CENTER for PRODUCE SAFETY

CPS 2015 RFP
FINAL PROJECT REPORT

Project Title
Pathogen physiological state has a greater effect on outcomes of challenge and validation studies than strain diversity

Project Period
January 1, 2016 – December 31, 2017

Principal Investigator
Martin Wiedmann
Cornell University
Department of Food Science
341 Stocking Hall
Ithaca, NY 14853
T: 607-254-2838
E: mw16@cornell.edu

Objectives
1. *Assemble a collection of Salmonella, L. monocytogenes, Shiga toxin–producing E. coli, and relevant surrogate, indicator, and index organisms representing national and international strain diversity associated with produce.*
2. *Evaluate the strains assembled in Obj. 1 for (i) growth on different produce types and (ii) survival of key produce relevant controls and intervention treatments (e.g., different sanitizers).*
3. *Expose selected pathogen strains (at least 5 per target pathogen) to different environmental and stress conditions (e.g., different growth phases and water activities; pH stress) and evaluate them for subsequent (i) growth on different produce and fruit products and (ii) survival of key produce relevant controls and intervention treatments.*
4. *Develop and publicize standard protocols for pathogen growth under produce relevant conditions and assemble standard produce pathogen strain sets for distribution to industry as well government and academic researchers.*

Funding for this project provided by the Center for Produce Safety through:
CDFA SCBGP grant# SCB15066

FINAL REPORT

Abstract

Effective control of foodborne pathogens on produce and in produce-associated environments requires science-based validation of interventions and control strategies. It has previously been shown that strains and/or genetic lineages of a pathogen may differ in their ability to survive different stress conditions. We hypothesize, however, that pre-growth conditions have a significantly greater effect on subsequent stress survival and stress response than genetic diversity within produce-relevant pathogens (e.g., linked to fresh produce outbreaks, recalls, and pre-harvest environment).

A proposed strain collection was evaluated by experts from academia (n=6), government (n=5), and industry (n=8), with at least 10 years of experience in food safety, for inclusion in the study. The final strain collection is comprised of *Salmonella enterica* (n=23)*, Listeria monocytogenes* (n=11)*, Escherichia coli* (n=13), and surrogate, indicator, and index organisms (n=8). Strain diversity was ensured by including the 10 most common serotypes (*Salmonella*), representation of lineages (*L. monocytogenes*), and inclusion of the "Big Six" (*E. coli*) (Objective 1). All isolates in the final collection were characterized by whole genome sequencing using the Illumina Miseq platform. Based on core single nucleotide polymorphism, a subset of representative strains for each pathogen group was selected.

To evaluate the phenotypic response, the selected representative pathogen strains were pre-grown under 7 different conditions (e.g., high salt, low pH, low water activity), followed by (i) inoculation on fresh produce items (i.e., tomato, cantaloupe rind, pre-cut romaine lettuce), and (ii) exposure to sanitizer (Objective 2 and 3). Briefly, strains were grown under 7 different pre-growth conditions, followed by exposure to 40 ppm of peroxyacetic acid (*Salmonella, E. coli*) or 60 ppm of peroxyacetic acid (*L. monocytogenes*) for 45 sec (Objective 2 and 3). Statistical analyses with a crossed random effects log-link binomial proportion model showed that pre-growth conditions have a larger effect on variation in log reduction due to sanitizer exposure than strain or strain and condition interaction; contributions of pre-growth conditions to variation were estimated to be 99.9% for *E. coli*, 87.6 % for *Listeria*, and 55.5% for *Salmonella*.

To evaluate growth and survival of strains, strains were pre-grown under 7 different conditions and inoculated on tomatoes, cantaloupe rind, and pre-cut romaine lettuce. Inoculated produce items were incubated for up to 7 days, and inoculated strains were enumerated on days 0, 1, 3, 4 and 7. Pre-growth conditions also showed a larger effect on the variation of growth and die-off as compared to the variation attributed to different strains grown under a given condition. For example, day 7 growth of *Salmonella* strains inoculated on tomatoes showed less strain variation within one condition (e.g., for pre-growth to stationary phase at 37°C, the recovery of strains ranged from 7.0 to 7.7 log at day 7; for pre-growth in NaCl, recovery of strains ranged from 8.0 to 8.4 log) as compared to variation between pre-growth conditions. For example, for a single *Salmonella* strain pre-grown under different conditions, the recovery on day 7 ranged from 5.1 to 8.2 log (for pre-growth at 21°C and in NaCl, respectively).

These data, along with the strain collection, which is publicly available (see https://foodsafety.foodscience.cornell.edu/research-and-publications/cps-strain-collection), will facilitate selection of appropriate strains and growth conditions for challenge studies. This study specifically indicates that pre-growth under different conditions has a larger impact on the outcome of challenge studies than strain diversity, suggesting that challenge studies should be conducted with strains pre-grown under different conditions, particularly if the conditions the strains are naturally exposed to are unknown or highly variable, as often is the case in natural contamination events in the produce industry.

Background

Effective control of foodborne pathogens on produce and in produce-associated environments requires science-based validation of interventions and control strategies. Typically, multiple strains are used for most microbial challenge studies in which strains are grown under conditions that most likely provide greatest resistance to a given challenge. This approach accounts for strain diversity by including multiple distinct strains, serotypes, subtypes, and specific food-associated strains. However, it has been well documented for key pathogens of concern that pathogen strains differ in their ability to survive stress conditions. Also, pre-growth conditions have a considerable effect on pathogen growth and survival under subsequent stresses, and increase the range of physiological responses of pathogens in challenge and validation studies. This project addresses the need for science-based challenge and validation studies by providing a comprehensive data set that allows the produce industry to justify use of (i) specific pre-growth conditions and (ii) specific pathogens, surrogate, indicator, and index strains. The key goal of this project was to determine whether strain diversity or different pre-growth conditions yield a larger variation in response to different intervention and control strategies, and which strains and conditions may yield the largest differences in response.

For this project, a produce-relevant strain collection was developed, including *Salmonella*, *Listeria monocytogenes*, and Shiga toxin–producing *E. coli* (STEC), and was evaluated genotypically as well as phenotypically. This strain collection and the associated phenotypic characterization data will be particularly valuable considering the changing food safety regulatory environment (including the Food Safety Modernization Act), which places an increasing emphasis on science-based approaches and scientific validations.

Research Methods

Survey for strain collection

An electronic survey was developed to identify the most relevant strains to include in the final strain collection. Isolates in the proposed strain collection were selected to represent isolates linked to fresh produce outbreaks, recalls, and pre-harvest environments. The survey was sent to 30 experts of which 19 responded, including experts in academia (n=6), government (n=5), and industry (n=8), all with at least 10 years of experience in food safety. Reviewers were asked to rate each strain on a scale of 1 to 5 (e.g., 1 – strain irrelevant, do not include; 2 – strain may be relevant; 3 – uncertain, strain may or may not be relevant; 4 – important, should be included; 5 – very important, must be included). Information provided on the proposed strain collection included ID numbers, isolate origin (e.g., associated outbreak), and reference. Reviewers also had the option to suggest inclusion of additional strains not included in the initial strain list. A strain was included when at least 50% of respondents considered the strain as important.

Bacterial strains

Isolates in the strain collection were organized into four groups: (i) *Salmonella enterica* (n=23), (ii) *Listeria monocytogenes* (n=11), (iii) *Escherichia coli* (n=13), and (iv) surrogate, indicator, and index organisms (n=8). Isolates to be included were selected based on the expert review detailed above, while also ensuring inclusion of the 10 most common serotypes (*Salmonella enterica*) (FoodNet 2015 Surveillance Report), representation of lineages (*L. monocytogenes*), and the “Big Six” (*E. coli*). Isolates were either part of a previously described set (e.g., ILSI NA Listeria strain collection [1], ATCC [2]), previously described by our group (e.g., [3, 4]) or obtained from various outside sources, e.g., Food and Drug Administration (FDA), Center for Disease Control and Prevention (CDC), United States Department of Agriculture (USDA), Wisconsin State Laboratory of Hygiene, Trade, and Consumer Protection, Texas Tech

University, STEC Center Michigan State University, University of California–Davis, and University of Florida. All bacterial strains were stored in brain heart infusion (BHI; Difco, Becton Dickinson, Sparks, MD) with 15% glycerol at −80°C.

Library preparation and whole genome sequencing

Isolates were streaked out from glycerol stocks onto BHI agar plates and incubated at 37°C for 24 h. An overnight culture was prepared by inoculating 5mL BHI broth with a single colony, followed by incubation at 37°C for 14 h. Then 2mL of overnight culture was pelleted and used for DNA extraction (DNeasy Blood and Tissue kit, Qiagen, Valencia, CA) following manufacturer instructions. Gram-positive bacteria were pre-treated in 200µL of lysis solution (4mg lysozyme, 4µL 1M Tris-HCl, 6.25µL 250mM EDTA, 2.4µL Triton X-100, add dH_2O to 200µL final volume). DNA was eluted in 50uL Tris-HCl at pH 7.5. DNA purity was assessed spectrophotometrically using NanoDrop 2000 (Thermo Scientific, Wilmington, DE). DNA was quantified with fluorescent nucleic acid dye (Qubit dsDNA HS Assay Kit, Thermo Fisher Scientific, Waltham, USA) and a Qubit 2.0 fluorometer (Thermo Fisher Scientific). Libraries were prepared for sequencing with the Nextera XT DNA sample preparation kit and associated Nextera XT Index kit with 96 indices (Illumina, Inc. San Diego, CA). Library preparation was conducted according to the PulseNet standard operation procedure for Illumina Miseq data quality control protocol (*https://www.cdc.gov/pulsenet/pdf/PNL32-MiSeq-Nextera-XT.pdf*). Pooled samples were sequenced on an Illumina MiSeq platform with 2×250bp paired-end reads.

