CPS 2015 RFP
FINAL PROJECT REPORT

Project Title
Comparative genomics analysis and physiological assessment of the avirulent *Salmonella* surrogate relevant to produce safety

Project Period
January 1, 2016 – December 31, 2017

Principal Investigator
Julie Meyer (January–December 2017)
University of Florida /IFAS
Soil and Water Sciences Department
Genetics Institute Rm330E
2033 Mowry Rd.
Gainesville, FL 32611
T: 352-273-8195
E: juliemeyer@ufl.edu

(Max Teplitski: January–December 2016)

Co-Principal Investigator
Marcos de Moraes
University of Florida /IFAS
Soil and Water Sciences Department
Gainesville, FL 32611
T: 352-273-8195
E: mhmoraes@ufl.edu

Objectives
1. Characterization of the outbreak *S. enterica* serovars under the conditions that simulate pre- and post-harvest environments and comparison to the avirulent *S. enterica Typhimurium* surrogate.
2. Identification of the core and accessory genomes of the serovars associated with produce-related outbreaks and comparison to the avirulent surrogate.
3. Functional characterization of the core and accessory genomes of the serovars associated with produce-linked outbreaks.

Funding for this project provided by the Center for Produce Safety through:
CDFA SCBGP grant# SCB15063
Abstract

Coliforms and generic *Escherichia coli* are poor predictors of behavior of human pathogens like *Salmonella*, pathogenic *E. coli*, and *Listeria* in the crop production environment. Mounting evidence suggests that accurate models of *Salmonella* behavior in the production environment will have to be built based on experiments conducted with *Salmonella*, and not based on data from distantly related surrogates like generic *E. coli*. This, however, necessitates the availability and careful characterization of “disarmed” strains of *Salmonella* that can be used for on-site research. A fully avirulent, non-immunogenic, non-transgenic, and antibiotic-sensitive surrogate of *Salmonella* was constructed with previous CPS funding. With the completion of this project, we have validated the use of the avirulent surrogate as a new tool in addressing FSMA regulations in the pre- and post-harvest production environments. Outbreaks of produce-borne salmonellosis have been disproportionately associated with only a few of the 2,500 *Salmonella* serovars. The comparative genomics analysis completed during this project revealed that this difference in ability to grow in produce is based in the genetic characteristics of particular strains.

Background

*Salmonella* and common indicator/index organisms. The need for suitable nonpathogenic indicators of microbiological safety of the environment has long been recognized. However, even after a 40-year-long search for an ideal non-pathogenic indicator, the behavior of which correlates with that of pathogens, it has not been found [1]. The search for a suitable index organism that could reliably predict the presence of *Salmonella* in the crop production environment has similarly been unsuccessful, likely because even in production fields where *Salmonella* has been isolated, its prevalence is very low and establishing correlations with the presence of potential index organisms is statistically challenging. In general, *Salmonella* is commonly accepted as being better able to survive in the environment than *E. coli* [2], although *Salmonella* is not as common in environmental samples as *E. coli* or coliforms. *Salmonella* can better persist in processed organic waste and amended soils compared with fecal coliforms, although its persistence appears to be strain- and/or site-dependent [3-5]. The persistence pattern of *Salmonella* on the surfaces of peaches, cantaloupes, lettuce, and bell pepper has been shown to be very different from that of generic *E. coli* (and/or select microbial indicators) [6-8]. These differences in survival and re-growth under field conditions and in the vegetable production chain raise important questions about the suitability of indicators for predicting the behavior of *Salmonella* in manure, irrigation water, and in association with plants. The remarkable genetic diversity of non-typhoidal *Salmonella*, reflected in its physiological adaptability and phenotypic plasticity, further complicate the search for an indicator organism that behaves like *Salmonella* under production conditions. It is almost certain that to accurately predict behavior of *Salmonella* under the conditions relevant to the produce industry, a suitable strain of *Salmonella* will have to be used.

