# Tracking The Spread of Carbapenemases Producing Organisms Using Genomic Information: From Raw Reads to Phylogenetic Trees

Baofeng Jia[1], Dan Fornika[2], Matthew Croxen[2], William Hsiao[2], Linda Hoang[2]

1. Simon Fraser University. 2. BC Centre for Disease Control.

INTRODUCTION

- **Antimicrobial resistance (AMR) and emergence of pan resistant superbugs** – >10 million deaths by 2015
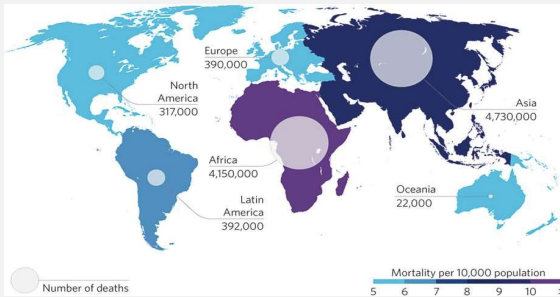
- **Carbapenemases producing organisms (CPO) -** Opportunistic pathogens found in humans, animals, environment; Most versatile extended spectrum beta-lactamases

- **CPO is nightmare for public health** - Highly transmissible AMR+ ubiquitously found + versatile resistance


- **Genomics driven CPO epidemiological investigation -**
  - Specialized expertise of scripting and command line is required for bioinformatics tools
  - Existing tools are not customized to suit the needs of CPO epidemiology
  - Lack of a proper way of characterizing and categorizing plasmids

Objective:

Produce an automated workflow to construct phylogenetic trees from raw Illumina reads that includes sequencing typing, AMR profile prediction and plasmid transmission information.
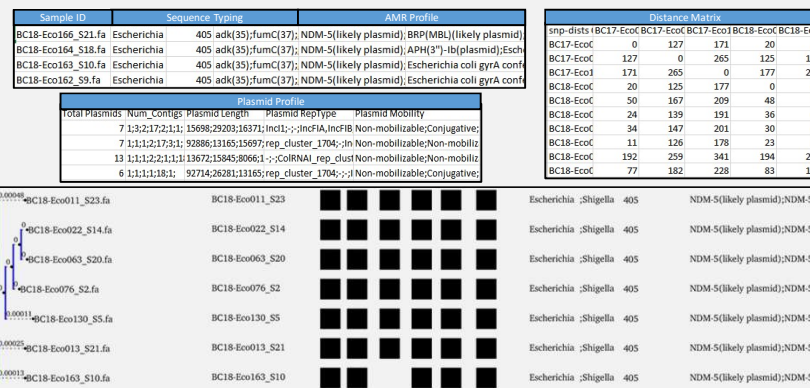
Method
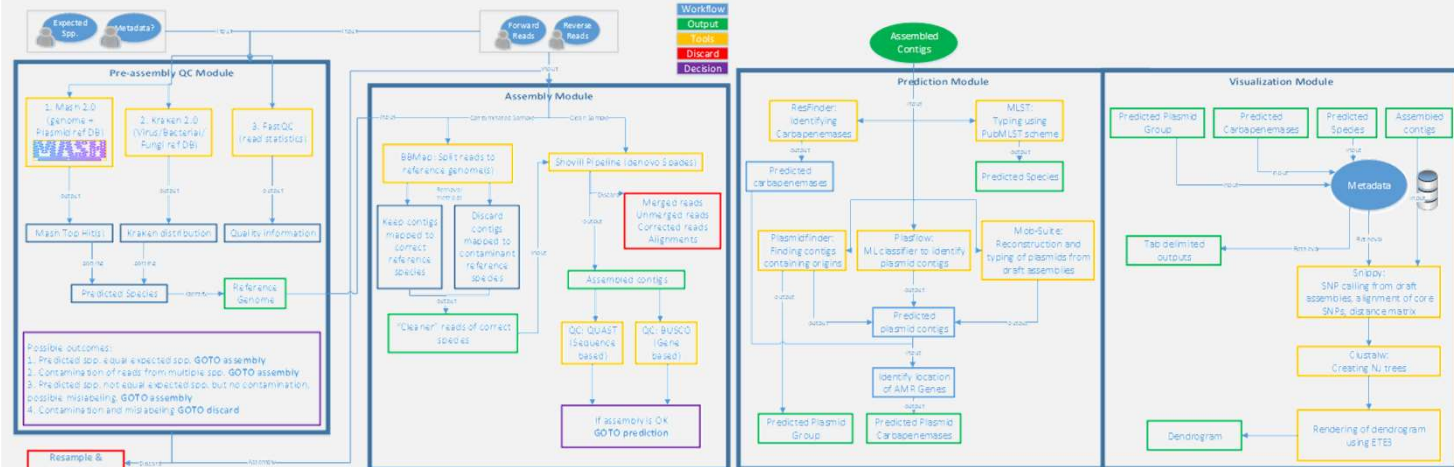
>$Python pipeline.py -i ID -f R1 -r R2 -o output -e "Klebsiella";