Genome assembly and phylogenetic analyses

Adapters were removed from sequences using Trimmomatic v 0.33 [5] followed by quality assessment using FastQC v 0.11.4 (*http://www.bioinformatics.babraham.ac.uk/projects/fastqc*). Sequences were assembled *de novo* with SPAdes v 3.8.0 [6]. Quality control of assembly was performed with QUAST v 3.2 [7] and average coverage determined using SAMtools v 1.4.1 [8]. Contigs smaller than 200bp were removed and contigs were blasted using Kraken [9] to confirm strain identity. Additionally, serotypes were confirmed with SeqSero (*http://www.denglab.info/SeqSero*) for *Salmonella* [10] and SeroTypeFinder for *E. coli* [11]. A standard set of 21 sequenced genomes of *Listeria* isolates, as previously described by Liao et al. [12], was used in a single nucleotide polymorphism (SNP)–based phylogenetic analysis to confirm lineages for *Listeria* isolates as part of the new strain collection. *Salmonella* isolates with acquired antibiotic resistance (e.g., rifampicin, nalidixic acid resistance) were compared to their wildtype parent to identify high quality SNPs using the CFSAN SNP Pipeline [13]. Identified SNPs were analyzed for synonymous and non-synonymous mutations. BLASTX was used to identify genes if SNPs were located in potential open reading frames (ORFs) [14]. The core-SNP based analysis was performed using kSNP v 3 with estimated optimal kmer size 13 using Kchooser for each pathogen group including: *Salmonella*, *E. coli*, and *Listeria* [15]. A maximum-likelihood phylogeny based on core genome SNP was generated with 1,000 bootstrap repetitions in RaxML for each pathogen group. Surrogate, indicator, and index sequences were integrated into phylogenetic analysis with their associated pathogen group, except for *E. faecium* (FSL R9-5275) [16]. Phylogenetic trees were edited using FigTree v 1.3.4 (*http://tree.bio.ed.ac.uk/software/figtree/*).

Pre-growth conditions

Bacterial strains were streaked out from glycerol stocks onto tryptic soy agar (TSA; Difco, Becton Dickinson) plates and incubated at 37°C for 24 h. Colonies for overnight cultures were picked from a plate held at 4°C for a minimum of 24 h and a maximum of 7 days and inoculated into 5mL of tryptic soy broth, followed by incubation at 37°C for 12-14 h. These cultures were used to inoculate, at a 1:1000 dilution, pre-warmed side-arm flasks (Nephelo Flasks/C38

300mL; Belco, Vineland, NJ) with either 30 or 100mL of growth medium. Bacterial cultures were grown to early stationary phase at 37°C except when the pre-growth condition was defined as mid-log phase or 21°C. Growth curves were generated for each strain in each condition to determine optical density (OD) for each growth phase. Pre-growth conditions used in this study included (i) low pH (pH 5.0 for *Salmonella* and *E. coli*; pH 5.5 for *Listeria*), (ii) high salt (4% NaCl), (iii) low water activity (0.96 for *Salmonella* and *E. coli*; 0.95 for *Listeria*), (iv) two different growth phases (mid-log phase, stationary phase), (v) minimal medium (M9 for *Salmonella* and *E. coli*, defined medium for *Listeria*), and (vi) two different temperatures (37°C, 21°C). Water activity was adjusted using glycerol at 15.6% (vol/vol) and 13% (vol/vol) to achieve a_w of 0.95 and 0.96, respectively. The pH was adjusted using lactic acid and the high salt environment was generated with additional 4% NaCl (wt/vol). Minimal medium M9 was prepared as previously described with 0.4% glucose (wt/vol) (*Salmonella*) or 0.3% fructose (wt/vol) (*E. coli*) [17]. The chemically defined medium for *Listeria* was prepared with 10mM glucose [18].

Sanitizer treatment

Bacterial cultures grown under different pre-growth conditions were exposed to peroxyacetic acid (Tsunami, Ecolab, St. Paul, MN). Peroxyacetic acid (PAA) concentration was measured using Reflectoquant (RQflex 10, Millipore, Darmstadt, Germany). For each treatment, 1mL of bacterial culture was added to 9mL PAA (in a 15-mL Falcon tube) for a final concentration of either 40ppm (*E. coli, Salmonella*) or 60ppm (*Listeria*), followed by mixing through four inversions and incubation for 45 sec. The sanitizer solution was inactivated by adding 100µL 50% $Na_2S_2O_3$ (wt/vol) immediately after the 45-sec incubation, followed by four inversions of the tube to ensure complete mixing. Control cultures were treated with phosphate-buffered saline solution (PBS) instead of PAA. Immediately afterwards, 50µL of the appropriate dilutions were plated on TSA using a spiral plater (Autoplate 4000; Advanced Instruments, Norwood, MA). Plates were incubated at 37°C for 24 h (*Salmonella, E. coli*) or 48 h (*Listeria*). Colonies were enumerated using Color Q-Count (Model 530, Advanced Instruments Inc., Norwood, MA).

Produce preparation and inoculation

Produce items were purchased from a local grocery store and stored at 4°C in their original packaging for up to 2 days, and pre-warmed at room temperature for 1 h prior to inoculation. In brief, for each combination of condition, strain and sampling day, the following inoculated products were prepared: (i) 3 grape tomatoes inoculated with *Salmonella*, (ii) 3g of pre-cut romaine lettuce inoculated with *E. coli*, and (iii) one cantaloupe rind prepared with a standardized cork-borer inoculated with *Listeria*.

Produce inoculation

Each tomato, lettuce, and cantaloupe rind section was inoculated in ten different places with 10µL of the appropriate bacterial culture. The inoculated produce was left to dry at room temperature for 1 h. The tomatoes were then transferred into Whirl-Pak filter membrane bags (Nasco, Fort Atkinson, WI) with 3 tomatoes per bag. The lettuce and cantaloupe rinds were inoculated in petri dishes, and sealed with parafilm (Parafilm M; Bemis Company, Neenah, WI) prior to incubation. Lettuce and cantaloupe were incubated at 7°C, while tomatoes were incubated at 21°C.

Bacterial enumeration of inoculated produce items

A 30mL aliquot of PBS was added to bags containing tomatoes, followed by gentle hand-rubbing for 1 min. This solution was diluted with PBS and spiral plated on xylose-lysine-deoxycholate (XLD) agar, followed by incubation at 30°C for 24 h. Lettuce was transferred to filter bags with 47mL PBS, followed by agitation in a stomacher (Stomacher 400 Circulator; Seward, Bohemia, NY) for 30 sec at 260 rpm, performed twice. Dilutions were plated on

MacConkey agar, followed by incubation at 37°C for 24 h. Each cantaloupe rind was transferred into a 50-mL conical tube with 10mL PBS, followed by vortexing for 1 min. Dilutions were plated on modified Oxford medium (MOX) followed by incubation at 35°C for 48 h. To assess aerobic plate counts and potential background contamination, negative control samples were plated on TSA, as well as Listeria monocytogenes Plating Medium (LMPM) and MOX for *Listeria*, on MacConkey agar for *E. coli*, and XLD for *Salmonella*, followed by incubation for 24 h at 37°C (TSA, MacConkey, XLD), and 48 h at 35°C (LMPM, MOX). All samples were plated in duplicate. Colony counts for each plate were determined by an automated colony counter (Q-Count; Advanced Instruments). All trials were conducted in triplicate.

Statistical analysis

Statistical analyses were performed in R (version 3.1, R Core Team, Vienna, Austria). A cross random-effects log-link binomial proportion model (with OLRE to correct for over dispersion) was applied. The outcome for the survival data of the sanitizer treatment was reported as log-reduction. In order to identify if strain diversity or pre-growth conditions contribute more to a larger range of responses, a random effects model was fitted using the lme4 package [19]. A fixed-effect linear model was fitted to the data set, and post-hoc comparisons of estimated means were performed with the lsmeans package to determine whether strain and condition means were significantly different [20].

Research Results

Objective 1

The final strain collection assembled includes 23 *Salmonella enterica,* 11 *Listeria monocytogenes,* and 13 *Escherichia coli* as well as 8 surrogate, indicator, and index organisms; details on these strains can be found at https://foodsafety.foodscience.cornell.edu/research-and-publications/cps-strain-collection. Details on the strain collection for the different organism groups are provided below.

