un codon stop. Le scroll vertical permet de faire un zoom sur la séquence et le scroll horizontal, de se déplacer sur la séquence. Vous remarquez que les positions depuis le début de la séquence sont indiquées.

La partie centrale, DNA view, présente une vue plus détaillée d'une région génomique. On y voit les deux brins d'ADN sur la ligne grise, et la traduction pour chacun des cadres de lecture du brin directe (au dessus) et du brin complémentaire (au dessous).

Vous pouvez sélectionner un acide aminé pour voir le codon correspondant. Remarquez que si vous sélectionner une région d'ADN, la séquence en acide aminé correspondante (dans le bon cadre de lecture) est sélectionnée aussi, ainsi que la portion correspondante dans la partie "Overview".

Si vous double-cliquez sur une région du panneau du haut, le panneau du milieu affiche automatiquement la vue détaillée pour cette région. Vous voyez d'ailleurs apparaître, sous le menu, dans la partie "Selected feature", les coordonnées de la portion sélectionnée.

Dans la vue "DNA view", les codons stops sont représentés par 3 symboles différents, lesquels ?

Quelle est la longueur de la séquence et comment l'avais vous trouvé?

Artemis permet de trouver automatiquement les ORFs. Afficher les ORFs de taille minimale 100 :

Menu Create > Mark Open Reading Frame

Artemis a créé une nouvelle entrée ORF_100+, qui est maintenant l'entrée active et contient toutes les ORFs.

Combien d'ORF trouvez-vous ?

3. Application dans l'annotation des séquences: Prédiction de la structure 3D des protéines avec l'outil Swiss-Model

Parmi les objectifs de l'annotation des séquences protéiques est de prédire la structure tertiaire et quaternaire des protéines. Un des modèles bioinformatiques disponibles, [Swiss-Model](#), utilise une méthode de modélisation par homologie. C'est l'un des modèles les plus précis et les plus utilisés.

L'actine est une protéine globulaire jouant un rôle important dans la structure et les mouvements cellulaires chez les eucaryotes. La plupart des cellules eucaryotes possèdent de nombreux gènes d'actine codant des protéines légèrement différentes.

Dans cet exemple, nous tentons de prédire la structure de l'actine alpha 1, l'une des isoformes d'actine identifiées chez l'homme, en utilisant Swiss-Model. La séquence de cette protéine est donnée ci dessous:

> Homo sapiens alpha skeletal muscle actin

```
MCDEDETTALVCDNGSGLVKAGFAGDDAPRAVFPISVGRPRHQGVMMVGMGQKDSYVGDEAQS1SKRGILTLKYP2IE
HGIITNWDDMEKIWHHTFYNELRVAP3EEHPTLLTEAPLNPKANREKMTQIMFETFNVPAMYVAIQAVLSLYASGR
TTGIVLDSGDGVTHNVWIYEVYALVHAIMRLDLAGRDLLDYL4MKILTERGYSFVTTAEREIVRDIKEKLCYVALD
FENEMATAASSSSLEKSYELPDGQVITIGNERFRCPETL5FQPSFIGMESAGIHETTYNSIMKCDIDIRKDLYANNVM
SGGTTMYPGIADRMQKEITALAPSTMKIKIIAPPERKYSVWIGGSILASLSTFQQMWITKQEYDEAGPSIVHRKCF
```

3-1-Lancez la modélisation sur Swiss Model :

3-2-Copiez la séquence FASTA dans la case correspondante et lancez la recherche de modèles (Templates)

3-3-Quel est le nombre de modèles qui correspondent à notre séquence?

Choisissez le modèle ayant donné le maximum de similarité pour prédire la structure 3D de l'actine alpha 1

3-4-D'après le résultat obtenus, quel est le % de similarité entre le modèle et l'actine alpha 1

Remarques:

- Tous les champs doivent être remplis ;

- Une fois que votre compte rendu est complété, sauvegardez le en appuyant sur "enregistrer" en haut à gauche de votre lecteur pdf (Veuillez changer le nom du fichier).

