
Projet L2D -Introduction à la bioinformatique

Manuel d'utilisation

Projet de Programmation

L2D1



Illustration 1: Image de présentation

Projet L2D -Introduction à la bioinformatique

Les informations d'identification du document :

Référence du document :
Version du document : 2
Date du document : 01/05/2021
Auteurs : Mehdi Hamiche Manal Boutajar Adelin Bodnar

Les éléments de vérification du document :

Validé par :
Validé le :
Soumis le :
Type de diffusion :
Confidentialité :

Les éléments d'authentification :

Maître d'ouvrage :	Chef de projet :
Date / Signature :	Date / Signature :

Projet L2D -Introduction à la bioinformatique

Sommaire

<u>1. Introduction</u>	5
1.1. Objectifs et méthodes	5
1.2. Documents de référence	6
<u>2. Guide de lecture</u>	7
2.1. Maîtrise d'œuvre	7
<i>2.1.1. Responsable</i>	7
<i>2.1.2. Personnel administratif</i>	7
<i>2.1.3. Personnel technique</i>	7
2.2. Maîtrise d'ouvrage	7
<i>2.2.1. Responsable</i>	8
<i>2.2.2. Personnel administratif</i>	8
<i>2.2.3. Personnel technique</i>	8
<u>3. Concepts de base</u>	9
<u>4. Mise en oeuvre</u>	11
<u>5. Liste des commandes</u>	16
<u>6. Messages d'erreur</u>	17

Projet L2D -Introduction à la bioinformatique

<u>7. Annexes</u>	18
<u>8. Glossaire</u>	19
<u>9. Références</u>	20
<u>10. Index</u>	21

1. Introduction

Le manuel d'utilisation est un document qui décrit l'ensemble des fonctionnalités de notre application ainsi que les concepts de base de notre logiciel avec les actions à réaliser afin de la mettre en œuvre. Cela doit permettre à l'utilisateur final d'exploiter ses fonctionnalités. Les rubriques du manuel d'utilisation sont les suivantes :

- **Mise en œuvre :** Indique la manière d'utiliser l'application pour réaliser les tâches de base
- **Liste des commandes :** Décrit l'ensemble des commandes de l'application
- **Messages d'erreur :** Décrit les erreurs pouvant survenir et les procédures permettant leur résolution

1.1. Objectifs et méthodes

Permettre à notre programme d'être plus simple à utiliser que ce soit dans la visibilité ou dans la façon d'utiliser.

- Alignement de séquences :

- ◆ protéiques
- ◆ nucléotidiques
 - Global
 - Multiple

- Interface graphique permettant de :

- ◆ voir les alignements globaux et multiples
- ◆ générer le logo

Modules

Projet L2D - Introduction à la bioinformatique

1. *Alignement global* - 2 séquences alignées via l'algorithme de Needleman - Wunsch
2. *Alignement multiple* - amélioration de l'alignement global (consensus) - alignement de trois séquences au moins
3. *Logo* - représentation graphique de l'alignement multiple
4. *Interface graphique* - relie les 3 modules et les présente

1.2. Documents de référence

Les documents du projet servant à l'élaboration du présent document :

- Le cahier des charges,
- Le cahier de recettes,
- La conception générale

2. Guide de lecture

2.1. Maîtrise d'œuvre

La maîtrise d'œuvre présente l'équipe du développement chargé du bon suivi du manuel d'utilisation et des besoins dont le maître d'ouvrage fait commande.

Elle représente l'équipe du développement :

- Adelin Bodnar
- Manal Boutajar
- Mehdi Hamiche

Cette équipe veillera au bon suivi du manuel d'utilisation coordonnées avec la conception générale représentant les besoins des enseignants encadrants.

2.1.1. Responsable

Il est conseillé pour le responsable de la maîtrise d'œuvre de lire le document dans sa totalité afin de prendre conscience de l'ensemble des éléments.

