

## Compte-rendu

**Présents :** Sophie LÈBRE, Thomas AYRIVIÉ, Mehdi BELKHITER, Jamila CHERKAOUI, Dina EL HIJJAWI, Magatte LO.

**Prochain rendez-vous :** Vendredi 13 Octobre 2023 9h30, Université Paul Valéry, bâtiment B.

### Travail à effectuer et répartition :

- Stats descriptives sur les longueurs de motifs JASPAR  
+ fréquences de nucléotides
- On peut représenter la distribution de la variabilité de la position de A, G, T, C... Qui sont fixes, qui sont variables...

### À aborder au prochain rendez-vous :

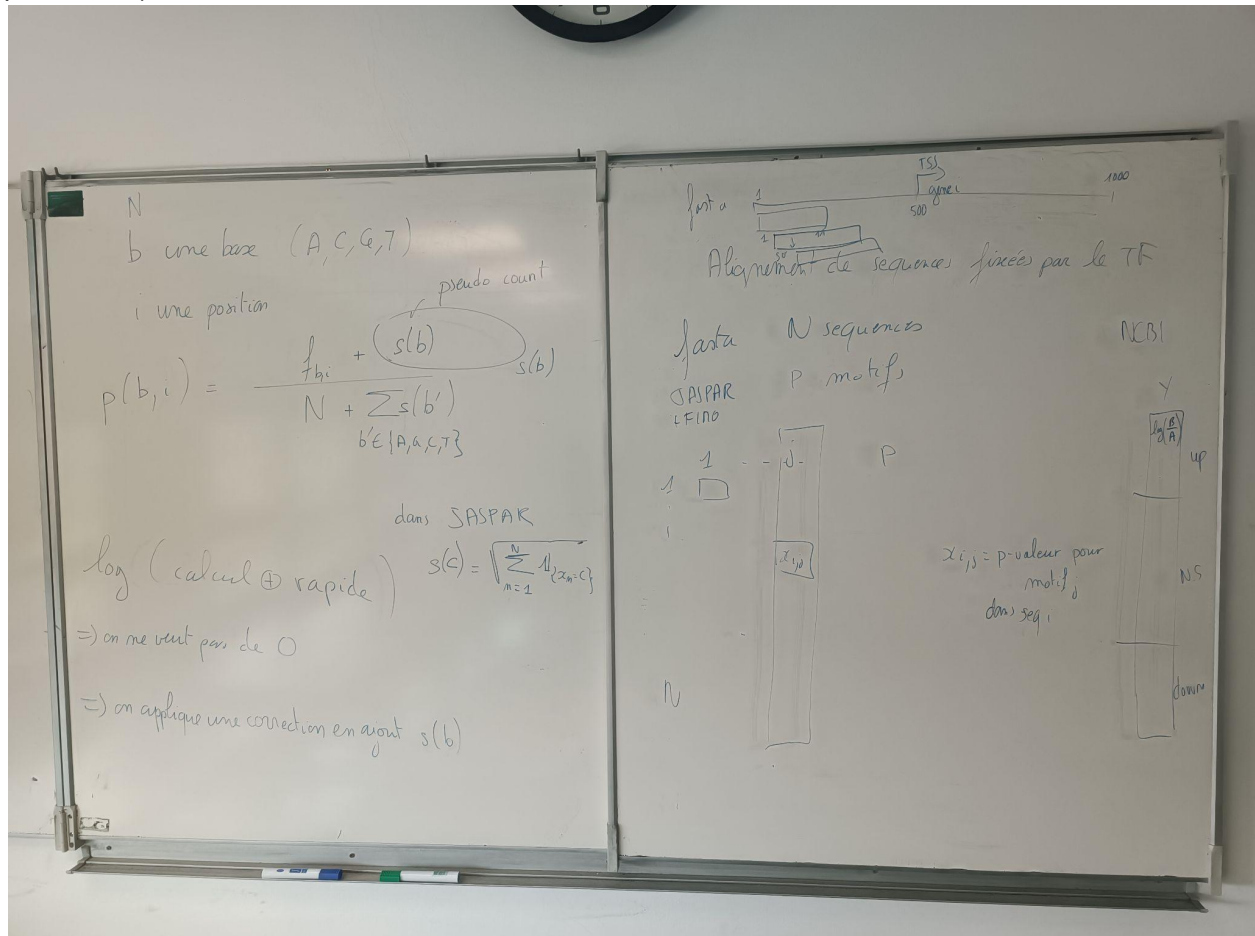
- s(b) de la correction
- Test qui suit le calcul du score. Qu'est-ce la p-valeur ?
- Lien entre le fichier NCBI et Jaspar/Fimo  
Faire la jointure (information à la rentrée)
- ACP, exploration des données
- Fichier sortie FIMO
- Planification des tâches
- Faire des diapos pour faire un résumé
- Découpage en classe, des y NCBI (0, -1, 1)

### Points abordés :

- Dernière ligne de commande, sort juste la liste des fréquences sur les 4 lettres  
Ordre 1 et Ordre 2 > enregistrer le fichier > le mettre en paramètre dans Fimo

Verbosity (choisir entre 1 et 5), Mathilde Robien choisit 1, mais on peut faire plus

Fréquences un peu corrigées avec bile sortie get (c'est pas puisqu'on l'a pas vu qu'il ne peut pas exister)



> on a plus de 0 donc on peut utiliser le LOG

Sinon on donne  $\frac{1}{4}$  à chaque lettre, pas forcément le cas dans la vraie vie, alors on peut faire de cette manière avec les fréquences observées dans les séquences (fichier fasta)

Wasserman explique bien l'analyse des

Fichiers des régions promotrices : consortium fantome qui a identifié chaque départ de gène on a extrait la séquences 500 avant et 500 après.