

TP1 : Exploration des banques de connaissances biologiques

Matériel du TP disponible ici : <http://helios.mi.parisdescartes.fr/~lomn/Cours/BI/>

Les bases de connaissance/**outils** les plus utiles à l'heure actuelle (Mars 2018) :

- **NCBI** (<http://www.ncbi.nlm.nih.gov/>) : National Center for Biotechnology Information, US) : **PubMed, PubMed Central**
- **EBI** (<https://www.ebi.ac.uk/>) avec **EMBL** (<http://www.embl.org/>) : The European Bioinformatics Institute, Part of the European Molecular Biology Laboratory, UK and Switzerland) (**Où est Hinxton?**)
- **STRING** (<http://string-db.org/>),
- **OMIM**® (<https://omim.org/>) : Online Mendelian Inheritance in Man® (**Qui est Mendel?**)
- La banque UniProtKB/Swiss-Prot (<http://www.uniprot.org>) pour les protéines
- **GeneCards**® (<http://www.genecards.org/>) : The Human Gene Compendium, Israël.
- **KEGG** (<http://www.genome.jp/kegg/>) : Kyoto Encyclopedia of Genes and Genomes
- etc. Remarque : vous trouverez parfois l'accès à un élément `NOM_BASE:CODE_ELEMENT` ex. SWISSPROT:P04156

Ce travail pratique est constitué de questions ouvertes qui doivent vous permettre de vous poser des questions : il faut gérer la complexité de vos réponses pour réaliser les exercices en un temps raisonnable. Prenez le temps d'observer, d'explorer, décrire, décortiquer ces bases de connaissance en ligne. Cet exercice préparera l'intervention de Prof. Frédéric Dardel et vous permettra d'appliquer les éléments de cours qu'il vous proposera.

Exercice 1 : La maladie de Creutzfeldt-Jakob (CJD Creutzfeldt-Jakob Disease) ou maladie de la vache folle (inspiré de collègues de Marseille, Campus de Luminy)

Avec l'outil **PubMed** (<http://www.ncbi.nlm.nih.gov/pubmed>), trouver **5 termes biologiques ou médicaux liés à la maladie de Creutzfeldt-Jakob** ? Faites la même recherche avec l'outil **PubMed Central** (<https://www.ncbi.nlm.nih.gov/pmc/>). Listez quelques bases de données sur lesquelles requête PubMed.

Interprétation/Commentaire de résultats : par exemple, qu'est-ce que PubMed versus les autres ressources (une compilation?) etc.

Avec l'outil **PubMed** (<http://www.ncbi.nlm.nih.gov/pubmed>),

- **Recherchez** les publications avec le mot clef **prion**
- Suivez le lien vers l'article écrit par Mabbott&Bradford publié en 2015 et de type Review.
- **L'article est-il consultable ? Si oui, sur quel site ?**

Interprétation/Commentaire de résultats : par exemple, disponibilité des ressources, openData ?

Exercice 2 : Visualisez des graphes d'interaction avec la protéine du prion PRNP ?

Recherchez les interactions d'autres protéines avec la protéine du prion PRNP (organisme Homo Sapiens) avec l'outil **STRING** (<http://string.embl.de/>) qui est visuel. Il vous proposera un graphe d'interaction entre gènes qui est interactif : essayez de déplacer les nœuds du graphe etc.

Demandez à ne visualiser que 5 interactions principales minimales ? Et avec un seuil de 0.9 combien obtient-on d'interactions ? Quel est le rôle de la protéine BCL2 ?

Sauvegardez les deux réseaux sous un format graphique, ainsi qu'une liste des protéines qui le constituent dans un fichier texte.

Interprétation/Commentaire de résultats : par exemple, comment l'algorithme extrait ces liens ou à quoi correspondent les différentes couleurs des liens ?

Exercice 3 : Interrogez l'outil OMIM® <http://omim.org/> sur la maladie de CJ (code 123400 pour CJD Creutzfeldt-Jakob Disease).

- Décrivez cet outil succinctement.
- Interrogez la base sur « Creutzfeld Jacob Disease »
- Suivre le lien vers le gène PRNP responsable de la CJD (fiche 176640)
- Explorez l'information disponible sur le gène :
 - Sur quel chromosome se situe ce gène PRNP?
 - Quels sont les gènes proches de PRNP (suivre le lien vers gène map locus: 20p13 avec l'application Mapping)?
- Suivre le lien vers la fiche **GeneCards®** (qui est un outil disponible à partir des liens externes sur la droite dans Genes Info). A quoi correspond cette base de connaissance ?
- Trouvez sur quel chromosome se trouve l'équivalent du gène PRNP chez la souris (vous verrez cette notion en cours on parlera d'orthologues)

