**Réseau d'interactions protéine-protéine - "*Single-Cell*"**

**Il est indispensable de connaître les cours.**

**Exercice 1**

Soit une cellule d'un volume de 103 µm3 avec une concentration totale de protéines = 5 µM :

Combien de molécules de protéines y a-t-il dans cette cellule ?

Combien de molécules de protéines d'un type donné y a-t-il si l'on considère qu'il y a 10.000 types différents de protéines dans cette cellule ?

Le nombre d'Avogadro (N) est le nombre d'atomes ou de molécules dans une mole de matière = 6,022 1023 mol-1.

**Exercice 2**

Aller au site "STRING exercises" : faire l'exercice 1.

**Exercice 3 - 1ère partie**

Aller à HPIDB.

"*Search Database By*" => choisir "*Keyword*". Menu "*Query Type*" : choisir "*Protein Accession / Name*" => entrer "Q16543".

Page de résultats : cliquer sur "*Network visualisation*". Choisir "*By taxon Names*" dans le menu déroulant.

a. Qu'est-ce que Q16543 ?

b. Quelle(s) protéine(s) du virus de la Dengue interagit/ssent avec l'homme ?

**Exercice 3 - 2ème partie**

Revenir à la page de résultats de HPIBD : cliquer sur "*Download unique proteins from Cytoscape (tab delimited)*"=> un fichier au format ".tsv" doit être téléchargé.

Lancer le programme Cytoscape (s'il est disponible) : Onglet "*File*" => menu "*Import*" => choisir "*Network from File ...*" : ouvrir le fichier récupéré de HPIBD.

Une fenêtre s'ouvre pour attribuer chaque colonne de ce fichier à un objet du graphe :

Colonne "*HPI\_Host\_Protein*" => attribuer "*Source Node*" (rond vert)

Colonne "*HPI\_Host\_Taxon*" => attribuer "*Not Imported*" (rond barré)

Colonne "*HPI\_Pathogen\_Protein*" => attribuer "*Target Node*" (rond rouge - point central)

Colonne "*HPI\_Pathogen\_Taxon*" => attribuer "*Target Node Attribute*" (fichier rouge)

a. Le réseau d'interactions créé avec Cytoscape a-t-il le même aspect (noeuds, arrêtes, attributs) que celui de HPIDB ?

b. Se familiariser avec certains outils du programme :

* La loupe du zoom.
* Cliquer sur un noeud quelconque puis sur l'icône "deux maison" ("*First Neighbors of Selected Nodes*").

c. Modifier l'aspect du réseau avec les options du menu général "*Layout*". Exemples :

* Modifier l'aspect avec les outils du sous-menu "*Layout tools*".
* Modifier le type de graphe avec "*Grid Layout*", "*Hierarchical Layout*", "*Circular Layout*", "*Stacked Node Layout*".
* Tester d'autres options à loisir ...

d. Voir le tutoriel Cytoscape pour modifier l'aspect des noeuds et des arêtes :

"*Set Node Fill Color*" (<https://cytoscape.org/cytoscape-tutorials/protocols/basic-data-visualization/#/5>).

* Entrer dans le sous-menu le menu "*Style*" de Cytoscape tout à gauche de la fenêtre.
* Modifier les couleurs des noeuds, modifier l'aspect des arêtes, ...

**Exercice 4**

Aller à l'exercice en ligne de l'EBI sur l'interactomique (URL ci-dessous).

<https://www.ebi.ac.uk/training/online/courses/network-analysis-of-protein-interaction-data-an-introduction/introduction-to-graph-theory/try-it-yourself-adjacency-matrices/>

Remplir la matrice d'ajacence du graphe proposé.

**Exercice 5**

Analyser le paragraphe "*Identification of existing drugs targeting SARS-CoV-2 human host factors*" et la Figure 5 de l'article Gordon et al. (2020) "*A SARS-CoV-2 protein interaction map reveals targets for drug-repurposing*" Nature 583, 459 – 468.