2.1.2. Personnel administratif

Il est conseillé pour le personnel administratif de lire la présentation du produit, mise en œuvre et sur la liste des commandes.

2.1.3. Personnel technique

Il est conseillé pour le personnel technique de prendre en compte la partie sur le concept de base, mise en œuvre et sur les messages d'erreur.

2.2. Maîtrise d'ouvrage

Projet L2D - Introduction à la bioinformatique

La maîtrise d'ouvrage représente dans notre cas le client du projet, c'est-à-dire les personnes dont les besoins permettent la conception du projet.

La maîtrise d'ouvrage est assistée par l'équipe de la maîtrise d'œuvre et donc ce rôle sera assuré par les enseignants encadrants Dragutin Jastrebic et Koviljka Lukic Jastrebic.

2.2.1. Responsable

Il est conseillé pour le responsable de la maîtrise d'ouvrage de lire le document dans toute sa totalité afin de prendre conscience de l'ensemble des documents.

2.2.2. Personnel administratif

Il est conseillé pour le personnel administratif de lire la présentation du produit, mise en œuvre et sur la liste des commandes.

2.2.3. Personnel technique

Il est conseillé pour le personnel technique de prendre en compte la partie sur le concept de base, mise en œuvre et sur les messages d'erreur.

3. Concepts de base

Pour une meilleure compréhension du document, voici quelques définitions :

- **Bio-informatique** : Discipline permettant d'analyser de l'information biologique qu'il y a dans les séquences nucléotidiques et protéiques. La bioinformatique est à la frontière de la biologie, de l'informatique et des mathématiques (bases de données volumineuses, algorithmes performants pour traiter les informations stockées).
- **Différents modules** :
 - ***Alignement global*** : recouvre les séquences alignées (par exemple 2 séquences) sur l'ensemble de leur longueurs,
 - *backtracking* - algorithmes qui permettent de résoudre des problèmes algorithmiques (de satisfaction de contraintes). Ils permettent de tester systématiquement l'ensemble des affectations potentielles du problème.
 - ***Alignement multiple*** : aligner un ensemble de séquences homologues, comme des séquences protéiques ou nucléiques qui assure des fonctions similaires dans différentes espèces vivantes,
 - *séquences consensus* - séquence nucléotidique ou peptidique la plus fréquente à chaque position d'un alignement de séquences.
 - ***Weblogo*** : application web conçue pour rendre la génération de logos de séquence aussi simple que possible. Les logos de séquences sont une représentation graphique d'un alignement de séquences multiples d'acide aminé ou d'acide nucléique,
 - ***Interface Graphique*** : application permettant de relier les 3 modules et les présente de manière simple et efficace.

Projet L2D - Introduction à la bioinformatique

- **Algorithme Needleman-Wunsch** : effectue un alignement global maximal de deux chaînes de caractères. Utilisé en bioinformatique pour aligner des séquences de protéines ou de nucléotides.
- **Format des séquences** : Fasta
- **Fasta** : format de fichier texte dans lesquels les séquences sont affichées par une suite de lettres (séquences protéiques et nucléotidiques). Elle permet de stocker les séquences protéiques et nucléotidiques.
- **Programmation orienté objet** : met en œuvre différents objets (instances de classes). Chaque objet associe des données et des méthodes (fonctions) agissant exclusivement sur les données.

4. Mise en oeuvre

Pour lancer l'interface graphique, il faut avoir un compilateur de langage java (Eclipse, Netbeans ...).