Serovar diversity and links to produce outbreaks. There are over 2,500 serovars of non-typhoidal *Salmonella*, however only a dozen of the serovars have been consistently associated with produce-borne outbreaks. This does not seem to be a stochastic association. Strains of *S. Newport* have been commonly associated with tomato-linked outbreaks of salmonellosis [9], while *S. Enteritidis* PT30 has been consistently isolated from almond orchards over a 5-year period and linked to several almond outbreaks [10, 11]. At least two hypotheses could be formulated to address this observation. One possibility is that certain serovars of *Salmonella* are
better adapted to persistence in the crop production environment (fields, processing facilities, in association with plants, etc.). In fact, when growth of different Salmonella serovars was compared, Newport (originally recovered from tomato fields on the Eastern Shore of Virginia [ESVA]) reached higher cell numbers in tomato pericarps, suggesting that this serovar is better adapted to survival within tomatoes [12]. The superior fitness of Newport under agricultural conditions also was found in another study [13], where this serovar was more fit in the tomato rhizosphere and phyllosphere. Salmonella serovars Newport, Montevideo, and Hadar were more adapted to growth within tomatoes, and Montevideo reached higher cell numbers in tomato fruits inoculated at the blossom stage [14]. Similarly, almond orchard isolates of S. Enteritidis generally had a metabolic profile that was distinct from clinical isolates [11], suggesting that despite the overall genotypic similarity to the clinical strains, almond isolates had unique metabolic adaptations. While outcomes of the studies from different labs consistently point to the possibility that certain serovars are better adapted to persistence in the production environment, little is known about physiological or genetic mechanisms underlying this adaptation. The second hypothesis is that some strains are more virulent and/or express their virulence genes during colonization of fruits and vegetables, which makes them more likely to cause an infection in humans when ingested. With this study, we addressed these two possibilities. Bridging this knowledge gap may improve approaches for managing human pathogens in the production environment.

Several tools can be employed to identify serovar-level adaptations. Our preliminary comparative genomics analysis to identify accessory and core genomes revealed large genomic islands that were unique to S. Newport and are involved in its persistence within tomatoes. We expect that a similar analysis of genomes of other produce outbreak-associated serovars may define genetic mechanisms that make these strains more persistent in the production environment. A similar approach has been successful in the analysis of serovar Weltevreden (responsible for recurring produce-borne outbreaks of salmonellosis in Europe) and led to the identification of the unique genomic islands that were expressed when this outbreak strain was grown on alfalfa sprouts [15]. We anticipate that our systematic analysis of the most common Salmonella outbreak strains will reveal their adaptations to persistence in soil or on/in plants, and these discoveries will be important for modeling behavior of pathogens in the production chain and devising strategies for their control.

Research Methods and Results

Objective 1. Characterization of the outbreak S. enterica serovars under conditions that simulate pre- and post-harvest environments and comparison to the avirulent S. enterica Typhimurium surrogate.

Objective 1 Methods:
Attachment to abiotic surfaces. Biofilm assays were performed on the avirulent S. enterica Typhimurium surrogate and the wild-type parental strain as described before [20] on polystyrene, vinyl, and stainless steel surfaces. Briefly, overnight cultures of bacteria were seeded in microtiter plates (polystyrene or vinyl) or stainless steel caps, incubated statically overnight at 22°C or at 4°C. Supernatant was decanted and the ability of bacteria to form biofilms was estimated using crystal violet staining, followed by quantification of the stained biofilms using a microtiter plate reader.
Thermal inactivation. Resistance to heat stress of the avirulent S. enterica Typhimurium surrogate and the wild-type parental strain was tested essentially as in [21]. Briefly, cultures were grown to exponential phase in LB, cells were pelleted, resuspended in 10 ml of PBS, and incubated at 50°C. Triplicate samples were collected after 1, 2, and 3 hours. A decrease in CFU
counts per ml was measured by serially diluting samples in PBS and plating them onto LB agar, followed by overnight incubation at 37°C.

