Composite-Sample Complex: Building a genomics model to understand the evolutionary history of antimicrobial resistance gene movement

Amy J. Mathers[1], Thomas Li[1,2], Hardik Parikh[1], Katie Barry[1], Chava Castaneda Barba[1], Sharvari Narendra[1], Shireen Kotay[1], Qijun He[2], Christian Reidys[2]

[1] School of Medicine, University of Virginia, Charlottesville, VA USA

[2] Biocomplexity Institute, University of Virginia, Charlottesville, VA USA

Background: Most emerging antimicrobial resistance (AMR) genes are located on mobile genetic elements (MGE), that can be shared between bacterial strains and species. However, the availability of validated MGE tools to understand evolutionary history often falls short in part because of incomplete sampling frames on which to verify performance. We have developed and validated a novel mathematical approach to capture gene movement within a real-world confined longitudinal environmental sampling frame.

Methods: Between Dec-2013 and Dec-2018, we cultured for *Klebsiella pneumoniae* carbapenemase producing organisms (KPCOs) from six ICU sinks from a single US hospital. Each sink drain was considered a longitudinal biofilm using dates and location as the associated metadata. Both Illumina and Oxford Nanopore sequencing platforms were applied to all isolates to generate hybrid assemblies and annotation PlasmidFinder and AMRFinder. The annotation results and metadata were assimilated into a simplicial complex model, which uses chromosome and all plasmids as potential sources and targets of *bla*$_{KPC}$ together in the model to capture movement and directionality of movement over time. Results were then combined with TETyper results from Tn*4401* transposition to assess and verify bla$_{KPC}$ movement patterns.

Results: Over the course of 286 samplings from sinks, we successfully sequenced 82 KPCOs. These isolates revealed 14 distinct strains across 10 species of bacteria, with *Serratia marcescens*, *Citrobacter freundii*, and *Raoultella ornithinolytica* being the most common. We identified 113 blaKPC-carrying plasmids across 15 unique plasmid replicon types, with many bacteria carrying multiple instances of *bla*$_{KPC}$. Using the Composite-Sample Complex approach on the isolates, we determined the movement of *bla*$_{KPC}$ via the transposition of Tn*4401* into the chromosome and other plasmids in several instances. In addition, we note that a “non-typeable” plasmid pKPC_UVA01 was shared across several species. By analyzing structural variations using TETyper and specific mutations, we were able to validate the mathematical approach and show that the movement of *bla*$_{KPC}$ into the chromosome from an original plasmid location was most likely within the sink biofilm in two of three instances.

Conclusions: Our study utilized closed genomes from a relatively stable real-world environmental niche over a five-year period to develop and validate a multimodal model. This model effectively captures the evolutionary history of *bla*$_{KPC}$ movement within and between strains and species of bacteria. Our approach revealed frequent AMR gene mobility, which was found to vary by genomic context and associated MGE. Evolutionary models that account for dynamic variation with associated MGE will be crucial for accurately tracking AMR movement, thereby aiding in developing effective strategies to combat the emergence of AMR.