Q1 :

Analyse pls :

PLS : Arial Latest Squares : expliquer la relation entre deux sous-ensembles de données continus (cherche combinaisons linéaires des variables de chaque ensemble de données

PLS est particulièrement efficace lorsque P + Q > N, où P est le nombre de variables dans le premier ensemble de données, Q est le nombre de variables dans le deuxième ensemble de données et N est le nombre d'échantillons dans chacun.

1. Analyse gène/protéine

mRNA et prot :

On remarque une bonne corrélation entre les m RNA et les prot

miRNA et prot

pas de corrélation miRNA et prot

miRNA et mRNA :

pas de corrélation miRNA et mRNA

1. Discussion nombre de composante à inclure : comp =2 (justifier grâce à ACP [jean])
2. Les graphes :

Graphe échantillons plot\_indiv

Graphe variables plot\_var

sPLS :

a-Faire une première analyse sPLS avec les gènes et les protéines (10 features/bloc composante 1, 5 pour composante 2) : c’est fait

b- les variables sélectionnés :

mRNA : 10 1

prot : 1 5

Leur rôle biologique :

14-3-3 epsilon

4E-BP1\_pT70,