Inactivation by disinfectant. Inactivation by Pheno-Tek II (Bio-Tek) was tested in the avirulent S. enterica Typhimurium surrogate and the wild-type parental strain. Overnight cultures were spun down and resuspended in PBS. Pheno-Tek II was diluted in PBS at 1:250 (application rate recommended by manufacturer), 1:500, 1:1,000, 1:5,000, or 1:50,000. A decrease in CFU counts per ml was measured by serially diluting samples in PBS and plating them onto LB agar, followed by overnight incubation at 37°C.

Proliferation within cantaloupes and red and green tomatoes. Shallow wounds on cantaloupe or tomato surfaces were seeded with ~100–1,000 cells of the outbreak serovars Typhimurium, Agona, Braenderup, Javiana, Montevideo, Newport, and Poona and the avirulent Salmonella surrogate. Fruits of two tomato varieties (red Solanum lycopersicum L. cv. Campari or green S. lycopersicum cv. Success) grown in the field or in the greenhouse under conventional production conditions were used for these experiments. Inoculated mature green tomatoes were incubated at 18–21°C, 90–95% relative humidity (to mimic standard ripening conditions) and 14–16°C (to mimic slow ripening in transit); inoculated red fruits and ripe cantaloupes were incubated under the conditions that mimic post-harvest storage at 7–10°C. Salmonella was recovered from internal tissues at the end of a week-long incubation period. Infected produce was macerated by stomaching in PBS and plating the dilutions onto a selective XLD medium.

Persistence on spinach leaves and cut romaine lettuce. Spinach plants, grown under organic agriculture-like conditions in the field (Archer, FL) were transplanted into individual pots and brought to the lab. Salmonella inoculum was prepared essentially as for inoculation of tomatoes and cantaloupes, and a suspension of cells was spotted onto the abaxial surfaces of spinach leaves (application site was labeled with permanent marker). Leaves were harvested after 1 or 3 days and the application site was excised from the leaf, ground with glass beads, and then supernatants were serially diluted and plated on XLD agar. Salmonella persistence on spinach leaves was expressed as the percentage of CFU recovered from the leaves in relation to total CFU inoculated. Bagged whole hearts of romaine were purchased from a local supermarket (Winter Haven, FL). Leaves were cut across the midrib in lengths of ~10 cm. Individual samples were spot inoculated onto the abaxial leaf surface. Salmonella populations were enumerated at 0, 1, 2, 5, 7, 14, and 21 days. At each time point, an individual leaf sample was combined with 10 ml of 0.1% peptone water. Samples were placed in a mechanical stomacher for 60 s. Serial dilutions were plated on XLD agar. To increase the sensitivity and lower the limit of detection to 1 log CFU/leaf, 1 ml of the initial dilution (10^-1) was spread over 4 plates (0.25 ml/plate). XLD plates were incubated overnight at 37°C; colonies were counted manually, and levels were reported as CFU counts per leaf.

Persistence in soil. Fitness of the outbreak strains and the avirulent surrogate was tested in envirotron chambers (available as BLS-1 and BSL-2 facilities at the UF Genetics Institute). Generally, the design of the study conformed to the Framework for Developing Research Protocols for Evaluation of Microbial Hazards for Untreated Soil Amendments [23]. The overall experimental design was similar to the one described previously [24]. Briefly, two types of soil (Gainesville loamy sand and Candler fine sand) were collected in production fields in Florida. Soil (~500 g) was weighed into a polyethylene bag and mixed with 5 ml of the inoculum. To mix the inoculum with soil, the bag was closed and rubbed by hand for 1 min. The contents of the bag were poured out and spread over two sheets of filter paper, briefly air-dried at ambient temperature and then pooled and mixed thoroughly. The pooled soil was divided into polyethylene bags, and sterile distilled water was added to achieve the desired moisture levels. Moistened inoculated soil (100 g) was placed in sterile screw-cap bottles and incubated at 20 ± 2°C for sampling at 0, 1, 2, 5, 7, 14, and 21 days. Salmonella was sampled on XLD medium.