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7431030>

**Exercice 6**

a. Analyser les 4 réseaux d’interactions de la figure 3 de l'article Reys & Labesse (2022) "*SLiMAn: An integrative Web server for exploring short linear motif-mediated interactions in interactomes*".

b. Aller à la base de données BioGRID : <https://thebiogrid.org/>

Combien de protéines interagissent avec FRAT2 de l'homme ?

Pourquoi le chiffre est-il différent de celui de l'article ?

**Exercice 7 - "Single-Cell"**

Analyser et décrire la figure 1 de l'article Clark et al. (2023) "*Microfluidics-free single-cell genomics with templated emulsification*" Nat. Biotechnol. 41, 1557 - 1566

<https://www.ncbi.nlm.nih.gov/pmc/articles/pmid/36879006>

**Exercice 8. Répondre par "VRAI" ou "FAUX" aux assertions suivantes.**

|  |  |
| --- | --- |
| **Assertions** | **V/F** |
| A1. Seules les protéines ont une dynamique conformationnelle. |  |
| A2. Toutes les protéines interagissent avec un gène. |  |
| A3. Toutes les méthodes d'étude des interactions protéine-protéine permettent de déterminer la cinétique d'association de ces protéines. |  |
| A4. Certaines protéines interagissent avec un gène. |  |
| A5. Toutes les méthodes d'étude des interactions protéine-protéine permettent de déterminer la constante de dissociation de ces protéines. |  |
| A6. Le pseudo-temps est le positionnement d'une cellule le long de la trajectoire qui quantifie la progression d'un processus biologique. |  |
| A7. Une technique utilisant une/des protéine(s) de fusion est rapide et simple à appliquer. |  |
| A8. La structure des macromolécules est l'élément central qui contrôlent leurs interactions. |  |
| A9. Les anticorps sont un outils important pour l'étude des interactions protéine-protéine. |  |
| A10. Les interactions protéines connues sont en très grande majorité issues de données expérimentales (biochimiques, biophysiques, …). |  |
| A11. La technique appelée double-hybrides est une technique à très haut débit. |  |
| A12. Tout comme les alignements de séquences et les arbres phylogénétiques, il n'y a pas de réseau d'interactions «*juste*» ou «*faux*». |  |
| A13. Certaines protéines interagissent avec plusieurs dizaines d'autres protéines. |  |
| A14. Toutes les protéines interagissent avec au moins une autre protéine. |  |
| A15. Plus le nombre de molécules d'un ligand se fixant sur les sites de fixation d'une protéine est élevé, plus la variation d'enthalpie de cette réaction est faible. |  |
| A16. Plus KD est petite, plus l'affinité de liaison du ligand pour son site de fixation est grande. |  |
| A17. La prédiction de trajectoire décrit l'évolution de chaque cellule en ordonnant ses états selon son processus de développement. |  |

**Exercice 9**

a. Analyser le paragraphe "*2.2. Protein Enrichment Analysis*" et la Figure 2 de l'article Chiaradia et al. (2019) "*Proteome Alterations in Equine Osteochondrotic Chondrocytes*" Int. J. Mol. Sci. 20, 6179

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6940994/>

b. Partie "*Materials and Methods*" => paragraphe "*4.4. Protein Enrichment Analysis*" : quelles sont les 2 bases de données et les 2 applications ("plugins") de Cytoscape utilisés pour analyser les réseaux regroupés sur la base de la fonction des protéines dérégulées dans OC ?

c. Quelles informations ontologiques et quelles ressources ont permis d'enrichir les réseaux ?

d. Quels sont les groupes fonctionnels les plus importants ?

**Exercice 10 - facultatif**

a. Aller aux réseaux analysables en ligne de NDEX : <https://www.ndexbio.org/iquery/>

Choisir l'exemple "*PANCREATIC BETA CELL*" : Analyser les résultats.

b. Ouvrir l'application Cytoscape si vous en disposez.

c. Revenir à la page de NDEX : cliquer sur l'icône orange (en haut à droite) pour ouvrir le réseau dans Cytoscape. Interpréter les différences d’aspects des réseaux.