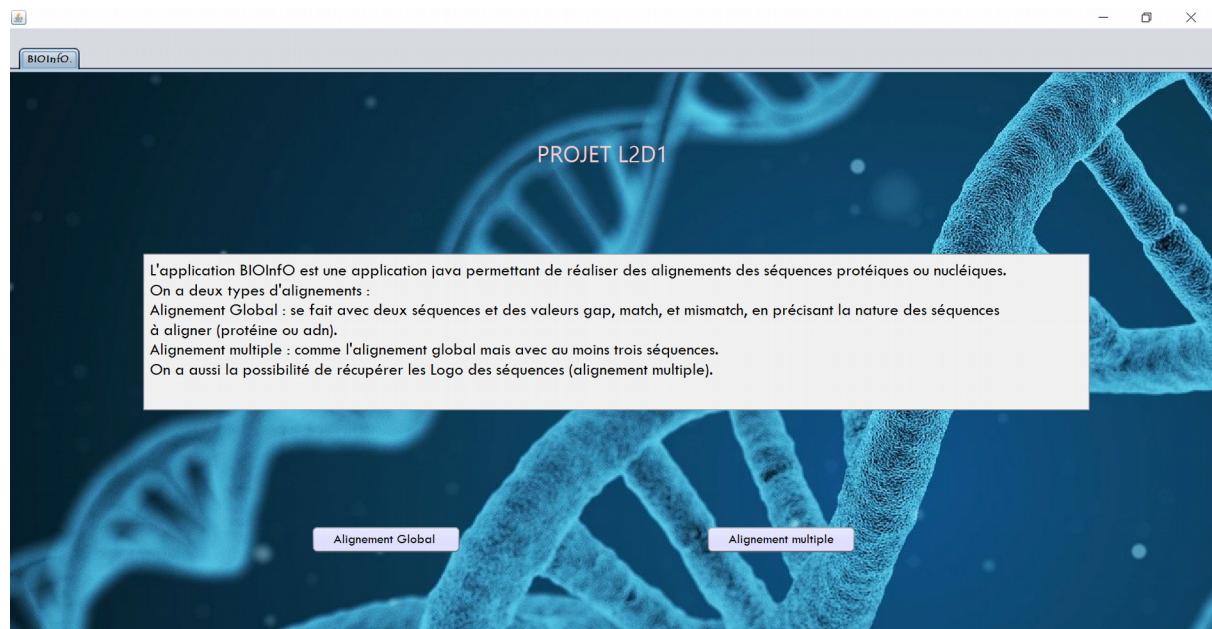


Illustration 2: Accueil de l'interface graphique

Une fois le programme exécuté, la fenêtre s'affichera :

Cette fenêtre contient deux parties :

1. La première est le bouton à gauche pour faire des alignements globaux de séquences protéiques ou nucléotidiques,
2. La deuxième est le bouton à droite pour faire l'alignement multiple des séquences saisies manuellement ou en insérant des fichiers .fasta et également pour récupérer les logo des séquences en format .txt.

Utilisation de l'interface graphique

Lors du choix de faire des alignements globaux, on se retrouve sur la fenêtre ci dessous :

