

# **AMR-TV: Antimicrobial Resistance Transmission Visualizer**

Ivan Gill

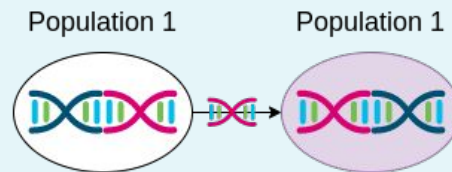
## Background

### AMR (antimicrobial resistance)

Bacterial populations with genes that provide resistance against antibiotics

### AMR transmission

Horizontal gene transfer (hgt) between populations



### Possible workflow for monitoring AMR transmission routes

1. **Identify potential routes by comparing AMR genes across populations**
  2. Validate potential routes using phylogenetic sequence comparison techniques
- **Hard to do when dealing with thousands of rows in tabular datasets** ↘

**GOAL: develop solution for finding and visualizing *potential* routes of AMR transmission from tabular data**

# 1. NCBI Isolates Browser (~700k data items!)

isolate	organism_group	create_date	location	isolation_source	host	amr_genotypes
PDT000873649.1	C. jejuni	2020-10-08	USA:FL	raw intact chicken		{blaOXA}
PDT000857925.1	Enterobacter	2020-10-30	USA: Midwest		Homo sapiens	{ant(2'')-Ia,blaOXA,fosA}
PDT000862009.1	C. jejuni	2020-10-19	USA:TN	chicken carcass		{blaOXA}
PDT000858659.1	Enterobacter	2020-10-08	Northeast	blood	Homo sapiens	{aph(3'')-Ib,aph(6)-Id,blaACT-17}

## 2. Derived network (DN)

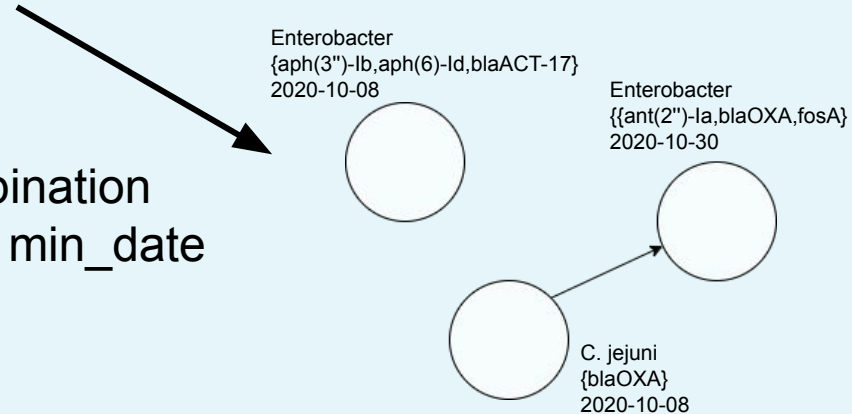
### Nodes

Unique organism\_group-amr\_genotypes combination

Attributes = organism\_group, amr\_genotypes, min\_date

Directed link between node A and node B if:

- A.min\_date < B.min\_date
- A.amr\_genotypes  $\subseteq$  B.amr\_genotypes



## **High-level tasks: summarize and identify species connected to an outbreak**

Filter NCBI Isolates Browser by create\_date range to form DN

Summarize # of links between DN nodes belonging to different organism\_groups

## **Mid-level tasks: explicitly visualize transmission routes**

Further filter DN by organism\_group values of interest

Explicitly visualize resulting DN topology, or node-link paths

## **Low-level tasks: gain further detail on each transmission event**

Map DN nodes back to NCBI Isolates Browser to:

- Identify sampled pathogens involved in potential AMR transmission routes
- Gain further context (location, host, isolation\_source)

**High-level tasks: summarize and identify species connected to an outbreak**

Adjacency matrix

**Mid-level tasks: explicitly visualize transmission routes**

Node-link (NL) diagram

**Low-level tasks: gain further detail on each transmission event**

Table

**DEMO TIME**

## Strengths

Easier to parse than tabular data alone

Novel solution to automatically visualizing hgt networks of this scale

## Weaknesses

Inconsistent data standards reduce context (location, isolation\_source, host)

DN nodes don't represent single populations (transmission w/in nodes?)

