Module 5: Genome annotation for antimicrobial resistance
Part 2: Pathogenwatch
Instructor: Rito Mikhari

PRACTICAL ASSIGNMENT

Task A:
For this part of the assignment, genomic raw reads of an unknown bacterial species have been extracted from a publicly available nucleotide database. We've named these ACORN75075_S10_R1_001.fastq.gz and ACORN75075_S10_R2_001.fastq.gz
 (https://wcs_data_transfer.cog.sanger.ac.uk/Module_5.zip). We are going to use pathogenwatch, a cost-free online tool, to identify and characterize the bacteria.

Instructions

- Open your web browser and navigate to https://pathogen.watch/
- Click on the "Upload" tab on the top right corner and sign in after the prompt.
- Select the "FASTQ" option as we will be looking at raw unassembled sequence data.
- Ensure that the files are named in the supported formats listed in the website.
- Drag and drop your sequences onto the website.
- The assembly will take a while, so at this point you can skip to Task B and return to answer Task A later.

Questions

1. What is the genus and species name of the bacteria you are characterizing?
2. Can you identify the sequence type and clonal complex of the pathogen?
3. Which database was searched for the MLST and cgMLST characterization of this pathogen?
4. Within this genus, only one other species is pathogenic. What is this other species?
5. Click on the taxonomy ID at the bottom of the genome report. Which phylum, class, order and family does this organism belong to?

Task B:

We will be looking at a dataset of *Salmonella Typhi* sequences with an accompanying metadata sheet to explore the collection view functions. You can download the sequences using the following

Instructions

- On Pathogenwatch, click on the Upload tab.
- Select "Single Genome FASTAs".
- Select all fasta files including the .csv file from your computer files, drag and drop to upload.
- Once complete, click on the "Genome" tab to find your genomes.
- Select all Salmonella Typhi genomes and click on "Selected Genomes" on the top right purple tab.
- Click on "Create Collection" and name your collection.
- Go to "Collections" and find your newly created collection of Salmonella Typhi genomes.
- Click on "View Collection" and generate tree.

Questions

1. First ensure that the phylogenetic tree fits to the panel, and take a snapshot of the tree. Paste the image below.
2. Zoom out on the map to the appropriate size of the global map and take a snapshot of the map. Paste the image below. Without looking at the metadata, how many countries do the genomes originate from? Name the countries.
3. Click on the query bar of the bottom panel and find the "Antibiotics" view. How many of the antibiotics are the samples not fully susceptible to?
4. All the samples are none-susceptible to one antibiotic. Click on this antibiotic and view distribution on the tree. Take a screenshot of the tree and paste below. Describe the distribution in your own words.
5. The variants responsible for the non-susceptibility mentioned in Question 4 can be analyzed in the "Variants" (query bar on the left). Describe the overall results in your own words.