Projet L2D - Introduction à la bioinformatique

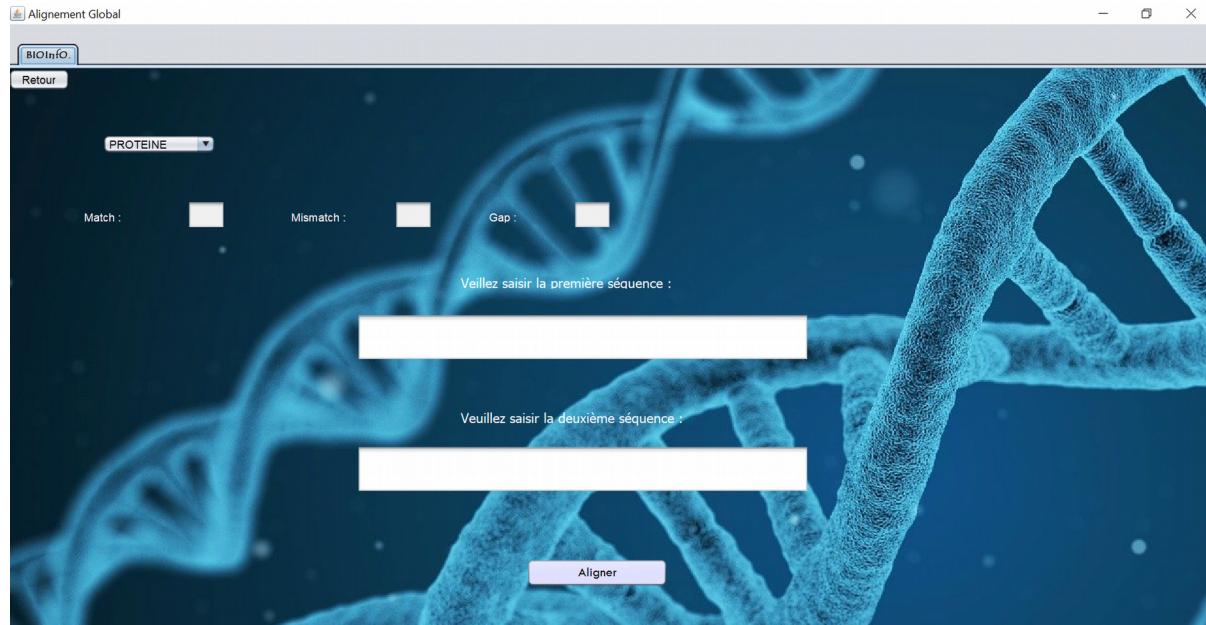


Illustration 3: Interface graphique - Page de l'Alignement Global

1. On remplit la liste qui nous permet de choisir le type de séquence à aligner
 1. “**PROTÉINE**” pour les séquences protéiques,
 2. “**ADN**” pour les séquences nucléotidiques.

Les champs match, mismatch et gap sont à remplir par l'utilisateur.

2. Ensuite, on a deux champs de saisie, un pour la première séquence, l'autre pour la deuxième. On récupère le résultat de l'alignement global sur une nouvelle fenêtre **JFrame** en appuyant sur le bouton “**Alignement Global**”.
3. On peut également faire un retour vers la page d'accueil en appuyant sur le bouton en haut à droite “**Retour**”.

Par exemple :

