PHENOTYPIC AND GENOTYPIC ANALYSIS OF PERSISTENT AND SPORADIC ISOLATES OF *LISTERIA MONOCYTOGENES*

BY

ALEXANDER JOSEPH TAYLOR

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Food Science and Human Nutrition
with a concentration in Food Science
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2018

Urbana, Illinois

Adviser:

Assistant Professor Matthew J. Stasiewicz

Abstract

The foodborne pathogen *Listeria monocytogenes* has the ability to persist for months to years within food-associated environments. These persistent strains maintain a constant risk of contamination and it is vital to identify and eliminate these persistent strains as soon as possible. Previous work has identified and characterized putative persistent and sporadic isolates from 30 retail delis in three regions of the U.S. over six months. Next generation sequencing allowed phylogenetic relationships to identify persistent strains based on the assumption that isolates of a persistent strain were more closely related than isolates of sporadic strains. Those studies provided the isolates and genomic data to investigate additional phenotypic and genotypic properties that may differentiate persistent and sporadic strains. In this present work, two questions were examined. (i) Do isolates of persistent strains contain distinct CRISPR spacers? (ii) Do isolates of persistent strains have higher growth rates or ability to grow in various stressful conditions or carbon sources associated with food environments?

CRISPRs, or clustered regularly interspaced short palindromic repeats, are comprised of pieces of foreign DNA (spacers) and are sandwiched between repeated sequences of host DNA (directed repeats). The foreign DNA is integrated into a bacteria's genome to act as a defense mechanism against foreign invading DNA, like bacteriophages. The integrated DNA, spacers, act like a logbook of which bacteriophages or foreign invaders have collided with a strain. Since persistent strains survive in environments for an extended period of time, and phages are common within food-associated environments, isolates of a persistent strain may contain more or different spacers than sporadic isolates.

Viewing the bioinformatic analysis of previously sequenced persistent and sporadic isolates from retail delis, 174 of 175 *L. monocytogenes* isolates contained CRISPR spacer arrays

that contained, at minimum, one CRISPR spacer within the array. Spacers that were found within isolates were conserved within the previously defined phylogenetic clades; these contiguous spacers were assigned to spacer patterns. While spacers were found within *L. monocytogenes* isolates, and isolates could be subtyped by their spacer patterns, there was no supporting data that persistent isolates could be differentiated from sporadic isolates based on their CRISPR spacer patterns.

Can isolates of persistent strains live for longer periods due to an increased ability to adapt to food-associated environmental stresses or carbon sources? The goal was to understand if persistent strains are more likely to grow and how well they grow in response to stress conditions, such as osmotic pressure, acidity, or sanitizer. Specifically, high-throughput growth assays were used to screen for isolates' ability to tolerate osmotic (5% or 10% NaCl), acidity (pH 5.2 or 9.2), or sub-lethal sanitizer (2 or 5 µg/mL benzalkonium chloride [BAC]) stresses, or to metabolize food-associated carbon sources (25 mM cellobiose, fructose, glucose, lactose, sucrose, or glycogen). At the end, persistent and sporadic strains did not differ from each other in either their ability to grow or their growth rate, if able to grow.

Therefore, taken together, these data suggest that there is no strain-specific phenotype that facilitates the persistence of *L. monocytogenes* in these retail deli environments. The data was able to show that CRISPRs are prevalent in *L. monocytogenes* and that *L. monocytogenes*' reactions to stressful treatments are consistent with previous literature, the persistence phenomenon is not distinguished through mobile genetic elements or responses to stress inducing environments. These data suggest that work should focus on identifying persistent strains of *L. monocytogenes* instead of understanding phenotypes or genotypes and look into improved environmental monitoring, cleaning and sanitation, or sanitary design.

Acknowledgements

I would like to thank my adviser, Dr. Matthew Stasiewicz, for choosing me to be one of his very first students, his support and guidance in this project, and the countless versions of drafts we have gone through. I would also like to thank Dr. Michael Miller for letting me work and be incorporated into his lab. With that, thank you to Dr. Yong-Su Jin and Dr. Dawn Bohn for taking the time to serve on my committee.

A big thank you goes out to my current labmates: Eric Xianbin Cheng and Shannon Rezac. As well as Dr. Miller's lab: Dr. Lili Zhang, Maxwell Holle, Luis Ibarra, Suneet Takhar, and Dr. Christine Xiaoji Liu for their support and making coming to lab incredibly fun and exciting. Along with my friends that I have met in FSHN, who have always been wonderful people that remind me why I chose Illinois.

I am thankful to the late Jonathan Baldwin Turner and to the JBT Fellowship Committee for allowing me the chance to study at Illinois without worry.

Thank you my family for supporting me through this part of my life. Thank you to my close friends for keeping my spirits up: Alex, Katie, Jory, Jeanette, Cameron, Dee, and my League team. Finally and most importantly, thank you to my significant other and partner-in-crime, Marissa Sylvester. Your smile and loving ways are what keep me going through each day. I would not be able to accomplish this without your help, your love, and your patience with me.

Table of Contents

Chapter 1: Introduction
1.1 Rationale and Significance

Listeria monocytogenes is recognized as a facultative intracellular pathogen that poses the highest risk for immunocompromised people, e.g. elderly, pregnant, infants, or people with weakened immunity.[5] Foods that are most frequently associated with *L. monocytogenes* are soft cheeses and dairy products, deli meats, smoked fish, leafy green vegetables, and, in general, ready-to-eat (RTE) products that are eaten without cooking or reheating.[9] *Listeria* organisms are widely found in natural and rural environments[6] and, thus, contaminate raw materials used in food processing plants. *Listeria* species are also well equipped to survive and grow in stressful conditions that most foodborne pathogens cannot. Able to grow in refrigerator temperatures, resistant to high osmotic pressure, tolerate pH ranges of 5.2 to 9.2, and slightly resistant to low-levels of sanitizers, *L. monocytogenes* is a serious threat to food safety.[1; 7]

Persistence has been researched extensively for how and why *L. monocytogenes* persists, and if persistence causes enhanced survival of *L. monocytogenes* in stress conditions or biofilm formation.[3] However, there is very little research on persistence outside of those areas. Recently, mobile genetic elements have been recently discovered to be able to differentiate *L. monocytogenes* strains.[10] Along with that study, CRISPRs, a type of mobile genetic element, have been shown to be useful for subtyping pathogenic bacteria[8], thus there may be a possibility to examine mobile genetic elements that are distinct to persistent strains. Some researchers have been able to identify CRISPRs within *L. monocytogenes*, but have not examined more about the sequences in the element called spacers.[2]

Most research on *L. monocytogenes* persistence has investigated the role of biofilm formation and sanitizer resistance.[4] However, very little data is on the effect of other stress-inducing conditions and carbohydrate sources on the growth rate and ability for persistent *L.*

monocytogenes to grow. Looking at the phenotypic responses of persistent strains to food-associated stress conditions and carbohydrate sources may be able to understand if the persistence phenomenon enhances the fitness and survivability of certain *L. monocytogenes* strains.

Thus, our ***rationale*** is that (i) by using CRISPRs as a subtyping tool, we can identify persistent strains of *L. monocytogenes*, which leads to seek-and-destroy tactics and (ii) by understanding persistent strains growth properties, we can know how to destroy persistent strains. The ***significance*** of this thesis is that we will be able to understand if there is an efficient genetic way to determine persistence or a phenotypic, biochemical way to determine persistence. By targeting persistent strains more effectively, food manufacturers can seek-and-destroy persistent strains more efficiently.

1.2 Objectives

Our objectives were as follows:

(i) Compare CRISPR spacers between persistent and sporadic strains of *L. monocytogenes*. We hypothesized that isolates of persistent strains would have differing spacers from sporadic strains and could be subtyped.

(ii) Analyze growth rates and ability to grow isolates from a persistent strain in food-associated stresses and various carbon sources. We hypothesized that persistent strains would have increased growth rate and more likely to grow in stress-inducing conditions and differing carbon sources.

1.3 References

(**1**) Buchanan, R. L., Gorris, L. G. M., Hayman, M. M., Jackson, T. C., & Whiting, R. C. (2017). A review of *Listeria monocytogenes*: An update on outbreaks, virulence, dose-response, ecology, and risk assessments. *Food Control, 75*, 1-13. doi:10.1016/j.foodcont.2016.12.016

(**2**) Di, H., Ye, L., Yan, H., Meng, H., Yamasak, S., & Shi, L. (2014). Comparative analysis of CRISPR loci in different *Listeria monocytogenes* lineages. *Biochem Biophys Res Commun, 454*(3), 399-403. doi:10.1016/j.bbrc.2014.10.018

(**3**) Ferreira, V., Wiedmann, M., Teixeira, P., & Stasiewicz, M. J. (2014). *Listeria monocytogenes* persistence in food-associated environments: epidemiology, strain characteristics, and implications for public health. *J Food Prot, 77*(1), 150-170. doi:10.4315/0362-028X.JFP-13-150

(**4**) Kremer, P. H., Lees, J. A., Koopmans, M. M., Ferwerda, B., Arends, A. W., Feller, M. M., et al. (2017). Benzalkonium tolerance genes and outcome in *Listeria monocytogenes* meningitis. *Clin Microbiol Infect, 23*(4), 265 e261-265 e267. doi:10.1016/j.cmi.2016.12.008

(**5**) McLauchlin, J. R., C. E. D. (2015). *Listeria**. *Bergey's Manual of Systematic of Archaea and Bacteria*, 1-29. doi:10.1002/9781118960608.gbm00547

(**6**) Orsi, R. H., & Wiedmann, M. (2016). Characteristics and distribution of *Listeria* spp., including *Listeria* species newly described since 2009. *Appl Microbiol Biotechnol, 100*(12), 5273-5287. doi:10.1007/s00253-016-7552-2

(**7**) Scallan, E., Hoekstra, R. M., Angulo, F. J., Tauxe, R. V., Widdowson, M. A., Roy, S. L., et al. (2011). Foodborne illness acquired in the United States--major pathogens. *Emerg Infect Dis, 17*(1), 7-15. doi:10.3201/eid1701.P11101

(**8**) Shariat, N., Sandt, C. H., DiMarzio, M. J., Barrangou, R., & Dudley, E. G. (2013). CRISPR-MVLST subtyping of *Salmonella enterica* subsp *enterica* serovars Typhimurium and Heidelberg and application in identifying outbreak isolates. *BMC Microbiology, 13*. doi:10.1186/1471-2180-13-254

(**9**) Vazquez-Boland, J. A., Kuhn, M., Berche, P., Chakraborty, T., Dominguez-Bernal, G., Goebel, W., et al. (2001). *Listeria* pathogenesis and molecular virulence determinants. *Clin Microbiol Rev, 14*.

(**10**) Wang, Q., Holmes, N., Martinez, E., Howard, P., Hill-Cawthorne, G., & Sintchenko, V. (2015). It is not all about single nucleotide polymorphisms: comparison of mobile genetic elements and deletions in *Listeria monocytogenes* genomes links cases of hospital-acquired listeriosis to the environmental source. *Journal of Clinical Microbiology, 53*(11), 3492-3500. doi:10.1128/JCM.00202-15

Chapter 2: Literature Review
2.1 *Listeria monocytogenes*

2.1.1 History of *L. monocytogenes*

Listeria is a genus of Gram-positive, short rods named after Lord Joseph Lister, who was an English surgeon and pioneer of antisepsis.[18] The genus *Listeria* currently includes 17 species: *L. aquatic, L. booriae, L. cornellensis, L. fleischmannii, L. floridensis, L. grandensis, L. grayi, L. innocua, L. ivanovii, L. marthii, L. monocytogenes, L. newyorkensis, L. riparia, L. rocourtiae, L. seeligeri, L. weihenstephanensis,* and *L. welshimeri.*[21] Of those species, *L. monocytogenes* and, more rarely, *L. ivanovii* are potentially pathogenic to humans. The first *Listeria* discovered, what we now know as *L. monocytogenes*, was isolated in 1926 by E.G.D. Murray, R.A. Webb, and M.B.R. Swann, from a rabbit, and originally named as *Bacterium monocytogenes.*[18] When another scientist, James Harvey Hunter Pirie, isolated the same bacterium, it was renamed to *Listerella* and then shortened to *Listeria*, as a fungus shared the *Listerella* name. The first cases of human listerosis were reported in 1929 in Denmark.[28] However, *L. monocytogenes* was not recognized as a foodborne pathogen until the 1980s when several large outbreaks occurred.[15]

2.1.2 Food-associated *L. monocytogenes* outbreaks

In 1985, one of the deadliest listerosis outbreaks occurred in California. A listerosis outbreak in Mexican-style soft cheese had a confirmation of 52 deaths, with 19 stillbirths and 10 infant children.[20] There was an immediate demand to track outbreaks and target *Listeria* spp. as soon as possible. Then, in 1987, the Food Safety and Inspection Service (FSIS) implemented regulatory microbial testing for *L. monocytogenes* for RTE meat and poultry products and labeled it as a "zero tolerance" pathogen.[22] An estimated 1,600 Americans every year become ill with listerosis and 260 fatalities occur per year.[4] With a fatality rate of 15.6%, *L. monocytogenes* is one

of the deadliest foodborne pathogens once acquired.[2] In a recently published source attribution estimate, the FSIS reported that nearly 90% of listerosis outbreaks were linked to fruits and dairy, with sprouts, turkey, and other foods being <10.0% of the outbreaks.[11] While the number of total *L. monocytogenes* outbreaks, in that report, was low (n=26) compared to the number of *Salmonella* outbreaks (n=638),[11] many recent recalls have been seen in biscuits, breads, apples, and other products.[17]

However, in the early 2000's, listerosis cases were attributed, majorly, to RTE deli meats.[5] Products that were sliced at retail delis, versus those at a processor, were 5 to 7 times more likely to be contaminated with *L. monocytogenes*.[2] Of the listerosis cases from RTE deli meat, 60% of them were attributed to retail delis.[8] Since *L. monocytogenes* is resistant to multiple environmental stresses, it is able to survive and replicate in retail deli facilities and leads to a phenomenon known as persistence.

2.2 Persistence

2.2.1 Concept of persistence

Persistence, in and of itself, is a simply understood, but complex problem, for the control of *L. monocytogenes*. The persistence of *L. monocytogenes* stems from the ability of the exact same strain of *L. monocytogenes* to be found for months up to years later within the same facility.[3] However, it is not fully known which the specific mechanisms allow these isolates to last and thrive for so long.[2] Mechanisms have been researched, and multiple papers have determined that persistence is either aided by the environment or the strain itself is fit for survival due to a resistance or other adaptation. By just existing, persistent strains increase the risk of contamination compared to non-persistent, sporadic, strains. While *L. monocytogenes* may enter a facility from raw materials, ineffective cleaning and sanitation, poor design or condition of food equipment, or

survival in niche areas of the facility add to why *L. monocytogenes* can persist. Strains that are identified as sporadic have been shown to transition to persistent later on.[7] In order to better understand mechanisms of persistence, understanding phenotypic and genotypic differentiations of persistent strains may have value.

2.2.2 Persistence: theoretical and empirical definitions

Before fully delve into the complexities of persistence, there need to be a clarification on a dissonance in literature between the definitions of persistence. There are two key definitions, theoretical and empirical, that are used throughout previous literature. The theoretical definition of persistence is that a strain is at least surviving in an environment, likely making clones of itself, and is therefore defined as persistent. The empirical definition is that isolates are identified as of a persistent strain if they are regarded as the same strain by a particular molecular subtyping method, and that strain is observed over a sufficiently long period, thus assumed to be persistent.[26]

2.2.3 Persistence and resistances to stress

Since *L. monocytogenes*, as a species, is known to be so resistant to multiple stresses, researchers were curious as to what additional phenotypic differences persistent strains happen to display. Researchers honed in on aspects that were consistently found within retail, manufacturing, or processing facilities, which were sanitizer resistance and biofilm formation. However, there was no conclusive data on if there phenotypes were the only ones that persistent strains held. More specifically, there is very little data on whether persistence can affect growth rate, ability to grow, or resistance to any stresses.[2] Pairing with this inconclusive data, the term of persistence has been used inconsistently. Some papers focus on finding the isolates a certain number of times within an environment. Some were basing persistence on previous genotypes

that were persistent. To get a clearer picture of persistence, there needs to be a clear definition with strong genetic relationships.

2.2.4 Persistence and subtyping tools

Many of the subtyping tools used to identify *L. monocytogenes* are described in detail in Section 2.3.1, but it should be of note that many of those tools are not able to accurately determine persistence. Pairing that with the fact that sporadic isolates can become, at some point, persistent strains makes identification incredibly difficult.[7] In fact, many times, to confirm persistence, multiple subtyping methods are used.[7] The main tool used was pulse-field gel electrophoresis (PFGE), which used the restriction enzymes *Apa*I and *Asc*I to digest (cut) the *L. monocytogenes* isolates' genomes. The digested DNA was then analyzed in an electrophoresis gel for the position and number of bands created. Isolates with matching bands would have the same PFGE subtype given to them and these genotypes.[13] However, PFGE subtyping needed multiple time point isolates to show if the isolate/strain was persistent or not. Then, one method was able to really help in discerning persistent isolates: Whole genome sequencing (WGS). WGS opened up the door to statistically identify and classify putatively persistent isolates that were based on single nucleotide polymorphisms (SNPs). By examining exact core genomic sequences, researchers could find SNP patterns that were unique to a single retail deli, unique to a single state, or spread across multiple states.[26]

2.3 Next generation sequencing

2.3.1 Previous subtyping and next generation sequencing technologies

Being able to identify and subtype *L. monocytogenes* strains is vital in differentiating an expansive number of *L. monocytogenes* strains. The last 20+ years has seen a major shift to molecular subtyping due to the specificity and sensitivity of genotypic approaches.[13] Techniques

like multi-locus sequence based typing (MLST) and PFGE were the gold standards for molecular subtyping. MLST focuses on seven loci from different housekeeping genes and are analyzed for nucleotide differences.[13] Whereas, PFGE, uses restriction enzymes, *Apa*I and *Asc*I, to cleave at specific nucleotide sites and, when ran on in a gel, produce a combination of bands that have patterns useful for subtyping.[14]

A specific type of next generation sequencing is whole genome sequencing (WGS). WGS is a tool used to determine the complete DNA sequence of an organism's genome. The idea is to look at specific nucleotide changes, single nucleotide polymorphisms (SNPs). WGS is usually done by breaking the genome into fragments. These fragments will have overhangs that align with other fragments, like a linear puzzle. Once all of the fragments are combined together, this becomes nearly the entire genome. This genome is then compared to a reference genome to align housekeeping genes and core genes. Once assembled, the genome can then be analyzed by researchers for genes, SNPs, etc. at incredibly high resolution and accuracy. By having such high resolution, researchers are able to investigate even further than before, as detailed in Section 2.3.3.

2.3.2 Advantages of next generation subtyping

There are many advantages of next generation subtyping compared to MLST/PFGE and other classical subtyping methods. WGS's cost has drastically gone down and WGS provides higher resolution SNP-level information. With that, WGS has the ability to be a high-throughput technique. These advances emboldened foodborne disease surveillance.

2.3.3 Follow-up analyses from WGS databases

Information acquired from next generation sequencing of microbes is of most of the entire genome. While WGS has clear advantages in subtyping, as denoted above, only a portion of the data gathered is used. With the rest of that data, having already been sequenced and readily

available, researchers are able to ask questions and investigate follow-up analyses besides subtyping.

Specifically, for *L. monocytogenes*, in two years of WGS application more numbers of listerosis cluster were detected, more were detected sooner, more outbreaks were solved, and drastically more number of cases were linked to a food source.[6;12;26] Two pairs of *L. monocytogenes* strains that had highly similar genomic compositions and PFGE profiles were subjected to WGS comparison. One pair of strains differed significantly from one another by their antibiotic and heavy metal stress resistance, their mobile genetic elements, and their pathogenic potential of an isolate.[6] Recently, an Australian paper, after WGSing isolates from a *L. monocytogenes* outbreak, found that specific mobile genetic elements within the genomes were linked to hospital-acquired listerosis.[29]

For example, multiple researchers have looked at a transporter genes, *qacH* and *bcrABC*, that belong to small multidrug resistance protein family and are associated with resistance to sanitizers commonly used in the food industry.[19] Another clear example would be to look into mobile genetic elements, as aforementioned by the finding of specific mobile genetic elements that were linked to a listerosis outbreak.[29] One such mobile genetic element, CRISPR, has been of significant importance in the past decade.[9]

2.4 CRISPRs

CRISPR is an acronym for clustered regularly interspaced short palindromic repeats, which are made of sequences known as spacers and directed repeats.[1] Spacers are short sequences of foreign DNA from bacteriophages or other harmful invaders to the bacteria.[1] Directed repeats are sequences of the host's genome that sandwich spacers on either side.[1] Therefore, a whole CRISPR array is a repeated motif of directed repeat – spacer – directed repeat – spacer … – directed repeat.

