**Représentation des interactions avec les protéines les plus proches contenant un mot-clé dans leur processus biologique**

Lorsque l’on étudie l’impact d’une protéine au sein d’une voie biologique, il est intéressant de trouver les protéines proches qui sont impliquées dans cette voie et les intéractions qui les relient à la protéine étudiée.

L’interactome humain utilisé provient de “STRING”, il contient 16948 protéines et 290849 intéractions, enregistré en format text (edges\_string.txt).

Les processus biologiques de chaque protéines ont été récupérés par l’API de QUICKGO et enregistrés en format JSON (bp\_string2.json).

Afin de sélectionner les protéines les plus proches, les nœuds du graph ont été transformés en vecteurs grâce à la librairie “node2vec”.

Le graph final représente le chemin le plus court entre la protéine de départ (rouge) et les protéines les plus proches contenant le mot clé dans leur processus biologiques

(bleu). Les protéines en vert permettent de comprendre l’interaction indirecte entre la protéine d’intérêt et les protéines ciblées La taille des nœuds du graphe représente la similarité des intéractions de la protéine avec celles de la protéine d’intérêt.

Les paramètres d'entrées sont le nom de la protéine étudiée, les mots clés recherchés dans les processus biologiques, ainsi que le nombre de protéines les plus proches à sélectionner.

Le programme fournit en sortie le graphe au format PNG, ainsi qu’un fichier texte contenant le nom, les processus biologiques et les fonctions moléculaires des protéines du graphe.