Projet L2D - Introduction à la bioinformatique

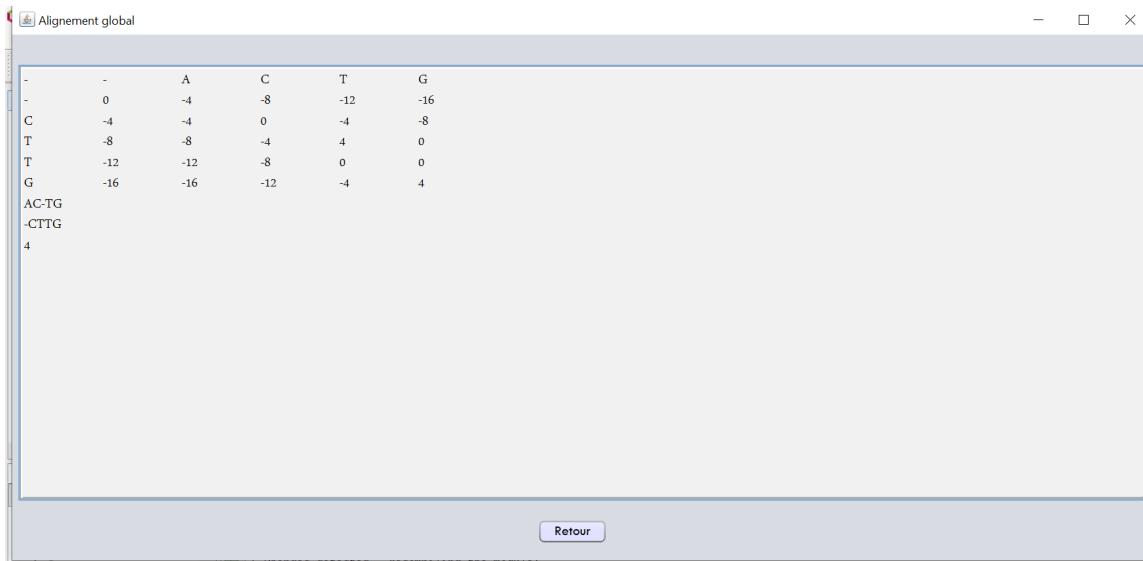


Illustration 4: Interface graphique – Exemple avec Alignement Global

4. Pour l’alignement multiple, on se retrouve sur cette fenêtre :

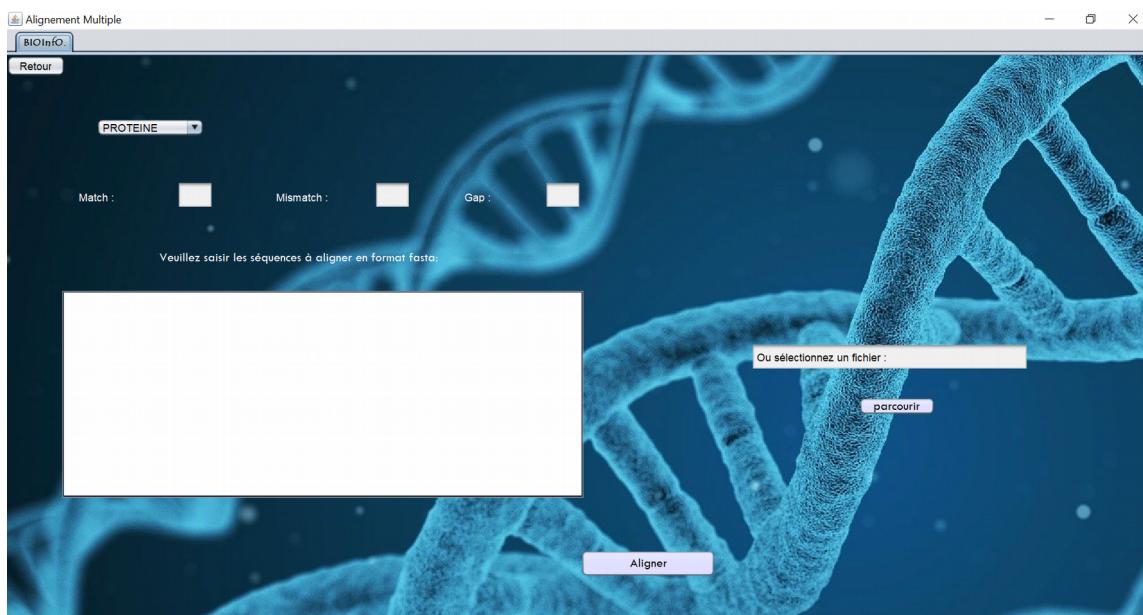


Illustration 5: Interface graphique - Page de l'Alignement multiple

5. on saisit de nouveau les valeurs **gap**, **mismatch**, **match** et on choisit le type de séquences qu’on va aligner (“**PROTÉINE**” ou “**ADN**”), puis on saisit les séquences soit manuellement sur le champs de saisie ou en appuyant sur le

Projet L2D - Introduction à la bioinformatique

bouton “**parcourir**” pour récupérer les séquences et les transmettre au programme qui permet de les aligner et de récupérer le résultat sur une nouvelle fenêtre. On a aussi la possibilité de faire un retour vers la page d'accueil en appuyant sur le bouton en haut à droite “**Retour**”.