CRISPRs have become a topic of interest due to their involvement in genome editing, but for bacteria, they are a defense mechanism to bacteriophages and to foreign bacteria.[10] In a way, CRISPRs can be considered as a logbook of what has tried to invade the bacteria of interest. Recently, there have been multiple papers that examine CRISPRs to subtype pathogenic bacteria[23;24;25] and there is evidence that CRISPRs are related to virulence in pathogens with Type IIA CRISPRs.[16;27] *L. monocytogenes* happens to have these Types IIA CRISPRs.[20] However, at the time of this thesis, it is not known what the exact role or relationship is between CRISPRs and *L. monocytogenes*' virulence. Also, little is known about CRISPRs and their relation to *L. monocytogenes*.

CRISPRs and *L. monocytogenes* are more thoroughly reviewed CRISPRs in the Chapter 3 introduction.

2.5 Phenotyping of persistent and sporadic *L. monocytogenes*

Phenotypes of persistent *L. monocytogenes* strains are more thoroughly reviewed in Chapter 4 introduction.

2.6 References

(**1**) Bolotin, A., Quinquis, B., Sorokin, A., & Ehrlich, S. D. (2005). Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology, 151*. doi:10.1099/mic.0.28048-0

(**2**) Buchanan, R. L., Gorris, L. G. M., Hayman, M. M., Jackson, T. C., & Whiting, R. C. (2017). A review of *Listeria monocytogenes*: An update on outbreaks, virulence, dose-response, ecology, and risk assessments. *Food Control, 75*, 1-13. doi:10.1016/j.foodcont.2016.12.016

(**3**) Carpentier, B., & Cerf, O. (2011). Review — Persistence of *Listeria monocytogenes* in food industry equipment and premises. *International Journal of Food Microbiology, 145*(1), 1-8. doi:10.1016/j.ijfoodmicro.2011.01.005

(**4**) CDC. (2017). *People at Risk*. Retrieved from Center for Disease Control: https://www.cdc.gov/listeria/risk.html

(**5**) CFSAN. (2003). *Quantitative assessment of the relative risk to public health from foodborne* Listeria monocytogenes *among selected categories of ready-to-eat foods*. Retrieved from https://www.fda.gov/downloads/food/foodscienceresearch/ucm197329.pdf

(**6**) Edward M. Fox, A. C., Kieran Jordan, Aidan Coffey, Cormac G.M. Gahan, Olivia McAuliffe. (2017). Whole genome sequence analysis; an improved technology that identifies underlying genotypic differences between closely related *Listeria monocytogenes* strains. *Innovative Food Science & Emerging Technologies, 44*, 89-96. doi:10.1016/j.ifset.2017.07.010

(**7**) Ferreira, V., Wiedmann, M., Teixeira, P., & Stasiewicz, M. J. (2014). *Listeria monocytogenes* persistence in food-associated environments: epidemiology, strain characteristics, and implications for public health. *J Food Prot, 77*(1), 150-170. doi:10.4315/0362-028X.JFP-13-150

(**8**) FSIS. (2010). *FSIS comparative risk assessment for* Listeria monocytogenes *in ready-to-eat meat and poultry deli meats report*. Retrieved from https://www.fsis.usda.gov/shared/PDF/Comparative_RA_Lm_Report_May2010.pdf

(**9**) Godde, J. S., & Bickerton, A. (2006). The repetitive DNA elements called CRISPRs and their associated genes: evidence of horizontal transfer among prokaryotes. *J Mol Evol, 62*. doi:10.1007/s00239-005-0223-z

(**10**) Grissa, I., Vergnaud, G., & Pourcel, C. (2007). The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. *BMC Bioinformatics, 8*(1). doi:10.1186/1471-2105-8-172

(**11**) Interagency Food Safety Analytics Collaboration. (2017). *Foodborne illness source attribution estimates for 2013 for* Salmonella, Escherichia coli *O157,* Listeria monocytogenes, *and* Campylobacter *using multi-year outbreak surveillance data*. Retrieved from United States, G.A. and D.C.:

(**12**) Jackson, B. R., Tarr, C., Strain, E., Jackson, K. A., Conrad, A., Carleton, H., et al. (2016). Implementation of nationwide real-time whole-genome sequencing to enhance listeriosis outbreak detection and investigation. *Clin Infect Dis, 63*(3), 380-386. doi:10.1093/cid/ciw242

(**13**) Jadhav, S., Bhave, M., & Palombo, E. A. (2012). Methods used for the detection and subtyping of *Listeria monocytogenes*. *J Microbiol Methods, 88*(3), 327-341. doi:10.1016/j.mimet.2012.01.002

(**14**) Laksanalamai, P., Joseph, L. A., Silk, B. J., Burall, L. S., L. Tarr, C., Gerner-Smidt, P., et al. (2012). Genomic characterization of *Listeria monocytogenes* strains involved in a multistate listeriosis outbreak associated with cantaloupe in US. *PLoS One, 7*(7), e42448. doi:10.1371/journal.pone.0042448

(**15**) Leclercq, A., Charlier, C., & Lecuit, M. (2014). Global burden of listeriosis: the tip of the iceberg. *Lancet Infect Dis, 14*(11), 1027-1028. doi:10.1016/S1473-3099(14)70903-X

(**16**) Louwen, R., Staals, R. H., Endtz, H. P., van Baarlen, P., & van der Oost, J. (2014). The role of CRISPR-Cas systems in virulence of pathogenic bacteria. *Microbiol Mol Biol Rev, 78*(1), 74-88. doi:10.1128/MMBR.00039-13

(**17**) Marler Clark. (2018). *Listeria* Blog. *Listeria Recalls.* Retrieved from https://www.listeriablog.com/listeria-recalls/

(**18**) McLauchlin, J. R., C. E. D. (2015). *Listeria**. *Bergey's Manual of Systematic of Archaea and Bacteria*, 1-29. doi:10.1002/9781118960608.gbm00547

(**19**) Moretro, T., Schirmer, B. C., Heir, E., Fagerlund, A., Hjemli, P., & Langsrud, S. (2017). Tolerance to quaternary ammonium compound disinfectants may enhance growth of *Listeria monocytogenes* in the food industry. *Int J Food Microbiol, 241*, 215-224. doi:10.1016/j.ijfoodmicro.2016.10.025

(**20**) Neuman, W. (2011, September 28, 2011). Deaths from cantaloupe *Listeria* rise. *New York Times*, p. B1. Retrieved from https://www.nytimes.com/2011/09/28/business/deaths-from-cantaloupe-listeria-rises.html?src=me&ref=general

(**21**) Orsi, R. H., & Wiedmann, M. (2016). Characteristics and distribution of *Listeria* spp., including *Listeria* species newly described since 2009. *Appl Microbiol Biotechnol, 100*(12), 5273-5287. doi:10.1007/s00253-016-7552-2

(**22**) Shank, F. R., Elliot, E. L., Wachsmuth, I. K., & Losikoff, M. E. (1996). US position on *Listeria monocytogenes* in foods. *Food Control, 7*(4), 229-234. doi:https://doi.org/10.1016/S0956-7135(96)00041-2

(**23**) Shariat, N., & Dudley, E. G. (2014). CRISPRs: molecular signatures used for pathogen subtyping. *Appl Environ Microbiol, 80*(2), 430-439. doi:10.1128/Aem.02790-13

(**24**) Shariat, N., Kirchner, M. K., Sandt, C. H., Trees, E., Barrangou, R., & Dudley, E. G. (2013). Subtyping of *Salmonella enterica* serovar Newport outbreak isolates by CRISPR-MVLST and determination of the relationship between CRISPR-MVLST and PFGE results. *Journal of Clinical Microbiology, 51*(7), 2328-2336. doi:10.1128/Jcm.00608-13

(**25**) Shariat, N., Sandt, C. H., DiMarzio, M. J., Barrangou, R., & Dudley, E. G. (2013). CRISPR-MVLST subtyping of *Salmonella enterica* subsp *enterica* serovars Typhimurium and Heidelberg and application in identifying outbreak isolates. *BMC Microbiology, 13*. doi:10.1186/1471-2180-13-254

(**26**) Stasiewicz, M. J., Oliver, H. F., Wiedmann, M., & den Bakker, H. C. (2015). Whole-genome sequencing allows for improved identification of persistent *Listeria monocytogenes* in food-associated environments. *Appl Environ Microbiol, 81*(17), 6024-6037. doi:10.1128/AEM.01049-15

(**27**) Stout, E., Klaenhammer, T., & Barrangou, R. (2017). CRISPR-cas technologies and applications in food bacteria. *Annu Rev Food Sci Technol, 8*, 413-437. doi:10.1146/annurev-food-072816-024723

(**28**) Vazquez-Boland, J. A., Kuhn, M., Berche, P., Chakraborty, T., Dominguez-Bernal, G., Goebel, W., et al. (2001). *Listeria* pathogenesis and molecular virulence determinants. *Clin Microbiol Rev, 14*.

(**29**) Wang, Q., Holmes, N., Martinez, E., Howard, P., Hill-Cawthorne, G., & Sintchenko, V. (2015). It is not all about single nucleotide polymorphisms: comparison of mobile genetic elements and deletions in *Listeria monocytogenes* genomes links cases of hospital-acquired listeriosis to the environmental source. *Journal of Clinical Microbiology, 53*(11), 3492-3500. doi:10.1128/JCM.00202-15

Chapter 3: CRISPR-based Subtyping Using Whole Genome Sequence Data does Not Improve Differentiation of Persistent and Sporadic *Listeria monocytogenes* Strains

3.1 Abstract

The foodborne pathogen *Listeria monocytogenes* can persist in food-associated environments for long periods. To identify persistent strains, the subtyping method pulse field gel electrophoresis (PFGE) is being replaced by whole genome sequence (WGS) -based subtyping. It was hypothesized that analyzing specific mobile genetic elements, CRISPR (Clustered Regularly Interspaced Short Palindromic Short Repeat) spacer arrays, extracted from WGS data, could differentiate persistent and sporadic isolates within WGS-based clades. To test this hypothesis, 175 *L. monocytogenes* isolates, from previously recovered from retail delis, were analyzed for CRISPR spacers using CRISPRFinder. These isolates represent 23 phylogenetic clades defined by WGS-based single nucleotide polymorphisms and closely related sporadic isolates. In 174/175 (99.4%) of isolates, at least one array with one spacer was identified. Numbers of spacers in a single array ranged from 1-28 spacers. Isolates were grouped into eighteen spacer patterns (SPs). SP variation was consistent with WGS-based clades forming patterns of (i) one SP to one clade, (ii) one SP across many clades, (iii) many SPs within one clade, and (iv) many SPs across many clades. Unfortunately, SPs did not appear to differentiate persistent from sporadic isolates within any WGS-based clade. Overall, the data shows that (i) CRISPR arrays are common in WGS data for these food-associated *L. monocytogenes* (ii) CRISPR arrays are conserved within phylogenetic clades of *L. monocytogenes*, and (iii) CRISPR subtyping cannot improve the identification of persistent or sporadic isolates.

3.2 Introduction

Listeria monocytogenes is a foodborne pathogen that causes the illness listeriosis. Listeriosis has one of the highest case fatality rates, 15.6%, of known foodborne pathogens.[8] The CDC estimates 1,600 Americans, annually, become ill with listeriosis and 260 fatalities occur per year.[9] In a study of 31 foodborne pathogens, *L. monocytogenes* had a hospitalization rate of 94%.[23] Many foods can harbor *L. monocytogenes*, such as ready-to-eat (RTE) foods like dairy products (ice cream, milk), fruits and vegetables (cantaloupes, spinach, lettuce), and deli meats.[18] While 86% of listeriosis illnesses were recently attributed to fruits and dairy foods[18], deli meats have long been identified as a source of listeriosis. Specifically, 90% of listeriosis cases in the U.S. in early 2000's were attributed to deli meats.[10] Products sliced at retail delis were estimated to be contaminated 5 to 7 times more frequently than products packaged at a processor[8] and meat sliced at retail is responsible for 60% of listeriosis cases from RTE deli meats.[13] The risk posted to retail deli meats is highlighted because this paper studies *L. monocytogenes* previously isolated from longitudinal studies in retail deli operations[26] and characterized by whole genome sequencing.[27]

3.2.1 Persistence

L. monocytogenes can persist in food-associated environments for months to years.[12] Persistent strains represent a risk of continual contamination within a manufacturing or processing environment. For this paper, persistence is defined as the continued presence of a clonal population of bacteria, over time, at a specific location. Persistence can then be empirically identified by isolating, on different dates, *L. monocytogenes* that are found to be of the same strain.

3.2.2 Subtyping

Subtyping methods for identifying persistent *L. monocytogenes* rely on the assumption persistent stains will be more genetically similar to each other than to sporadic strains.[27]

Investigators use the best available subtyping methods to identify strains at the sub-species level, with pulse field gel electrophoresis (PFGE) being the previous gold-standard for *L. monocytogenes.*[19] Next generation whole genome sequencing (WGS) has improved foodborne disease epidemiology[21] and has been specifically applied to improve the identification of persistent *L. monocytogenes* in food-associated environments.[27] WGS-based subtyping methods currently use either single nucleotide polymorphisms (SNPs)[27] or whole genome MLST.[19] Both methods actively exclude the impact of mobile genetic elements that can be acquired by mechanisms other than vertical transmission; yet raw WGS data allows for analysis of mobile genetic elements, if one chooses to do so.

3.2.3 CRISPRs as a Subtyping Tool

Clustered regularly inter-spaced short palindromic repeats (CRISPRs) are one such mobile genetic element. Nearly half of all bacterial species have CRISPR systems[14;15] and bacteria use them to incorporate short sequences of invading genetic elements, virus or plasmid, into the host's genome. They are also guides for a multifunctional protein complex to cleave other foreign genetic materials.[3] CRISPRs are segments of DNA comprised of spacers and directed repeats (DRs) that form an array.[29] Spacers are short sequences in the array, usually between 21 to 47 base pairs, of inserted viral or invading DNA sandwiched between repeated, consensus sequences, DRs.[4] Spacers can act as a logbook for what bacteriophages infect the microbe, as spacers can be passed onto progeny. Therefore, isolates with similar spacers or spacer patterns are, in theory, more closely related than isolates with different spacers. This relationship is the conceptual foundation for CRISPR-based subtyping.

CRISPR-based subtyping has shown to be useful for industrial starter cultures, probiotic cultures and strains, and animal commensal species.[1;5;7;16] CRISPR spacers have also been shown

to subtype other pathogens, such as *Erwinia amylovora*[24], *Yersinia* spp., Shiga-toxin producing *Escherichia coli*, *Campylobacter jejuni*, *Staphylococcus aureus*, *Mycobacterium tuberculosis*, and *Salmonella enterica*.[2] Each of these CRISPR-based subtypings have opened new paths to understanding the microbes' short-term evolution.[2] To the best of the authors' knowledge, there have been no reports of using CRISPR-based subtyping with *L. monocytogenes*.

3.2.4 Hypothesis

Previous research has identified CRISPR systems within *L. monocytogenes* genomes, and have shown different CRISPR loci associated with different *L. monocytogenes* lineages.[11] The hypothesis in this study was that differences in CRISPR spacer arrays may improve differentiation of persistent from sporadic *L. monocytogenes* strains beyond what is possible using WGS SNP-based subtyping. The conceptual model was that if functional CRISPR systems are able to log bacteriophage infection, then strains that persist within a given food-associated environment with a given phage pressure would acquire more similar spacers than otherwise similar strains from other environments. In this way, CRISPR arrays would have increased diversity, independent of core genome SNP analysis.

3.3 Materials and Methods

3.3.1 Isolate Selection and Sequence Status

To test this hypothesis, genomes of 175 *L. monocytogenes* isolates were examined. These isolates were collected from twenty retail delis in three U.S. states from early 2010 to late 2011[26], and were previously classified as persistent or sporadic by WGS analysis.[27] Briefly, the 2015 study conducted WGS SNP-based phylogenetic analysis to identify epidemiologically relevant clades, compared these phylogenetic clades to the associated MLST and PFGE subtypes, and used a SNP difference metric to identify persistent strains. In this study, CRISPR spacer arrays were

identified within all these well-characterized isolates. By subtyping using CRISPR spacers, comparisons can be made between CRISPR arrays and the previous WGS-SNP and PFGE-based subtyping results. The comparisons represent both the isolates' population as a whole and the differentiation of persistent and sporadic strains. The isolates represent 23 phylogenetic clades, in which each clade has at least one putative persistence event. As a starting point for this project, *de novo* assembled contigs were retrieved, from the BioProject accession number PRJNA245909, as input data for CRISPR analysis.

3.3.2 Bioinformatic CRISPR Identification

To identify CRISPR arrays, the web-based tool CRISPRFinder was utilized.[15] Draft genome data was inputted into CRISPRFinder, specifically contigs, and then the tool identifies CRISPR arrays by finding the short, palindromic repeats. The output for each genome contains the whole CRISPR Array (spacers, directed repeats), the location (node) of the CRISPR, the length of the array, the DR consensus sequence, and the individual spacer sequences. This work focuses on the specific spacer sequences. As the CRISPRFinder database is updated infrequently, it should be of note that the final analysis used the database updated as of May 9, 2017.

There are important consequences to using the default options in this software i) the parameters for the repeats and spacers are, by default, 23-55 bp for DRs and 25-60 bp for spacers, which means that DRs and spacers above or below those sizes will be excluded and ii) spacers are identified as "questionable" if the identified CRISPR array has only two or three DRs and therefore less than three spacers in the array.[14] Questionable spacers were kept in the analysis because as all CRISPR-based variations were examined, specifically for differences between persistent and sporadic strains of *L. monocytogenes*.

3.3.3 Spacer Pattern Assignment

Spacers were grouped into patterns to reduce redundancy in the data. First, letters were assigned to each unique spacer. If the specific spacer was present in more than one isolate (matching in both nucleotide composition and length), each instance would receive the same letter. Second, certain spacers were noticed to always identify as a consistent set of the same individual spacers, in a single node, in the same order. This single set, found in one or more isolates, was given a single Spacer Pattern number (e.g. SP 1). Then isolates with multiple spacer patterns in different nodes were identified by composite spacer pattern, e.g. SP 1-3, each number referring to a unique spacer pattern and the dash indicating a genomic distance between arrays. The process of assigning unique SP identifiers across isolates is illustrated in Figure 3.1.

3.3.4 Data Analysis

The discriminatory powers of previously identified PFGE, phylogenetic clade subtyping, and newly applied Spacer Pattern subtyping were evaluated according to Simpson's Index of Diversity.[17] Subsequent analysis was focused on the comparisons of SPs with isolates of various clades, PFGE subtypes, locations, and persistence statuses. Description analysis was sufficient to show the value of CRISPR spacers to subtype isolates, therefore there was no need to test for statistical differences between groups. All data analyses were carried out in Excel.

3.4 Results

CRISPR spacer arrays were identified in whole genome sequencing data for 175 persistent and sporadic isolates of *L. monocytogenes*. In total, there were 1,572 spacers identified within 175 isolates and those spacers were present in one of 18 unique spacer patterns.

3.4.1 CRISPR Spacers are Identified in Most *L. monocytogenes* Isolates

A total of 1,572 spacers were identified, representing 160 unique spacers (Table S3.2). The individual spacers ranged from 30 to 73 base pairs (bp) long, with an average length of 37 bp long. These 160 unique spacers were initially assigned to one of 24 SPs. Six of the SPs were found to be reverse compliments of other spacer patterns, so these were combined to leave 18 individual, (no dashes) final SPs. The number of spacers represented in each final SP ranged from 1 to 28 (Figure 3.2).

Of the 175 isolate genomes analyzed, 174 contained at least one CRISPR spacer throughout their genomes (Figure 3.3). The most common number of spacers identified within an isolate was 3 spacers, in 51% (89/175) of the isolates. The greatest number of spacers identified in an isolate was 41 spacers, identified in seven isolates. Of the 174 isolates, 89 isolates contained one or more questionable SPs (patterns with 1-2 spacers). The other 85 isolates all had CRISPR arrays of 3 or more spacers in their SPs, denoted as "confirmed." Of the 89 isolates with questionable spacers, a portion of the isolates (24) were found to have other confirmed SPs. If these isolates are included as isolates without questionable spacers, then 110 (63%) out of 175 isolates contained CRISPR arrays with three or more spacers. This means that even if isolates with only questionable spacers are subtracted, CRISPR spacers are identified in the majority of the *L. monocytogenes* isolates.

