Phylogenomic analysis of *Pseudomonas aeruginosa* isolates from bacteremia cases at CHOP

Alena Nag[1], Bianca Galis[2], Steven Jones[2] Stephanie Higgins[2], Erin Theiller[2], Paul J. Planet[2,3], Joseph P. Zackular[2,3], Ahmed M. Moustafa[2,3]

[1]College of Arts and Sciences (COL 2025), University of Pennsylvania.[2] microbial ARchive and Cryo-collection (microbialARC), CHOP Microbiome Center, Children's Hospital of Philadelphia, Philadelphia, PA,USA. [3]Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA,USA

BACKGROUND

The microbial ARchive and Cryo-collection (microbialARC) is a newly formed initiative under the PennCHOP Microbiome Program umbrella and supported by the CHOP Microbiome Center. It supports the collection, biobanking, and whole genome sequencing (WGS) of commensal and pathogenic bacteria, viruses and fungi, making them widely available for clinical and basic researchers, and provides whole genome sequences and computational tools for clinical, translational, and basic investigation. microbialARC aims to bring personalized patient care and infection control at CHOP through the power of genomics.

As part of microbialARC mission, we have biobanked pathogens from every patient with bacteremia at CHOP in collaboration with the CHOP Infectious Disease and Diagnostics Lab (IDDL) with a goal of better patient care and cutting-edge discoveries. In this study, we investigate bacteremia isolates of *Pseudomonas aeruginosa* species. Using comparative genomics techniques, we are interested in understanding the relationship between these different isolates.

AIMS

1. Determine whether there is one single strain of *Pseudomonas aeruginosa* causing a bacteremia outbreak at CHOP.
2. Determine whether *Pseudomonas aeruginosa* isolates from the same bacteremia subject will be closely related using a maximum likelihood phylogenetic tree.

METHODS: 3-STEP WORKFLOW

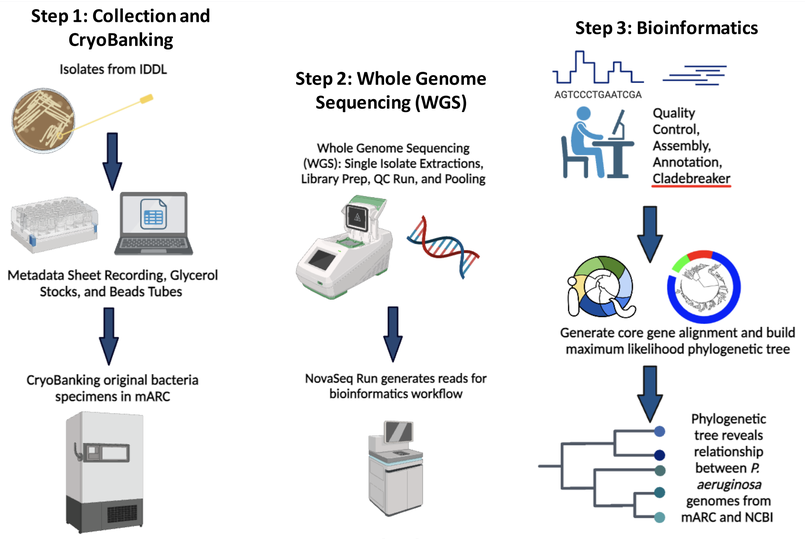

Fig 1a. **Fig 1b.** **Fig 1c.**

METHODS (CONT.)

Figure 1a. Bacteremia samples for microbial Archive and Cryo-collection (microbialARC) are collected from our ongoing studies and our collaborators across campus at Penn and CHOP, as well as the CHOP IDDL. Once collected, the laboratory workflow includes biobanking at the CHOP BioRC.

Figure 1b. Next, the workflow involves DNA extractions, library prep, and whole genome sequencing. Then, a NovaSeq machine generates reads to pass data to the bioinformatics team for analysis.

Figure 1c. The bioinformatics workflow starts with extensive quality control checks (FastQC[1] and MultiQC[2]) before assembling (Shovill[4]) and annotating (Prokka[5]) the genomes. Once assembled, we generate core gene alignment and build a maximum likelihood phylogenetic tree to understand the relationship between the different isolates from different subjects at CHOP.

RESULTS

Fig 2: Phylogenetic analysis of *P. aeruginosa* bacteremia isolates and closely related genomes from public database isolates.