Salmonella enterica

The 23 *Salmonella* strains in the strain collection include isolates from a number of produce-relevant outbreaks associated with *Salmonella*, including outbreaks linked to cantaloupe (2000-2002 and 2008), tomatoes (2002 and 2005), orange juice (2005), jalapeno peppers (2008), cucumbers (2015), alfalfa sprouts (2016), and bean sprouts (2014) as well as isolates associated with outbreaks linked to almonds (2000-2001) and peanut butter (2006-2007 and 2008-2009). Other isolates were included to ensure serotype diversity and inclusion of strains that are already in use for validation studies. Surrogate organisms for *Salmonella* included in the strain collection are *E. faecium* (FSL R9-5275) and avirulent *Salmonella* Typhimurium (FSL R9-6232, FSL R9-6231). Isolates clustered based on core SNP in a maximum-likelihood tree with their serotypes except for I 4,[5],12:i:-, which as expected clustered closely within Typhimurium, and one Newport isolate, which clustered closely with serotype Litchfield (**Figure 1**), consistent with previous data that Newport is polyphyletic and represents multiple lineages [21–23]. For two *Salmonella* isolates in the initial strain collection, the serotype predicted by whole genome sequencing did not match the reported serotype that was expected based on the outbreak a given isolate was associated with. These two strains were newly acquired from another source; whole genome sequencing of these new isolates confirmed that these strains represented the correct serotype. Among all *Salmonella* isolates (excluding antibiotic resistant strains derived from a given strain), the number of core SNP differences ranged from 51 to 13,436 SNPs. The strain collection also includes two *Salmonella* isolates resistant to rifampicin (FSL R9-6567, FSL R9-5251) and one strain resistant to nalidixic acid (FSL R9-5220). High quality SNP analysis

confirmed single point mutations in the RNA polymerase subunit B gene (*rpoB*) (FSL R9-6567) and DNA gyrase subunit A (*gyrA*) genes (FSL R9-5220). Isolate FSL R9-6567 had an additional non-synonymous mutation in a flagellar gene (*flhE*), which encodes a regulator in the flagellar system [24]. Relative to its parent strain, isolate FSL R9-5251 includes non-synonymous SNPs in genes encoding the fatty acid oxidation complex subunit alpha (*fadJ*) [25], the transcriptional regulator for fatty acid degradation (*fadR*), the glutamate-a-semialdehyde aminotransferase (*hemL),* a transcriptional repressor of sugar metabolism (*glpR*), the regulatory protein CsrD (*yhdA*) that controls degradation of CsrB and CsrC RNA [26], as well as two mutations in RNA polymerase subunit beta' (*rpoC*) (**Table 1**).

Listeria monocytogenes
L. monocytogenes isolates in the strain collection include isolates linked to listeriosis outbreaks associated with cantaloupe (2011), packaged salad (2016), and sprouts (2014). Other isolates were included for lineage and serotype representation, including isolates representing lineage I (n=6), lineage II (n=3), lineage III (n=1), and lineage IV (n=I). A *L. innocua* isolate (FSL C2-0008) is included in the collection as a surrogate organism. Isolates clustered based on core SNP in a maximum-likelihood tree as expected with their lineages (**Figure 2**). The maximum number of core SNP differences within lineage I isolates was 359 SNPs, as compared to 1,055 SNPs within lineage II. Lineage I isolates FSL J1-0107 and FSL J1-0108 showed no SNP differences in the core SNP differences matrix, consistent with the fact that both isolates were obtained from the same outbreak. In contrast, when running high quality SNP analysis, 14 SNP differences were identified between these two isolates.

Escherichia coli
Escherichia coli isolates in the strain collection include isolates associated with outbreaks linked to baby spinach (2006), shredded romaine lettuce (2010), sprouts (2009, 2012), clover sprouts (2012, 2014), cabbage salad (2014), and alfalfa sprouts (2016). Other isolates were included to ensure representation of the "Big-Six" non-O157 STEC serogroups. Surrogate organisms for pathogenic *E. coli* that are part of the strain collection include environmental "generic" *E. coli* isolates from water, plant, and soil (FSL R9-4077, FSL R9-4078, FSL R9-4079) as well as a naturally occurring non-pathogenic attenuated O157:H7 isolate (FSL R9-3467). In a core SNP phylogeny (**Figure 3**), all *E. coli* isolates clustered as expected according to their serogroup. The minimum number of core SNP differences among all *E. coli* isolates in the strain collection is 22 SNPs, whereas the maximum number of core SNP difference is 1,174 SNPs.

Objective 2 and Objective 3
Objectives 2 and 3 included establishing standardized growth conditions and protocols for bacterial strains; these protocols were used to pre-grow the selected strains prior to (i) sanitizer treatment or (ii) produce inoculation. Results for these two objectives are reported together to allow for better comparisons between the effects of stain diversity (Obj. 2) and pre-growth conditions (Obj. 3) on sanitizer-treatment survival and growth on relevant produce items.

<u>Effect of strain diversity and pre-growth conditions on sanitizer-treatment survival</u>
Survival after peroxyacetic acid (PAA) exposure showed limited variation among different strains when comparing strains grown under the same conditions. For example, *Salmonella* strains in mid-log phase showed log reductions (after 45-sec exposure to PAA) ranging from 5.6 to 5.7 (representing almost a million-fold reduction). In contrast, strains showed considerable differences in log reduction when grown under different conditions. For example, reductions of the 4 *Salmonella* strains after pre-growth to mid-log phase ranged from 7.2 to 8.0 log (representing a 10- to 100-million-fold reduction), while the reductions after pre-growth under salt stress ranged from 3.0 to 4.5 log (representing only about a 1,000- to 30,000-fold reduction

after PAA treatment) (**Figure 4**). Similarly, *E. coli* strains grown to mid-log phase showed a consistent 8.1-log reduction after PAA exposure (with virtually no strain differences), while *E. coli* strains pre-grown under salt stress showed only a 3.5-log reduction after exposure to PAA (**Figure 5**). For *Listeria,* strains grown in defined medium (DM) showed little variation in log reduction after PAA exposure, with a range of 5.4- to 6.2-log reduction (representing approx. a million-fold reduction), while *L. monocytogenes* pre-grown under salt stress showed less than 2-log reduction (100-fold reduction) after PAA exposure (with a range of 0.5- to 1.4-log reduction with different strains). These data show that pre-growth conditions have a considerable effect on sanitizer-treatment survival (**Figure 6**). Formal statistical analyses using a crossed random effects log-link binomial proportion model specifically showed that pre-growth conditions have a larger effect on variation in log reduction due to PAA than strain or strain and condition interaction; contributions of pre-growth conditions to variation were estimated to be 99.9% for *E. coli*, 87.6 % for *Listeria*, and 55.5% for *Salmonella*.

Effect of strain diversity and pre-growth conditions on growth on selected produce commodities
Overall, pre-growth conditions of bacterial inocula also had a considerable effect on subsequent pathogen growth on the selected produce items. For example, *Salmonella* pre-grown under salt-stress did not show a decline in bacterial numbers after inoculation onto tomatoes and maintained the inoculated populations over 7 days storage, while *Salmonella* pre-grown at 21°C showed an initial drop in numbers of up to 5 log (100,000-fold) at day 0 after inoculation of tomatoes, with slow regrowth over 7 days (**Figure 7**). Over the first 24 h of incubation, condition also showed a significantly larger contribution to variation in *Salmonella* numbers than strain diversity. *Salmonella* numbers after pre-growth conditions of 21°C, glycerol, pH5, and stationary phase did not differ significantly, but differed from numbers for *Salmonella* after pre-growth conditions of M9, NaCl, and mid-log phase. No significant difference was detected for *Salmonella* strains used for inoculation (FSL R9-5272, FSL R9-5244, FSL R9-5400, FSL R9-5502), including one avirulent *Salmonella* surrogate (FSL R9-6232) (**Figure 8**). For *Listeria*, pre-growth under different conditions before inoculation onto cantaloupe rind affected the length of lag phase. After 7 days, *Listeria* consistently reached a final concentration of 7 log CFU/plug, regardless of pre-growth conditions (**Figure 9**). However, for the first 24 h of incubation, *Listeria* strains showed a significantly larger contribution to variation of *Listeria* numbers than conditions. The surrogate organism *L. innocua* (FSL C2-0008) showed similar numbers, over the first 24 h, as compared with FSL R9-0506 and FSL R9-5506 but differed from FSL R9-5411 (**Figure 10**). For *E. coli*, which was tested for survival on pre-cut romaine lettuce incubated at 7°C, bacterial numbers decreased over 7 days when bacteria were pre-grown at 37°C but remained constant with no die-off when bacteria were pre-grown under 21°C (**Figure 11**). For the first 24 h of incubation, conditions showed significantly larger contribution to variation in bacterial numbers than strain diversity. *E. coli* pre-grown in glycerol, M9, and pH5 showed no significant difference in bacterial numbers. However, *E. coli* numbers differed significantly when pre-grown at 21°C compared to numbers when pre-grown to mid-log phase. No significant difference in numbers was detected, after 24 h of incubation, for *E. coli* strains used for inoculation (FSL R9-5271, FSL R9-5516, FSL R9-5517, FSL R9-5515, FSL R9-5258), including surrogate isolate FSL R9-4077 (**Figure 12**).