In addition, there were cases where multiple SPs were identified in separate nodes or with genomic space between the arrays within an isolate's genome (e.g., see Figure 3.1). To capture this point of interest in subsequent data analysis, the SP labels were hyphenated for a given isolate to indicate the separate SPs. For example, an isolate with the CRISPR arrays SP 4 and SP 5 identified in separate nodes would be labeled as having SP 4-5.

3.4.2 CRISPR-based Subtyping

While there has been focus in standard CRISPR-based subtyping, it was not the main interest of this study. Isolates were subtyped based on their CRISPR SPs alone (Table S3.1). However, as groupings were arbitrarily labelled, it should be recognized that these groupings do not fully determine the isolates' phylogenetic relationships. From this generic subtyping, 13 composite (e.g. 4-5), distinct groupings were observed.

3.4.3 Spacer Diversity is Conserved within Phylogenetic Clades

When comparing the distribution of SPs represented among isolates within WGS-based clades (Table 3.1), four relationship types were observed. Those are (i) one SP to one clade, such as SP 21-22-23 representing all isolates in clade X and only isolates in clade X; (ii) one SP over multiple clades, such as SP 3 representing all isolates with spacers in clades D, E, and F; (iii) multiple SPs in one clade, such as SP 4-6-7 and 4-6-7-12 representing all isolates in clade B; and (iv) many SPs across many clades, such as SPs 4-5 and 4-5-10-11 representing isolates in clades H through K.

The distribution of SPs were also assessed as they were represented among isolates with various PFGE subtypes. From the previously mentioned survey[27] the isolates were separated into 12 PFGE subtypes (Table 3.1). Upon the first pass, the isolates, mostly maintained a one clade to one PFGE to one SP relationship. As was described above, there were four distinct relationships (Table 3.1): (i) One PFGE subtype to one SP, such as Clade A, which is the only time SP9 was observed with CU-294-321; (ii) One PFGE subtype to multiple SPs were observed in isolates of Clade B, which isolates are of the subtype CU-57-267, but could contain SPs 4-6-7 or 4-6-7-12; (iii) Many PFGE subtypes to one SP was observed in Clades N through V, which all contain SP

1; (iv) Finally, many PFGE subtypes to many SPs were seen in Clades G through K, which all contain SP 4-5 and 4-5-10-11, but have 3 unique PFGE subtypes associated with those clades. Overall, the limited variability observed in SPs, within previously established SNP or PFGE-based subtypes, suggest that these CRISPR arrays are more likely vertically transmitted through mechanisms consistent with the transmission of SNPs and PFGE-based subtypes, rather than suggesting novel CRISPR spacers are being acquired or lost over the course of the original sampling study.

Two other features of the SP distribution are worth noting. First, these results show that when there was more than one SP represented among isolates of a given clade, often these SPs contained a common SP between them. For example, Clade C isolates contain SPs 7-12, 7-19, and 7-19-20, all three SPs contains SP 7. This SP was called a "backbone" SP. SPs 1, 4-5, and 7 are examples of backbone SPs.

Second, these results display a case in which SP 1-10-11, seen in Clade M (Table 3.1), is a blend of SPs observed in other clades. Specifically, SPs 10 and 11 are only seen in Clades H through K and SP 1 is only observed in Clades N through V. This blend of SPs is consistent with the phylogenetic clade relationships identified in the WGS study of the isolates.[27] From that study, those Clade M isolates are phylogenetically more closely related to Clades N through V than to isolates from Clades H to K. Isolates in Clade M are the only time that this blend of SPs was observed and bridged SPs between two groups of isolates.

There are two other interesting observations within this data. Firstly, two SPs, SPs 3 & 4, which only contained one spacer in each, varied from each other by one SNP. Due to the previously defined emphasis on spacer sequences being exact matches, these two separate SPs were considered as individual SPs, rather than the same. The second point is that there was one

single isolate in Clade A that contained 21/28 of the original spacers in SP 9. Not only was it missing 7 spacers, but it was missing 7 spacers from the middle of the SP, rather than at an end, suggesting that the array lost those spacers. Since no other isolates were missing spacers from within a CRISPR array, and that this isolate still contained a majority of SP 9, this isolate was labelled as SP 9^{r}. While not an entirely different SP, it seemed worthy of being separated from other isolates of SP 9.

3.4.4 CRISPR-based Subtyping does not Improve Discrimination of Isolates by Store

Isolates with the same SP were recovered from multiple retail deli stores (Table 3.2). In these data, there are still the same correspondence relationships of one store to one SP type (SP 15-16-17 and Store 19), many stores to one SP (SP 3 to Stores 4, 7, 10, 22, 23, 24, etc.), and many stores to many SPs (SP 1, SP 3, SP 4-5, SP 4-5-10-11, SP 4-6-7, and SP 4-6-7-12 all come from Stores 4, 7, 10, 16, 23, etc.). Therefore, it does not seem that SPs can discriminate based on isolation location, i.e. store.

3.4.5 CRISPR-based Subtyping does not Improve Discrimination of Persistent and Sporadic Isolates

Isolates were then categorized by their persistence or sporadic status, to compare CRISPR SPs within clades that represent putative persistent strains (Table 3.3). Isolates in some clades presented only one SP (e.g. Clades N-V, Table 3.3); it is clear that CRISPR SPs did not differentiate the persistent or sporadic isolates within those clade, or between clades. In other cases, there were isolates from clades with multiple SPs represented (e.g. Clade C), but there were no isolates of sporadic status. However, when a clade had multiple SPs and isolates of persistent and sporadic status, isolates did not vary in SPs. For example, Clades R through V, excluding U,

all had persistent and sporadic isolates. They also all had SP 1, persistent or sporadic. Therefore, the data does not suggest that CRISPR SPs improve identification of persistent strains.

3.4.6 Simpson's Index of Discrimination

The previous methods of subtyping, WGS SNP-based clades and PFGE, gave indices of 0.94 and 0.82, respectively, when applied to these isolates. The CRISPR-based SP subtyping, reported here, gave an index of 0.77. The lower index for CRISPR SPs than for both WGS SNPs and PFGE supports that SPs are not an improvement over PFGE or WGS-SNP based subtyping.

3.5 Discussion

Overall, this work was to apply CRISPR-based subtyping to food-associated *L. monocytogenes* isolates identified CRISPR spacers within nearly all (174/175) analyzed *L. monocytogenes* genomes. Also, *L. monocytogenes* CRISPR spacers are highly conserved across phylogenetic clades. Apart from SPs 3 and 4, which were different by one SNP, the other SPs all contain unique, identical spacers. Finally, CRISPR subtyping does not seem to differentiate persistent and sporadic *L. monocytogenes* isolates.

To the authors' knowledge, no other researchers have reported CRISPR subtyping methods for *L. monocytogenes* or to differentiate persistence with CRISPRs. One related study has focused on characterizing CRISPR loci by amplifying whole loci with PCR.[11] Di et al. did report on the spacer sequences found within their isolates, but as supplemental information. Another study has characterized the CRISPR-cas9 systems within *L. monocytogenes*.[22] Neither paper has reported on *L. monocytogenes* CRISPRs, their ability to subtype, and their relationship to persistence. Thus, the authors believe that they are the first to report on CRISPRs as a subtyping method for *L. monocytogenes* and explore the ability of CRISPRs to discriminate persistent strains.

3.5.1 CRISPRs in *L. monocytogenes* vary mainly as whole arrays

CRISPR-based subtyping of *Salmonella* species has been previously reported.[26] Shariat et al. brings forward how *Salmonella* CRISPR arrays differed based on SNPs within spacers. These results contrast how *Salmonella* and other microbes have been reported to gain or lose individual spacers, and have similar spacers that vary by individual SNPs.[25] Within the data, there were the two SPs that varied by one SNP (SPs 3 & 4), but did not have additions/deletions of entire other sequences. The only time that there were deletions was in SP 9^{r}, where the spacers were only 21/28 of the original SP 9. Instead, addition/deletion of whole arrays rather than individual spacers were observed more. Isolates could vary within clades, as they could have a different number of CRISPR arrays, but still be related to one another based on their core genomes. Further research is needed to determine if whole arrays are actively being acquired or lost as units, if the observed patterns are artifacts of the bioinformatic analysis, or if differences in array presence are due to past events that have been vertically transmitted.

3.5.2 CRISPR-based subtyping is possible for *L. monocytogenes*

CRISPR-based subtyping has become practical due to next generation sequencing and new bioinformatic analyses and tools. In some cases, CRISPR analyses require less time and money than subtyping by WGS.[24] However, not all bacteria contain CRISPRs, roughly 48% of all bacteria[14], and thus not all bacteria are optimal candidates for CRISPR-based subtyping. However, for *L. monocytogenes*, according to this data, CRISPR-based subtyping does seem possible. While it was not the main focus of this paper, the 174 CRISPR-containing isolates can form 13 groups arranged by SPs (Table S3.1). However, like PFGE, these groupings do not show phylogenetic relationships. This method turned out to be not as discriminatory as WGS-SNP subtyping and is very similar in discrimination as PFGE subtyping. Therefore CRISPR-based subtyping does not

provide improved clarity or resolution over other methods. As a whole, these results help understand the capacity to subtype *L. monocytogenes* using CRISPRs.

3.6 Conclusion for *L. monocytogenes* and subtyping persistence by CRISPR spacers

An honest assessment of these results suggests that CRISPR spacers are not useful to improve the discrimination of persistent and sporadic *L. monocytogenes* strains beyond what could already be provided by WGS-based SNP analysis (or PFGE). With that, not all CRISPRs are appropriate for molecular subtyping.[24] This suggests that the underlying assumption that persistent strains may acquire spacers during the time-period of persistence within a food-associated environment is also not supported by these data. While it is possible that the duration of the persistence events (<1 year) is not long enough for acquisition to occur, if longer persistence times are necessary for this subtyping method, then it is not relevant as an identification tool to the food industry.

There have been a variety of recent papers that state CRISPRs may have different purposes than just a defense mechanism.[6;28;29] For example, Type II CRISPR-cas genes have been associated with pathogens and pathogenicity.[20] Two such concepts for what the role CRISPRs play are: (i) Can *L. monocytogenes* acquire new spacers and incorporate them into isolates' genomes? If *L. monocytogenes* cannot acquire new spacers, (ii) what is the purpose or role of CRISPRs in this and other food-associated pathogens?

Another scenario is that the CRISPR systems in *L. monocytogenes* no longer function as an adaptive immune system, as was recently suggested for *Salmonella*.[25] If this last explanation is true, then it is worth noting that this data of CRISPR spacers showed a remarkable degree of conservation, only one SNP difference between a pair of SPs, and no other differences within a single SP, suggesting they may serve some other function.

3.7 Tables

Table 3.1 – Comparison of subtyping by previously assigned WGS-based clades and PFGE subtypes to CRISPR Spacer Patterns.

WGS-based Clade	PFGE Pattern Type	CRISPR Spacer Pattern (no. isolates)
A	CU-294-321	**9** (10), **9**[r](1)
B	CU-57-267	**4-6-7** (5), **4-6-7-12** (7)
C	CU-55-266	**7-12** (1), **7-19** (2), **7-19-20** (1)
D	CU-SNP2	**3** (20)
E	CU-258-69	**3** (2)
F	CU-258-69	**-**[n] (1), **3** (15)
G	CU-262-318	**4-5** (4)
H	CU-SNP3	**4-5** (1), **4-5-10-11** (1)
I	CU-SNP3	**4-5-10-11** (16)
J	CU-262-334	**4-5-10-11** (2)
K	CU-SNP3	**4-5** (4), **4-5-10-11** (2)
M	CU-11-282	**1-10-11** (4)
N	CU-11-326	**1** (3)
O	CU-8-340	**1** (5)
P	CU-SNP1	**1** (4)
Q	CU-SNP1	**1** (8)
R	CU-SNP1	**1** (7)
S	CU-SNP1	**1** (11)
T	CU-SNP1	**1** (12)
U	CU-SNP1	**1** (10)
V	CU-SNP1	**1** (12)
X	CU-296-330	**21-22-23** (2)
Y	CU-182-173	**15-16-17** (2)

[r] - isolate genome that contains 21/28 spacers from SP 9, either as in SP 9 or as reverse compliments.
[n] – No spacers were identified in this isolate's genome

Table 3.2 – Locations (stores) of isolates organized by phylogenetic clades and Spacer Patterns

WGS-based Clade	Spacer Pattern	Store (No. of Isolates)
A		
	9	2 (**1**), 4 (**1**), 7 (**1**), 10 (**3**), 16 (**2**), 21 (**2**)
	9[r]	12 (**1**)
B		
	4-6-7	4 (**3**), 7 (**1**), 23 (**1**)
	4-6-7-12	10 (**3**), 16 (**2**), 23 (**1**), 28 (**1**)
C		
	7-12	2 (**1**)
	7-19	2 (**2**)
	7-19-20	2 (**1**)
D		
	3	10 (**1**), 21 (**1**), 23 (**18**)
E		
	3	23 (**2**)
F		
	3	4 (**1**), 7 (**8**), 10 (**1**), 22 (**1**), 24 (**2**), 27[p] (**1**), 28[p] (**1**)
	-[n]	7[p] (**1**)
G		
	4-5	21 (**3**), 26 (**1**)
H		
	4-5	29 (**1**)
	4-5-10-11	29 (**1**)
I		
	4-5	21 (**1**), 23 (**1**), 28[p] (**6**)
	4-5-10-11	23 (**1**), 24 (**1**), 28[p] (**6**)
J		
	4-5-10-11	23 (**2**)
K		
	4-5	2 (**4**)
	4-5-10-11	2 (**2**)
M		
	1-10-11	13 (**4**)
N		
	1	8 (**2**), 11 (**1**)
O		
	1	8 (**5**)
P		
	1	13 (**2**), 23 (**1**), 29 (**1**)
Q		
	1	12[p] (**1**), 17[p] (**1**), 21[p] (**6**)
R		
	1	7[p] (**1**), 16[p] (**5**), 28 (**1**)
S		
	1	16[p] (**1**), 24[p] (**10**)
T		
	1	10[p] (**10**), 23 (**1**), 25[p] (**1**)
U		
	1	1 (**10**)
V		
	1	2[p] (**1**), 3[p] (**1**), 8[p] (**1**), 11[p] (**1**), 18[p] (**3**), 26[p] (**1**), 29[p] (**4**)
X		
	21-22-23	23 (**2**)
Y		
	15-16-17	19 (**2**)

[r] - Isolate genome that contains 21/28 spacers from SP 9, either as in SP 9 or as reverse compliments

[n] - No spacers were identified in this isolate's genome

[p] – Isolates from these stores are putatively persistent

Table 3.3 - Comparison of Spacer Patterns in persistent and sporadic isolates of the same clade

Clade	Spacer Pattern	Persistent	Sporadic	Total No.
A				
	9	7	3	10
	9^{r}	0	1	1
B				
	4-6-7	4	1	5
	4-6-7-12	6	1	7
C				
	7-12	1	0	1
	7-19	2	0	2
	7-19-20	1	0	1
D				
	3	19	1	20
E				
	3	2	0	2
F				
	3	11	4	15
	$-^{n}$	1	0	1
G				
	4-5	3	1	4
H				
	4-5	1	0	1
	4-5-10-11	1	0	1
I				
	4-5	7	1	8
	4-5-10-11	7	1	8
J				
	4-5-10-11	2	0	2
K				
	4-5	4	0	4
	4-5-10-11	2	0	2
M				
	1-10-11	4	0	4
N				
	1	2	1	3
O				
	1	5	0	5
P				
	1	3	1	4
Q				
	1	6	2	8
R				
	1	6	1	7
S				
	1	10	1	11
T				
	1	10	2	12
U				
	1	10	0	10
V				
	1	6	6	12
X				
	21-22-23	2	0	2
Y				
	15-16-17	2	0	2
Total No.		147	28	175

3.8 Figures

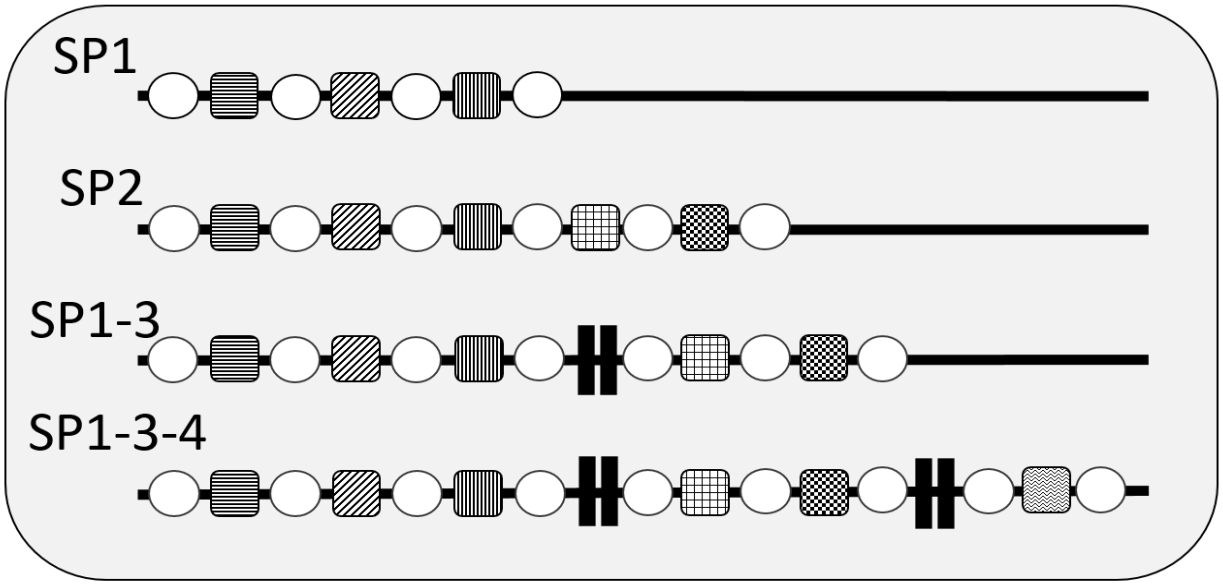

Figure 3.1. Four different cases that illustrate Spacer Patterns (SPs) assignment across the 175 isolates. Boxes with the same pattern are representative of the same spacer. Case 1 (SP 1): A Spacer Pattern with only one CRISPR array of three spacers (**boxes**) between four direct repeats (**circles**). Case 2 (SP 2): A different SP is assigned to this single CRISPR array with five spacers, even though some spacers are shared with SP 1. Case 3 (SP 1-3): Two CRISPR arrays, which are separated by genomic distance (**II**) for which separate spacers patterns are assigned to each array (and separated by the '-'); because the first array matches SP 1, that number is recycled. Case 4 (SP 1-3-4): Three CRISPR arrays, two previously identified and one novel, containing a variety of spacers in each pattern.

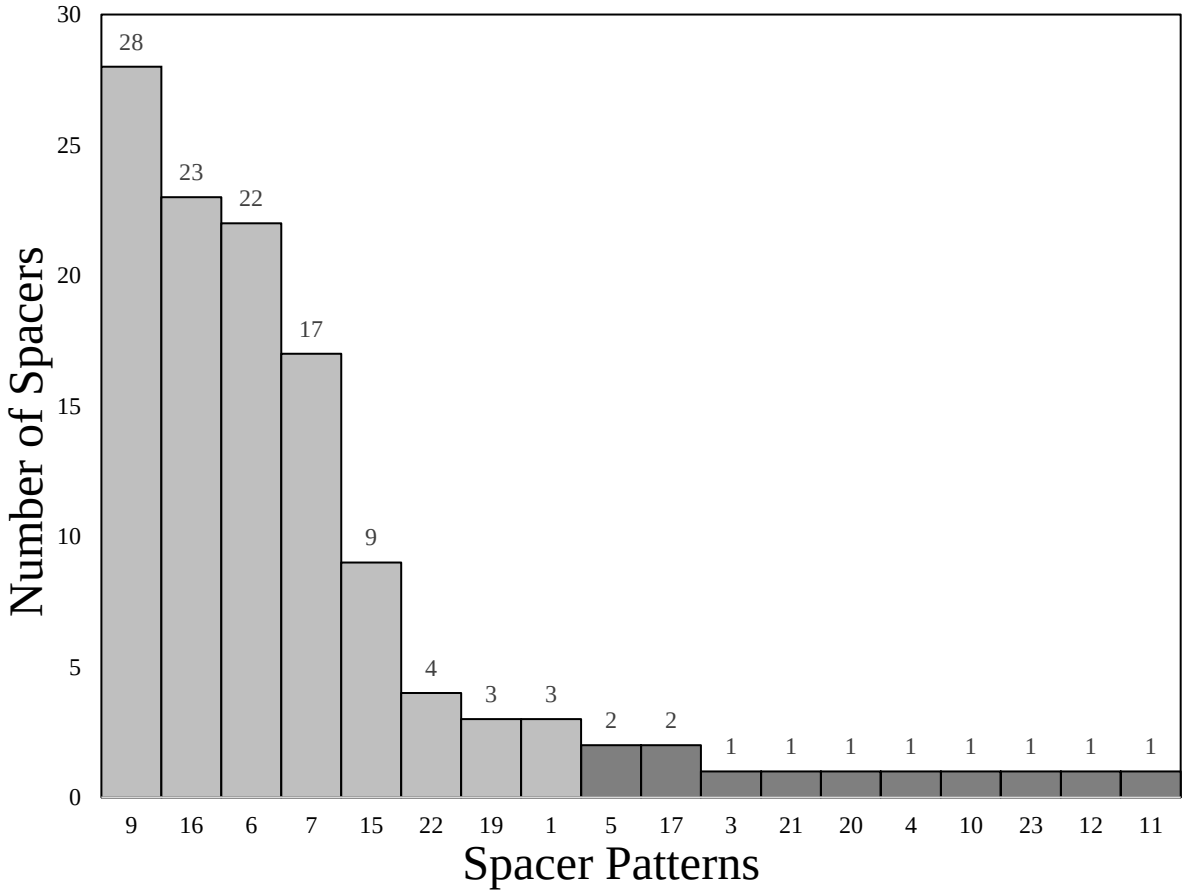

Figure 3.2 - Number of individual CRISPR spacers in every unique Spacer Pattern (SP). Any SP that had less than three spacers in one array were classified as questionable spacers by the CRISPR finder algorithm and are dark gray in this histogram.