Available As Galaxy Workflow

The four modules of the CPO_Prediction workflow
The workflow is split into 4 modules responsible for quality check, assembly, genomic characterization and visualization. Each step of the workflow is highlighted in blue, tools used highlighted in yellow, input and outputs highlighted in green

Results

Sample ID	Sequence Typing		AMR Profile
BC18-Eco166_S21.fa	Escherichia	405 adk(35);fumC(37);	NDM-5(likely plasmid); BRP(MBL)(likely plasmid);
BC18-Eco164_S18.fa	Escherichia	405 adk(35);fumC(37);	NDM-5(likely plasmid); APH(3")-Ib(plasmid);Esch
BC18-Eco163_S10.fa	Escherichia	405 adk(35);fumC(37);	NDM-5(likely plasmid); Escherichia coli gyrA conf
BC18-Eco162_S9.fa	Escherichia	405 adk(35);fumC(37);	NDM-5(likely plasmid); Escherichia coli gyrA conf

		Distance Matrix			
snp-dists	BC17-EcoC	BC17-EcoC	BC17-Eco1	BC18-EcoC	BC18-EcoC
BC17-EcoC	0	127	171	20	50
BC17-EcoC	127	0	265	125	167
BC17-Eco1	171	265	0	177	209
BC18-EcoC	20	125	177	0	48
BC18-EcoC	50	167	209	48	0
BC18-EcoC	24	139	191	36	60
BC18-EcoC	34	147	201	30	40
BC18-EcoC	11	126	178	23	53
BC18-EcoC	192	259	341	194	236
BC18-EcoC	77	182	228	83	101

Plasmid Profile				
Total Plasmids	Num_Contigs	Plasmid Length	Plasmid RepType	Plasmid Mobility
7	1;3;2;17;2;1;1;	15698;29203;16371;	Incl1;-;-;IncFIA,IncFIB	Non-mobilizable;Conjugative;
7	1;1;1;2;17;3;1;	92886;13165;15697;	rep_cluster_1704;-;-;In	Non-mobilizable;Non-mobiliz
13	1;1;1;2;2;1;1;1	13672;15845;8066;1-;-;	colRNAI_rep_clust	Non-mobilizable;Non-mobiliz
6	1;1;1;1;18;1;	92714;26281;13165;	rep_cluster_1704;-;-;I	Non-mobilizable;Conjugative;

Example output of the CPO_prediction workflow for an *E.coli* outbreak cluster

Future Directions

1) Validation of workflow with existing methods

2) Incorporation of sensitive patient data

3) Better streamlining module 1 of the workflow to remove the need for command line

4) Recombination filtering for SNV phylogenomics

Reference

1. Dugden R, Kelly R, Davies, S. Combatting antimicrobial resistance globally. Nature Microbiology. 2016; 1(16187)

Baofeng (Justin) Jia
M.Sc. Candidate, Simon Fraser University
BC CDC Bioinformatics Intern
Contact: bja20@sfu.ca