Persistence in irrigation water. Unchlorinated well-water was drawn from a private well in
Archer, FL. Surface water was collected from Lake Alice on the Gainesville campus of the University of Florida and filter-sterilized or autoclaved for 30 min. Outbreak strains and the avirulent surrogate were inoculated into the water samples and incubated with gentle agitation for up to 21 days at 15–20°C and 30°C (temperatures that approximate ground and surface water temperatures during the production season). Salmonella colony forming units were counted after dilution plating on XLD medium.

**Objective 1 Results:**
Attachment to abiotic surfaces. Biofilm formation by the wild-type strain S. Typhimurium 14028 was more extensive on polypropylene than on polystyrene. The avirulent surrogate formed approximately twice as much biofilm as the parental wild-type strain on the two plastics; however, it showed no difference in biofilm formation from the wild-type on stainless steel. These results suggest that the surrogate may exhibit improved survival in response to treatments or conditions where biofilm is protective, which may lead to underestimation (but not overestimation) of effects.

Inactivation by heat or disinfectant. The avirulent surrogate was similarly affected by heat stress and disinfection with Pheno-Tek II as the parental wild-type strain. Heat stress for 1, 2, or 3 h induced a 100-fold decrease in CFU per ml. Pheno-Tek II at 1:250, 1:500, and 1:1,000 dilution levels inhibited all growth of Salmonella.

Proliferation within cantaloupes and red and green tomatoes. Salmonella levels increased by about 10,000-fold in tomatoes or cantaloupes at 3 days or 5 days, respectively. The avirulent surrogate strain had somewhat decreased growth in green, but not red, tomatoes, although the difference was not statistically significant. The growth of the avirulent surrogate was indistinguishable from the wild type in cantaloupes. Proliferation of Salmonella serovars Typhimurium, Agona, Braenderup, Javiana, Montevideo, Newport, and Poona in red tomatoes exhibited the most variability, with the highest growth in Agona and Braenderup. Typhimurium grew significantly better in green tomatoes than did all other tested serovars. Montevideo and Javiana achieved the highest levels of growth in cantaloupes, although results between trials were variable.

Persistence on spinach leaves and cut romaine lettuce. On leaf surfaces, culturable populations of Salmonella declined over time. On spinach leaves, both the wild-type S. Typhimurium and the avirulent surrogate demonstrated a 91% decrease at 1 day after inoculation and further decreases 3 days after inoculation.

Persistence in soil and water. All tested Salmonella strains, including the wild-type strain S. Typhimurium, the avirulent surrogate, and the produce-associated serovars Agona, Braenderup, Javiana, Montevideo, Newport, and Poona declined in abundance but were still detectable in soil for 14 days and in water for 7 days, the duration of experimental test periods. Salmonella abundance decreased 100-fold within 2 days in water and decreased another 100-fold after 7 days in water. Overall, these results indicate that Salmonella virulence genes located on pathogenicity islands 1 to 5 and/or the virulence plasmid do not contribute to the persistence of this organism outside of animal hosts.

**Objective 2.** Identification of the core and accessory genomes of the serovars associated with produce-related outbreaks and comparison to the avirulent surrogate.

**Objective 2 Methods:**
Genome retrieval and gene prediction. Raw reads from Illumina sequencing of Salmonella isolates were retrieved from the Sequence Read Archive hosted by the National Center for Biotechnology Information, and assembled de novo with SPAdes [25] to remove bias associated with different assembly methods. The quality control of genome assembly was done using
CheckM [26]; genomes with more than 1% of contamination and less than 99% completeness were excluded. Gene predictions for the remaining high quality assembled genomes were performed with Prokka [27] and a pangenome matrix was created with Roary [28].