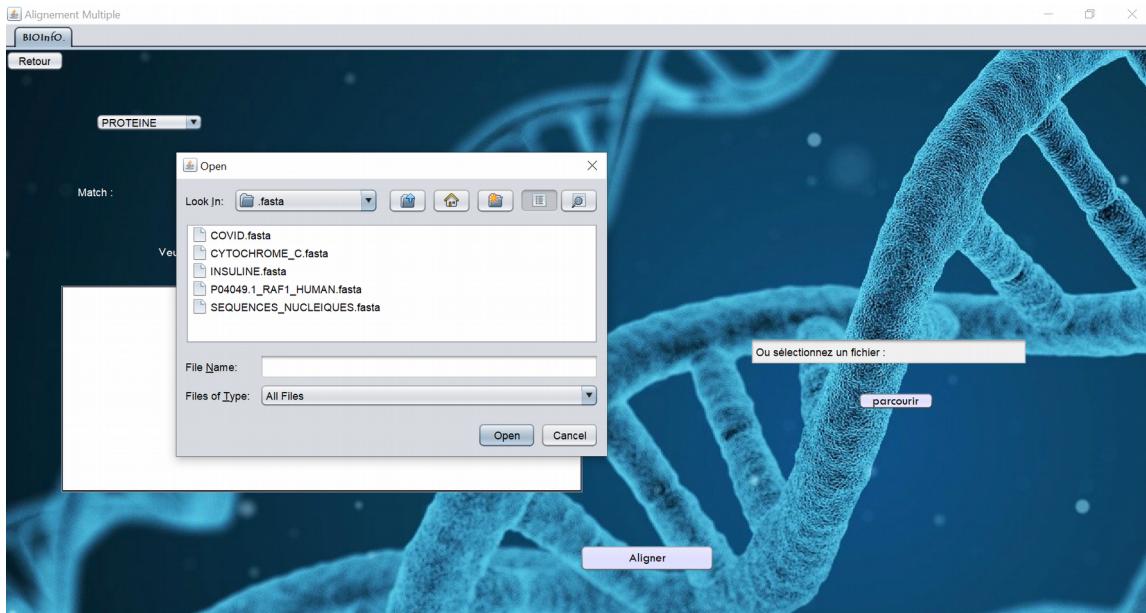


Illustration 6: Interface graphique - Choix de fichier Alignement multiple

Comme dans cet exemple :

Projet L2D - Introduction à la bioinformatique



Illustration 7: Interface graphique - Alignement multiple Cytochrome_C

- Pour le logo, on a le bouton « **LOGO** » en dessous du résultat de l’alignement multiple permettant l’affichage :