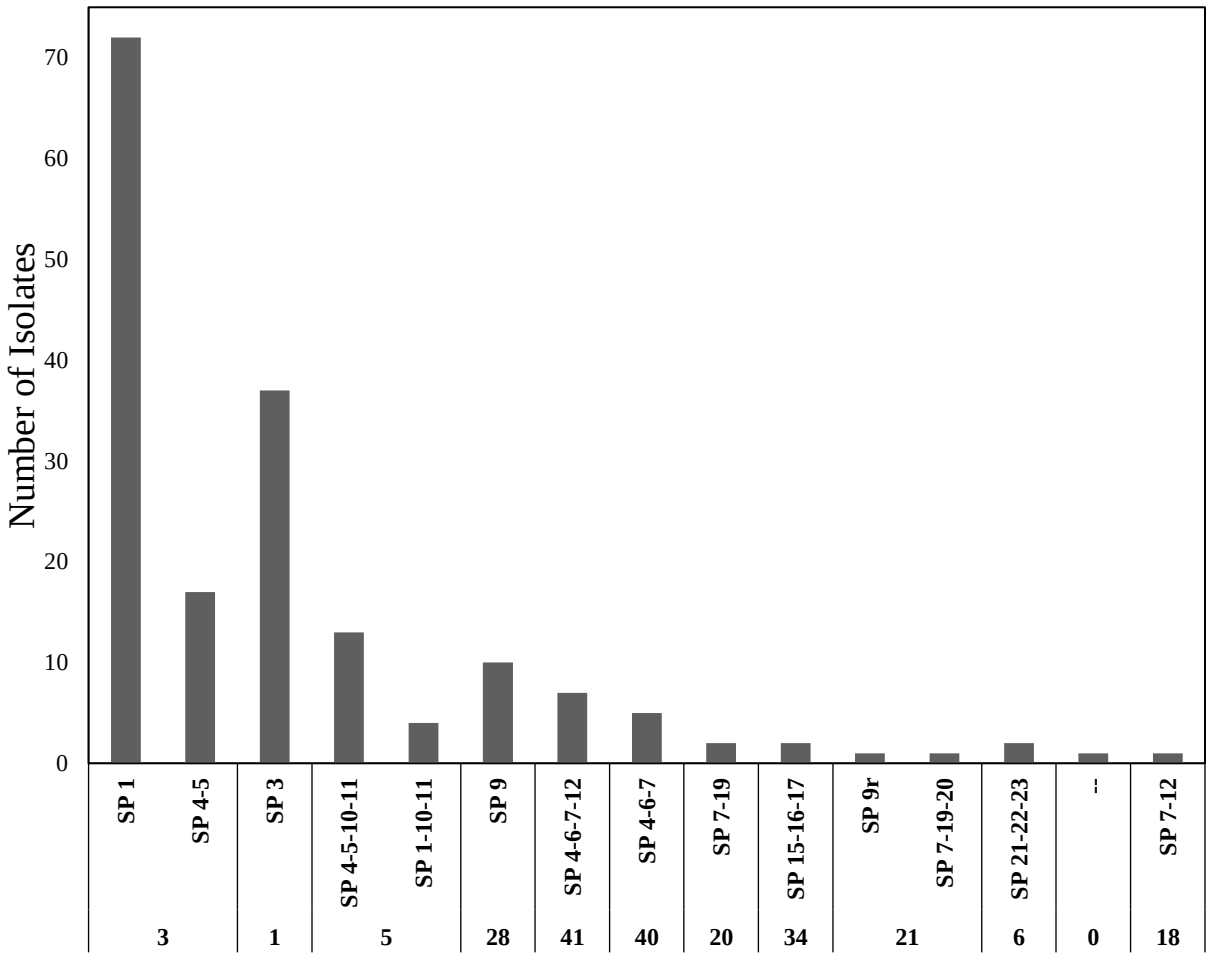

Figure 3.3 - Number of isolates with a given number of individual CRISPR spacers. (-) represents multiple Spacer Patterns (SPs) found within an isolate. X-axis displays SPs grouped by the number of spacers in each SP.

3.9 References

(**1**) Barrangou, R., Briczinski, E. P., Traeger, L. L., Loquasto, J. R., Richards, M., Horvath, P., et al. (2009). Comparison of the complete genome sequences of *Bifidobacterium animalis* subsp. *lactis* DSM 10140 and Bl-04. *J Bacteriol, 191*(13), 4144-4151. doi:10.1128/JB.00155-09

(**2**) Barrangou, R., & Dudley, E. G. (2016). CRISPR-based typing and next-generation tracking technologies. *Annu Rev Food Sci Technol, 7*, 395-411. doi:10.1146/annurev-food-022814-015729

(**3**) Bhaya, D., Davison, M., & Barrangou, R. (2011). CRISPR-cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation. *Annu Rev Genet, 45*, 273-297. doi:10.1146/annurev-genet-110410-132430

(**4**) Bland, C., Ramsey, T. L., Sabree, F., Lowe, M., Brown, K., Kyrpides, N. C., et al. (2007). CRISPR recognition tool (CRT): a tool for automatic detection of clustered regularly interspaced palindromic repeats. *BMC Bioinforma, 8*. doi:10.1186/1471-2105-8-209

(**5**) Briner, A. E., & Barrangou, R. (2014). *Lactobacillus buchneri* genotyping on the basis of clustered regularly interspaced short palindromic repeat (CRISPR) locus diversity. *Appl Environ Microbiol, 80*(3), 994-1001. doi:10.1128/AEM.03015-13

(**6**) Briner, A. E., & Barrangou, R. (2016). Deciphering and shaping bacterial diversity through CRISPR. *Curr Opin Microbiol, 31*, 101-108. doi:10.1016/j.mib.2016.03.006

(**7**) Briner, A. E., Lugli, G. A., Milani, C., Duranti, S., Turroni, F., Gueimonde, M., et al. (2015). Occurrence and diversity of CRISPR-cas systems in the genus *Bifidobacterium*. *PLoS One, 10*(7). doi:10.1371/journal.pone.0133661

(**8**) Buchanan, R. L., Gorris, L. G. M., Hayman, M. M., Jackson, T. C., & Whiting, R. C. (2017). A review of *Listeria monocytogenes*: An update on outbreaks, virulence, dose-response, ecology, and risk assessments. *Food Control, 75*, 1-13. doi:10.1016/j.foodcont.2016.12.016

(**9**) CDC. (2017). *People at Risk*. Retrieved from Center for Disease Control: https://www.cdc.gov/listeria/risk.html

(**10**) CFSAN. (2003). *Quantitative assessment of the relative risk to public health from foodborne* Listeria monocytogenes *among selected categories of ready-to-eat foods*. Retrieved from https://www.fda.gov/downloads/food/foodscienceresearch/ucm197329.pdf

(**11**) Di, H., Ye, L., Yan, H., Meng, H., Yamasak, S., & Shi, L. (2014). Comparative analysis of CRISPR loci in different *Listeria monocytogenes* lineages. *Biochem Biophys Res Commun, 454*(3), 399-403. doi:10.1016/j.bbrc.2014.10.018

(**12**) Ferreira, V., Wiedmann, M., Teixeira, P., & Stasiewicz, M. J. (2014). *Listeria monocytogenes* persistence in food-associated environments: epidemiology, strain characteristics, and implications for public health. *J Food Prot, 77*(1), 150-170. doi:10.4315/0362-028X.JFP-13-150

(**13**) FSIS. (2010). *FSIS comparative risk assessment for* Listeria monocytogenes *in ready-to-eat meat and poultry deli meats report*. Retrieved from https://www.fsis.usda.gov/shared/PDF/Comparative_RA_Lm_Report_May2010.pdf

(**14**) Grissa, I., Vergnaud, G., & Pourcel, C. (2007). The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. *BMC Bioinformatics, 8*(1). doi:10.1186/1471-2105-8-172

(**15**) Grissa, I., Vergnaud, G., & Pourcel, C. (2007). CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res, 35*(Web Server issue), W52-57. doi:10.1093/nar/gkm360

(**16**) Horvath, P., Romero, D. A., Coute-Monvoisin, A. C., Richards, M., Deveau, H., Moineau, S., et al. (2008). Diversity, activity, and evolution of CRISPR loci in *Streptococcus thermophilus*. *J Bacteriol, 190*(4), 1401-1412. doi:10.1128/JB.01415-07

(**17**) Hunter, P. R., & Gaston, M. A. (1988). Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity. *J Clin Microbiol, 26*(11), 2465-2466.

(**18**) Interagency Food Safety Analytics Collaboration. (2017). *Foodborne illness source attribution estimates for 2013 for* Salmonella, Escherichia coli *O157,* Listeria monocytogenes, *and* Campylobacter *using multi-year outbreak surveillance data*. Retrieved from United States, G.A. and D.C.:

(**19**) Jadhav, S., Bhave, M., & Palombo, E. A. (2012). Methods used for the detection and subtyping of *Listeria monocytogenes*. *J Microbiol Methods, 88*(3), 327-341. doi:10.1016/j.mimet.2012.01.002

(**20**) Louwen, R., Staals, R. H., Endtz, H. P., van Baarlen, P., & van der Oost, J. (2014). The role of CRISPR-Cas systems in virulence of pathogenic bacteria. *Microbiol Mol Biol Rev, 78*(1), 74-88. doi:10.1128/MMBR.00039-13

(**21**) Moura, A., Criscuolo, A., Pouseele, H., Maury, M. M., Leclercq, A., Tarr, C., et al. (2016). Whole genome-based population biology and epidemiological surveillance of *Listeria monocytogenes*. *Nat Microbiol, 2*, 16185. doi:10.1038/nmicrobiol.2016.185

(**22**) Rauch, B. J., Silvis, M. R., Hultquist, J. F., Waters, C. S., McGregor, M. J., Krogan, N. J., et al. (2016). Inhibition of CRISPR-cas9 with bacteriophage proteins. *Cell, 168*(1), 150-158.e110. doi:10.1016/j.cell.2016.12.009

(**23**) Scallan, E., Hoekstra, R. M., Angulo, F. J., Tauxe, R. V., Widdowson, M. A., Roy, S. L., et al. (2011). Foodborne illness acquired in the United States--major pathogens. *Emerg Infect Dis, 17*(1), 7-15. doi:10.3201/eid1701.P11101

(**24**) Shariat, N., & Dudley, E. G. (2014). CRISPRs: molecular signatures used for pathogen subtyping. *Appl Environ Microbiol, 80*(2), 430-439. doi:10.1128/Aem.02790-13

(**25**) Shariat, N., Timme, R. E., Pettengill, J. B., Barrangou, R., & Dudley, E. G. (2015). Characterization and evolution of *Salmonella* CRISPR-cas systems. *Microbiology, 161*, 374-386. doi:10.1099/mic.0.000005

(**26**) Simmons, C., Stasiewicz, M. J., Wright, E., Warchocki, S., Roof, S., Kause, J. R., et al. (2014). *Listeria monocytogenes* and *Listeria* spp. contamination patterns in retail delicatessen establishments in three U.S. states. *J Food Prot, 77*(11), 1929-1939. doi:10.4315/0362-028X.JFP-14-183

(**27**) Stasiewicz, M. J., Oliver, H. F., Wiedmann, M., & den Bakker, H. C. (2015). Whole-genome sequencing allows for improved identification of persistent *Listeria monocytogenes* in food-associated environments. *Appl Environ Microbiol, 81*(17), 6024-6037. doi:10.1128/AEM.01049-15

(**28**) Stern, A., Keren, L., Wurtzel, O., Amitai, G., & Sorek, R. (2010). Self-targeting by CRISPR: gene regulation or autoimmunity? *Trends Genet, 26*. doi:10.1016/j.tig.2010.05.008

(**29**) Stout, E., Klaenhammer, T., & Barrangou, R. (2017). CRISPR-cas technologies and applications in food bacteria. *Annu Rev Food Sci Technol*, *8*, 413-437. doi:10.1146/annurev-food-072816-024723

Chapter 4: Persistent and Sporadic Isolates of *Listeria monocytogenes* from Retail Do Not Differ in Growth in the Presence of Food-Associated Stress or Carbon Sources

4.1 Abstract

The foodborne pathogen, *Listeria monocytogenes*, causes the lethal disease listerosis. Within food-associated environments, *L. monocytogenes* can persist for long periods and increase the risk of contamination by its presence in processing and manufacturing facilities or other food-associated environments. Most published literature has looked at *L. monocytogenes*' biofilm formation and sanitizer resistance, but have not explored other phenotypic responses to food-associated stresses or carbon sources. Therefore, this study aimed to explore if persistent *L. monocytogenes* show adaptations to stresses by growth rate and ability to grow. It was hypothesized that isolates of persistent strains have increased fitness compared to closely related sporadic isolates in stress-inducing, food-associated environments and food-associated carbon sources. To test this hypothesis, 95 isolates (74 persistent and 21 sporadic), from a previous longitudinal study, were grown under different (i) stress conditions: salt levels (0, 5%, and 10% NaCl), varying pHs (5.2, 7.2, and 9.2), and sanitizer conditions (benzalkonium chloride [BAC], 0, 2, and 5 µg/mL) and (ii) carbon sources: 25 mM glucose, cellobiose, glycogen, fructose, and sucrose. Growth rate and the ability to grow of 95 isolates were tested, at 37°C, using high-throughput, optical-density growth curves. Throughout all isolates, all stress conditions reduced growth rates compared to control ($p < 0.05$, all comparisons). In addition, growth varied by carbon source, with more strains likely able to consistently grow using cellobiose ($p = 0.052$) than the control, glucose, and fewer strains able to grow using glycogen ($p = 0.02$), lactose ($p = 2.2x10^{-16}$), and sucrose ($p = 2.2x10^{-16}$). With that, there was at least one strain that was able to grow in each of the carbon sources. Yet, this hypothesis must be rejected, as there was not a significant

difference in growth rate or ability to grow for isolates of persistent strains compared to sporadic isolates for the stress conditions or carbon sources utilization tests. Therefore, these data suggest that persistence is not determined by a phenotype unique to persistent strains isolated from these retail deli environments.

4.2 Introduction

4.2.1 *Listeria monocytogenes*

Listeria monocytogenes is a foodborne pathogen that causes listeriosis and is estimated to cause 1,600 illnesses and 260 deaths per year in the US.[6] *L. monocytogenes* is found commonly in natural and rural environments, where it can contaminate raw food materials directly and also be transferred into processing facilities.[31] Local environments can lead to the formation of niches or places that are difficult for sanitizers to reach.[22] Usually, water and organic materials are readily available, which allow the *L. monocytogenes* to survive and multiply.[3;9] Once in the niche, the pathogen can form a biofilm, which can make it even harder to kill the strain, as the disinfectant cannot reach bactericidal concentrations.[26;33]

Pair biofilm formation with *L. monocytogenes* inherent resistances and this pathogen is more durable to numerous environmental stresses than most other foodborne pathogens. *L. monocytogenes* can survive and grow at low pHs (>4.7) and high pHs (<9.2)[12], high salt concentrations (10% w/v)[35], and in small amounts of antimicrobial solutions or sanitizers (amount varies per sanitizer).[13;39] Resistance to these food-associated stresses likely contribute to its survival in processing environments (where pH, osmotic, and sanitizer stresses are common), and survival represents a risk for cross-contaminating food products produced in those environments.

4.2.2 Persistence

L. monocytogenes can persist in food processing plants for years to decades.[12;17] Persistent strains represent a risk of continual contamination within a manufacturing or processing environment. For this paper, persistence is defined as the continued presence of a clonal population of bacteria, over time, at a specific location. Many researchers believe that niches within the food environment[4], biofilm formation[29], and *L. monocytogenes*' inherent resistance to sanitizers[18] and other food-associated stresses[19], contribute to strain persistence. While the contribution of niches and biofilm formation have been discussed as components of comprehensive reviews on *L. monocytogenes* persistence[5;12], this study will focus on a gap in the literature on food-associated stress response and metabolic phenotypes.

4.2.3 Stress Response Phenotypes of Persistent Strains

There are multiple publications comparing persistent and sporadic strains for differences in responses to biofilm and sanitizer conditions. However, there is only a handful of reports on salt and acidity and persistent *L. monocytogenes* strains.[4] One recent paper has found that persistent strains from a cheese-processing facility were better adapted than sporadic strains to grow in 2.5%, 4%, and 8% NaCl and acidic, pH 5, conditions.[24] However, an earlier paper compared acid tolerance in 17 persistent to 23 non-persistent strains from three meat-processing plants.[22] No difference was observed in log reduction after the acid stress, but the authors noted that two non-persistent strains were the most acid-sensitive strains. Another paper[32] found that 14 persistent isolates from two separate pork-processing plants had BAC resistant genes. These strains related back to the MLST sequence type 121 (ST121)[32], which has been known to be both persistent and

contains the BAC resistant transposon Tn*6188*.[31;32] However, not all persistent strains contain this transposon or BAC resistant genes. Other authors have found no relationship between persistent strains and resistance of commonly used sanitizers between isolates of persistent and sporadic strains.[12] Overall, there is a lack of consistency in the literature on whether persistent strains are more resistant to particular stress conditions compared to sporadic isolates from similar sources.

4.2.4 Classification of Persistent Strains and Relationship to Phenotype work

One potential explanation for inconsistency in phenotyping study results is that each study has used different empirical rules to identify persistent and sporadic comparison groups. Persistent strains are typically empirically identified by isolating, on different dates, *L. monocytogenes* that are found to be indistinguishable, or otherwise of the same stain, by the best subtyping method available to the researchers. In particularly, most literature[2;4;21;23;34;37] has used pulse-field gel electrophoresis (PFGE) for subtyping. While PFGE has been the gold standard for assessing genetic relationships between *L. monocytogenes* isolates, this technique has recently been replaced whole genome sequencing (WGS) which has improved listeriosis outbreak investigation.[6;30] As costs lower, WGS is becoming a viable alternative for distinguishing isolates and investigating contamination in food processing plants.[11;15] WGS has also been used to improve the differentiation of persistent and sporadic isolates from retail delis in multiple regions of the US.[40] Yet, the authors' are unaware of work to compare phenotypes of persistent and sporadic strains using isolates classified by modern, WGS-based methods.

4.2.5 Motivation and Hypothesis

This study wanted to compare phenotypes relevant to food-associated environmental survival between persistent and sporadic strains of *L. monocytogenes* classified by the best available subtyping methods. To do so, isolates were from a previous, longitudinal study of 30

retail delis across the US that had been previously subtyped by PFGE[36] and WGS-based methods, and persistent strains were identified by peer-reviewed SNP-difference metrics.[38] With reliably classified isolates representing 31 individual, putative persistence events, and closely related sporadic strains, a panel of strains was assembled. The panel had the statistical power to rigorously test if persistent and sporadic strains differ in growth responses to osmotic pressure, acidic and alkali conditions, sanitizers, and food-associated carbon sources. The hypothesis was that if persistent strains have fitness advantages over closely related, but sporadic strains, they would show significantly greater growth rates, or ability to grow, in the presence of these food-associated stress conditions and carbon sources.

4.3 Materials and Methods

4.3.1 Isolate Selection

In this study, 95 isolates (74 [77%] persistent and 21 [23%] sporadic) were analyzed, from a chosen study that used WGS based phylogenetics to identify persistent strains from retail delis.[40] Critical metadata for all isolates selected for sequencing are found in Table S4.1 and additional metadata are stored in the database www.foodmicrobetracker.com.

