

Projet 748 : Comparaison de paires de génomes bactériens par dotplot

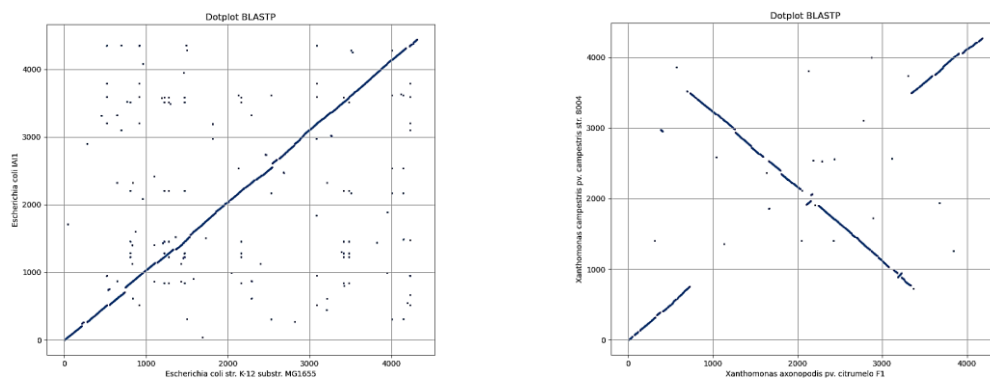
I – Objectifs du projet et stratégie adoptée

Le but de ce projet vise à développer une interface graphique connectée à une base de données de protéomes bactériens, afin d'analyser la synténie entre deux protéomes. Cette analyse est réalisée en visualisant la répartition des gènes homologues à travers un dotplot. D'un point de vue biologique, on anticiperait une réduction des blocs de synténie à mesure que les espèces s'éloignent sur le plan taxonomique, ce qui correspond à une diminution de la conservation des répertoires génomiques. Nous nous focaliserons sur quelques exemples de protéomes pour illustrer ce phénomène de perte de conservation.

Pour ce faire, nous avons conçu une interface graphique en utilisant Tkinter associée une base de données PostgreSQL. Cette application offre des fonctionnalités interactives pour sélectionner les protéomes, ajuster les seuils de e-value, et exécuter les BLASTP et les CD Search. Les résultats sont représentés par des dotplots en scatter, affichés dans le canevas de l'interface, permettant ainsi la visualisation de la synténie entre les gènes homologues. De plus, l'application propose des options supplémentaires telles que l'élimination du bruit et des fonctions de zoom pour mieux visualiser les détails du dotplot.

II – Résultats et interprétation

Voici un exemple de résultats de BLASTP entre 2 organismes taxonomiquement proches (à gauche) et taxonomiquement plus éloignés (à droite), avec un seuil d'e-value à $1e-6$ et élimination du bruit :



La figure de gauche illustre le dotplot entre deux souches d'une même espèce bactérienne (*Escherichia coli*). On peut observer une conservation significative de l'ordre des gènes avec un grand bloc de synténie, entrecoupé par quelques duplications discernables lorsqu'on zoome sur le dotplot. La figure de droite représente le dotplot entre *Xanthomonas axonopodis* pv. *citrumelo* F1 et *Xanthomonas axonopodis* pv. *campestris* str. 8004, toutes deux appartenant au genre *Xanthomonas*. Dans ce cas, des blocs de synténie discontinus sont visibles au début et à la fin du dotplot. Une inversion majeure du protéome, affectant environ les deux tiers de celui-ci (soit environ 3000 gènes), est également observée. Ces résultats suggèrent une conservation moins importante entre ces deux espèces par rapport au dotplot précédent.

III – Conclusions et perspectives

Ainsi, au sein d'une même espèce, l'ordre des gènes est généralement plus conservé que celui observé entre des espèces appartenant à un même genre, et encore moins entre des genres différents. Bien que ce résumé ne puisse couvrir tous les résultats possibles de BLASTP et de CD Search, il est intéressant de noter que même entre des espèces de genres différents, quelques blocs de synténie peuvent être observés. Cette approche met en lumière le potentiel des blocs de synténie pour retracer l'évolution des génomes bactériens.

Parmi les nombreuses améliorations envisageables, une piste intéressante serait de cibler plus précisément les blocs de synténie et d'extraire les régions homologues pour d'autres types d'analyses.