**Objective 2 Results:**
The pangenome analysis of 2,098 high quality *Salmonella* genomes identified 31,675 gene orthologues. While serovar Typhimurium exhibited a small standard deviation in the number of genes per genome, the serovar Newport contained genomes that had up to 2,000 additional genes due to prophage and plasmids. *Salmonella* serovar Typhimurium had an average of 4,628 genes per genome, while serovar Newport Group II had 4,554 genes per genome and Newport Group III had 4,413 genes per genome. The characterization of the presence/absence profile of genes across *Salmonella* genomes allowed the identification of the 724 genes shared in the Newport Group III shell and 84 genes unique to its core. Since we expected that the adaptation to plants is associated with this group, those genes involved in this phenotype should be exclusive or more frequent in this group. The pangenome analysis also confirmed that the *S. enterica* Newport strain C4.2 used for transposon mutant screening in Objective 3 was a typical representative of the genome content of Newport Group III serovars.

**Objective 3.** Functional characterization of the core and accessory genomes of the serovars associated with produce-linked outbreaks.

**Objective 3 Methods:**
Libraries of *S. enterica* Typhimurium and *S. enterica* Newport C4.2 Tn5 insertion mutants were constructed and sequenced by paired 150-bp reads with Illumina. Mutant sequences were mapped to reference genomes with Bowtie 2 [29]. 230,000 transposon insertions were screened for growth in tomato pericarps to identify genes associated with reduced fitness, followed by validation of the mutant screening results using competition assays of isogenic mutants lacking genes involved in amino acid biosynthesis (*thrC, metA, ilvD, trpC, glnG, glnA*) against the wild type.

**Objective 3 Results:**
The massive mutant screen to identify functions that the outbreak strains use to colonize tomatoes revealed several interesting results. First, *Salmonella* strains require two key functions for survival in both animals and plants: lipopolysaccharide (LPS) and nucleotide biosynthesis. The production of LPS results in a coating over the outer membrane of pathogen cells, which essentially hides the pathogen from the host’s defense systems because most immunity is based on the detection of unique identifying molecules on the exterior of the pathogen cell. In addition, it appears that the functions required by *Salmonella* for proliferation in tomatoes have little overlap with essential functions in phytopathogens. However, both *Salmonella* and phytopathogens require multiple genes for the biosynthesis of amino acids in order to grow within plant tissues that have low availability of amino acids, compared to growth in animal hosts. Isogenic mutants lacking genes involved in amino acid biosynthesis (*thrC, metA, ilvD, trpC, glnG, glnA*) had severe defects in fitness against the parental strain, confirming the transposon mutant screening results.

Most genes in *Salmonella* serovar Newport that were selected for during persistence in tomatoes were shared with and similarly selected in *S. Typhimurium*. Many of these functions are involved in basal metabolism, including amino acid biosynthetic pathways, iron acquisition, and maintenance of cell structure. One exception was a greater need for core genes involved in purine metabolism in Typhimurium than in Newport. A new gene, *papA*, which was unique to
serovar Newport, also contributed to the fitness of a strain of S. Newport in tomatoes. The \textit{papA} gene present in 71 Newport Group III genomes was generally absent from other \textit{Salmonella} genomes, but a homolog was detected in members of the \textit{Enterobacteriaceae} family that can colonize both plants and animals.

\textbf{Outcomes and Accomplishments}

The avirulent \textit{Salmonella} surrogate developed for this project (\textit{Salmonella enterica} subsp. \textit{enterica} serovar Typhimurium strain MHM112) is publicly available from the American Tissue and Culture Collection (ATCC), accession #BAA-2828, and its genome is available from GenBank, accession #LONA00000000.

In addition, the findings of this project, including the validation of the avirulent surrogate and the \textit{Salmonella} pangenome analysis, have resulted in four peer-reviewed journal articles (see Appendices). One of the published articles (de Moraes et al., 2017, \textit{AEM} 83:e03028) was featured on the blog of the American Society for Microbiology in an article titled “How to be a successful foodborne pathogen: step 1, grow on food” (https://tinyurl.com/y93dgby2).

