

# Genomic Island Analysis Clean

*Cameron Strachan*

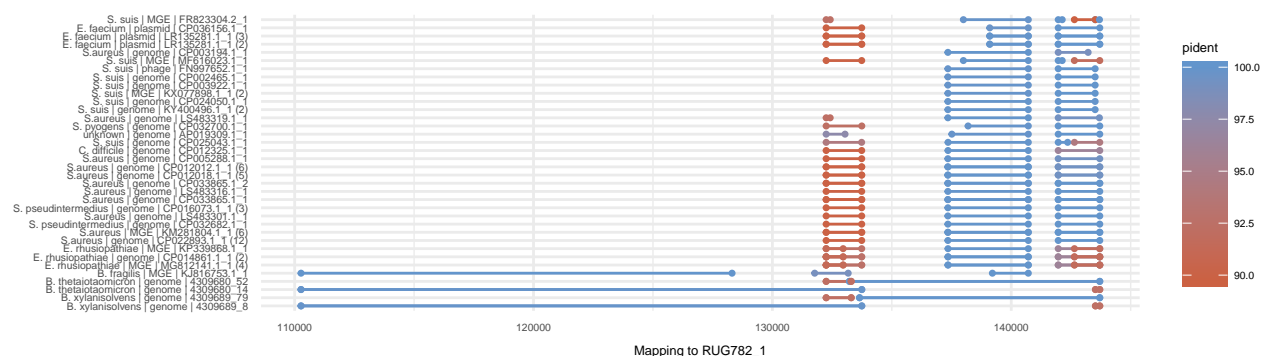
2019-04-04

## Summary

When blasting the CARD database at the protein level against the rumen genomes, I noticed that there were 2 proteins with 100% identify both on the same contig for 3 genomes. I decided to take these 3 cotigs and blast them on NCBI at the DNA level, which seemed to hit many pathogens. I took the top 50 hits for each of the 3 contigs, extracted the regions that contained all the alignments and then re-blasted these against the 3 rumen contigs. These revealed the structure of the island. I also checked if this island is in anyt other rumen genomes, but it doesn't appear to be.

After doing the online blast and then the local blast against the rumen contigs, I have used a combined length cutoff (of all alignments) of 4000 with a percent identity of 80%. This only removed a single contig, which only spanned part of the island from the online blast. Therefore, these should be all the island from NCBI and using the same cut offs, I confirmed that the island is not in any further rumen genomes.

## Analysis steps



## Gene diagrams of representative sequences

