Detecting within-host diversity of *Staphylococcus aureus* during colonization

Vishnu Raghuram[1,]
@Vishnu_Raghuram
vraghu2@emory.edu
@VishnuRaghuram94
Jessica J. Gunoskey[3] Natasia F. Jacko[3] Katrina S. Hofstetter[2] Timothy D. Read[2] Michael Z. David[3]

[1] *Microbiology and Molecular Genetics, Emory University*
[2] *Division of Infectious Diseases, Emory University*
[3] *Division of Infectious Diseases, University of Pennsylvania*

Introduction

While the cost of whole genome sequencing has decreased, the process of DNA extraction, library preparation, physical sample storage, and subsequent bioinformatic analysis have not scaled down at the same rate. We believe there are open questions about the optimal genomic workflow design to minimize time and labor costs while maximizing the amount of information obtained from each clinical sample. In this work, we use *S. aureus* as a model system to compare the advantages of sequencing individual colonies versus pooled populations from clinical samples. We analyzed both single colonies as well as pooled colonies from humans who have been identified to have MRSA (Methicillin resistant *S. aureus*) positive skin and soft tissue infections (SSTI)

Methods

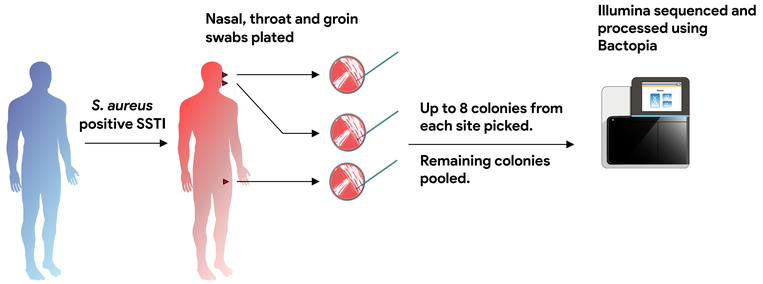

Figure 1: The patients submit nasal, throat and inguinal skin swabs at enrollment and every three months since enrollment for up to one year. Each swab is streaked out and 8 individual colonies (singles) are collected and sequenced. The remaining colonies were pooled and sequenced (pools). For each collection, reads from all eight singles/four random singles /two random singles were combined in silico to generate artificial pools.

If you work with **clinical samples** from **microbial infections**, please **sequence** the entire **population**!

EMERGENT Lab
EMORY UNIVERSITY

I am graduating soon!
Scan for more info

Results

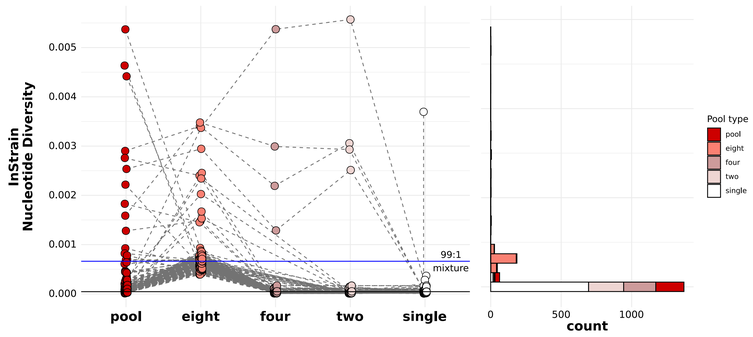

*Figure 2: **Most patients have relatively homogeneous pools - a single colony has as much diversity as the pool.** The nucleotide diversity as measured by InStrain between single colonies, two/four/eight singles combined, and the pools suggest that most multi-colony samples have a diversity value equal to the average diversity expected in a pure single colony*

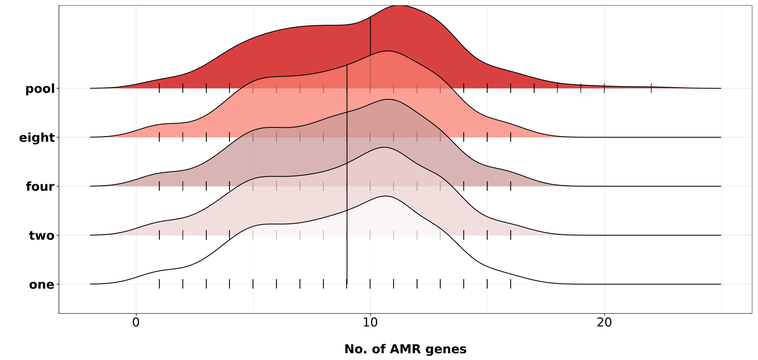

*Figure 3: **Though rare, there can be antibiotic resistance genes in the pool that are not found in individual colonies**. The number of AMR genes as detected by AMRFinderPlus in single colonies, two/four/eight singles combined, and the corresponding pools show that in rare cases, more AMR genes are dectected in the pools compared to the singles or artificial pools.*

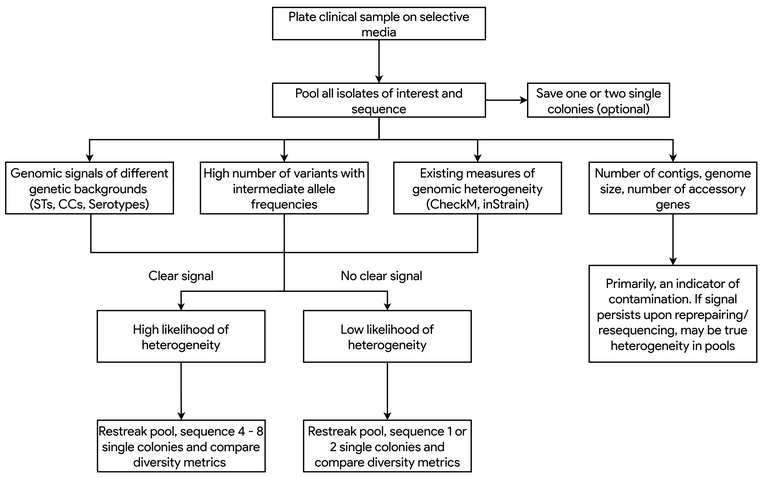

Figure 4: Measuring genetic diversity metrics from pooled population sequences can dictate further sampling protocols. Low pool diversity = fewer singles needed to represent total population. Following this workflow can reduce time and labour costs associated with clinical sample analysis.