



Université d'Angers
Faculté des Sciences
Module : *perl*

Département Informatique
ISSBA, LTE3, 07/08
Jean-Michel Richer

TD BIOPERL

1 Recherche de motif sur les glutamate déshydrogénase

La spécificité fonctionnelle de la GDH4¹ (Glutamate Déshydrogénase EC 1.4.1.4) est la formation du glutamate en utilisant NADPH. Bien que la GDH4 soit bien caractérisée pour les animaux et les prokaryotes, on dispose de moins de connaissances pour les plantes.

Exercice 1 - Dans un premier temps, on désire récupérer un ensemble de séquences GenBank dont les numéros d'accèsion sont les suivants (fichier `glutamate.acc`) :

```
# fichier glutamate
CAA41635
CAA41636
CAA09478
AAL36888
S28829
AAL51413
E75362
```

Ecrire un programme Perl nommé `bio_get_glutamate.pl` qui permet de récupérer les séquences GenBank qui correspondent aux numéros d'accèsion et stocke les séquences obtenues au format `fasta` dans le fichier `glutamate.fasta`. **Remarque** : le fichier des numéros d'accèsion peut contenir des lignes de commentaires qui commencent par le caractère `#`.

Exercice 2 - Dans un second temps on désire rechercher si le motif suivant, au format PROSITE, existe pour chacune des séquences du fichier `glutamate.fasta` :

```
K-G-G-x-R-x(12,23)-L-x(6)-K-x(4,6)-P-x-G-G-x-K
```

Ecrire un programme Perl nommé `bio_find_pattern.pl` qui indique quelles séquences contiennent ce motif et donnent le motif complet trouvé.

Exercice 3 - Modifier le programme précédent afin d'exhiber le motif consensus (2 versions).

Exercice 4 - Ecrire un programme `bio_align.pl` qui permet de réaliser un alignement multiple des séquences du fichier `glutamate.fasta` avec Clustalw. On stockera le résultat de l'alignement dans le fichier `align1.ms`. Visualiser l'alignement obtenu avec le programme `cinema`. Retrouve-t-on le motif initial ?

Exercice 5 - Avec l'alignement obtenu calculer et afficher la matrice des distances en utilisant le programme `ProtDist` du package logiciel Phylip.

¹Ce travail s'appuie sur l'article d'Emmanuel Jaspard, UMR PMS, Université d'Angers : A computational analysis of the three isoforms of glutamate dehydrogenase reveals structural features of the isoform EC 1.4.1.4 supporting a key role in ammonium assimilation by plants, *Biology Direct*, 2006, (1) :38