
Projet BioInformatique

Master SDSC 2020

SALAS VEGA Dominique - MOREL Louis - BONNETAIN Baptiste - LÉNON Lucas - MONET Anaïs
SDIA TPS

Rapport :

Big data en Bioinformatique - Statistiques sur les trinuéotides dans les gènes

Instructions et librairies permettant l'exécution du programme avec Eclipse

Nous utilisons les librairies Java de base (qui s'ajoutent automatiquement à la création d'un projet dans Eclipse) et Jxl qui permet de traiter les fichiers excel.

Informations sur l'arborescence locale et fichiers nécessaires

L'arborescence se met à jour et est composée d'un dossier Results qui contient 4 autres dossiers Eucaryotes, Procaryotes, Plasmids et Viruses, contenant les différents types d'être vivant. Le niveau inférieur de l'arborescence contient les sous types des êtres vivants.

Enfin nous avons les fichiers excel (un par être vivant et par groupes). Lors de la création de l'arborescence par l'interface graphique, les fichiers .DS_Store apparaissent. Ces fichiers sont invisibles dans l'arborescence locale. Nous n'avons pas réussi à les enlever.

La mise à jour automatique de l'arborescence lors de la création d'un nouveau fichier fonctionne. Le code couleur ne marche pas car nous ne faisons pas la mise à jour des fichiers, nous les créons seulement.

Classe contenant main

La classe contenant le main est *interface.java* .

Remarques :

Le parsing fonctionne correctement et sur toutes les classes, et les statistiques nous semble correct, les fichiers xls d'en tête sont correctement mise à jours. Nous utilisons 4 threads qui travaillent indépendamment sur chaque groupe (Eukaryotes, Prokaryotes, Plasmids et Viruses), ce qui permet d'accélérer le processus et aussi de vérifier le parsing pour chaque domaine. Le programme n'est pas capable de mettre à jour un fichier existant, il ne fait que crée les fichiers des espèces qui ne sont pas encore dans « Results ».La barre de chargement fonctionne.

Lorsque nous cliquons sur le bouton **démarrer** de l'interface, le programme se lance. Les fichiers sont en xls et nonxlsx car nous utilisons une librairie de java un peu ancienne qui ne génère pas correctement des nouveaux fichier xlsx, cependant les fichiers xls sont parfaitement ouvrable sous windows avec excel. La mise à jour de l'arborescence fonctionne, mais elle n'est pas complètement optimisée, lorsqu'un fichier est ajouté toute l'arborescence se referme, il faut donc l'ouvrir à nouveau.