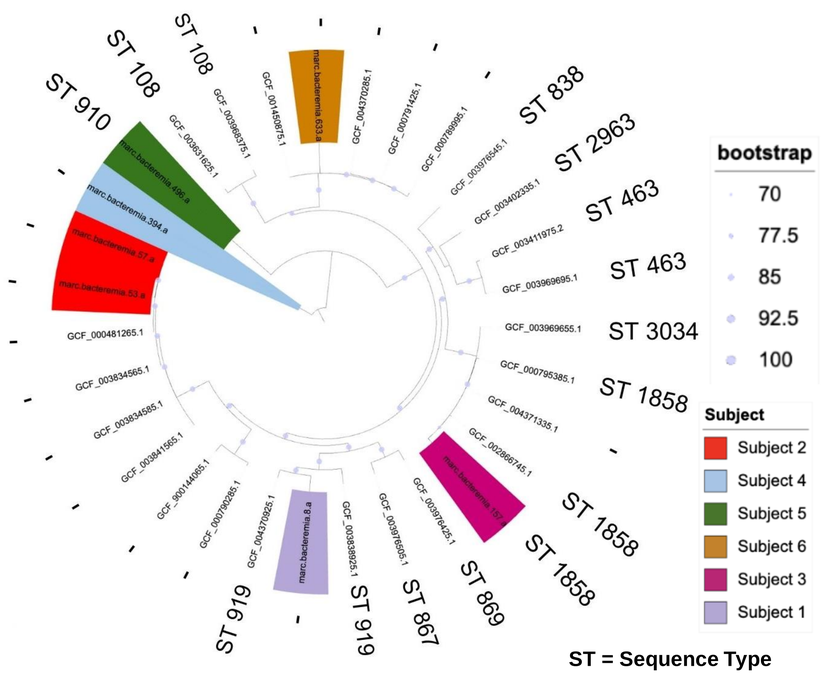

Fig 2: Maximum likelihood tree of core genes alignment for *P. aeruginosa* bacteremia isolates collected at CHOP and reference genomes from public databases (e.g. NCBI) using the Cladebreaker approach.

ST = Sequence Type
Characterizes isolates of the species based on differences in 7 housekeeping genes

March - October, 2022
Collection of isolates

FINDINGS

1. The CHOP bacteremia genomes did not segregate into a distinct clade apart from the NCBI genomic sequences, indicating that there was no single *P. aeruginosa* strain responsible for a bacteremia outbreak in the hospital.

2. The isolates of *P. aeruginosa* from the same individual were closely positioned on the maximum likelihood tree generated from the alignment of core genes, suggesting that the individual was infected only once.

3. High bootstrap values on the tree indicate that most of the clades are strongly supported.

FUTURE DIRECTIONS

In this initial endeavor by microbialARC into real-time pathogen tracking, numerous research possibilities with *P. aeruginosa* and other pathogens have been uncovered. Our primary aim is to better understand the circulating lineages of *P. aeruginosa* in bacteremia patients, as real-time tracking is crucial for improving patient care and catalyzing more extensive research in the future.

- ❖ Examine more *Pseudomonas aeruginosa* genomes from bacteremia isolates
- ❖ Look into genomes with potential recombination (marc.bacteremia.394.a)
- ❖ Possible mutations that are common in all of the bacteremia genomes
- ❖ Common accessory genes across all bacteremia isolates

REFERENCES

1 – FastQC: https://github.com/s-andrews/FastQC. Accessed 26 August 2023.
2 – MultiQC: https://github.com/ewels/MultiQC. Accessed 26 August 2023.
3 - NCBI BLAST+: https://www.ncbi.nlm.nih.gov/books/NBK279684/. Accessed 26 August 2023.
4 – Shovill: https://github.com/tseemann/shovill. Accessed 26 August 2023.
5 – Prokka: https://github.com/tseemann/prokka. Accessed 26 August 2023.

ACKNOWLEDGEMENTS

I would like to thank Dr. Ahmed Moustafa from the CHOP Microbiome Center for his support. Additionally, I would like to thank the CHOP Microbiome Sequencing Core lab members (Bianca, Erin, Stephanie, and Steven) and other members of the CHOP Microbiome Center for their guidance, and Dr. Ann Vernon-Grey at the Penn Center for Undergraduate Research and Fellowship (CURF) for providing me this opportunity through the Penn Undergraduate Research Mentoring Program (PURM) grant.