Objective 4
A www-based resource (*https://foodsafety.foodscience.cornell.edu/research-and-publications/cps-strain-collection)* was created, which will provide access to the strain collection, including all relevant information (e.g., whole genome sequencing data). Highly specific and detailed strain records and data are available through a www-based database (FoodMicrobeTracker), which can be searched using the standard strain ID designations (e.g., FSL R9-5251) and ensures that users will continue to have access to up-to-date strain data.

Outcomes and Accomplishments

Overall this project provides critical data on how environmental conditions affect the behavior of foodborne pathogens in produce as well as how environmental conditions affect the response of the different organisms to control strategies such as sanitizers. Importantly, this study (i) provides industry with a standard strain collection for future challenge and validation studies, and (ii) indicates that pre-growth conditions have a larger effect on pathogen growth and survival as compared to strain differences, including specific data on pre-growth conditions that may yield the most stress resistant cultures. Specifically, a key outcome of this study is the development of a produce-relevant standard strain collection with a total of 55 strains, including (i) 23 *Salmonella enterica*; (ii) 11 *Listeria monocytogenes*, (iii) 13 Shiga toxin–producing *E. coli* (STEC), and (iv) 8 surrogate, indicator, and index organisms. All strains have been characterized by whole genome sequencing, and all sequencing data have been submitted to NCBI and will be made publicly available. Information about the CPS strain collection is accessible through https://foodsafety.foodscience.cornell.edu/research-and-publications/cps-strain-collection, with updated strain information available through FoodMicrobeTracker. Phenotypic characterization of selected strains by assessing growth and survival on (i) inoculated fresh produce items (i.e., tomato, cantaloupe rind, pre-cut romaine lettuce) and (ii) exposure to sanitizer, demonstrated that pre-growth conditions have a greater effect on pathogen survival and growth in challenge and validation studies than does strain diversity. Consequently, another key outcome of this study is new information, which scientifically supports that, at least in some situations, strains grown under different conditions should be included in cocktails used for validation and challenge studies. Finally, this study also provides reference data for the behavior of different surrogate strains relative to actual pathogens.

Summary of Findings and Recommendations

- A produce-relevant strain collection with a total of 55 strains, including (i) 23 *Salmonella enterica*; (ii) 11 *Listeria monocytogenes*, (iii) 13 Shiga toxin–producing *E. coli* (STEC), and (iv) 8 surrogate, indicator, and index organisms provides a valuable resource for industry.
- Pre-growth conditions have a larger effect on variation in challenge studies than strain diversity or strain and condition interaction, suggesting that, at least in some situations, strains grown under different conditions should be included in cocktails used for validation and challenge studies.
- Specific growth conditions yield more “stress resistant” strains; hence data on the effect for growth conditions on subsequent survival and growth will allow industry, academic and government researchers to better select strains and pre-growth conditions for (i) validation of different interventions and (ii) assessment of growth and survival parameters for specific pathogens on different produce types.
- Most surrogate strains have a similar “stress resistance” than the corresponding pathogens. However, *Enterococcus faecium*, a surrogate organism for *Salmonella,* consistently shows a higher “stress resistance” when exposed to peroxyacetic acid.

APPENDICES

Publications and Presentations

Wiedmann, M. *Pathogen physiological state has a greater effect on outcomes of challenge and validation studies than strain diversity*.

- Poster presented at CPS Research Symposium, 28–29 June 2016, Seattle, WA
- Oral presentation at CPS Research Symposium, 21June 2017, Denver, CO

In addition, initial results from the project were presented in 2017 in a poster entitled *Characterization of a Pathogen Strain Collection to Allow for Improved Validation of Sanitizer Efficacy in the Produce Industry* at the annual meeting of the International Association for Food Protection and the annual conference of the New York State Association for Food Protection.

Budget Summary

Total funds allocated = $274,030; funds utilized as of 1/30/18:

Salaries and Wages	148,352.31
Employee Benefits	49,068.16
Travel	4,003.72
Materials and Supplies	55,692.44
Services	7,278.30
Other Direct Expenses	485.90
Indirect Costs	8,809.96
Total:	**$273,690.79**

Suggestion to CPS

We would suggest that the strain collection developed through this project is highlighted in a future newsletter and on the CPS www-page, to ensure industry use of this resource.

Table 1 and Figures 1–12

Table 1: Number of SNPs and positions for antibiotic resistant strains of *Salmonella* isolates

FSL ID	No. of SNP	Position	Strand	Gene	Amino acid change
FSL R9-6567	2	NODE 12 (pos. 68512)	+	*flhE*	Leu → Ser
		NODE 54 (pos. 6996)	-	*rpoB*	Arg → His
FSL R9-5220	1	NODE 6 (pos. 364027)	+	*gyrA*	Ser → Stop
FSL R9-5251	9	NODE 4 (pos. 145900)	-	*fadJ*	Leu → Pro
		NODE 6 (pos. 288419)	+	*fadR*	Val → Ile
		NODE 6 (pos. 319958)	-	hypothetical protein	Ile → Val
		NODE 9 (pos. 51000)	+	*hemL*	Pro → Thr
		NODE 11 (pos. 52140)	-	*glpR*	Asn → Lys
		NODE 13 (pos. 109858)	+	hypothetical protein	Ile → Thr
		NODE 19 (pos. 25052)	-	*yhdA*	Stop → Tyr
		NODE 20 (pos. 19872)	-	*rpoC*	Arg → Cys
		NODE 20 (pos. 21554)	-	*rpoC*	Val → Ala

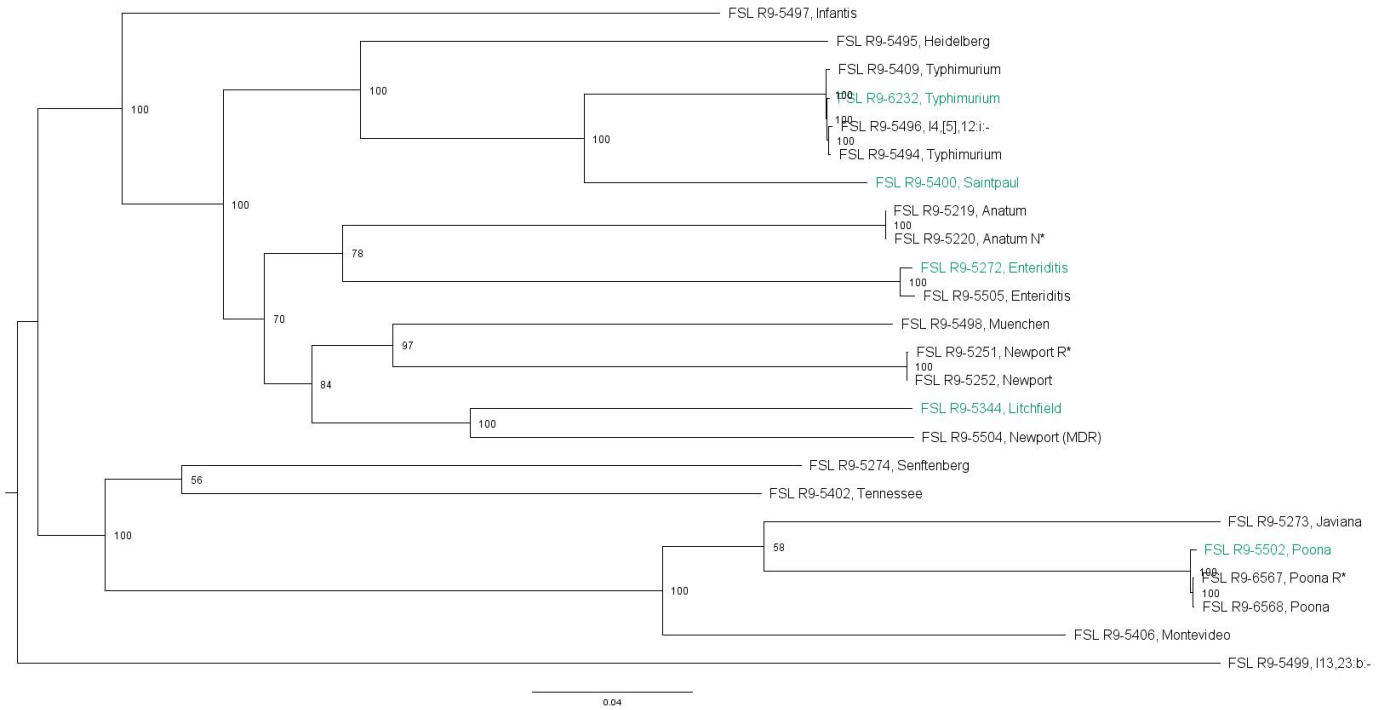

Figure 1: Maximum-likelihood tree based on core SNP analysis for *Salmonella* isolates. Strains marked in green were selected as representative strains for phylogenetic analysis.