While the referenced study analyzed 175 isolates, the authors selected 95 of those 175 as this number is appropriate for high-throughput analysis in microtiter plates. The 95 chosen isolates were systematically selected. First, only isolates that were associated with statistically-significant WGS SNP evidence for persistence were selected.[40] Second, the panel included sporadic isolates closely related to the persistence events. Third, only isolates physically available from Dr. Oliver's lab at Purdue University were acquired. This consolidated 175 isolates to 105 candidate isolates for the phenotyping panel. Of the 105 candidates, some putative persistence events were more represented than others were, so 10 randomly selected isolates were discarded from events that

already had sufficient representation. Isolates were maintained at -80ºC in 12.5% v/v glycerol-brain-heart infusion (BHI) media in 96-well microplate format (Corning Clear Polystyrene 96-Well Microplates 360µL, Corning, Tewksbury, MA). Before being assayed in the Bioscreen C, isolates were resuscitated from frozen stocks by pre-growth in control media (BHI or DM Glucose, described below) at 37ºC, for optimal growth, for 18 to 24 hours (Overnight, O/N) with shaking at 150 rpm, again in the 96-well microplates.

4.3.2 Treatment Media

There were multiple treatment media used in this project (Table 4.1). This project used nutritive media to create stress conditions and chemically defined media to assay growth in different food-associated carbon sources. The control medias were either BHI (Sigma-Aldrich, St. Louis, MO) or a chemically defined media (DM), specifically formulated for *Listeria* species[1] which uses glucose as a control carbon source. As the tested *L. monocytogenes* isolates come from retail delis, three different food-associated stresses were tested, as they are likely present in delis – osmotic pressure, acidic and alkali pH, and sanitizers. Therefore, the following medias were made: BHI with 5% or 10% w/v NaCl; BHI at pH 5.2 or 9.2 BHI (adjusted with 3M HCl or 3M NaOH); BHI with 2 or 5µg/mL benzalkonium chloride (BAC), a quaternary ammonium compound (QAC) and common industry sanitizer.[29]

Food-associated carbon sources were also assayed, such as glucose (control), cellobiose, fructose, glycogen, lactose, and sucrose. DM was used to limit the carbohydrate sources compared to, e.g. those present in BHI, and test other food-associated carbon sources. The DM carbon sources were substituted at the same initial concentration (25mM) as directed in previous literature.[1] Each carbon source was chosen to represent an energy source the pathogen may encounter within a food environment. Cellobiose was for observing *L. monocytogenes*' ability to

grow on plant matter (vegetables). Fructose was representative of fruit sugars (fruits). Glycogen was representative of muscle tissue (meats). Lactose was representative of milk sugars (dairy products). Sucrose was representative of refined sugar (sweets).

4.3.3 Growth Assay

O/N cultures were transferred from a 96-well plate to a 100-well Honeycomb Bioscreen Plate (Growth Curves USA, Piscataway Township, NJ) in the treatment specific media (20µL O/N culture with 180µL of fresh, treatment media). These were then grown for 24 hours, at 37ºC, in the Bioscreen C (Growth Curves USA, Piscataway Township, NJ) Automated Growth Curve Analysis System with the Bioscreener software measuring the OD_{600} of each of the 100 wells at 15 minute intervals, and shaking at medium-intensity 15 seconds before each reading.

Growth data was analyzed using an open-source regression tool specifically adapted to fitting food microbiological growth models to OD data.[14] The tool fits a Baranyi Roberts growth curve to the OD_{600} data. Curves were only fit to data where growth was observed, which is defined as $\Delta OD_{600} > 0.1$. The plots would display the lag time, growth rate, doubling time, and ΔOD_{600} of the isolates grown in the treatments. Within the initial analysis, prominence was placed on the isolates' lag time, ΔOD, and growth rate. However, focus was shifted to the isolates' growth rates, as the lag time was inversely related to growth rate and ΔOD was directly proportional to growth rate. As each isolate was grown at a minimum of three times and a maximum of six times. For each treatment, the growth rates were averaged for the treatments.

4.3.4 Data Analysis

Growth parameter data was analyzed to compare both if strains were able to grow, and growth rate if growth was observed. As for the isolates' ability to grow, isolates were given the designations of "Growth," (G) "Variable," (VAR) or "No Growth" (NG) if they either grew in

($\Delta OD_{600} \geq 0.1$) every replicate of a treatment, grew in at least one replicate but not all, or did not grow in any of the replicates of a treatment, respectively. Significant differences were tested in the number of strains for each growth category for each treatment compared to its respective control (Control [BHI] and DM Control [DM Glucose], for stress response and carbon source utilization, respectively) using χ^2 tests (or Fisher's Exact tests if any cell had <5 counts).

To analyze growth rate data, data was aggregated across replicates by calculating the mean lag time, growth rate, and ΔOD, for each strain for each treatment where growth was observed ($\Delta OD_{600} \geq 0.1$). Then, it was tested for the effects of treatment, persistence, and their interaction on growth parameters using Analysis of Variance (ANOVA). Plotting and further statistical analysis were performed in JMP (JMP Pro 13.0.0, SAS Inc., Cary, NC).

4.4 Results and Discussion

Ninety-five *L. monocytogenes* isolates, comprised of 74 persistent and 21 sporadic isolates, were tested for their ability to grow in the presence of food-associated stress conditions and utilization of food-associated carbon sources (Table 4.1). These isolates were collected from a previous, longitudinal study that analyzed persistence based on WGS SNPs.[40]

4.4.1 Growth responses to food-associated environmental stresses are consistent with previous literature

4.4.1.1 Results. Environmental Stress

To represent isolate growth ability in the presence of osmotic, pH, and sanitizer stress conditions, isolates were counted if they were able to grow in all (Growth), some (Variable), or no (No Growth) replicates of each treatment (Table 4.2). Isolate ability to grow was not significantly different from control for the 5% NaCl and pH 5.2 & 9.2 conditions. Isolates had significantly reduced ability for growth in 10% NaCl and 2 & 5 μg/mL BAC. BAC 5μg/mL media was least

likely to support growth, with just above 50% (49/95) of isolates having variable growth and the rest not growing at all. The five isolates with variable growth at pH 5.2 and pH 9.2 were not the same isolates.

Overall, stress conditions decrease growth rate among *L. monocytogenes* strains that were able to grow (Figure 4.1, overall ANOVA treatment effect $p < 0.001$). When comparing all food-associated stress conditions, all treatment means were significantly lower than the control of normal BHI (Tukey's HSD, p-value = 0.05 threshold). The conditions of 5% NaCl, BAC 2µg/mL, and pH 9.2 were all indistinguishable (Figure 4.1), with remaining treatments showing even lower growth rates. Stress condition treatments were separated into three individual groups: salt, pH, and sanitizer. Within each group, growth rates are significantly different (i.e. BAC 2µg/mL results are significantly different from BAC 5µg/mL, and so on).

4.4.1.2 Discussion. Environmental Stress

As a species, *L. monocytogenes* is relatively resistant to many environmental stresses.[4;27] It is not surprising that all isolates were always able to grow in five percent salt, and all showed at least variable growth in 10% salt, as *L. monocytogenes* is known to grow at high salt concentrations (up to 10% w/v).[27;35] Similarly, most isolates were always able to grow in both acid (pH 5.2) and alkali (pH 9.2) conditions, and it is known *L. monocytogenes* can survive and grow at low pHs (≥ 4.7) and high pHs (≤ 9.2).[12;27] It is interesting *L. monocytogenes* isolates grew significantly faster at pH 9.2 than at pH 5.2 because each treatment deviated from neutral pH by approximately the same magnitude.

As for the BAC data, treatment with 2-5 µg/mL allowed, at best, variable growth with significantly reduced growth rates. While industry uses a variety of different sanitizers, the concentration of BAC needed for complete inhibition of growth is around 60 µg/mL[20], a level

consistent with the observation that some, but not all, strains are able to grow when exposed to lower concentration. Overall, these data align with what has been seen already in literature for treatment effects of salt, pH, and sanitizer stress on growth of *L. monocytogenes* isolates, and this study adds substantial data on strain-to-strain variability.

4.4.1.3 Results. Carbon Source Utilization

L. monocytogenes was also examined for its ability to grow on various carbon sources in chemically defined media (DM, Table 4.2). It was surprising to see the control condition data, DM Glucose, was split between consistent (n=46) and variable (n=47) growth with two isolates that never grew. This is surprising because the media was developed using glucose as the sole carbon source. Comparatively, DM Cellobiose maintained more consistent growth of isolates (n=60) than any other treatment. Only DM Cellobiose and Fructose conditions had zero No Growths; DM Lactose and Sucrose had the most No Growths. DM Glycogen, Lactose, and Sucrose were the only treatments that had significantly different growth distributions than control ($p < 0.05$ for all), all with reduced ability to support growth. Cellobiose showed a trend towards supporting more growth than control (p-value = 0.052).

Growth rates of the *L. monocytogenes* isolates did not vary as much in carbon sources (Figure 4.3) compared to stress-inducing treatments (Figure 4.1). When comparing the treatments to the control (DM Glucose), only DM Lactose and Sucrose gave statistically significant differences in growth rates (by Tukey's HSD test, p=0.05 threshold) and growth rate was reduced. These two treatments, with significantly reduced growth rates among isolates that were able to grow, are also the same treatments that are least likely to support growth (Table 4.2). One important note, in these analyses, is that the definition of growth is a given change in OD over

time. This a created a growth rate limit of detection of $\Delta OD_{600} > 0.1 / 24$ h = 0.004 ΔOD_{600} / h; therefore, any outlying strains with growth below this threshold are excluded from the analysis.

These data confirm that *L. monocytogenes* can grow in multiple food-associated carbon sources. Isolates were found to grow especially well on cellobiose and not as likely to grow on lactose or sucrose as the sole carbon source.

4.4.1.4 Discussion. Carbon Source Utilization

Defined Media and Supported Growth. The results for which carbon sources support the growth of *L. monocytogenes* are mostly consistent with previous studies in chemically defined media (DM), with this work testing a larger panel of carbon sources and strains. Most isolates were able to grow on glucose, cellobiose, fructose, and glycogen, whereas lactose and sucrose only rarely supported growth. The DM formula used in this study was a version of the formula used by Amezaga et al., the only difference was the use of different carbohydrates. Amezaga et al. stated that their media supported *L. monocytogenes* growth on glucose; however, no other carbohydrates were tested.[1] A similar DM formula, developed by Premaratne et al., supported growth on fructose, cellobiose, and a few other sugars not tested here, but not on lactose, sucrose, and other sugars not tested here; glycogen was never reported on in Premaratne et al.[36] The major differences between these two DM formulae is that Amezaga et al. had added trace metals, added α-lipoic acid in ethanol, and had different phosphate salts.[1] While both DMs supported growth, only Premaratne et al. looked at other carbon sources besides glucose. The Premaratne formula used 10.0 g/L of glucose (equivalent to 55.5 mM) and did not specifically state the concentrations of the other tested sugars.[36] Thus, it is assumed that 10.0g/L was used for all of the tested sugars. In contrast, all media in this study were formulated with 25mM of a sole carbon source.

The results presented in this study are consistent with other studies that show growth is supported by glucose, cellobiose and fructose[1;36;38], but the literature varies on if lactose and sucrose support *L. monocytogenes* growth, and glycogen has not been extensively studied. In addition, Bergey's Manual states that under aerobic conditions lactose supports growth, but sucrose does not.[27] In this study there was at least one strain (of 95 total) that consistently grew in each carbon source, and a handful of variable growth strains; it is possible that previous research reported specific carbon sources, e.g. lactose or sucrose, did not support growth due to a bias of small sample sizes. For example, Amezaga et al. tested only one unique strain and Premaratne et al. tested 5 names strains and 28 unlisted diary strains.

Furthering the point of sample size effects, in this study isolates seemed to grow more consistently on cellobiose than on the control condition of glucose, and showed a trend towards faster growth rates. Both other DM formulation papers discussed above used glucose as their main carbon source and reported consistent growth. Specifically, growth for three replicates of the common lab strain ATCC 23704[1] and unknown replicates for growth of strains Scott A, V7, CA, OH, ATCC 19115, and the other 28 dairy isolates (unlisted).[36] Given that this study tested a larger panel of isolates, these results suggest that cellobiose may be a better sole carbon source for formulating DM to support the growth of a wide range of *L. monocytogenes* isolates.

Cellobiose. There are a few intriguing implications of the possibly increased ability of cellobiose to support growth over glucose. Since cellobiose is comprised of two glucose molecules, also known as a dimer, one could assume growth on cellobiose should be similar to glucose. However, as there is a slight difference favoring cellobiose, there are at least three possible explanations for this difference. First, *L. monocytogenes* can be found in many different environments, but is most abundant within the soil as a saprotroph.[16] Thus, *L. monocytogenes*

likely has adapted to grow better using the carbohydrate that is more abundant in its natural environment.

Second, it is possible that cellobiose is more energetically favorable compared to glucose metabolism. A few studies have found that in the presence of cellobiose, the main transcriptional activator of virulence genes, *prfA*, is down-regulated.[16;25;38;41] Summarizing many authors, the pathogen favors its saprophytic lifestyle by not expressing virulence genes that would not confer a fitness benefit.[10] However, at low concentrations (5 to 7mM) glucose can also inhibit *prfA* activity.[10] The relationship between carbon catabolite repression of *prfA* activity and other signaling pathways is complex. For example, glucose is present in blood at similarly low concentrations[10] and when *L. monocytogenes* EGD-e was inoculated into donated, healthy-human blood, the virulence gene locus was highly overexpressed.[42] When *L. monocytogenes* utilizes either cellobiose or glucose, that contributes to virulence gene repression and it is not clear which carbohydrate would be more energetically favorable in DM. More importantly, evaluating the energetic differences between various carbohydrate sources, in conditions mimicking natural environments of these energy sources, could be a focus of future research.

Finally, the cellobiose treatment may have provided more gross energy simply due to formulating media on a mM basis. The implication of formulating our media on a mM basis is that there was an equal concentration of cellobiose and glucose molecules in each media. *L. monocytogenes* carbon catabolism is usually associated with phosphotransferase system (PTS) of transport[10;38], which requires expenditure of energy for uptake of carbohydrates like glucose and cellobiose.[10] In importing cellobiose, the PTS system phosphorylates the cellobiose.[10] It is then hydrolyzed into glucose and glucose-6-phosphate by the enzyme, 6-P-β-glucosidase. The resulting glucose is also phosphorylated by a glucokinase or hexokinase.[10] Since cellobiose is

effectively broken down into two phosphorylated glucoses, it may be possible that cellobiose supported more growth because it effectively became twice as much available glucose.

Glycogen. The DM data suggest some *L. monocytogenes* can grow on media with glycogen as the sole carbon source, which was not previously reported by papers developing chemically defined media. Total growth is weak, usually around a ΔOD_{600} of 0.15. Still, according to Bergey's manual of 2015, *L. monocytogenes* is known to not have any acid production from glycogen (their proxy for growth).[27] This discrepancy may be due to different methods to determine growth. The data collected here did not test or evaluate acid production from carbohydrate sources. However, *L. monocytogenes* metabolism of glycogen is suggested by a pathogenesis paper where triglycerides and glycogen storages of *Drosophila melanogaster* deplete over time when infected by *L. monocytogenes.*[8] Another caveat is that 10/95 isolates never grew in glycogen treated media, and all were from a single PFGE type, suggesting there may be sub-populations of *L. monocytogenes* that differ in glycogen utilization.

Lactose and Sucrose. While the DM data confirmed that most *L. monocytogenes* isolates are unable to grow with lactose or sucrose as the sole carbon source, there is intriguing variability in these phenotypes. Specifically, at least 1 isolate was able to consistently grow on each of these sugars, and a few more isolates showed variable growth. This variability of growth is particularly interesting for lactose, as *L. monocytogenes* can be isolated from dairy products[4], and unpasteurized dairy products have long been identified as high-risk foods for listeriosis.[7] However, the DM Lactose data suggests that *L. monocytogenes* cannot grow well on lactose, by itself, in chemically defined media. It would be interesting to compare these results to growth of these same strains on whole dairy foods, and to compare growth of *L. monocytogenes* from dairy

sources in lactose-supplemented DM, to assess if this variability is more a function of strain or media components.

4.4.2 Persistent and sporadic isolates from deli sources do not differ in food-associated stress tolerance or carbon sources utilization

To test if persist and sporadic strains differ in relevant phenotypes the growth rate and ability to grow data were reanalyzed, separating isolates by persistence status. Figures 4.2 and 4.4 are expanded versions of Figures 4.1 and 4.3, respectively, with the added persistence factor. While there were visual differences in the boxplots of mean growth rate by treatment, the means did not differ systematically or statistically between isolates of persistent and sporadic strains for any treatments ($p > 0.05$ in all cases, t-test). In addition, overall tests in stress-inducing conditions and carbon source utilization were non-significant for the persistence factor (ANOVA, $p = 0.82$ & $p = 0.22$, respectively) and the interaction of persistence and treatment (ANOVA, $p = 0.79$ & 0.92, respectively). This suggests that there is no interaction between treatment and persistence of *L. monocytogenes* in its growth rate.

Isolates' ability to grow were also examined by their persistence status (Table 4.3). None of the treatments showed a significant difference in the proportion of isolates able to grow, or not, when comparing persistent and sporadic isolates ($p > 0.05$ in all cases that were testable). From both of these assays, it is unlikely that isolates of persistent strains are able to grow better in stress inducing environments or on food-associated carbon sources than otherwise similar sporadic isolates that are also from retail delis.

4.4.3 Discussion of the differences between persistent and sporadic *L. monocytogenes* strains in their growth rates and their ability to grow

Previous reports have shown persistent and non-persistent *L. monocytogenes* isolates and how they react to varying salt concentrations, acidity conditions, and QAC concentrations.[22;24;28;32;30] Recently, a report by Magalhães et al. looked at BAC resistance, osmotic pressure, and pH conditions on growth kinetics, in 41 persistent and non-persistent isolates from three cheese processing plants classified by PFGE subtyping. They concluded that there were not significant differences in lag time or growth rate between persistent and non-persistent in 50µg/mL of BAC. The BAC data confers with this section of their report.[24] As for the other two stresses, the osmotic pressure and acid data was in conflict with Magalhães et al.'s data. While they were able to find that lag time was shorter for persistent isolates at 2.5%, 4%, 8% NaCl and pH 5, there were no significant differences between persistent and sporadic isolates under similar conditions. For persistence and acid tolerance, Lundén et al., that showed 17 persistent strains, from three meat-processing plants, had higher tolerance, less log reduction, to acidic (pH 2.4 for 2 hours) conditions than the 23 non-persistent strains.[22] Growth under sub-lethal acid stress was not tested.

There could be two possible contributions to the differing results of persistent strain growth under food-associated stresses: classification methods and small sample sizes. The Lundén study identified 34 different PFGE subtypes (which they called RFLP types), of which 12 were persistent and 22 were non-persistent. This means that their comparisons of persistent and non-persistent strains are of strains with unclear relation to one another. To determine persistence, they only labelled strains that they found 5+ times in a span of 3 months as persistent. Persistent strains isolates less frequently might have been mislabeled as non-persistent, or sporadic. In addition, these studies have relatively smaller sample sizes ($n<50$) of isolates, again with unclear phylogenetic relationships. This study uses well-defined, putatively persistent isolates that are

phylogenetically related to one another. The data presented here is a strong case for representing both persistent strains and closely related sporadic isolates. Within the full data set, it does not seem that persistent isolates of *L. monocytogenes* grow faster or have a better ability to grow than sporadic isolates. It seems that, more likely, persistent strains rely on permissive environmental conditions rather than phenotypic adaptations.[4]

4.5 Conclusion for persistence and growth rates and ability to grow of *L. monocytogenes*

This study set out to characterize different phenotypic responses of isolates from persistent and sporadic strains. Stress-inducing conditions and various carbohydrate sources have significant effects on *L. monocytogenes*' ability to grow and growth rate. However, when comparing growth between persistent and sporadic isolates, there does not seem to be difference in the ability to grow or growth rate. These results indicate that *L. monocytogenes* strains are not likely persistent due to strain specific phenotypes in food-associated environments. Rather, persistence is likely a combination of environmental conditions and factors. From this conclusion, future research should be focused on the control of persistent *L. monocytogenes* would be better focused on improving environmental-based monitoring and seek-and-destroy strategies[21] to eliminate harborage sites, which are known to contribute to persistence. While there was no shown stark difference in persistence, there was the addition to *Listeria* metabolism by finding evidence for strain-to-strain variability of *L. monocytogenes*' carbon source utilization, particularly glycogen, lactose, and sucrose.