## Lessons learned

I made two unrealistic domain assumptions when making DN links:

1. Bacteria must transmit all their AMR genes during hgt
2. hgt directionality can be inferred through date of detection

To get around these, sequence comparison **MUST** be more intimately involved in the visualization process, rather than used for post-hoc validation of NL diagram

## Future work

Allow users to upload their own files:

Common vocabulary to standardize ALL fields

- No missing context
- Use host and location information when mapping nodes

Ditch the faulty assumptions from the last slide:

Validate/filter links with sequence comparisons BEFORE visualization

Why?

- Considering events where bacteria may have transferred a subset of their AMR genes drastically increases DN size (hard to visualize without filtering)
- Need new indicator of directionality

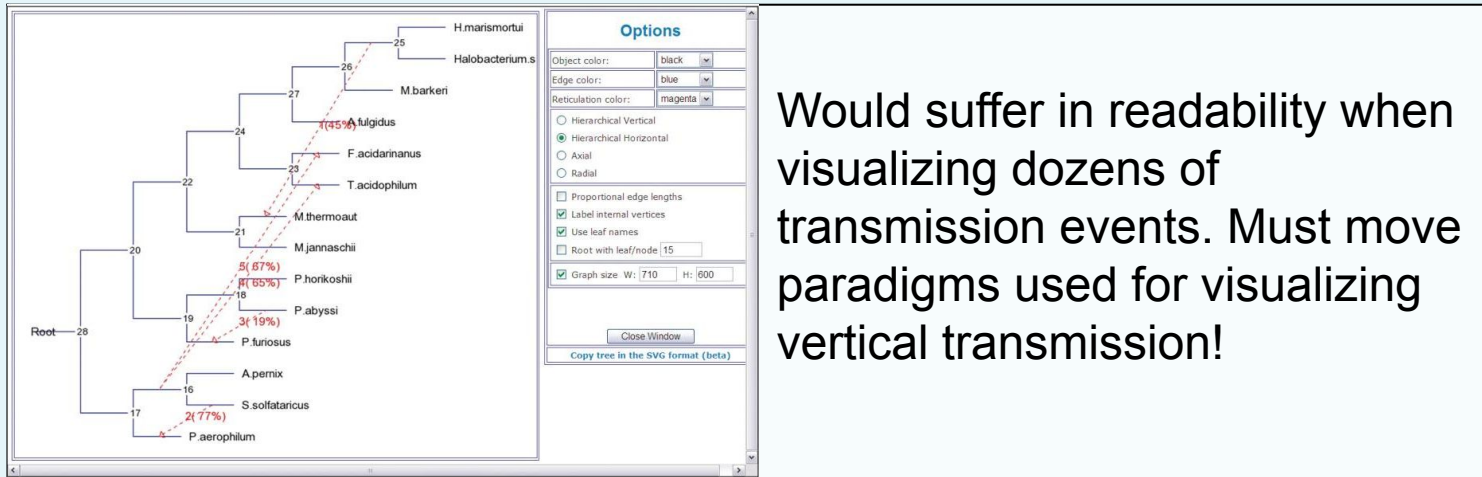
**Questions?**



# Monitoring AMR transmission

Difficult to parse from tabular data (visualization solution needed)

Current visualization solutions ignore hgt or inadequate:



Goal: create scalable visualization solution specifically for horizontal transmission

## DN Cardinality

Filters	# of nodes	Avg # of edges per node
None	~77k	Unknown
Oct 2020	~5k	~17
Oct 2020; <i>Enterobacter</i>	190	~5
Oct 2020; <i>Enterobacter</i> and <i>C. jejuni</i>	384	~8
Oct 1 2020; <i>Salmonella Enterica</i>	439	~20

## \_\_\_ # of NCBI Isolate Browser data items mapped to DN nodes

Min	1
Max	93
Avg	~9