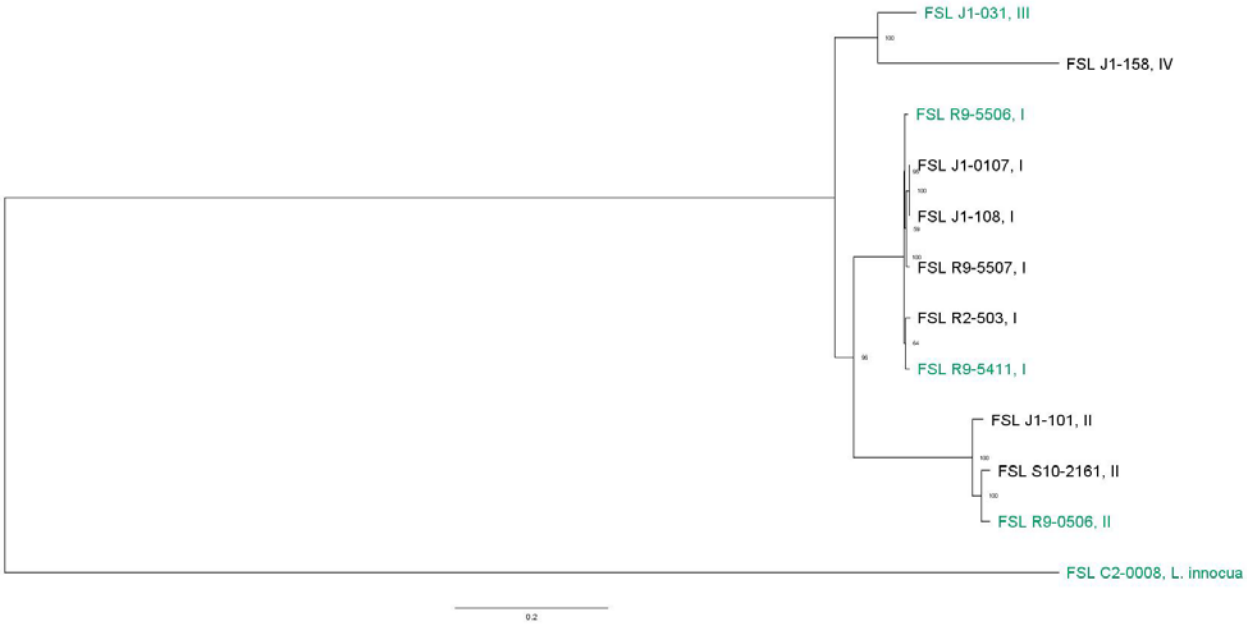

Figure 2: Maximum-likelihood tree based on core SNP analysis for *Listeria* isolates. Strains marked in green were selected as representative strains for phylogenetic analysis.

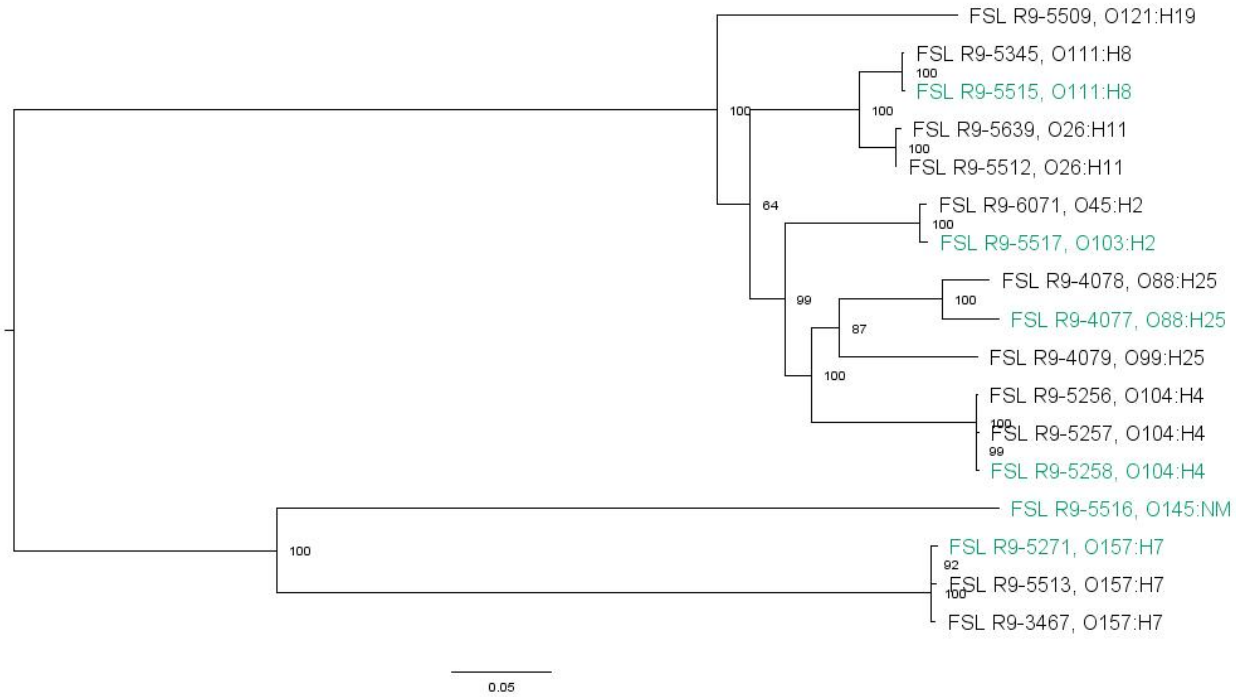

Figure 3: Maximum-likelihood tree based on core SNP analysis for *E. coli* isolates. Strains marked in green were selected as representative strains for phylogenetic analysis.

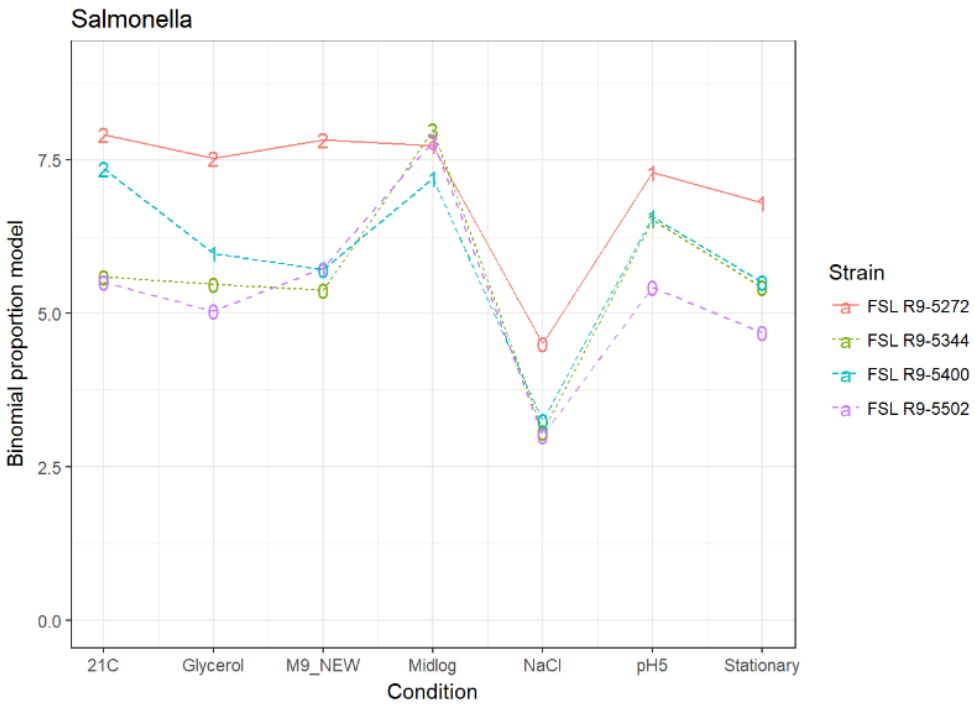

Figure 4: Crossed random effects log-link binomial proportion model showing the log reduction of *Salmonella* strains (x-axis) when pre-grown under different conditions (y-axis) prior to sanitizer exposure. Numbers in the graphs indicate the number of replicates (among a total of three) that yielded no survivors after PAA treatment (i.e., <200 CFU/mL).

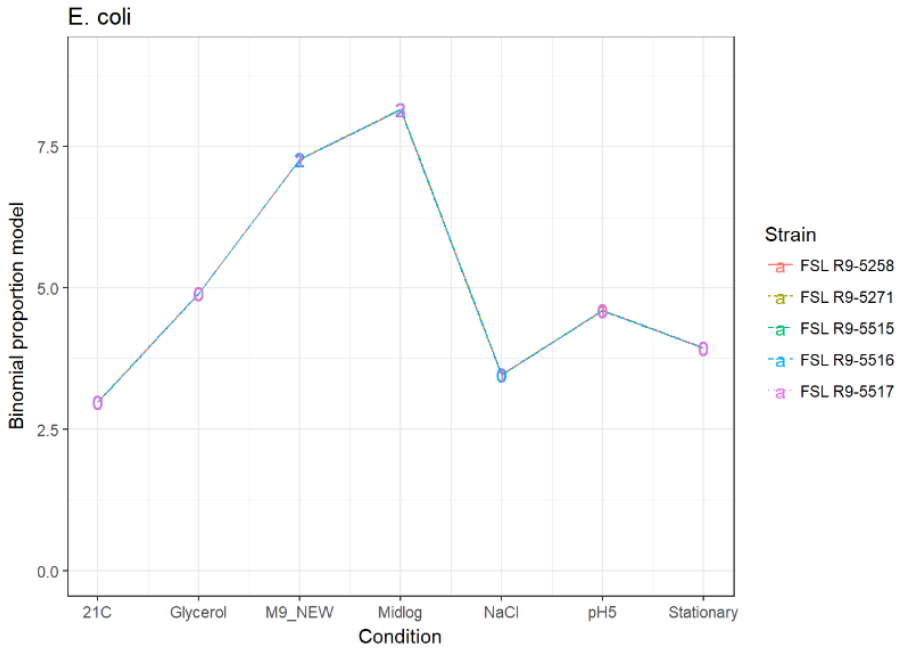

Figure 5: Crossed random effects log-link binomial proportion model showing the log reduction of *E. coli* strains (x-axis) when pre-grown under different conditions (y-axis) prior to sanitizer exposure. Numbers in the graphs indicate the number of replicates (among a total of three) that yielded no survivors after PAA treatment (i.e., <200 CFU/mL).