Interprétation/Commentaire de résultats :

Exercice 4 : les serveurs d'outils génériques European Bioinformatics Institute (EBI)

Sur le serveur de l'EBI <https://www.ebi.ac.uk/services> qui répertorie les outils populaires à l'instant présent, à l'aide de l'outil ENSEMBL (<http://www.ensembl.org/>) :

1. Cherchez dans la banque EMBL (European Molecular Biology Laboratory), le mot clef **prion**.
 - Combien de séquences de gènes trouvez-vous?
 - Cherchez les prions chez **l'homme (homo sapiens)**.
 - **Quelles données trouve-t-on dans les fiches EMBL?**
 - Cherchez les transcrits (ARN) du prion *chez l'homme ou la souris*.

Interprétation/Commentaires de résultats :

Exercice 7 : Genome Browser

- Trouvez la séquence du gène prion PRNP et le visualiser avec l'outil **Genome Browser** <http://genome.ucsc.edu>

Interprétation des résultats :

Ouverture

- Vous pourrez retester ces bases quand on vous parlera pendant le cours du récepteur humain de la vitamine D (Fig. 4.3 du cours de Prof. Frédéric Dardel, VDR_HUMAN, P11473).
- **Vous pouvez explorer** le projet d'exploration de fonds marin organisé à Luminy, Marseille. <http://annotathon.org> qui est un projet d'envergure et intéressant sollicitant les étudiants.

Annexe

Le gène Prnp codant pour la protéine PrP

« *Major prion protein (PrP, for prion protein or protease-resistant protein), also known as CD230 (cluster of differentiation 230), is the only known example of a prion protein in animals. In humans, it is encoded by the PRNP gene (PRioN Protein)* »

https://fr.wikipedia.org/wiki/Prion_%28prot%C3%A9ine%29 :

« Un prion est un type d'agent pathogène de nature protéique (constitué d'une protéine ayant adopté une conformation ou un repliement anormal) qui au contraire des agents infectieux conventionnels tels que les virus, les bactéries ou encore les parasites, est exempt d'acide nucléique (ADN et ARN) comme support de l'information infectieuse. Ce terme fut introduit pour la première fois en 1982 par Stanley Prusiner et correspond à l'acronyme de *PRoteinaceous Infectious ONly particle* (particule protéique infectieuse).

On distingue les prions de mammifères qui infectent l'humain et différentes espèces animales, des prions retrouvés chez les champignons comme chez Saccharomyces cerevisiae (levure de boulanger). Les prions de mammifères sont les agents causaux responsables des encéphalopathies spongiformes transmissibles (EST) ou maladies à prion. Parmi les EST les plus connues, on peut citer chez l'homme, les différentes formes de la maladie de Creutzfeldt-Jakob, l'insomnie fatale familiale (IFF), le syndrome de Gerstmann-Sträussler-Scheinker (SGSS), le Kuru et chez l'animal, la tremblante du mouton et de la chèvre, l'encéphalopathie spongiforme bovine (ESB), l'encéphalopathie spongiforme féline, l'encéphalopathie spongiforme du vison et le dépérissement chronique du cervidé (CWD pour Chronic Wasting Disease). L'ensemble de ces maladies se caractérise par une dégénérescence du système nerveux central (cerveau et moelle épinière) liée à la propagation ou multiplication des prions chez l'hôte infecté. D'un point de vue anatomopathologique, on observe ainsi au niveau de l'encéphale la formation de vacuoles (donnant un aspect spongieux au cerveau, d'où le nom de spongiforme dans EST), une mort des neurones, une gliose (multiplication des astrocytes et de la microglie) et l'accumulation d'une protéine de l'hôte, la PrP^C, sous une conformation anormale (ou mal repliée) alors dénommée PrP^{Sc}.

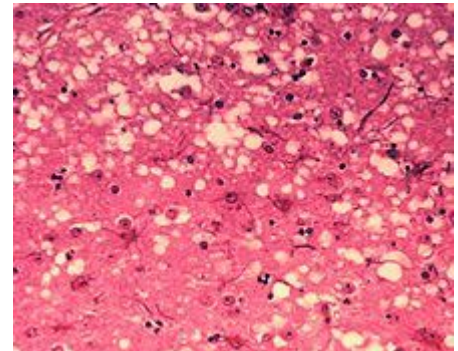


Illustration 1: Les « trous » microscopiques sont caractéristiques des tissus infectés de prions, leur donnant une consistance spongieuse.

Des prions atypiques, apparemment également pathogènes, mais probablement sporadiques, ont aussi été récemment découverts (publication 2011 par une équipe franco-espagnole 1). »

Quelques slides intéressants sur la maladie ovine : <http://slideplayer.fr/slide/1477652/>