4.6 Tables

Table 4.1: Treatment formulations for the food-associated stress and carbon source medias

Treatments	Replicates	Pre-growth media	Carbon Source	Salt (%)	pH levels	BAC[a] (µg/mL)
Stress Conditions (tested in nutritive media [BHI])						
Control	3	BHI	Dextrose	0.5	7.20	0
5% NaCl	4	BHI	Dextrose	**5**	7.20	0
10% NaCl	4	BHI	Dextrose	**10**	7.20	0
pH of 5.2	3	BHI	Dextrose	0.5	**5.20**	0
pH of 9.2	3	BHI	Dextrose	0.5	**9.20**	0
BAC 2ug/mL	6	BHI	Dextrose	0.5	7.20	**2**
BAC 5ug/mL	6	BHI	Dextrose	0.5	7.20	**5**
Carbon Source Utilization (tested in chemically defined media [DM])						
DM Control	3	DM Glucose	Glucose	0.5	6.75	0
DM Cellobiose	6	DM Glucose	**Cellobiose**	0.5	6.75	0
DM Fructose	4	DM Glucose	**Fructose**	0.5	6.75	0
DM Glycogen	6	DM Glucose	**Glycogen**	0.5	6.75	0
DM Lactose	4	DM Glucose	**Lactose**	0.5	6.75	0
DM Sucrose	6	DM Glucose	**Sucrose**	0.5	6.75	0

[a] – BAC, Benzalkonium Chloride

Table 4.2: Count of growth response for 95 *L. monocytogenes* isolates

Treatments[a]	Growth	Variable	No Growth	p-value[b,c]
Stress Conditions (tested in nutritive media [BHI])				
Control	95	0	0	--
5% NaCl BHI	95	0	0	No test
10% NaCl BHI	51	44	0	**$2.2x10^{-16}$**
pH of 5.2	90	5	0	0.059
pH of 9.2	90	5	0	0.059
BAC 2µg/mL	0	95	0	No test
BAC 5µg/mL	0	49	46	**$2.2x10^{-16}$**
Carbon Source Utilization (tested in chemically defined media [DM])				
DM Control	46	47	2	--
DM Cellobiose	60	35	0	0.052
DM Fructose	47	48	0	0.62
DM Glycogen	33	52	10	**0.020**
DM Lactose	1	11	83	**$2.2x10^{-16}$**
DM Sucrose	1	18	76	**$2.2x10^{-16}$**

[a] – See Table 4.1 for formulations of each treatment

[b] – p-values test for difference in growth response category between control and individual treatments in the same stress condition or carbon source test category

[c] – p-values were determined by Chi Square Test, or Fisher's exact test if any cells contained less than five isolates

Bolded p-values are statistically significant at $p < 0.05$

Table 4.3: Comparison of observed growth ability of persistent and sporadic isolates.

	Growth		Variable		No Growth		
Treatment[a]	Persistent	Sporadic	Persistent	Sporadic	Persistent	Sporadic	p-value[b]
Stress Conditions in Nutritive Media (BHI)							
Control	74	21	0	0	0	0	No test
5% NaCl BHI	74	21	0	0	0	0	No test
10% NaCl BHI	41	10	33	11	0	0	0.70
pH of 5.2	69	21	5	0	0	0	0.58
pH of 9.2	70	20	4	1	0	0	1.0
BAC 2µg/mL	0	0	74	21	0	0	No test
BAC 5µg/mL	0	0	41	8	33	13	0.25
Carbon Source Utilization (tested in chemically defined media [DM])							
DM Glucose	38	8	34	13	2	0	0.46
DM Cellobiose	50	10	24	11	0	0	0.16
DM Fructose	38	9	36	12	0	0	0.66
DM Glycogen	24	9	41	11	9	1	0.50
DM Lactose	1	0	9	2	64	19	1.0
DM Sucrose	1	0	15	3	58	18	0.81

[a] – See Table 4.1 for treatment formulations

[b] – p-values were determined by Chi Square Test, or Fisher's exact test if any cells contained less than five isolates

4.7 Figures

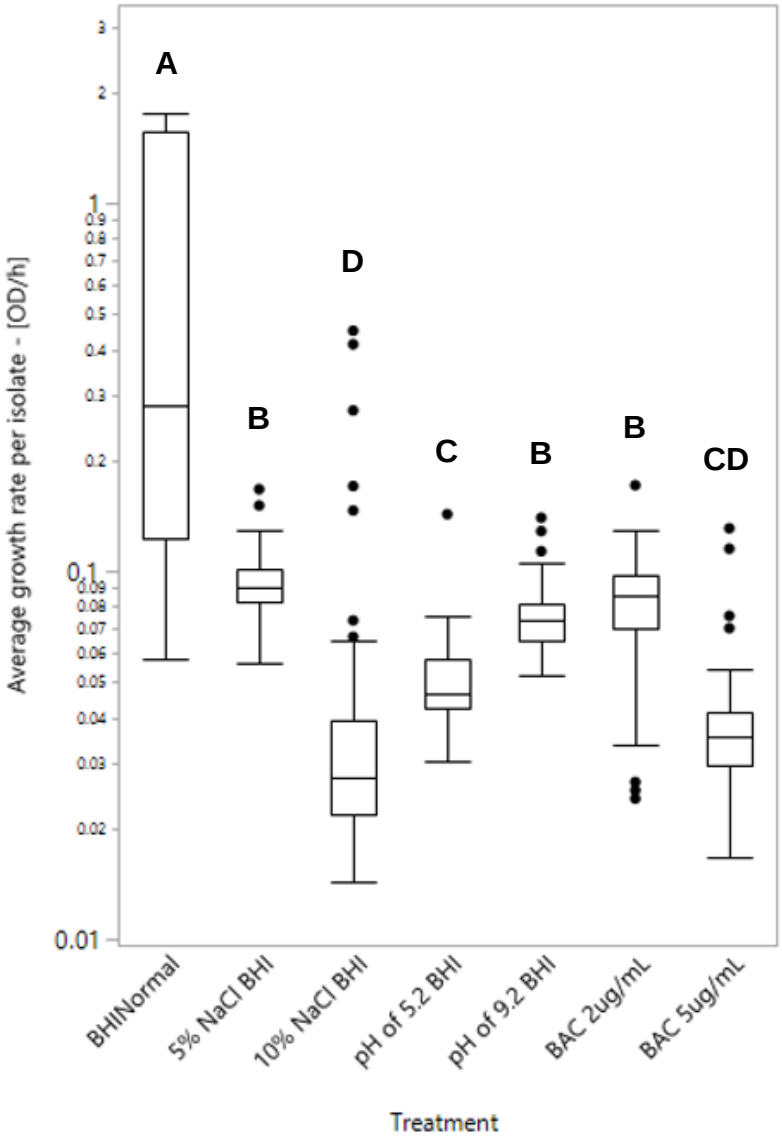

Figure 4.1: Box plots of the average growth rate for *L. monocytogenes* isolates exposed to stress conditions. Data includes the average of only replicates whose growth was observed (ΔOD_{600} $\geq$0.1), in log scale. The box represents the interquartile range (IQR), the line represents the median of the treatment, whiskers are drawn to the furthest point within 1.5 x IQR from the box, and the points are outliers of the data. Groups were created by use of Tukey's HSD, where the same letters indicate means that are not different from each other.

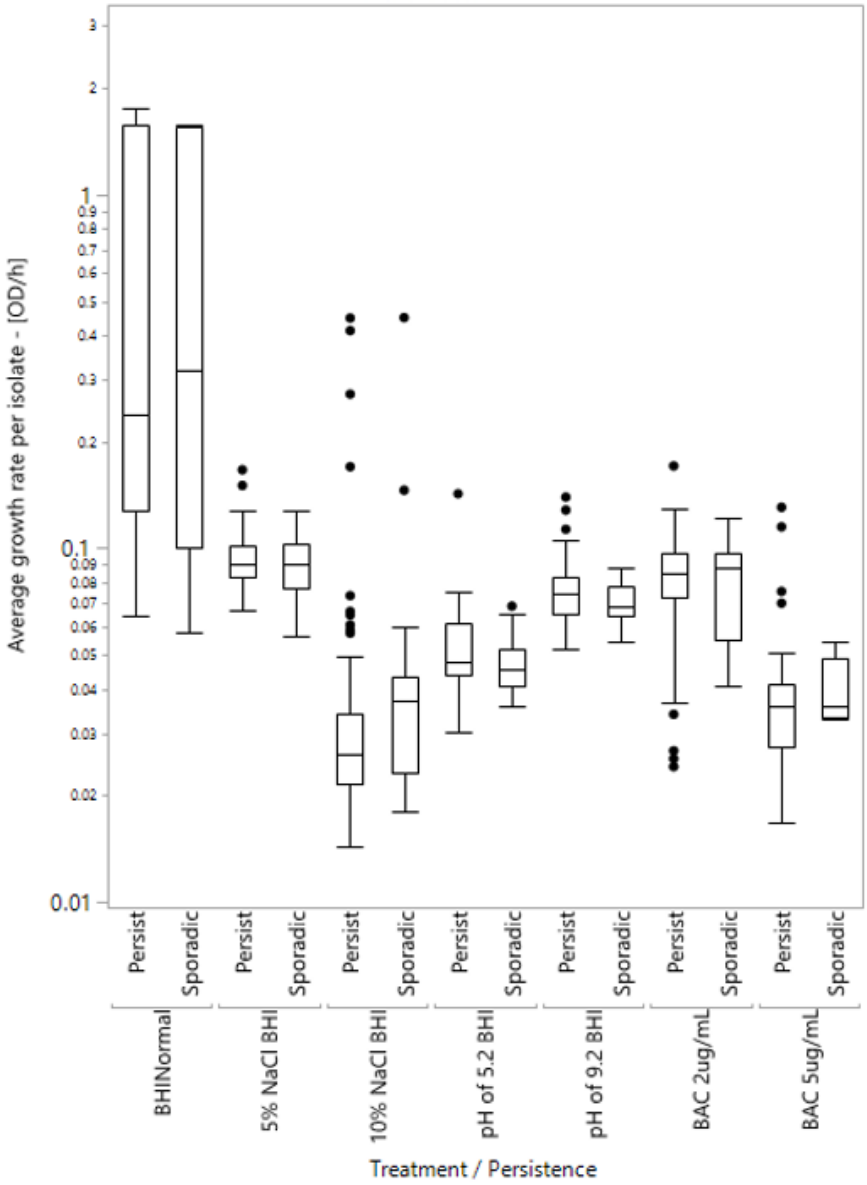

Figure 4.2: Box plots of average growth rate for *L. monocytogenes* isolates exposed to stress conditions, separated by their persistence factor. Data includes the average of only replicates where growth was observed ($\Delta OD_{600} \geq 0.1$), in log scale. No significant differences were observed in average growth rate between persistent and sporadic isolates for any treatment.

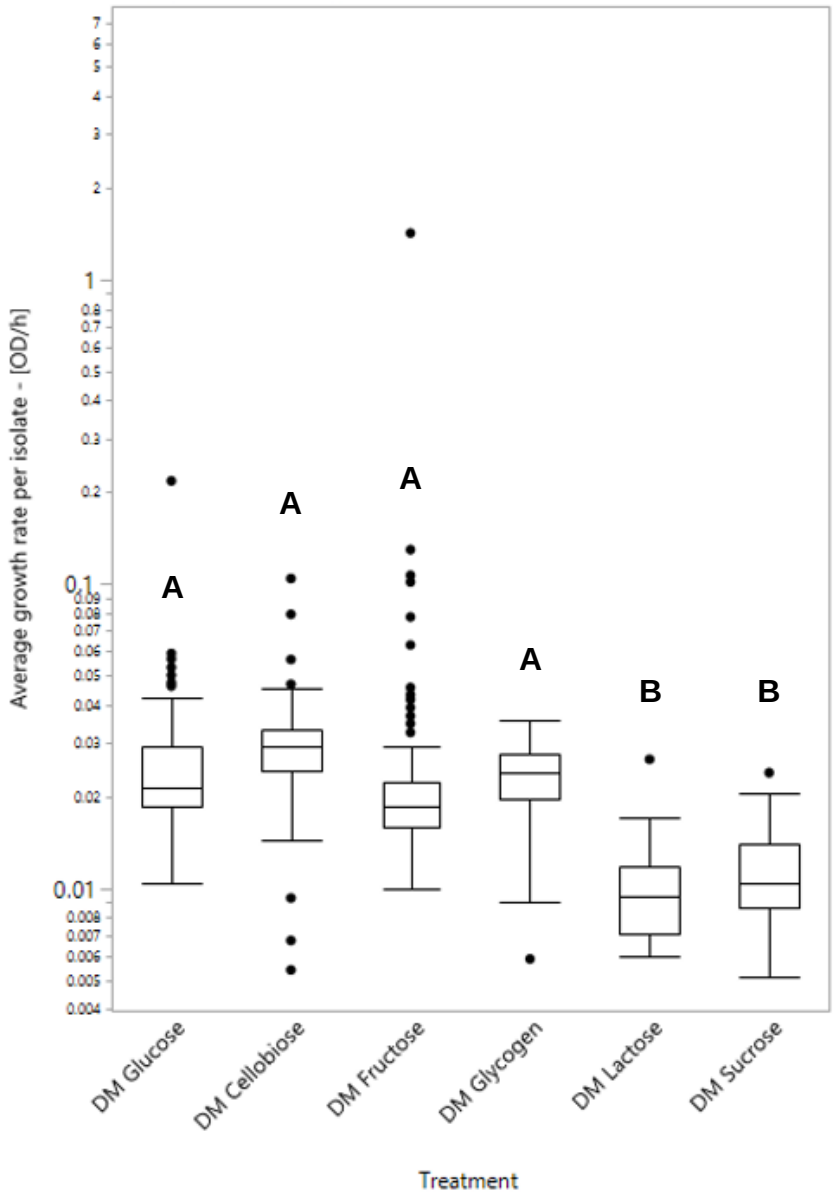

Figure 4.3: Box plots of average growth rate for *L. monocytogenes* isolates grown in chemically defined media (DM) at 25mM concentrations of each carbon source. Data includes averages of only replicates whose growth was observed ($\Delta OD_{600} \geq 0.1$), in log scale. Groups were created by use of Tukey's HSD, where the same letters indicate means that are not different from each other.

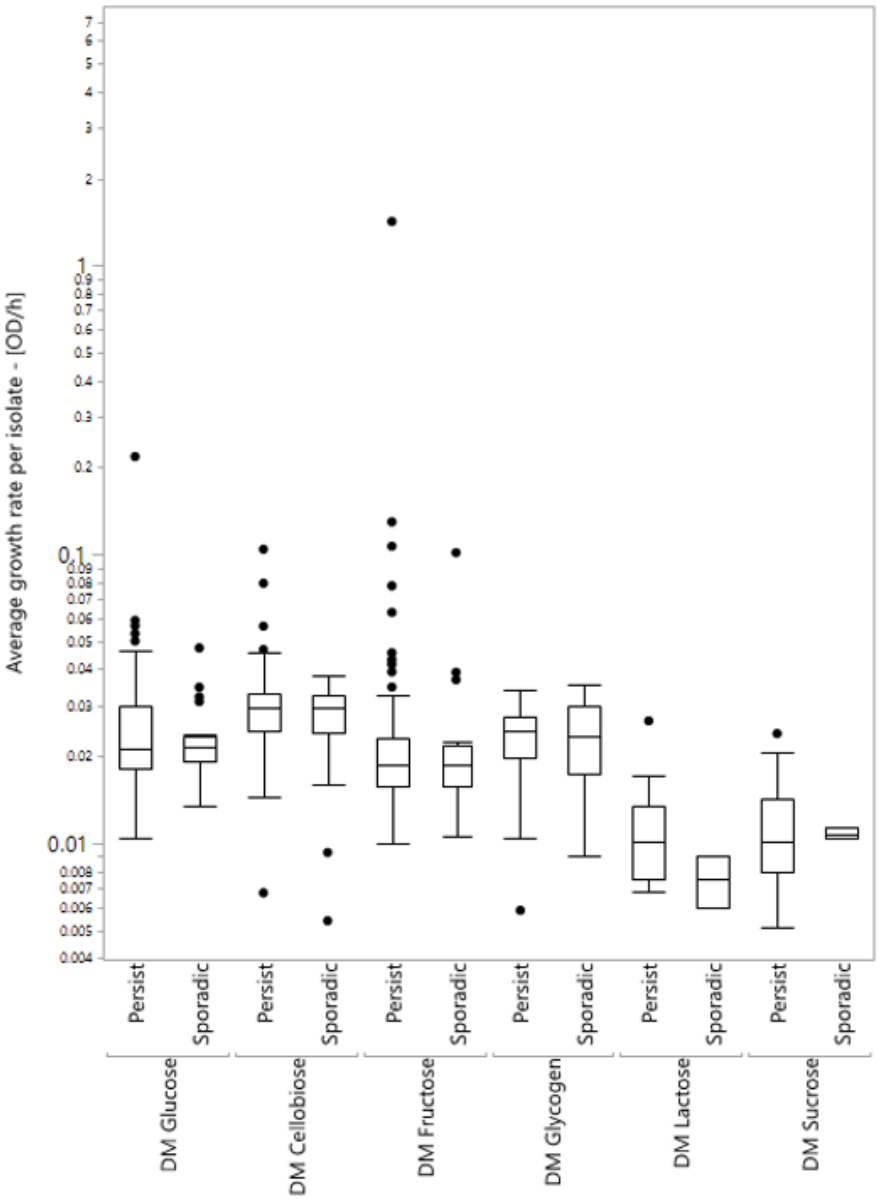

Figure 4.4: Box plots of average growth rate for *L. monocytogenes* isolates grown in chemically defined media (DM), separated by their persistence factor. Data only includes replicates whose growth was observed ($\Delta OD_{600} \geq 0.1$), in log scale. No significant differences were observed in average growth rate between persistent and sporadic isolates for any treatment.

4.8 References

(**1**) Amezaga, M. R., Davidson, I., McLaggan, D., Verheul, A., Abee, T., & Booth, I. R. (1995). The role of peptide metabolism in the growth of *Listeria monocytogenes* ATCC 23074 at high osmolarity. *Microbiology, 141 (Pt 1)*, 41-49. doi:10.1099/00221287-141-1-41

(**2**) Autio, T., Keto-Timonen, R., Lundén, J., Björkroth, J., & Korkeala, H. (2003). Characterisation of persistent and sporadic *Listeria monocytogenes* strains by Pulsed-Field Gel Electrophoresis (PFGE) and Amplified Fragment Length Polymorphism (AFLP). *Systematic and Applied Microbiology, 26*(4), 539-545. doi:10.1078/072320203770865846

(**3**) Bolocan, A. S., Pennone, V., O'Connor, P. M., Coffey, A., Nicolau, A. I., McAuliffe, O., et al. (2017). Inhibition of *Listeria monocytogenes* biofilms by bacteriocin-producing bacteria isolated from mushroom substrate. *J Appl Microbiol, 122*(1), 279-293. doi:10.1111/jam.13337

(**4**) Buchanan, R. L., Gorris, L. G. M., Hayman, M. M., Jackson, T. C., & Whiting, R. C. (2017). A review of *Listeria monocytogenes*: An update on outbreaks, virulence, dose-response, ecology, and risk assessments. *Food Control, 75*, 1-13. doi:10.1016/j.foodcont.2016.12.016

(**5**) Carpentier, B., & Cerf, O. (2011). Review — Persistence of *Listeria monocytogenes* in food industry equipment and premises. *International Journal of Food Microbiology, 145*(1), 1-8. doi:10.1016/j.ijfoodmicro.2011.01.005

(**6**) CDC. (2017). *People at Risk*. Retrieved from Center for Disease Control: https://www.cdc.gov/listeria/risk.html

(**7**) CFSAN. (2003). *Quantitative assessment of the relative risk to public health from foodborne* Listeria monocytogenes *among selected categories of ready-to-eat foods*. Retrieved from https://www.fda.gov/downloads/food/foodscienceresearch/ucm197329.pdf

(**8**) Chambers, M. C., Song, K. H., & Schneider, D. S. (2012). *Listeria monocytogenes* infection causes metabolic shifts in *Drosophila melanogaster*. *PLoS One, 7*(12), e50679. doi:10.1371/journal.pone.0050679

(**9**) da Silva, E. P., & De Martinis, E. C. (2013). Current knowledge and perspectives on biofilm formation: the case of *Listeria monocytogenes*. *Appl Microbiol Biotechnol, 97*(3), 957-968. doi:10.1007/s00253-012-4611-1

(**10**) Deutscher, J., Moussan Désirée Aké, F., Zebre, A., Nguyen Cao, T., Kentache, T., Mai Ma Pham, Q., et al. (2014). Carbohydrate utilization by *Listeria monocytogenes* and its influence on virulence gene expression. In E. C. Hambrick (Ed.), Listeria monocytogenes*: Food Sources, Prevalence and Management Strategies* (pp. 49-76): Nova science publishers.