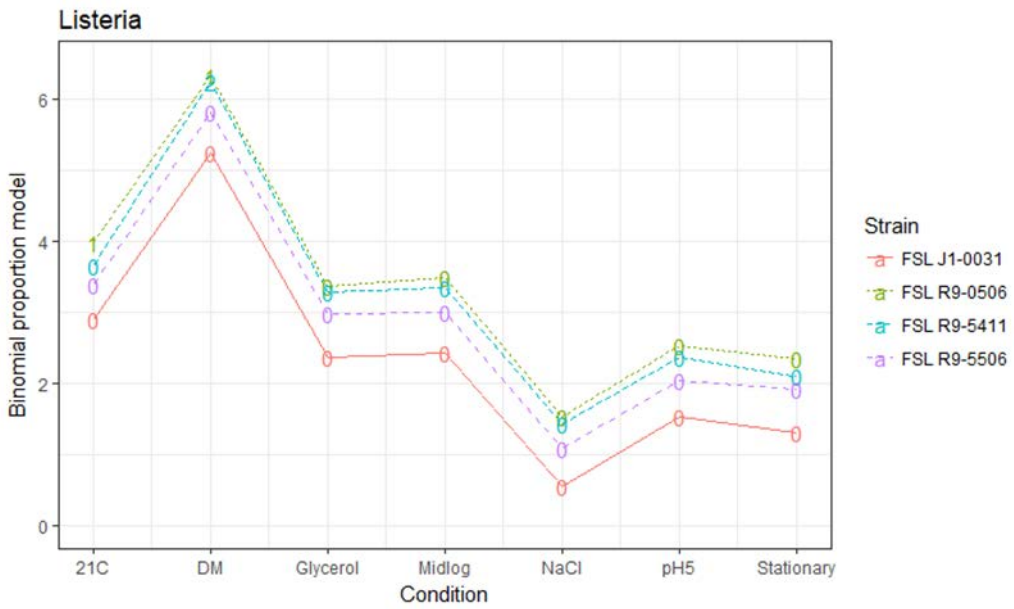

Figure 6: Crossed random effects log-link binomial proportion model showing the log reduction of *Listeria* strains (x-axis) when pre-grown under different conditions (y-axis) prior to sanitizer exposure. Numbers in the graphs indicate the number of replicates (among a total of three) that yielded no survivors after PAA treatment (i.e., <200 CFU/mL).

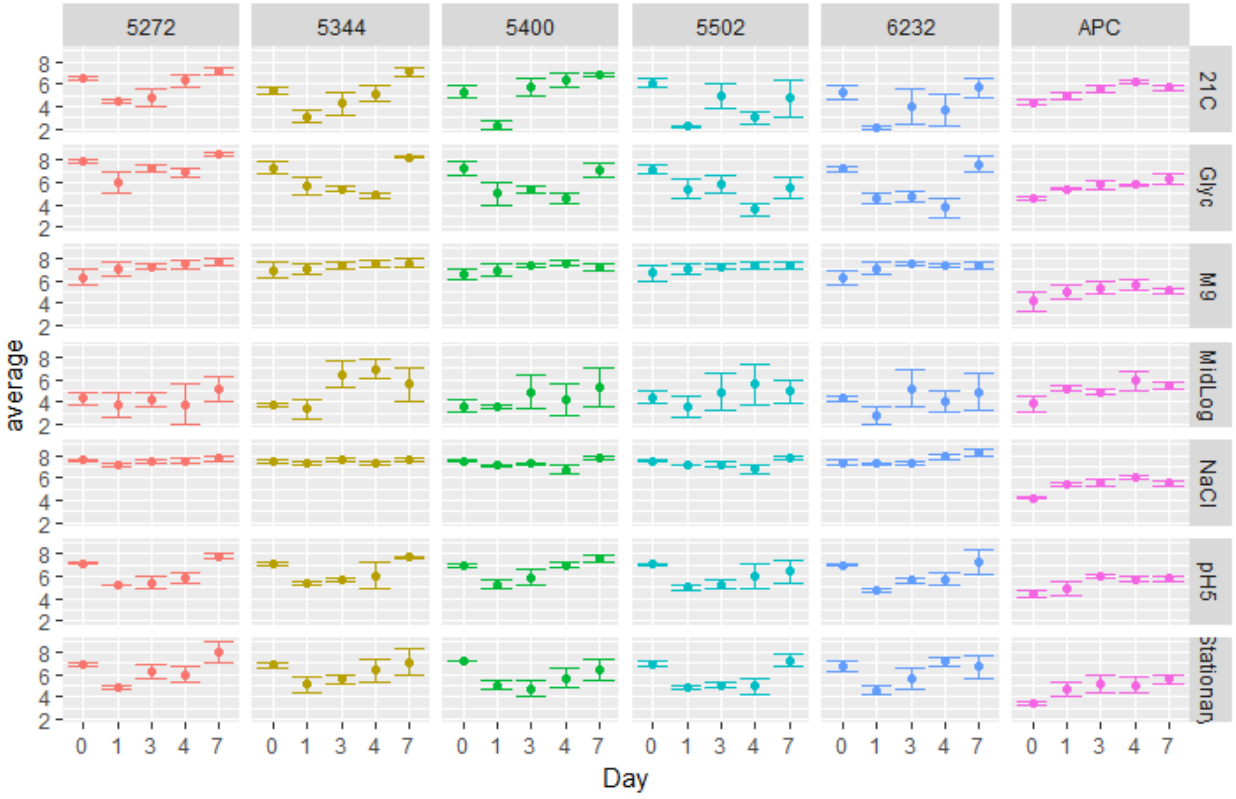

Figure 7: Growth curves for *Salmonella* strains inoculated on grape tomatoes. Y-axis represents average of three trials in log CFU/tomato; x-axis shows days of bacterial enumeration. Facets are divided into pre-growth conditions and strains. APC stands for aerobic plate count; strains are abbreviated, e.g., 5272 stands for FSL R9-5272.

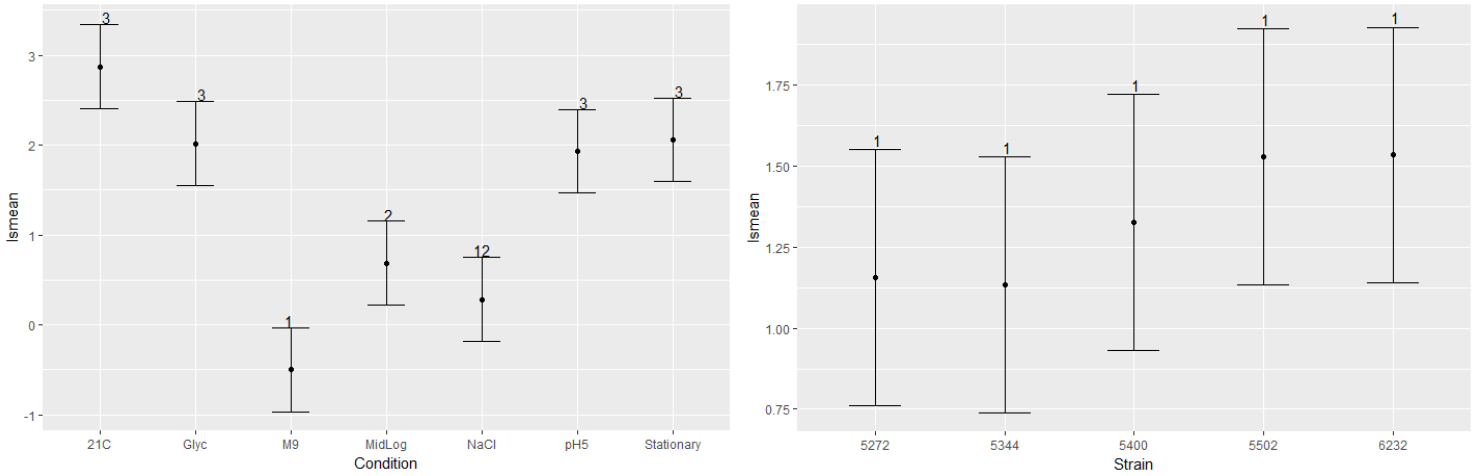

Figure 8: Fixed-effect linear model and post-hoc comparison of estimated means for *Salmonella* strains, including surrogate, indicator, and index organisms for the first 24 h of incubation. Y-axis represents estimated means; x-axis on the left-hand graph shows condition and on the right-hand graph shows strains. Numbers indicate groups based on Tukey Honest Significant Difference method.