\textbf{Summary of Findings and Recommendations}

Previous funding from CPS supported the construction of a non-virulent surrogate strain of \textit{Salmonella} that is suitable for on-site experiments. The current project demonstrated that this avirulent surrogate behaves like the parental virulent strain during growth in produce, persistence in soil and water, and in its sensitivity to heat and disinfectants. Thus, we have validated the use of this fully avirulent, non-immunogenic, non-transgenic, and antibiotic-sensitive surrogate as a new tool in addressing FSMA regulations.

In addition, the pangenome analysis of over 2,000 serovars of \textit{Salmonella} revealed high diversity in the genomic composition across strains. Genomes of the \textit{Salmonella} serovar Newport, which has been associated with several produce-linked foodborne illness outbreaks, had a set of unique core genes that may be involved in adaptation to growth in plants. A massive parallel screening of mutants revealed that growth of \textit{Salmonella} serovars Typhimurium and Newport required the production of lipopolysaccharide and nucleotide biosynthesis, regardless of the host. However, growth in plants required genes involved in the biosynthesis of amino acids, due to the low availability of amino acids in plant tissues. In addition, the novel gene \textit{papA} (for \textit{plant} associated \textit{protein A}) was identified in \textit{Salmonella} Newport and contributed to fitness of the strain in tomatoes. While the function of \textit{papA} is currently unknown, homologues of this gene were found in \textit{Enterobacteriaceae} capable of growth in both plants and animals.

Collectively, the pangenome analysis supports the idea that only a handful of \textit{Salmonella} serovars can grow in produce because their genotype contains genes for adaptation to growth in plants. Furthermore, the discovery of this novel plant-associated \textit{Salmonella} gene could be used as a marker gene to detect \textit{Salmonella} that flourish in both plants and animals.
Literature Cited
APPENDICES

Publications and Presentations

Publications


Presentations

Marcos de Moraes, 2016 CPS Research Symposium, Lightning Round presentation: “Genome-wide functional and comparative analysis reveal adaptations of Salmonella sv Newport to proliferation in tomatoes.” June 2016, Seattle, WA.


Budget Summary

<table>
<thead>
<tr>
<th>Category</th>
<th>Meyer 2016CPS09 Years 1 &amp; 2 (Research)</th>
<th>Total spent</th>
<th>Remaining budget</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personnel – Salary</td>
<td>$48,720.00</td>
<td>$17,248.80</td>
<td>$31,471.20</td>
</tr>
<tr>
<td>Fringe Benefits</td>
<td>$14,177.00</td>
<td>$875.07</td>
<td>$13,301.93</td>
</tr>
<tr>
<td>Domestic Travel</td>
<td>$3,316.00</td>
<td>$3,316.00</td>
<td>$-</td>
</tr>
<tr>
<td>Materials &amp; Supplies</td>
<td>$29,823.00</td>
<td>$17,699.82</td>
<td>$12,123.18</td>
</tr>
<tr>
<td>Other Expenses</td>
<td>$8,600.00</td>
<td>$366.25</td>
<td>$8,233.75</td>
</tr>
<tr>
<td>Indirect Costs</td>
<td>$3,145.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>$107,781.00</td>
<td>$107,781.00</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Category</th>
<th>Meyer 2016CPS09 Year 3 (Publications &amp; Presentation)</th>
<th>Total spent</th>
<th>Remaining budget</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personnel – Salary</td>
<td>$-</td>
<td>$-</td>
<td>$-</td>
</tr>
<tr>
<td>Fringe Benefits</td>
<td>$-</td>
<td>$-</td>
<td>$-</td>
</tr>
<tr>
<td>Domestic Travel</td>
<td>$1,386.00</td>
<td></td>
<td>$1,386.00</td>
</tr>
<tr>
<td>Materials &amp; Supplies</td>
<td>$-</td>
<td>$-</td>
<td>$-</td>
</tr>
<tr>
<td>Other Expenses</td>
<td>$2,950.00</td>
<td></td>
<td>$2,950.00</td>
</tr>
<tr>
<td>Total</td>
<td>$4,336.00</td>
<td></td>
<td>$4,336.00</td>
</tr>
</tbody>
</table>