***Analyse de séquence Marco et Base de données***

**2014-2015**

Svitlana :

on m'a demandé d'expliquer l'algorithme de Needleman-Wunsch et comment l'adapter pour les différents types de problèmes. Puis je crois que j'ai eu des questions sur comment interpréter les données dans PDB.  
  
Pour bases de données, c'était principalement des questions sur le diagramme relationnel car le prof ne comprend pas grand chose au reste. Il m'a quand même demandé quelles difficultés j'ai eu lors de l'import des données et des trucs du genre.

**2015-2016**

Alizée

**Database**: Il interroge sur les trucs qui ne lui ont pas semblé clair dans le rapport. Les limitations de la DB et comment on pourrait les résoudre.

**Macromol :** On pioche une question au hasard.

Question n4 : Recherche d’une séquence dans une DB

Expliquer en gros FASTA et BLAST (il ne demande ni d’expliquer l’automate à stade fini ni le suffix tree et les tables index :p) il veut vraiment globalement la différence entre les 2.

Comment fait on pour retrouver des séquences homologues ? Recirpocal Best Hit. Comment ca marche et les limitations.

Il pose une question sur l’ago de N-W a tout le monde je crois. Moi j’ai eu quel est l’avantage de la programmation dynamique.

Voilà il aide vraiment beaucoup quand on ne sait pas, il est sympa ☺

Florian

**Database**: il pose des questions sur ton rapport et c’est plus une justification des raisons de ce que tu as écrit

**Macro :** Questionnaire 5 : alignement multiples. Pourquoi on ne peut utiliser needleman pour le faire ? Comment évolue le temps de calcule en fonction de la longueur et du nombre de séquences.

Expliquer différentes programmes, les limitations ? Si on peut comparer tout un génome ou pas ? Si translocation, comment observer 🡪 dot plot.

Charlotte

**Database :** pleins de petites questions sur ce qu’il n’a pas compris dans le rapport.

**Macro :** Questionnaire 4 : expliquer dynamical programming, les avantages et pourquoi est-ce qu’on l’utilise.

Expliquer FASTA et BLAST, différence (alignement global vs alignement local), qu’est-ce qu’on peut changer dans BLAST et expliquer les paramètres

Joel

**Database :** pleins de petites questions sur des petits détails, souvent du détail inutile comme des formulations de phrases ou pourquoi avoir fait la query ainsi et pas comme ça. En gros, il est chiant à la « Richelle », tous ceux qui étaient en cours sauront de quoi je parle.

**Macro :** (Question 2) En gros expliquer N et W. Dans mon cas il avait vu mon implémentation dans R, donc ça jouait en ma faveur (pas besoin d’expliquer chaque détail). Je devais discuter comment j’implémenterais des gaps étendus (opening + extending). Puis il m’a demandé comment ont score les alignements : donc fallait se lancer dans PAM et BLOSUM. Bien expliquer comment chaque type de matrix est généré. Puis, étant donné des séquences supposé avoir divergé il y a longtemps, comment choisirai-je les matrix : PAM250 (ou haut indice, car divergé il y a longtemps) soit BlOSUM30 (indice bas car faible pourcentage de similarité attendu).  
Finalement, comme BLOSUM utilise du MSA, j’ai dû toucher quelques mots sur les alignements multiples.

Benoît

**Database : no comment**

**Macro :** question numéro 6. Ici j’ai dû plutôt détailler les différents types d’algorithmes utilisés en alignements multiples (programation dynamique améliorée, alignement progressif et algo génétique avec un exemple de programmes donc dans l’ordre : lipmann-carillo, clustalw et saga). Puis j’ai dû parler des limitations (ne tiennent pas forcément compte du fait que on puisse avoir des inversions, etc). Et enfin j’ai dû dire la spécificité de COFFEE -> tient compte des infos des alignements en pair-wise pour optimiser les alignements progressifs.

Gilles

**Database : questions sur quelles sont les étapes pour faire une database**

**Macromol :** Questionnaire 1 : Comment savoir si le score obtenu avec N-W est relevant ? ( parlé de la P-value, comment la calculer pour ce cas? ) Comment trouver des gènes dans un génome ?

2eme sess  
Q2 Pattern matching : D'abord N-W global puis local

Trouver un motf dans une sequence, parler de PFFM , back ground model, ....

Comment trouvé des séquences homologues ? Il m’a parlé de programmes comme gims/consenus mais je ne sais pas ce qu’il voulait.

Nancy :

**Macro :** Pleins de questions sur les dot plots et aussi globalement expliquer les etapes dans la prediction des genes.