

## TP 3 - Indexation de séquences ADN

Damien STENGEL, Benjamin PELLIEUX, Paul-Elie VALLEE

### Premier défi

L'objectif du premier défi était de retrouver une certaine chaîne d'acides aminés dans une séquence ADN. Pour ce faire, nous avons premièrement défini une structure d'arbre, afin d'indexer les différentes chaînes présentes dans la séquence d'ADN fournie. Le noeud final étant le noeud de la dernière lettre de la chaîne de 5 caractères, il contient la liste chaînée permettant de retrouver l'index des motifs dans la séquence. En réutilisant les méthodes du TP Wordblitz et liste chaînée, nous avons donc modifier les structures d'arbres et de liste chaînée pour qu'elles conviennent au problème (cf code).

### Deuxième défi

Nous avons rajouté un champ séquence dans la structure Noeud (cf code) Pour trouver si un motif est commun au 5 séquences, nous parcourons l'arbre par ses feuilles, et nous nous intéressons à toutes les feuilles de profondeurs 5. Ces noeuds finaux symbolisant une chaîne de 5 acides, nous pouvons dès lors regarder si ce noeuds comprends des chaînes contenu dans les différentes séquences. Nous pouvons alors faire une liste des chaînes respectant ces conditions.

Nous sommes actuellement entrain de faire la méthode permettant de trouver les noeuds de profondeur 5 contenant des références à chaque séquence. Pour cela nous avons écrit une méthode permettant de parcourir l'arbre de façon récursive. Lorsque nous trouvons un noeud indiquant une correspondance. On renvoie de façon récursive le résultat pour indiquer aux appels précédents d'afficher la lettre correspondant à l'index du fils qu'ils ont appelé. Nous n'avons pas fini cette méthode