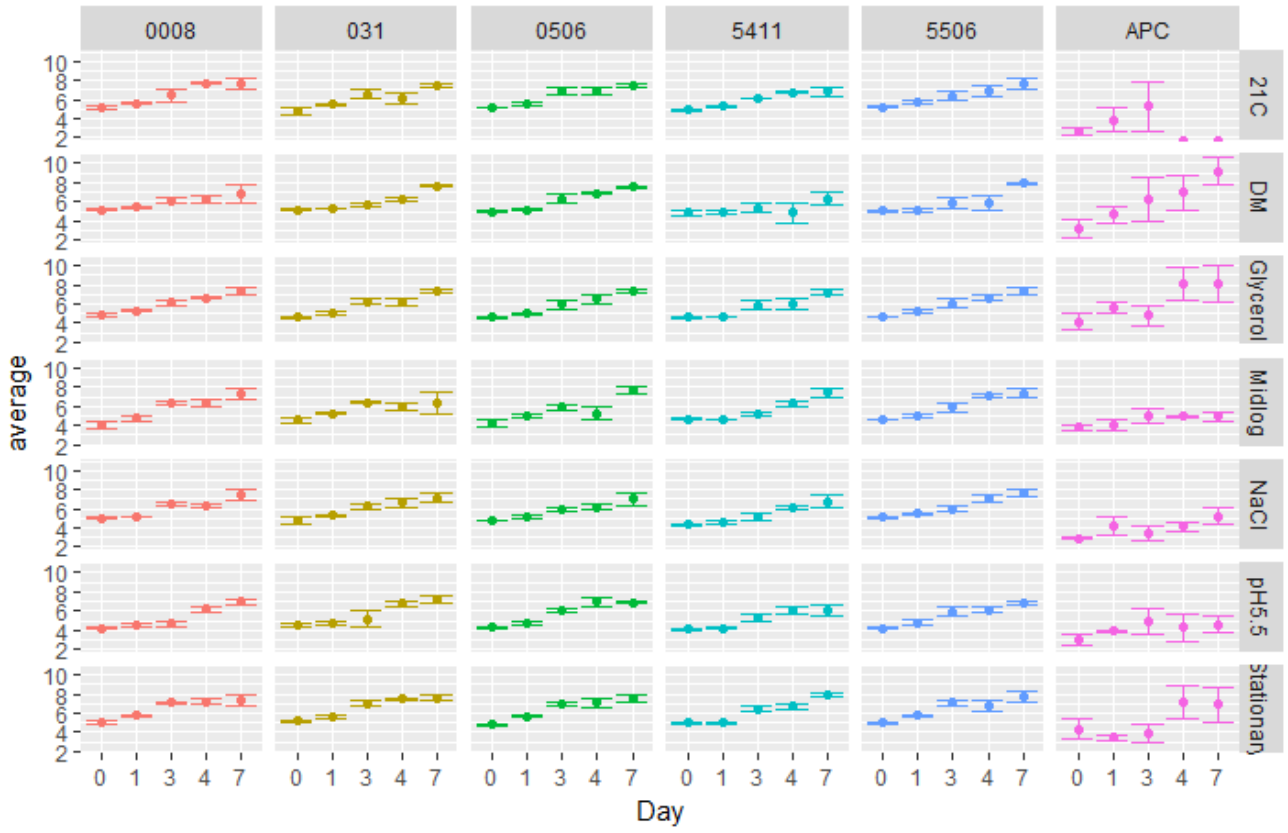

Figure 9: Growth curves for *Listeria* strains inoculated on cantaloupe rind. Y-axis represents average of three trials in log CFU/rind; x-axis shows days of bacterial enumeration. Facets are divided into pre-growth conditions and strains. APC stands for aerobic plate count. Strains are abbreviated: 5411, 5506, and 0506 stand for FSL R9-5411, FSL R9-5506, and FSL R9-0506; 0008 stands for FSL C2-0008; and 031 stands for FSL J1-031.

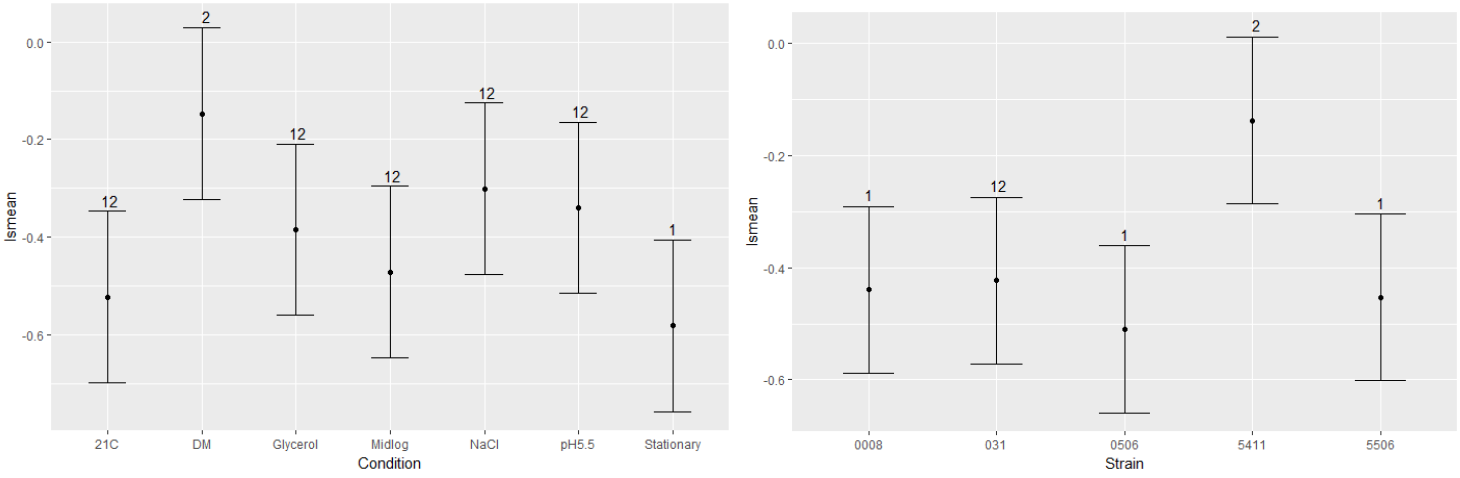

Figure 10: Fixed-effect linear model and post-hoc comparison of estimated means for *Listeria* strains, including surrogate, indicator, and index organisms for the first 24 h of incubation. Y-axis represents estimated means; x-axis on the left-hand graph shows condition and on the right-hand graph shows strains. Numbers indicate groups based on Tukey Honest Significant Difference method.

