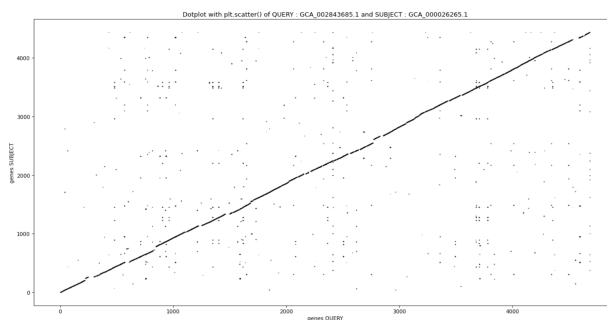
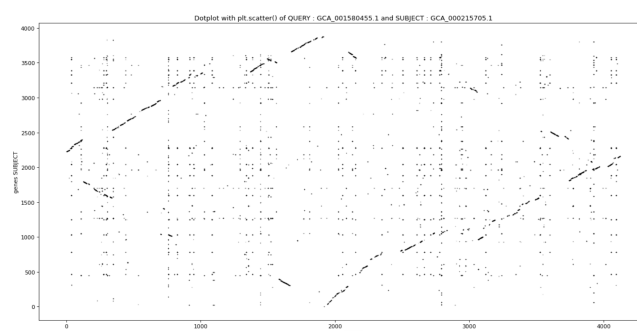


L'objectif de ce projet était de développer une interface graphique permettant, pour une paire de génomes bactériens complets choisie par l'utilisateur et dont les informations seront stockées dans une base de données relationnelles, de représenter par un dotplot les gènes homologues et leur positions dans ces génomes. La stratégie adoptée a d'abord été de créer une base de données avec 3 tables principales organismes, gène, blast et une table intermédiaire blastinterm. Puis, j'ai rempli la base avec les données fournies sur moodle. J'ai réfléchi à comment présenter mon interface graphique pour qu'elle soit claire et simple d'utilisation. Les combobox permettent de choisir les assembly des séquences query et subject. Il faut ensuite entrer la valeur pour le seuil de e-value, la valeur pour la fenêtre, pour le pas et ainsi que le seuil de stringence. Ensuite, la personne peut lancer l'analyse. L'analyse prend un certain temps puis le dotplot apparaît. Matplotlib permet d'enregistrer le graphique. Search_Org permet d'obtenir l'assembly grâce au nom précis d'une espèce. L'utilisation de fenêtre permet d'avoir moins de bruit et de mieux visualiser les régions de synténie.



Sur le graphique ci-dessus, nous observons une diagonale montante. Lorsqu'il y a une diagonale montante centrée cela signifie qu'il y a un match entre les gènes et que par conséquent que les régions géniques sont très conservées. Ces deux gènes appartiennent sûrement à la même espèce, peut-être de souche différentes.



Sur ce graphique ci-dessus, nous observons une sorte de losange. Les diagonales inversées peuvent correspondre à des régions géniques inversées et répétées. On observe que la diagonale n'est pas correctement centrée. On suppose qu'il y a dû avoir des translocations de gènes entre les deux génomes. Les deux organismes sont de la même espèce mais de souches différentes. Elles sont plus éloignées que les *E.coli* du dessus.