**INTRO :**

Nous allons parler du contrôle du taux de faux positifs (False Discovery Rate) dans le cas des Volcano plot pour les données omiques.

Dans le cadre d'un test d'hypothèse, il y a toujours une chance de rejeter à tort l'hypothèse nulle, qui est connu sous le nom de d’erreur de type I. Dans le cadre des données omiques, il y a énormément de données. Nous introduisons donc le FDR qui est une méthode de conceptualisation du taux d'erreurs de type I dans les tests d'hypothèses nulles lors de comparaisons multiples. Le FDR est l’espérance de la proportion de faux positif (False Discovery Proportion), c’est-à-dire l’espérance du nombre de faux positif divisé par le nombre total de positif.

La procédure la plus populaire de contrôle du FDR est la procédure de Benjamini-Hochberg en ajustant les p-valeurs. La procédure s’applique comme ceci. Elle classe d’abord les p-valeurs par ordre croissant en leur assignant un rang. Puis, elle calcule la valeur critique de BH pour chaque p-valeur qui est le rang de la p-valeur sur le nombre total de test multiplié par la part de faux positifs à un niveau alpha, 5% en général. Enfin, la procédure trouve la plus grande p-valeur qui est inférieure à cette valeur critique, notée pi chapeau. Et désigne comme étant significative toutes les p-valeurs inférieures à pi chapeau.

Lorsque trop de gènes sont sélectionnés, on peut appliquer la procédure de double filtrage qui consiste à ne conserver que les p-valeurs ajustées qui ont une grande taille d’effet (le fold change). Cette procédure peut être représentée par un Volcano plot qui est un diagramme de dispersion des −log10 des p-valeurs ajusté par BH par rapport au log2 de la taille de l’effet.

**CONCLU :**

Ces méthodes peuvent s’appliquer dans le cadre de données réelle. Comme nous pouvons le voir dans le poster, la méthode de closed testing est une méthode conservatrice qui permet d’avoir une meilleure estimation du FDP.

Le problème du contrôle du FDR survient lorsque des procédures telles que Benjamini-Hochberg sont combinées avec la procédure de double filtrage. Le contrôle du FDR sur un ensemble n'implique pas le contrôle du FDR sur une sous-partie de cet ensemble. L'inflation du FDR est élevée lorsque la variance des gènes différentiellement exprimée est inférieure à celle des gènes non différentiellement exprimés. Cependant, elle est moins importante lorsqu’il y a de fortes corrélations ou une faible proportion de gènes nuls. Deux méthodes permettent un double filtrage tout en contrôlant le FDR, le closed testing et le focused BH.