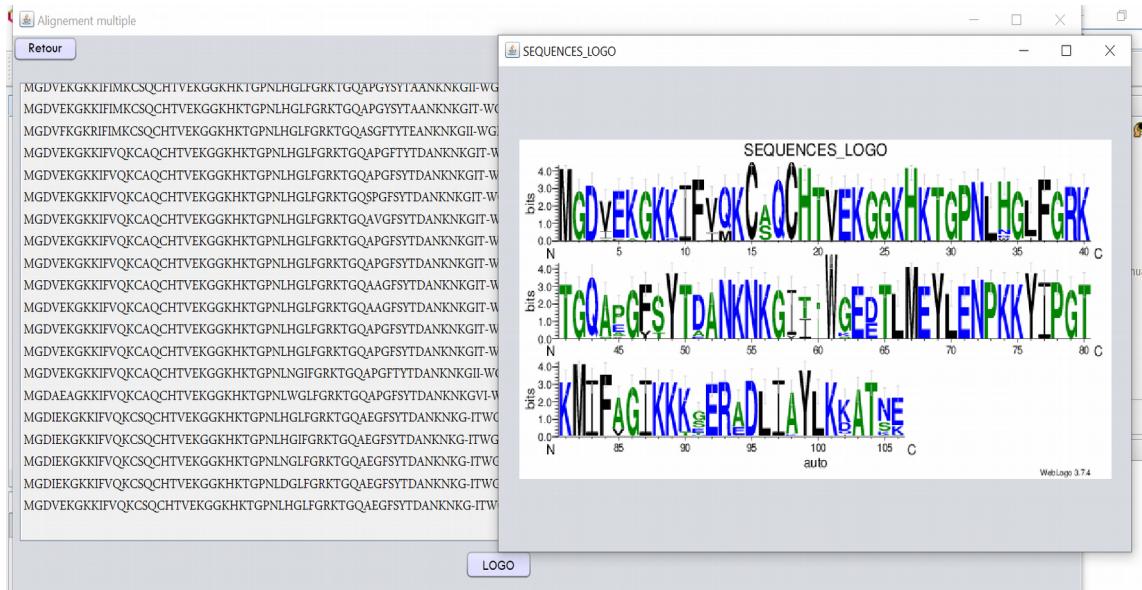


Illustration 8: Interface graphique - Logo Cytochrome_C

5. Liste des commandes

• CTRL + C	➔ Copier
• CTRL + V	➔ Coller
• CTRL + X	➔ Couper
• CTRL + Z	➔ Annuler dernière action
• Souris	➔ Copier - coller possible également
• Touches de l'interface	➔ Exécution des tâches (alignement global, alignement multiple, gap, match, mismatch, type des séquences)

6. Messages d'erreur

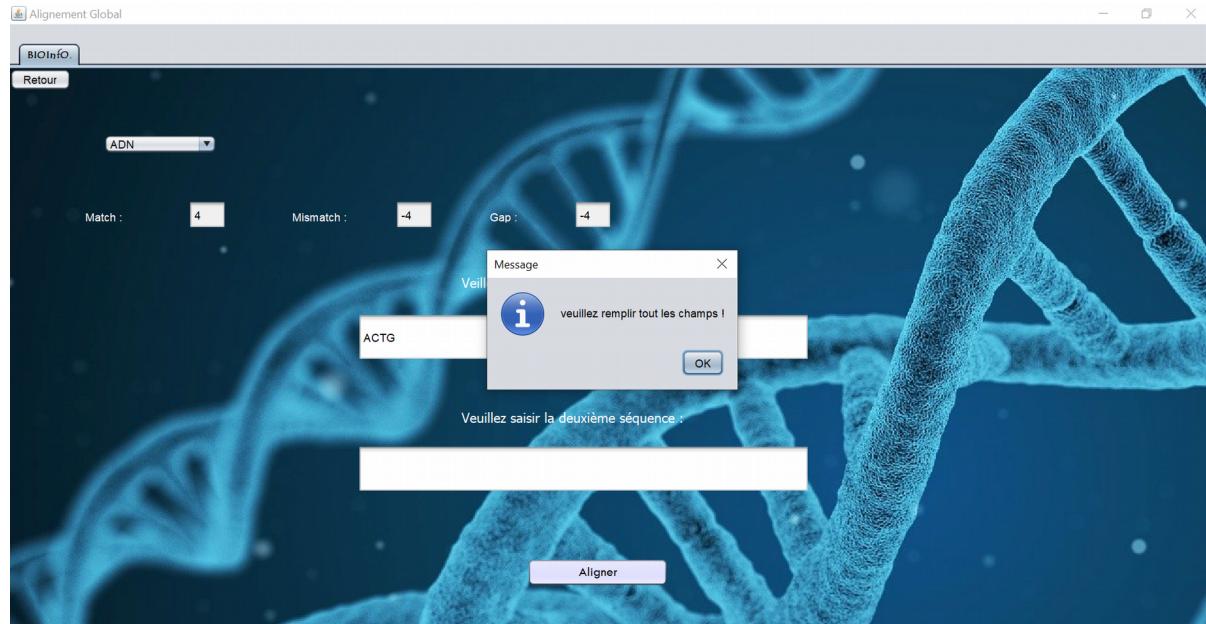


Illustration 9: Interface graphique - Messages d'erreur "veuillez remplir tout les champs"

7. Annexes

- Le cahier des charges
- Le cahier de recettes
- La conception détaillée

8. Glossaire

- ★ **BIO-INFORMATIQUE** : discipline permettant d'analyser de l'information biologique qu'il y a dans les séquences nucléotidiques et protéiques
- ★ **PROGRAMMATION ORIENTÉE OBJET** : met en œuvre différents objets (instances de classes). Chaque objet associe des données et des méthodes (fonctions) agissant exclusivement sur les données
- ★ **ALGORITHME NEEDLEMAN - WUNSCH** : effectue un alignement global maximal de deux chaînes de caractères

