MANUEL DES FONCTIONS D’ANALYSE DE DONNEES DE CRIBLAGES

L’outil d’analyse de données de criblages est composé de fonctions codées en R.

Il permet de traiter des données de criblages secondaires contenus dans une fichier CSV via des outils exécutables en Terminal permettant de déterminer les paramètres de la courbe et de tracer les graphiques correspondants. 5 scriptes avec chacun une fonction précise ont été développés.

Le fichier soumis peut être normalisé ou non, une des fonctions étant à disposition pour le normaliser.

Les données peuvent être issues de monoplica, de duplica ou de triplica. Pour ces deux dernières possibilités les fichiers doivent être remodelés par les fonctions à disposition.

Ces fonctions étant exécutables en terminal, chaque commande est précédée de « R –vanilla ».

Vous trouverez ci-dessous les cinq fonctions mises à disposition et la manière de les exécuter.

1. Normaliser les résultats

R --vanilla fichier.csv fichier\_de\_sortie.csv < normalization.R

La fonction normalization.R permet de normaliser les données issues de criblages monoplica, duplica et triplica. Le fichier doit être de type csv, la décimale doit être exprimé par un point et les valeurs séparés par des points-virgules.

Il doit être renseigné en premier argument le nom de fichier et en second argument la destination suivie du nom de sortie séparé d’un slash, et se terminer par « .csv ».

1. Formater des données de duplica et de triplica

R --vanilla fichier.csv fichier\_de\_sortie.csv < make\_duplica.R

R --vanilla fichier.csv fichier\_de\_sortie.csv < make\_triplica.R

Ces deux fonctions traitent respectivement des données de duplica et de triplica.

Ils permettent de réunir les colonnes issues de la même molécule en une seule pour que les données soit exécutable par le script multi\_llogistic.R. Pour les contrôles il effectue une moyenne des valeurs.

Le fichier doit être de type csv, la décimale doit être exprimé par un point et les valeurs séparés par des points-virgules.

Il doit être renseigné en premier argument le nom de fichier et en second argument la destination suivie du nom de sortie séparé d’un slash, et se terminer par « .csv ».

1. Paramètres et courbes

R --vanilla fichier.csv MIN MAX destination/reference < mono\_llogistic.R

R --vanilla fichier.csv MIN MAX destination/reference < multi\_llogistic.R

Les fonctions mono\_llogistic.R et multi\_llogistic.R permettent de déterminer les paramètres de la fonction log-logistic et de tracer la courbe pour chaque molécule.

Quatre argmuments sont à indiquer. Le premier argument et le nom du fichier à exécuter. Les deuxième et troisième argument sont les valeurs « minimum » et « maximum » à prédéfinir au besoin, dans le cas contraire il est demandé d’indiquer « NA ».

Le troisième argument et la destination suivi du début de nom des fichiers. Les fonctions génèrent un fichier csv contenant les quatre paramètres (dans l’ordre, min, max, EC50 et Hill) et un graphique pour chaque molécule. Multi\_llogistic génère l’intervalle de confiance de la courbe.

Le nom de chaque graphique contiendra le nom de la molécule.