(**11**) Edward M. Fox, A. C., Kieran Jordan, Aidan Coffey, Cormac G.M. Gahan, Olivia McAuliffe. (2017). Whole genome sequence analysis; an improved technology that identifies underlying genotypic differences between closely related *Listeria monocytogenes* strains. *Innovative Food Science & Emerging Technologies, 44*, 89-96. doi:10.1016/j.ifset.2017.07.010

(**12**) Ferreira, V., Wiedmann, M., Teixeira, P., & Stasiewicz, M. J. (2014). *Listeria monocytogenes* persistence in food-associated environments: epidemiology, strain characteristics, and implications for public health. *J Food Prot, 77*(1), 150-170. doi:10.4315/0362-028X.JFP-13-150

(**13**) Gerald McDonnell, A. D. R. (1999). Antiseptics and disinfectants: activity, action, and resistance. [Jan 2001]. *Clin Microbiol Rev, 12*(1), 147-149.

(**14**) Hoeflinger, J. L., Hoeflinger, D. E., & Miller, M. J. (2017). A dynamic regression analysis tool for quantitative assessment of bacterial growth written in Python. *J Microbiol Methods, 132*, 83-85. doi:10.1016/j.mimet.2016.11.015

(**15**) Jadhav, S., Bhave, M., & Palombo, E. A. (2012). Methods used for the detection and subtyping of *Listeria monocytogenes*. *J Microbiol Methods, 88*(3), 327-341. doi:10.1016/j.mimet.2012.01.002

(**16**) Kentache, T., Milohanic, E., Cao, T. N., Mokhtari, A., Ake, F. M., Ma Pham, Q. M., et al. (2016). Transport and catabolism of pentitols by *Listeria monocytogenes*. *J Mol Microbiol Biotechnol, 26*(6), 369-380. doi:10.1159/000447774

(**17**) Keto-Timonen, R., Tolvanen, R., Lunden, J., & Korkeala, H. (2007). An 8-year surveillance of the diversity and persistence of *Listeria monocytogenes* in a chilled food processing plant analyzed by amplified fragment length polymorphism. *J Food Prot, 70*(8), 1866-1873.

(**18**) Komora, N., Bruschi, C., Magalhães, R., Ferreira, V., & Teixeira, P. (2017). Survival of *Listeria monocytogenes* with different antibiotic resistance patterns to food-associated stresses. *International Journal of Food Microbiology, 245*, 79-87. doi:10.1016/j.ijfoodmicro.2017.01.013

(**19**) Kovacevic, J., Ziegler, J., Wałecka-Zacharska, E., Reimer, A., Kitts, D. D., & Gilmour, M. W. (2016). Tolerance of *Listeria monocytogenes* to quaternary ammonium sanitizers is mediated by a novel efflux pump encoded by *emrE*. *Appl Environ Microbiol, 82*(3), 939-953. doi:10.1128/AEM.03741-15

(**20**) Kremer, P. H., Lees, J. A., Koopmans, M. M., Ferwerda, B., Arends, A. W., Feller, M. M., et al. (2017). Benzalkonium tolerance genes and outcome in *Listeria monocytogenes* meningitis. *Clin Microbiol Infect, 23*(4), 265 e261-265 e267. doi:10.1016/j.cmi.2016.12.008

(**21**) Laksanalamai, P., Joseph, L. A., Silk, B. J., Burall, L. S., L. Tarr, C., Gerner-Smidt, P., et al. (2012). Genomic characterization of *Listeria monocytogenes* strains involved in a multistate listeriosis outbreak associated with cantaloupe in US. *PLoS One, 7*(7), e42448. doi:10.1371/journal.pone.0042448

(**22**) Lunden, J., Tolvanen, R., & Korkeala, H. (2008). Acid and heat tolerance of persistent and nonpersistent *Listeria monocytogenes* food plant strains. *Lett Appl Microbiol, 46*(2), 276-280. doi:10.1111/j.1472-765X.2007.02305.x

(**23**) Luo, L., Zhang, Z., Wang, H., Wang, P., Lan, R., Deng, J., et al. (2017). A 12-month longitudinal study of *Listeria monocytogenes* contamination and persistence in pork retail markets in China. *Food Control, 76*, 66-73. doi:10.1016/j.foodcont.2016.12.037

(**24**) Magalhaes, R., Ferreira, V., Brandao, T. R., Palencia, R. C., Almeida, G., & Teixeira, P. (2016). Persistent and non-persistent strains of *Listeria monocytogenes*: A focus on growth kinetics under different temperature, salt, and pH conditions and their sensitivity to sanitizers. *Food Microbiol, 57*, 103-108. doi:10.1016/j.fm.2016.02.005

(**25**) Marr, A. K., Joseph, B., Mertins, S., Ecke, R., Muller-Altrock, S., & Goebel, W. (2006). Overexpression of PrfA leads to growth inhibition of *Listeria monocytogenes* in glucose-containing culture media by interfering with glucose uptake. *J Bacteriol, 188*(11), 3887-3901. doi:10.1128/JB.01978-05

(**26**) Martinez-Suarez, J. V., Ortiz, S., & Lopez-Alonso, V. (2016). Potential Impact of the resistance to quaternary ammonium disinfectants on the persistence of *Listeria monocytogenes* in food processing environments. *Front Microbiol, 7*, 638. doi:10.3389/fmicb.2016.00638

(**27**) McLauchlin, J. R., C. E. D. (2015). *Listeria**. *Bergey's Manual of Systematic of Archaea and Bacteria*, 1-29. doi:10.1002/9781118960608.gbm00547

(**28**) Mereghetti, L., Quentin, R., Marquet-Van Der Mee, N., & Audurier, A. (2000). Low sensitivity of *Listeria monocytogenes* to quaternary ammonium compounds. *Appl Environ Microbiol, 66*(11), 5083-5086.

(**29**) Moretro, T., Schirmer, B. C., Heir, E., Fagerlund, A., Hjemli, P., & Langsrud, S. (2017). Tolerance to quaternary ammonium compound disinfectants may enhance growth of *Listeria monocytogenes* in the food industry. *Int J Food Microbiol, 241*, 215-224. doi:10.1016/j.ijfoodmicro.2016.10.025

(**30**) Moura, A., Criscuolo, A., Pouseele, H., Maury, M. M., Leclercq, A., Tarr, C., et al. (2016). Whole genome-based population biology and epidemiological surveillance of *Listeria monocytogenes*. *Nat Microbiol, 2*, 16185. doi:10.1038/nmicrobiol.2016.185

(**31**) Müller, A., Rychli, K., Muhterem-Uyar, M., Zaiser, A., Stessl, B., Guinane, C. M., et al. (2013). Tn6188 - A novel transposon in *Listeria monocytogenes* responsible for tolerance to benzalkonium chloride. *PLoS One, 8*(10), e76835. doi:10.1371/journal.pone.0076835

(**32**) Muller, A., Rychli, K., Zaiser, A., Wieser, C., Wagner, M., & Schmitz-Esser, S. (2014). The *Listeria monocytogenes* transposon Tn6188 provides increased tolerance to various quaternary ammonium compounds and ethidium bromide. *FEMS Microbiol Lett, 361*(2), 166-173. doi:10.1111/1574-6968.12626

(**33**) Nakamura, H., Takakura, K., Sone, Y., Itano, Y., & Nishikawa, Y. (2013). Biofilm formation and resistance to benzalkonium chloride in *Listeria monocytogenes* isolated from a fish processing plant. *J Food Prot, 76*(7), 1179-1186. doi:10.4315/0362-028X.JFP-12-225

(**34**) Ortiz, S., Lopez-Alonso, V., Rodriguez, P., & Martinez-Suarez, J. V. (2015). The connection between persistent, disinfectant-resistant *Listeria monocytogenes* strains from two geographically separate Iberian pork processing plants: evidence from comparative genome analysis. *Appl Environ Microbiol, 82*(1), 308-317. doi:10.1128/AEM.02824-15

(**35**) P.J. McClure, T. A. R., P. Otto Oguru. (1989). Comparison of the effects of sodium chloride, pH and temperature on the growth of *Listeria monocytogenes* on gradient plates and in liquid medium. *Letters in Applied Microbiology, 9*(3), 95-99. doi:10.1111/j.1472-765X.1989.tb00299.x

(**36**) Premaratne, R. J., Lin, W. J., & Johnson, E. A. (1991). Development of an improved chemically defined minimal medium for *Listeria monocytogenes*. *Appl Environ Microbiol, 57*(10), 3046-3048.

(**37**) Simmons, C., Stasiewicz, M. J., Wright, E., Warchocki, S., Roof, S., Kause, J. R., et al. (2014). *Listeria monocytogenes* and *Listeria* spp. contamination patterns in retail delicatessen establishments in three U.S. states. *J Food Prot, 77*(11), 1929-1939. doi:10.4315/0362-028X.JFP-14-183

(**38**) Slaghuis, J., Joseph, B., & Goebel, W. (2007). Metabolism and physiology of *Listeria monocytogenes*. In H. Goldfine & H. Shen (Eds.), *Listeria monocytogenes: Pathogenesis and Host Response* (pp. 63-80). Boston, MA: Springer US.

(**39**) Stanga, M. (2010). *Sanitation: cleaning and disinfection in the food industry*: John Wiley & Sons, Inc.

(**40**) Stasiewicz, M. J., Oliver, H. F., Wiedmann, M., & den Bakker, H. C. (2015). Whole-genome sequencing allows for improved identification of persistent *Listeria monocytogenes* in food-associated environments. *Appl Environ Microbiol, 81*(17), 6024-6037. doi:10.1128/AEM.01049-15

(**41**) Stoll, R., Mertins, S., Joseph, B., Muller-Altrock, S., & Goebel, W. (2008). Modulation of PrfA activity in *Listeria monocytogenes* upon growth in different culture media. *Microbiology, 154*(Pt 12), 3856-3876. doi:10.1099/mic.0.2008/018283-0

(**42**) Toledo-Arana, A., Dussurget, O., Nikitas, G., Sesto, N., Guet-Revillet, H., Balestrino, D., et al. (2009). The *Listeria* transcriptional landscape from saprophytism to virulence. *Nature, 459*(7249), 950-956. doi:10.1038/nature08080

Chapter 5: Conclusions

5.1 Research Summary

Throughout this work, genotypic and phenotypic properties of persistent and sporadic strains of *L. monocytogenes* were examined. It was hypothesized that (i) CRISPR spacers might provide evidence that there is a genotype associated with persistence and that (ii) persistent strains would have higher growth rates or abilities to grow than sporadic isolates from retail deli environments. In both cases the hypothesis was refuted, as it is clear that the persistence cannot be identified by CRISPRs or destroyed by phenotypic means. However, this research is novel as it is the first to examine CRISPR subtyping within *L. monocytogenes*, identify conserved Spacer Patterns in *L. monocytogenes* isolates, look at persistence within mobile genetic elements, and test a distinct panel of persistent *L. monocytogenes* in food-associated stress environments and carbon sources. The importance of seek-and-destroy strategies to eliminate these persistent *L. monocytogenes* strains in food-associated facilities is strongly stressed due to this work.

5.2 Future Work

There are a few key suggestions we want to make towards future work. The first is that CRISPRs within *L. monocytogenes* and pathogens, in general, must be more understood. There is a looming question on if CRISPRs are still necessary within current *L. monocytogenes* strains or if they are remnants of previous phage infections. This could be looked at through acquisition and integration of new spacers or by PCR amplification of current CRISPR arrays. Secondly, *L. monocytogenes*' resistance to stresses is well explored, but its carbon source utilization is not. More classical papers in biology have examined part of this inquiry, but there is little research that has explored more on the biochemistry behind food-associated carbon source utilization. Finally, it is strongly suggested that research continues on persistence. Whether this is by examining persistence's relationship to sanitizer resistance or if persistence can be determined by other

genotypic / mobile genetic elements. We have begun some work on both areas of this last section and will continue to pursue identification and elimination tools to persistent strains of *L. monocytogenes*.

Appendix A – Supplemental Tables for Both Projects

Table A.1. CRISPR-based Subtyping of *L. monocytogenes*

Isolate ID	Spacer Patterns	WGS-Based Clade	Persistence Status	WGS-Corrected PFGE Subtype	Store
FSL R8-5112	1	N	Persist	CU-11-326	8
FSL R8-5119	1	N	Sporadic	CU-11-326	11
FSL R8-5584	1	N	Persist	CU-11-326	8
FSL R8-5108	1	O	Persist	CU-8-340	8
FSL R8-5559	1	O	Persist	CU-8-340	8
FSL R8-5576	1	O	Persist	CU-8-340	8
FSL R8-6257	1	O	Persist	CU-8-340	8
FSL R8-7107	1	O	Persist	CU-8-340	8
FSL R8-5809	1	P	Sporadic	CU-SNP1	23
FSL R8-6283	1	P	Persist	CU-SNP1	13
FSL R8-7121	1	P	Persist	CU-SNP1	29
FSL R8-7161	1	P	Persist	CU-SNP1	13
FSL R8-5610	1	Q	Persist	CU-SNP1	21
FSL R8-6208	1	Q	Persist	CU-SNP1	21
FSL R8-6617	1	Q	Persist	CU-SNP1	21
FSL R8-7020	1	Q	Sporadic	CU-SNP1	12
FSL R8-7198	1	Q	Sporadic	CU-SNP1	17
FSL R8-7571	1	Q	Persist	CU-SNP1	21
FSL R8-7786	1	Q	Persist	CU-SNP1	21
FSL R8-9398	1	Q	Persist	CU-SNP1	21
FSL R8-5786	1	R	Persist	CU-SNP1	16
FSL R8-6112	1	R	Persist	CU-SNP1	16
FSL R8-6359	1	R	Sporadic	CU-SNP1	7
FSL R8-6603	1	R	Persist	CU-SNP1	16
FSL R8-6936	1	R	Persist	CU-SNP1	16
FSL R8-7534	1	R	Persist	CU-SNP1	16
FSL R8-8816	1	R	Persist	CU-SNP1	28
FSL R8-6232	1	S	Persist	CU-SNP1	24
FSL R8-6244	1	S	Persist	CU-SNP1	24
FSL R8-6791	1	S	Persist	CU-SNP1	24
FSL R8-6848	1	S	Persist	CU-SNP1	24
FSL R8-6870	1	S	Persist	CU-SNP1	24
FSL R8-7554	1	S	Sporadic	CU-SNP1	16
FSL R8-7603	1	S	Persist	CU-SNP1	24
FSL R8-7635	1	S	Persist	CU-SNP1	24
FSL R8-7860	1	S	Persist	CU-SNP1	24
FSL R8-7866	1	S	Persist	CU-SNP1	24
FSL R8-7870	1	S	Persist	CU-SNP1	24
FSL R8-5318	1	T	Persist	CU-SNP1	10
FSL R8-5410	1	T	Persist	CU-SNP1	10
FSL R8-5686	1	T	Sporadic	CU-SNP1	25
FSL R8-6045	1	T	Persist	CU-SNP1	10
FSL R8-6341	1	T	Sporadic	CU-SNP1	10
FSL R8-6345	1	T	Persist	CU-SNP1	10
FSL R8-6721	1	T	Persist	CU-SNP1	10
FSL R8-6725	1	T	Persist	CU-SNP1	10
FSL R8-6739	1	T	Persist	CU-SNP1	23
FSL R8-7047	1	T	Persist	CU-SNP1	10
FSL R8-7460	1	T	Persist	CU-SNP1	10
FSL R8-9402	1	T	Persist	CU-SNP1	10
FSL R8-5081	1	U	Persist	CU-SNP1	2
FSL R8-5088	1	U	Persist	CU-SNP1	2
FSL R8-5233	1	U	Persist	CU-SNP1	2
FSL R8-5463	1	U	Persist	CU-SNP1	2
FSL R8-6036	1	U	Persist	CU-SNP1	2
FSL R8-6305	1	U	Persist	CU-SNP1	2
FSL R8-6317	1	U	Persist	CU-SNP1	2
FSL R8-6649	1	U	Persist	CU-SNP1	2
FSL R8-6661	1	U	Persist	CU-SNP1	2
FSL R8-7133	1	U	Persist	CU-SNP1	2
FSL R8-5543	1	V	Sporadic	CU-SNP1	3
FSL R8-5748	1	V	Persist	CU-SNP1	29
FSL R8-5797	1	V	Persist	CU-SNP1	18
FSL R8-6271	1	V	Persist	CU-SNP1	29
FSL R8-6446	1	V	Persist	CU-SNP1	18

Table A.1 Cont

FSL R8-6488	1	V	Sporadic	CU-SNP1	8
FSL R8-6641	1	V	Sporadic	CU-SNP1	2
FSL R8-7474	1	V	Sporadic	CU-SNP1	11
FSL R8-7489	1	V	Sporadic	CU-SNP1	26
FSL R8-7493	1	V	Persist	CU-SNP1	29
FSL R8-7694	1	V	Sporadic	CU-SNP1	18
FSL R8-7716	1	V	Persist	CU-SNP1	29
FSL R8-5124	1-10-11	M	Persist	CU-11-282	13
FSL R8-6480	1-10-11	M	Persist	CU-11-282	13
FSL R8-7153	1-10-11	M	Persist	CU-11-282	13
FSL R8-7478	1-10-11	M	Persist	CU-11-282	13
FSL R8-7722	4-5-10-11	H	Persist	CU-SNP3	29
FSL R8-5878	4-5-10-11	I	Persist	CU-SNP3	28
FSL R8-6176	4-5-10-11	I	Persist	CU-SNP3	28
FSL R8-6180	4-5-10-11	I	Persist	CU-SNP3	28
FSL R8-6557	4-5-10-11	I	Persist	CU-SNP3	28
FSL R8-6836	4-5-10-11	I	Sporadic	CU-SNP3	24
FSL R8-6963	4-5-10-11	I	Persist	CU-SNP3	28
FSL R8-6967	4-5-10-11	I	Persist	CU-SNP3	28
FSL R8-7833	4-5-10-11	I	Persist	CU-SNP3	23
FSL R8-6761	4-5-10-11	J	Persist	CU-262-334	23
FSL R8-6900	4-5-10-11	J	Persist	CU-262-334	23
FSL R8-5449	4-5-10-11	K	Persist	CU-SNP3	2
FSL R8-6313	4-5-10-11	K	Persist	CU-SNP3	2
FSL R8-6220	4-5	G	Persist	CU-262-318	21
FSL R8-6456	4-5	G	Sporadic	CU-262-318	26
FSL R8-6828	4-5	G	Persist	CU-262-318	21
FSL R8-7559	4-5	G	Persist	CU-262-318	21
FSL R8-6426	4-5	H	Persist	CU-SNP3	29
FSL R8-5884	4-5	I	Persist	CU-SNP3	28
FSL R8-6168	4-5	I	Persist	CU-SNP3	23
FSL R8-6553	4-5	I	Persist	CU-SNP3	28
FSL R8-6808	4-5	I	Sporadic	CU-SNP3	21
FSL R8-7653	4-5	I	Persist	CU-SNP3	28
FSL R8-7676	4-5	I	Persist	CU-SNP3	28
FSL R8-7881	4-5	I	Persist	CU-SNP3	28
FSL R8-7902	4-5	I	Persist	CU-SNP3	28
FSL R8-5230	4-5	K	Persist	CU-SNP3	2
FSL R8-5257	4-5	K	Persist	CU-SNP3	2
FSL R8-6024	4-5	K	Persist	CU-SNP3	2
FSL R8-7381	4-5	K	Persist	CU-SNP3	2
FSL R8-5487	4-6-7	B	Sporadic	CU-57-267	7
FSL R8-6625	4-6-7	B	Persist	CU-57-267	4
FSL R8-6735	4-6-7	B	Persist	CU-57-267	23
FSL R8-7149	4-6-7	B	Persist	CU-57-267	4
FSL R8-7399	4-6-7	B	Persist	CU-57-267	4
FSL R8-5775	4-6-7-12	B	Persist	CU-57-267	16
FSL R8-6321	4-6-7-12	B	Persist	CU-57-267	10
FSL R8-6592	4-6-7-12	B	Persist	CU-57-267	16
FSL R8-6717	4-6-7-12	B	Persist	CU-57-267	10
FSL R8-7043	4-6-7-12	B	Persist	CU-57-267	10
FSL R8-7806	4-6-7-12	B	Persist	CU-57-267	23
FSL R8-7914	4-6-7-12	B	Sporadic	CU-57-267	28
FSL R8-6637	7-12	C	Persist	CU-55-266	2
FSL R8-5436	7-19	C	Persist	CU-55-266	2
FSL R8-9399	7-19	C	Persist	CU-55-266	2
FSL R8-6033	7-19-20	C	Persist	CU-55-266	2
FSL R8-5805	3	D	Persist	CU-SNP2	23
FSL R8-5833	3	D	Persist	CU-SNP2	23
FSL R8-5844	3	D	Persist	CU-SNP2	23
FSL R8-6037	3	D	Sporadic	CU-SNP2	10
FSL R8-6122	3	D	Persist	CU-SNP2	23
FSL R8-6138	3	D	Persist	CU-SNP2	23
FSL R8-6160	3	D	Persist	CU-SNP2	23
FSL R8-6743	3	D	Persist	CU-SNP2	23
FSL R8-6765	3	D	Persist	CU-SNP2	23
FSL R8-6908	3	D	Persist	CU-SNP2	23
FSL R8-6910	3	D	Persist	CU-SNP2	23
FSL R8-6914	3	D	Persist	CU-SNP2	23