- ★ **ALIGNEMENT GLOBAL** : recouvre les séquences alignées (par exemple 2 séquences) sur l'ensemble de leur longueur
- ★ **ALIGNEMENT MULTIPLE** : aligner un ensemble de séquences homologues, comme des séquences protéiques ou nucléiques qui assure des fonctions similaires dans différentes espèces vivantes
- ★ **WEBLOGO** : application web conçue pour rendre la génération de logos de séquence aussi simple que possible
- ★ **INTERFACE GRAPHIQUE** : relie les 3 modules et les présente de manière simple et efficace

- ★ **GAP, MATCH, MISMATCH** : valeurs pour alignement global
- ★ **FASTA** : format de fichier texte dans lesquels les séquences sont affichées par une suite de lettres

9. Références

[Alignement de séquences](#)

[Alignement global](#)

[Alignement multiple](#)

[JAVA - Main \(site 1\)](#)

[JAVA - Main \(site 2\)](#)

[Chaînes de caractères](#)

[Tableau](#)

[Needleman et Wunsch – Java](#)

[Needleman et Wunsch – Python](#)

[Interface graphique en Java](#)

[Interface graphique avec SWING](#)

10. Index

Index des figures

<i>Illustration 1: Image de présentation</i>	1
<i>Illustration 2: Accueil de l'interface graphique</i>	11
<i>Illustration 3: Interface graphique - Page de l'Alignement Global</i>	12
<i>Illustration 4: Interface graphique – Exemple avec Alignement Global</i>	13
<i>Illustration 5: Interface graphique - Page de l'Alignement multiple</i>	13
<i>Illustration 6: Interface graphique - Choix de fichier Alignement multiple</i>	14
<i>Illustration 7: Interface graphique - Alignement multiple Cytochrome_C</i>	15
<i>Illustration 8: Interface graphique - Logo Cytochrome_C</i>	15
<i>Illustration 9: Interface graphique - Messages d'erreur "veuillez remplir tout les champs"</i>	17

Index lexical

ADN	12, 14
alignement de séquences	9
Alignement de séquences	5, 20
alignement g	6
alignement global	10, 12, 19

Projet L2D - Introduction à la bioinformatique

Alignement global	6, 9, 20
Alignement Global	12
ALIGNEMENT GLOBAL	19
alignement multiple	6, 11, 13, 16
Alignement multiple	6, 9, 20
ALIGNEMENT MULTIPLE	19
application	5, 9, 19
backtracking	9
bio-informatique	9
Bio-informatique	9
BIO-INFORMATIQUE	19
bioinformatique	10
consensus	6, 9
Fasta	10
FASTA	19
fasta et également	11
gap	12, 14, 16
GAP	19

Projet L2D - Introduction à la bioinformatique

interface graphique	11
Interface graphique	5 sv, 9, 11, 19 sv
Interface Graphique	9
INTERFACE GRAPHIQUE	19
JFrame	12
alignement global	16
liste des commandes	7 sv
Liste des commandes	5, 16
logo	5, 9, 11, 15, 19
Logo	6
LOGO	15, 19
manuel d'utilisation	5, 7
Manuel d'utilisation	1
match	12, 14, 16
MATCH	19
messages d'erreur	7 sv
Messages d'erreur	5, 17
Mise en œuvre :	5

Projet L2D - Introduction à la bioinformatique

mismatch	12, 14, 16
MISMATCH	19
Needleman - Wunsch	6
NEEDLEMAN - WUNSCH	19
nucléotidiques	9, 19
parcourir	14
Programmation orienté objet	10
protéine	10
PROTÉINE	12, 14
protéiques	5, 9 sv, 19
retour	12, 14
Retour	12, 14
Weblogo	9
WEBLOGO	19