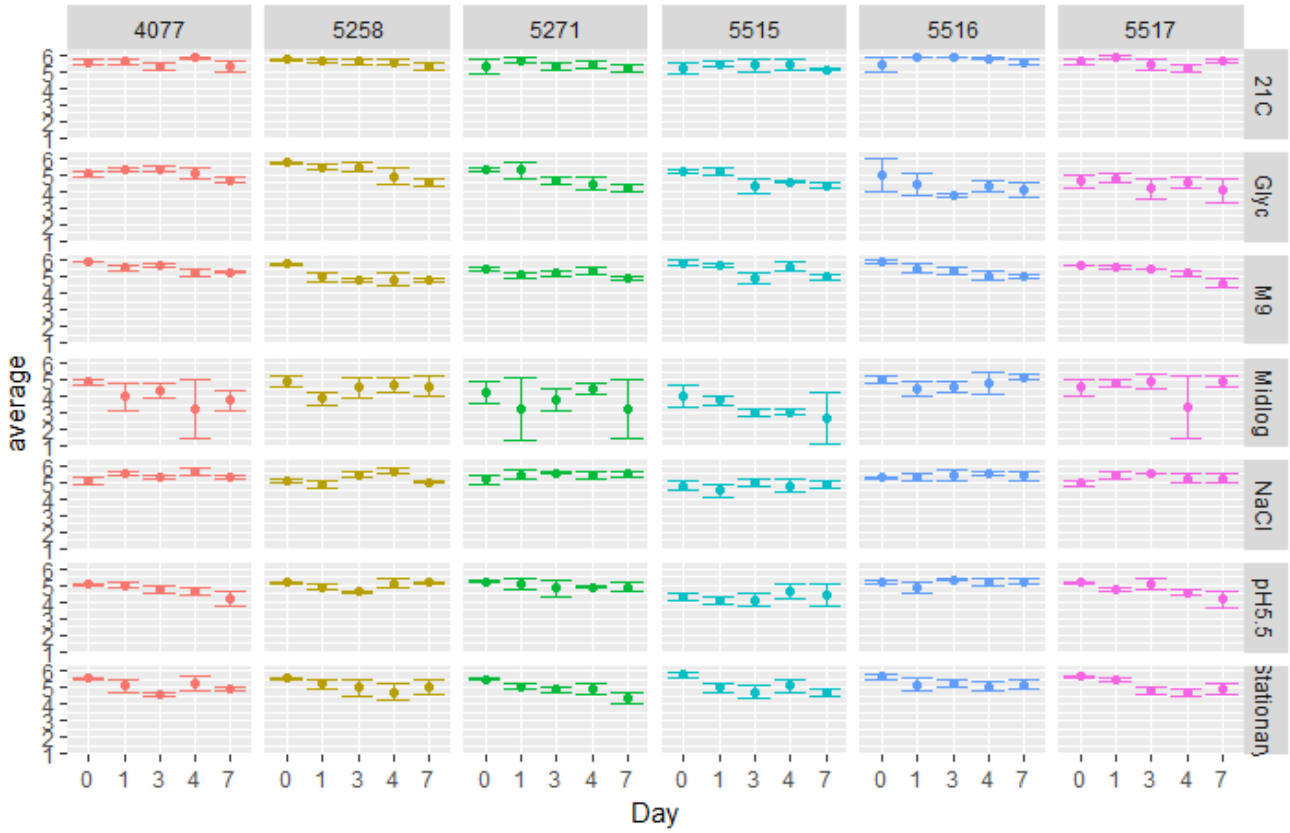

Figure 11: Growth curves for *E. coli* strains inoculated on pre-cut romaine lettuce. Y-axis represents average of three trials in log CFU/g; x-axis shows days of bacterial enumeration. Facets are divided into pre-growth conditions and strains. APC stands for aerobic plate count; strains are abbreviated, e.g., 4077 stands for FSL R9-4077.

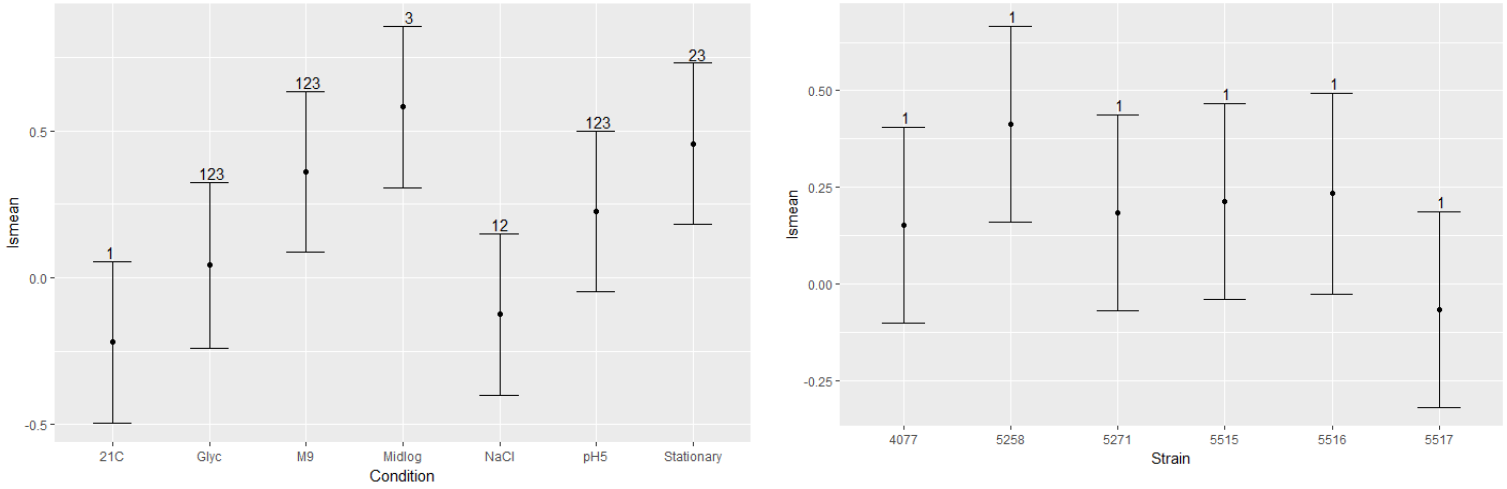

Figure 12: Fixed-effect linear model and post-hoc comparison of estimated means for *E. coli* strains, including surrogate, indicator, and index organisms for the first 24 h of incubation. Y-axis represents estimated means; x-axis on the left-hand graph shows condition and on the right-hand graph shows strains. Numbers indicate groups based on Tukey Honest Significant Difference method.

Bibliography

1. Fugett E, Fortes E, Nnoka C, Wiedmann M. International Life Sciences Institute North America Listeria monocytogenes strain collection: development of standard Listeria monocytogenes strain sets for research and validation studies. Journal of food protection. 2006;69(12):2929-38.
2. Clark WA, Geary DH. The story of the American Type Culture Collection--its history and development (1899-1973). Advances in applied microbiology. 1974;17(0):295-309.
3. Nightingale K, Bovell L, Grajczyk A, Wiedmann M. Combined sigB allelic typing and multiplex PCR provide improved discriminatory power and reliability for Listeria monocytogenes molecular serotyping. Journal of microbiological methods. 2007;68(1):52-9.
4. Weller D, Wiedmann M, Strawn LK. Spatial and Temporal Factors Associated with an Increased Prevalence of Listeria monocytogenes in Spinach Fields in New York State. Appl Environ Microbiol. 2015;81(17):6059-69.
5. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics (Oxford, England). 2014;30(15):2114-20.
6. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. Journal of computational biology : a journal of computational molecular cell biology. 2012;19(5):455-77.
7. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. Bioinformatics (Oxford, England). 2013;29(8):1072-5.
8. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics (Oxford, England). 2009;25(16):2078-9.
9. Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence classification using exact alignments. Genome Biology. 2014;15(3):R46.
10. Zhang S, Yin Y, Jones MB, Zhang Z, Deatherage Kaiser BL, Dinsmore BA, et al. Salmonella serotype determination utilizing high-throughput genome sequencing data. Journal of clinical microbiology. 2015;53(5):1685-92.
11. Joensen KG, Tetzschner AM, Iguchi A, Aarestrup FM, Scheutz F. Rapid and Easy In Silico Serotyping of Escherichia coli Isolates by Use of Whole-Genome Sequencing Data. Journal of clinical microbiology. 2015;53(8):2410-26.
12. Liao J, Wiedmann M, Kovac J. Genetic Stability and Evolution of the sigB Allele, Used for Listeria Sensu Stricto Subtyping and Phylogenetic Inference. Appl Environ Microbiol. 2017;83(12).
13. Davis S, Pettengill JB, Luo Y, Payne J, Shpuntoff A, Rand H, et al. CFSAN SNP Pipeline: an automated method for constructing SNP matrices from next-generation sequence data. PeerJ Computer Science. 2015;1:e20.
14. States DJ, Gish W. Combined use of sequence similarity and codon bias for coding region identification. Journal of computational biology : a journal of computational molecular cell biology. 1994;1(1):39-50.
15. Gardner SN, Slezak T, Hall BG. kSNP3.0: SNP detection and phylogenetic analysis of genomes without genome alignment or reference genome. Bioinformatics (Oxford, England). 2015;31(17):2877-8.
16. Stamatakis A. Using RAxML to Infer Phylogenies. Current protocols in bioinformatics. 2015;51:6.14.1-.
17. Green MR, Sambrook, Joseph. Molecular cloning : a laboratory manual / Michael R. Green, Joseph Sambrook. Cold Spring Harbor, NY Cold Spring Harbor Laboratory Press, [2012]. 2012.
18. Premaratne RJ, Lin WJ, Johnson EA. Development of an improved chemically defined minimal medium for Listeria monocytogenes. Appl Environ Microbiol. 1991;57(10):3046-8.
19. Bates D, Mächler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Using lme4. 2015. 2015;67(1):48.
20. Lenth RV. Least-Squares Means: The R Package lsmeans. 2016. 2016;69(1):33.

21. Toboldt A, Tietze E, Helmuth R, Junker E, Fruth A, Malorny B. Population structure of Salmonella enterica serovar 4,[5],12:b:- strains and likely sources of human infection. Appl Environ Microbiol. 2013;79(17):5121-9.
22. Timme RE, Pettengill JB, Allard MW, Strain E, Barrangou R, Wehnes C, et al. Phylogenetic Diversity of the Enteric Pathogen Salmonella enterica subsp. enterica Inferred from Genome-Wide Reference-Free SNP Characters. Genome Biology and Evolution. 2013;5(11):2109-23.
23. Zheng J, Luo Y, Reed E, Bell R, Brown EW, Hoffmann M. Whole-Genome Comparative Analysis of Salmonella enterica Serovar Newport Strains Reveals Lineage-Specific Divergence. Genome Biology and Evolution. 2017;9(4):1047-50.
24. Lee J, Monzingo AF, Keatinge-Clay AT, Harshey RM. Structure of Salmonella FlhE, conserved member of a flagellar Type III secretion operon. Journal of molecular biology. 2015;427(6 0 0):1254-62.
25. Iram SH, Cronan JE. The beta-oxidation systems of Escherichia coli and Salmonella enterica are not functionally equivalent. Journal of bacteriology. 2006;188(2):599-608.
26. Suzuki K, Babitzke P, Kushner SR, Romeo T. Identification of a novel regulatory protein (CsrD) that targets the global regulatory RNAs CsrB and CsrC for degradation by RNase E. Genes & development. 2006;20(18):2605-17.