RESEARCH IN PROCESS

Illuminating the role of whole genome sequencing in produce safety

SUMMARY:

Whole genome sequencing (WGS) is rapidly becoming the gold standard for foodborne outbreak investigations by public health agencies around the world. Source-tracking investigations during an outbreak can often come down to the epidemiological data and a difference of just a few single nucleotide polymorphisms (SNPs) in the foodborne pathogen's genome (**Figure 1**). Investigations can be made even more challenging with highly clonal pathogens that lack high levels of genetic diversity (**Table 1**). Therefore, the goal of this project is to provide data to public health agencies to help improve the use of WGS as an outbreak investigation tool. Furthermore, this project can be used by the produce industry to implement WGS for internal source tracking to identify "resident" versus "transient" pathogens, sources of contamination for either, and a better overall understanding of the breakdowns or gaps in prevention, thus improving produce safety by closing these gaps.

OBJECTIVE:

1. Determine mutational rates of pathogens during persistent colonization in agricultural soil and irrigation water under distinctive geographical environmental conditions.

METHODS:

Irrigation water and soil samples will be collected from romaine lettuce fields in Arizona and California and transferred back to the Cooper laboratory. Individual agricultural samples along with buffered peptone water control will be inoculated with either *Salmonella*, *Listeria*, or *Escherichia coli* O157:H7, and then maintained under environmental conditions (light, humidity and temperature) for Yuma, AZ and Salinas, CA. Every two weeks for the first three months, and then monthly afterwards, each of the inoculated samples will be cultured for the corresponding inoculated pathogen, and five colonies of the pathogen will be selected for genome sequencing. Each of the five selected colonies will be whole genome sequenced and compared to the genome of the original pathogen inoculum to determine the number of mutations that have developed since the initial inoculation. The project will maintain the inoculated samples for up to one year to establish an annual mutation rate (**Figure 2**).

RESULTS TO DATE:

We are in the process of obtaining soil and irrigation water samples from the leafy greens fields in Arizona and California, and will be inoculating each agricultural sample with *Salmonella*, *Listeria*, and *Escherichia coli* O157:H7 for the long-term evolution studies under the different environmental conditions prior to the CPS Research Symposium. Due to the long-term nature of this study, initial mutational rates for the various pathogens will not be available until at least November or December 2019.

Table 1. *Salmonella* Typhimurium strains from different sources and temporal isolation demonstrating the difficulty of source tracking

Strain	Source	State or country	Isolation date	# of SNPs compared to reference strain 14028s[1]	# of SNPs compared to RM13677[2]
14028s (Reference strain)	Poultry	N/A	1960	N/A	13
RM6835	Celery	Washington	2002	9	4
RM6837	Cantaloupe	Washington	2002	9	4
RM10602	Pond Water	California	01/23/2009	12	5
RM13670	Clinical	Oregon	08/04/2009	11	4
RM13671	Clinical	Oregon	08/08/2009	10	3
RM13672	Clinical	Oregon	08/08/2009	9	4
RM13673	Clinical	Oregon	08/13/2009	10	3
RM13674*	Clinical	Oregon	08/14/2009	10	3
RM13675	Clinical	Oregon	08/22/2009	11	4
RM13676*	Clinical	Oregon	09/15/2009	10	3
RM13677	Clinical	Oregon	10/21/2009	13	N/A
RM14512	Clinical	Maine	08/06/2009	9	4
RM14513	Clinical	Pennsylvania	08/08/2009	10	3
LT2 (Outlier strain)	N/A	Sweden	1940s	394+	380+

*Samples isolated from the same person
[1]Number of SNPs in non-repeat regions compared to the complete reference genome of *S.* Typhimurium strain 14028s
[2]Number of SNPs in non-repeat regions compared to most distantly related of the 18 strains *S.* Typhimurium strain RM13677
+SNPs for *S.* Typhimurium str. LT2 are only the chromosome

BENEFITS TO THE INDUSTRY:

This project has direct and indirect benefits to the produce industry. The indirect benefits include an improvement of WGS as a tool for outbreak investigations, thus benefiting the industry by allowing for more accurate and rapid outbreak investigations. The project helps support the further development of the source tracking database, GenomeTrakr, especially the aspect of speeding up investigations using geographical or regional pathogen data. This benefits the industry by decreasing outbreak investigation times, removing contaminated products faster, preventing additional cases, and ultimately getting an outbreak declared over faster. The direct benefit of this project is that it improves the use of WGS as a tool that can be applied for internal source tracking of contamination in a facility. Moreover, using these mutation rates can identify the source of "resident" or long-term pathogen contamination by tracking the pathogen through a produce facility, as has been applied in other food industries, and identify and close gaps in prevention.

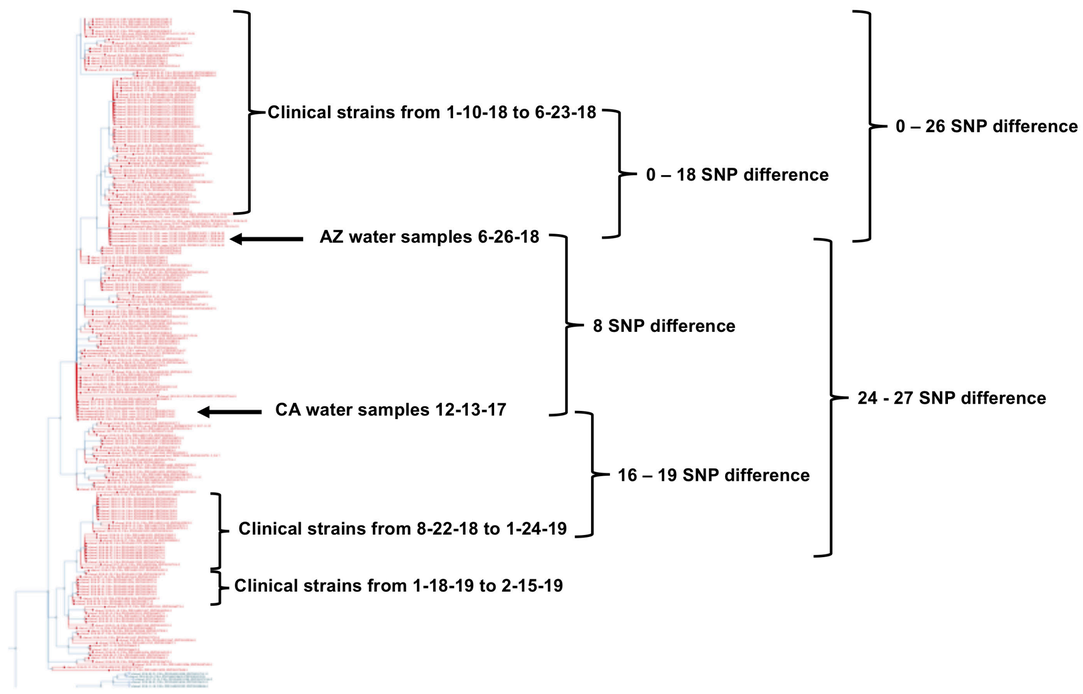

Figure 1. Phylogenic tree from the National Center for Biotechnology Information (NCBI) Pathogen Detection database, representing the complexity of WGS and source tracking. This tree shows single nucleotide polymorphism (SNP) differences among clinical and environmental *Escherichia coli* O157:H7 associated with the 2018 Arizona romaine lettuce outbreak. Results show up to 26 SNP differences between clinical strains and irrigation water strains that were isolated in Arizona during June 2018, whereas up to 27 SNP differences between irrigation water samples and clinical strains from as recently as January 2019. This example demonstrates the importance of the epidemiological data obtained during the outbreak investigation, while also the need to refine WGS as a tool to assist investigations.

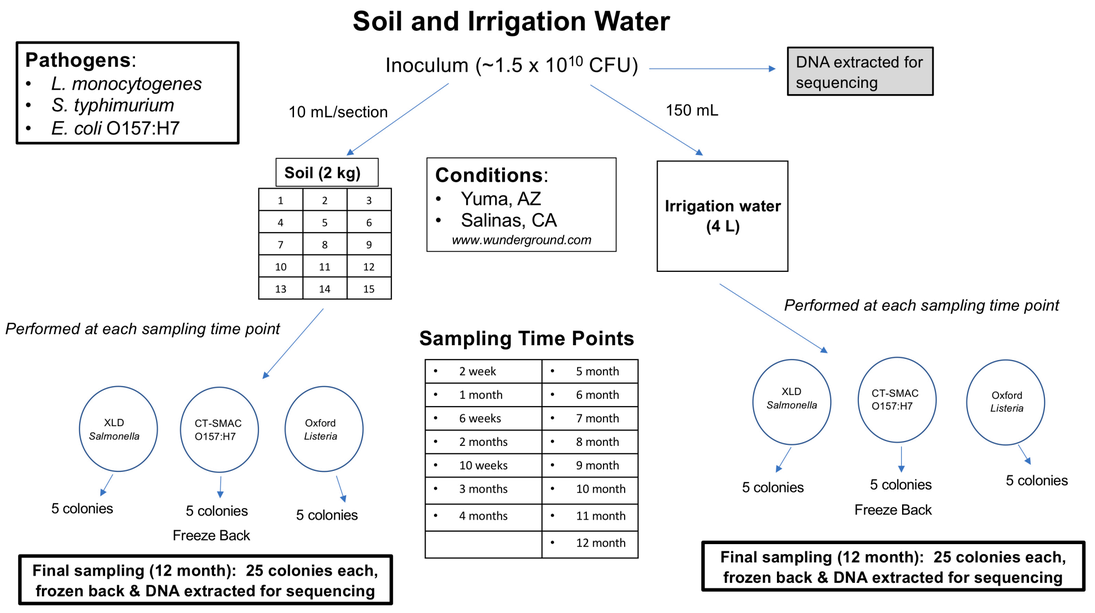

Figure 2. Flowchart of irrigation water and soil experiments for the entire project.

CONTACT

Dr. Kerry Cooper
The University of Arizona
E: kcooper@email.arizona.edu

AUTHORS

Kerry Cooper

ACKNOWLEDGEMENTS:

Many thanks to my industry collaborator, Megan Chedwick, Church Brothers Farms and True Leaf Farms for continued support in numerous aspects of the project, and to Courtney Sams for her technical assistance in accomplishing the objectives of the project.

LENGTH OF FUNDING

January 1, 2019 – December 31, 2020

CPS CENTER *for* PRODUCE SAFETY