Table A.1 Cont

FSL R8-7585	3	D	Persist	CU-SNP2	23
FSL R8-7591	3	D	Persist	CU-SNP2	23
FSL R8-7599	3	D	Persist	CU-SNP2	23
FSL R8-7812	3	D	Persist	CU-SNP2	23
FSL R8-7820	3	D	Persist	CU-SNP2	23
FSL R8-7825	3	D	Persist	CU-SNP2	23
FSL R8-7842	3	D	Persist	CU-SNP2	23
FSL R8-8476	3	D	Persist	CU-SNP2	21
FSL R8-6896	3	E	Persist	CU-258-69	23
FSL R8-7829	3	E	Persist	CU-258-69	23
FSL R8-5095	3	F	Persist	CU-258-69	7
FSL R8-5104	3	F	Sporadic	CU-258-69	10
FSL R8-5237	3	F	Sporadic	CU-258-69	4
FSL R8-5249	3	F	Persist	CU-258-69	7
FSL R8-5309	3	F	Persist	CU-258-69	7
FSL R8-5528	3	F	Persist	CU-258-69	7
FSL R8-5646	3	F	Sporadic	CU-258-69	22
FSL R8-5726	3	F	Sporadic	CU-258-69	27
FSL R8-5854	3	F	Persist	CU-258-69	24
FSL R8-5975	3	F	Persist	CU-258-69	7
FSL R8-7097	3	F	Persist	CU-258-69	7
FSL R8-7427	3	F	Persist	CU-258-69	7
FSL R8-7846	3	F	Persist	CU-258-69	24
FSL R8-8446	3	F	Persist	CU-258-69	28
FSL R8-8511	3	F	Persist	CU-258-69	7
FSL R8-5402	9	A	Persist	CU-294-321	10
FSL R8-5760	9	A	Persist	CU-294-321	21
FSL R8-6046	9	A	Persist	CU-294-321	10
FSL R8-6607	9	A	Persist	CU-294-321	21
FSL R8-6918	9	A	Persist	CU-294-321	16
FSL R8-7057	9	A	Sporadic	CU-294-321	7
FSL R8-7348	9	A	Sporadic	CU-294-321	2
FSL R8-7389	9	A	Sporadic	CU-294-321	4
FSL R8-7450	9	A	Persist	CU-294-321	10
FSL R8-7761	9	A	Persist	CU-294-321	16
FSL R8-7006	9[r]	A	Sporadic	CU-294-321	12
FSL R8-6691	--	F	Persist	CU-258-69	7
FSL R8-6373	15-16-17	Y	Persist	CU-182-173	19
FSL R8-7926	15-16-17	Y	Persist	CU-182-173	19
FSL R8-6134	21-22-23	X	Persist	CU-296-330	23
FSL R8-6892	21-22-23	X	Persist	CU-296-330	23

Table A.2. Initial Spacer Sequences and Pattern Assignments

Spacers Sequences	Individual Pattern	Spacer Patterns
AAAGGTACTTAAATCTAGCTCTTCAAGACTAGTACAATTTTGAAACATTCTATT	A	1
CCAATTGCTTACATCCAAAATCTCCAGTGATCCACAACCCGAAAACATGGAATA	B	1
AAAATGACTTAAATCTAACTCTTCAAGACCAGAGCAAAATTGAAACATGCCATA	C	1
TATGGCATGTTTCAATTTTGCTCTGGTCTTGAAGAGTTAGATTTAAGTCATTTT	**D – revcom C**	**1**
TATTCCATGTTTTCGGGTTGTGGATCACTGGAGATTTTGGATGTAAGCAATTGG	**E – revcom B**	**1**
AATAGAATGTTTCAAAATTGTACTAGTCTTGAAGAGCTAGATTTAAGTACCTTT	**F – revcom A**	**1**
TTCGAATGAGCTCGGTTCTGATGTAGCAGGAGCATTAAAACCAAGAAGCATCGGT	G	3
ACCGATGCTTCTTGGTTTTAATGCTCCTGCTACATCGGAACCGAGCTCATTCGAA	H	4
GTCTAGCTCTTCAAAACTAGTACATCCTTCAAACATACTATTCATGCGAGTTACT	I	5
ATTCAGTTTCTCAAGCGATTTACAATATTGAAACATCATCTGCATAAAAGTAACA	J	5
GGCGATTCTGTTGAAACTGCAACGAAACTT	K	6
CCACGTCAGACCAAGAACTTAAAACTATGC	L	6
CTAAATTCATGTTGCGGGATGTTGTGGATG	M	6
GAACTTCTTCACCATCCTTCATTTCTGTTT	N	7
TCAGACTTCTATATCCACAATAAAAGCCCT	O	7
GCGAGTCAATTCATCAAACCCAATCAGAAA	P	7
GTGTACACGATAGTCCAAGTCGGTATTTCC	Q	7
CTCACGTTGTCAAGGTCAAATTTTAGATATG	R	7
GAGGTTCGTGGGGGTAATCCCGGCTTGCGA	S	7
AGAGGCAACGAGACTACAGACGCACTTAGA	T	7
TTTGGTCGAATTCTTCCGCATCTTTAAACT	U	7
TTAATGAAGAACTAGAAACAATTGAAAACA	V	7
GCACGTATGATCGTCTTCTTTGATGACCTC	W	7
TACTCGTATAAAAATCAAGAGAGACGTATT	X	7
CTTCCGCGTTACATGTTTGACATCAAGTAA	Y	7
GGTTGTGCCGTCACCCGTTGGTGGTAGGCT	Z	6
TTTTCTTTTTCACGATGCGATCGAACGTTT	AA	6
AGTGCAATATTTAGCGTACACATCTTTTAC	AB	6
TTGGTCAATAGATGACTGTATTTCTTGCTC	AC	6
TGAAACACACAACATTTCAGAATGGCTTAG	AD	6
ACTTACTGAACAACATTGATTACCACAGTT	AE	6
TAATAAACAAGAAATATTACTTCATGAATC	AF	6
CACTATCCACTACAGTGATTTGTATTGTGC	AG	6
GTAATCCCAATTAACCCCGCAGAGGGTGTA	AH	6
TGTCATGGCGAAAGGTAAAACGGTCGATTG	AI	6
AGTGTTGGAGACTGCAAGGATTTCCGGATT	AJ	6
AACCTGCAGGTGCTGTGTTCACGTCAGCAA	AK	6
TTGATGAGAATCTATACAGGTACTTAACCG	AL	6
TGTTGTCAAAGATGGTAATAAATGGGTGAC	AM	6
CATCGAATTGATACTTTTCGAGTGAAGCAA	AN	6
GTGGGAAACGTTAAATATTATAAAACAGAT	AO	6
GCATCGTACCCCAGTTCATGAAGCGCGGTA	AP	6
ACAAAACTCTCTAATTCAATTGCTCCATCA	AQ	6
TTTATAAAGAATACTTGCGGGGCATAAATG	AR	6
GTCTAAACTTGGAATATAACTTAGGTCTTG	AS	7
ACTGATGAAAGTATTTCTCCAATGACAAGT	AT	7
TTAAATACGATGGAATTAATCTATCGATGC	AU	7
TCAAAGAGTTTATACCTGTTTTGATTGAGT	AV	7
TAATATTCTTCCCCCCCATCGCTAACAGGG	AW	7
ACCGATGCTTCTTGGTTTTAATGCTCCTGCTACATCAGAACCGAGCTCATTCGAA	**AX – revcom G**	**3**
CTCTTCTCGTGTCGCTTGTTTTGCTCATTCCGCCAC	AY	9
ATAAATTTCGCTACTTTATCAATCATTCCAAATCTCC	AZ	9
ATTACATTTATTAAACAATGTACAGGTTGAATTAA	BA	9
CTTTTTATATATTGTGCGCTTAACTTAGTTGTTAAGA	BB	9
TTTTTTGTAGCGGAGGTGTACAATGCCGAAAGTGAC	BC	9
CCCTGTAGGATTAATATACAAAGGTTCTCGTTTTTC	BD	9
AATGCCGAAAGGTTCGCTTTCATTTCGCTATTCAC	BE	9
TTGATAACATCTTTTACTGAAAGACCTGCCTTTAA	BF	9
TTGGATGCAAAAACTTTAATAGCCTCTGCATCAA	BG	9
TTATCTAACCAACCCCGCAAAACAGTATCAACAGT	BH	9
TTGGTTTATCAAGATGAATGCAACACACTTGCGCCAT	BI	9
AAGAAACGCAGGAAGAAGCTCCACGCCGCAAGCGTC	BJ	9
GTTGTTAATGAAGAAGGAAAAGGTATTTTGAAATTC	BK	9
ATTTTATTACGGCTGGTGAGTTGTCCGCCGCACTA	BL	9
GTCGATACAATCAACAAAGAAGCGCTAGATATGA	BM	9
ATGAGTGCATCATAACGTTTTGCTACATAAGAAAT	BN	9

Table A.2 Cont

AAAAGTTCAAGAATTTGAACCATCCATTTAGCCCCC	BO	9
ATAGACCGTTTTTTCAAAGCCATTTATCAACTG	BP	9
ACTTTCTGTCAACGTGACAGGCATCGGAGATTTAAT	BQ	9
ATAGCTTATCAATCCGTTGGACAGATGAACCTTT	BR	9
TTCGTTAGCATTAACGCTGAAAATCAACTTCATGT	BS	9
TTTGTTCGATTTCGCTAATTTGGTAGTTTGTGACTTC	BT	9
GTGGTTGGCAAGAAACCTTTTTTGGCAGGTAATTTTT	BU	9
AAAATAATCATTAATGCAAACGCAATACAAATGACTG	BV	9
GTGGTAACAGCGTTCTTATTTCCTGGCGGAGCGCTTCC	BW	9
TTGGTCCTTCCTATCCCTCCACCTCTTTGTTTTT	BX	9
TCGTTCGACGGCTTAGACAACCATCACGCTTCTTTAA	BY	9
CCTAATACACCTAAACTACTCTTATATATAAAAACA	BZ	9
TTGTCAATTAGTTACTACCTTAATTGCTAATTAGTTACTACCTTAATTCGTTACTACCTTA ATTATTTTGGCT	CA	10
TTGCCGATTTTAGCCTATTAGTTACTACCTTAATTTGTTACTACCTTAATTTTTTTGACT	CB	11
GATTGGCTGTATTTAATAATTTGTGCTCCCCTTTTTGAGTTGTTCGTCCACT	CC	12
TTCGAATGAGCTCGGTTCCGATGTAGCAGGAGCATTAAAACCAAGAAGCATCGGT	CD	4
TGTTTTTATATATAAGAGTAGTTTAGGTGTATTAGG	**CE – revcom BZ**	**9**
TTAAAGAAGCGTGATGGTTGTCTAAGCCGTCGAACGA	**CF –revcom BY**	**9**
AAAAACAAAGAGGTGGAGGGATAGGAAGGACCAA	**CG –revcom BX**	**9**
GGAAGCGCTCCGCCAGGAAATAAGAACGCTGTTACCAC	**CH – rvcom BW**	**9**
CAGTCATTTGTATTGCGTTTGCATTAATGATTATTTT	**CI – revcom BV**	**9**
AAAAATTACCTGCCAAAAAAGGTTTCTTGCCAACCAC	**CJ – revcom BU**	**9**
GAAGTCACAAACTACCAAATTAGCGAAATCGAACAAA	**CK – revcom BT**	**9**
ACATGAAGTTGATTTTCAGCGTTAATGCTAACGAA	**CL – revcom BS**	**9**
AAAGGTTCATCTGTCCAACGGATTGATAAGCTAT	**CM – revcom BR**	**9**
ATTAAATCTCCGATGCCTGTCACGTTGACAGAAAGT	**CN – revcom BQ**	**9**
CAGTTGATAAATGGCTTTGAAAAAACGGTCTAT	**CO – revcom BP**	**9**
GGGGGCTAAATGGATGGTTCAAATTCTTGAACTTTT	**CP – revcom BO**	**9**
ATTTCTTATGTAGCAAAACGTTATGATGCACTCAT	**CQ – revcom BN**	**9**
TCATATCTAGCGCTTCTTTGTTGATTGTATCGAC	**CR – revcom BM**	**9**
TAGTGCGGCGGACAACTCACCAGCCGTAATAAAAT	**CS – revcom BL**	**9**
GAATTTCAAAATACCTTTTCCTTCTTCATTAACAAC	**CT – revcom BK**	**9**
GACGCTTGCGGCGTGGAGCTTCTTCCTGCGTTTCTT	**CU – revcom BJ**	**9**
ATGGCGCAAGTGTGTTGCATTCATCTTGATAAACCAA	**CV – revcom BI**	**9**
ACTGTTGATACTGTTTTGCGGGGTTGGTTAGATAA	**CW – revcom BH**	**9**
TTGATGCAGAGGCTATTAAAGTTTTTGCATCCAA	**CX – revcom BG**	**9**
TTAAAGGCAGGTCTTTCAGTAAAAGATGTTATCAA	**CY – revcom BF**	**9**
GTGAATAGCGAAATGAAAGCGAACCTTTCGGCATT	**CZ – revcom BE**	**9**
GAAAAACGAGAACCTTTGTATATTAATCCTACAGGG	**DA – revcom BD**	**9**
GTCACTTTCGGCATTGTACACCTCCGCTACAAAAAA	**DB – revcom BC**	**9**
TCTTAACAACTAAGTTAAGCGCACAATATATAAAAAG	**DC – revcom BB**	**9**
TTAATTCAACCTGTACATTGTTTAATAAATGTAAT	**DD – revcom BA**	**9**
GGAGATTTGGAATGATTGATAAAGTAGCGAAATTTAT	**DE – revcom AZ**	**9**
GTGGCGGAATGAGCAAAACAAGCGACACGAGAAGAG	**DF – revcom AY**	**9**
TCTAAAAATCAAAGGAGGATGTTAGAAATGAATGAAC	DG	15
TCTGCCGTTTTGATATCATTTTTTGCAATAACATC	DH	15
GAAGCCTTACTAGTTGAATTATTGCCGCCGGAAGAT	DI	15
ACGCCTATCTGTGCGTCTTCCCAAACCAAGTCGTGT	DJ	15
AAAAGAGTTGATTGACCAAATCAGAAACACACTCAT	DK	15
AATTAACAGAAGCGGAAATTAAAGGCATGCAAAAA	DL	15
TAGATAGTTTAGACTTTTCGACGTAGTACTTCCGCTA	DM	15
GGCACAGTCTTCACGATTTCGGGGATGGGGAAAACT	DN	15
TAAACAAGTGCTTCTGGCAATACCAACGTTCCATC	DO	15
CGAGCAATAGACTCTCGAACATCAATACCATACA	DP	16
AGAAGTTTGCCCCCTGCGCCTAAGAGTTTTCCTAGC	DQ	16
CAATTTGACGAAAGAAACAATGTGAAGTGGTATTCT	DR	16
CCTAAAGACTTACATCCACTTGTTTTCCAGAGTTAT	DS	16
TTGCATTAATACTCATTTCTAAGCTTCTAAGTTCATTCA	DT	16
CAAAATAAATAATATTGAATAAGCTGAAATATTTTT	DU	16
TTGACACATCAAATGTAAATCCGCTAAACAATTCTT	DV	16
TCCGACACCACCCAAAAAGCCAATGCAGTGCTTCAAT	DW	16
GAAGTAGATGTAATAATTTTCAAAGAAGCTTTGA	DX	16
ACATTTTGTACGCACCACGCTTGACGGTCTGAAACT	DY	16
AGCTAAACAAGTTCTTGACTGCTCTTCTTAATTCC	DZ	16
GTTATATCTTCTATTTGCGTCGTTTTTCCGTCATAAA	EA	16
CGAGTTAGGGTTAGCTACTCCCGTTAGTCTTTTTA	EB	16
ATAATTCTGTGAACAAATACTCTTTTTCCTTTAATAG	EC	16
ATTTTCGATGGCGGGCATGAAACGACACTAGCTTTT	ED	16

Table A.2 Cont

CAATACCCGCTGGGTTTAGACCTCTAGCAGGTAATCA	EE	16
GATGGAAGTATTCGTGTGAGTGTTCTGGCATACG	EF	16
TCAGCTCCATCAATGCTTTTTGATAGCACTTCTAAA	EG	16
GTACTCCATCGGCAATGATAATCACATCCGAAGTACCG	EH	16
GAAAATAAAAGAGATCGAAAAACTAAACGAACAAGT	EI	16
AACCATCTGATGGATCAATGGGGAATGGAGTTTTTAC	EJ	16
TTGTAGTCTGCTAGGTCATCAAACGAATCAATACCA	EK	16
AAGTTAGTTTCAAGGTGGATCGACTTCTTGAAATGG	EL	16
TGTATTGGTTAATAAATTTGCATTTAAATCTAAAAATACGA	EM	17
AATAGTTGTCAATTGGTTAGAACTTAGATCTAACGTAATAA	EN	17
TGTTACTTTTATGCAGATGATGTTTCAATATTGTAAATCGCTTGAGAAACTGAAT	**EO – revcom J**	**5**
AGTAACTCGCATGAATAGTATGTTTGAAGGATGTACTAGTTTTGAAGAGCTAGAC	**EP – revcom I**	**5**
ACGTATTTTGAAATCTGGATGGCATTTTCA	EQ	19
TCCGCTTAGCCGATGAAATTCCAAGCAATA	ER	19
TTTATTCACCGGACAAGAAACGACCTCGTC	ES	19
TCCCTATATCTATATATCGCTAAACATGTTAACTATAAGAGTTACCTGAGCAATGTTAAT TACAGAT	ET	20
CGTCTCTGGCAAAGCAAGTCTTAACTCTTCATCCGTAGGTGGTGGTGGAAATTCGAAT	EU	21
CGATAAAAAAGATATTAGTAATTTATTTGCAATTTA	EV	22
TGCTCTCTTTTGTGTAAAGCACATCAAGCATGTAGC	EW	22
GCGATTTTTGTCAAAGGGACAGCGATGGGTTACAA	EX	22
TTACCACCAAAGTCCCTACACTCAATACCACCAAAGC	EY	22
GAAAAAATAAAAAACCAATCGTATTCGA	EZ	23
GCTTTGGTGGTATTGAGTGTAGGGACTTTGGTGGTAA	**FA – revcom EY**	**22**
TTGTAACCCATCGCTGTCCCTTTGACAAAAATCGC	**FB – revcom EX**	**22**
GCTACATGCTTGATGTGCTTTACACAAAAGAGAGCA	**FC – revcom EW**	**22**
TAAATTGCAAATAAATTACTAATATCTTTTTTATCG	**FD – revcom EV**	**22**

Bold letters and sequences indicate reverse compliments of spacers that were identified previously in our data. Their reverse compliments are noted with "revcom – __".

Table A.3. Metadata of the 95 Isolates Analyzed

Isolates	96 Well Plate location	BioC Well	Persistence	Corrected PFGE	Store no.[a]
FSL R8-5081	A1	101	Persistent	CU-SNP1	2
FSL R8-5088	A2	102	Persistent	CU-SNP1	2
FSL R8-5095	A3	103	Persistent	CU-258-69	7
FSL R8-5230	A4	104	Persistent	CU-SNP3	2
FSL R8-5233	A5	105	Persistent	CU-SNP1	2
FSL R8-5449	A6	106	Persistent	CU-SNP3	2
FSL R8-5487	A7	107	Sporadic	CU-57-267	7
FSL R8-5543	A8	108	Sporadic	CU-SNP1	3
FSL R8-5584	A9	109	Persistent	CU-11-326	8
FSL R8-5805	A10	110	Persistent	CU-SNP2	23
FSL R8-5797	A11	111	Persistent	CU-SNP1	18
FSL R8-5844	A12	112	Persistent	CU-SNP2	23
FSL R8-6046	B1	113	Persistent	CU-294-321	10
FSL R8-6160	B2	114	Persistent	CU-SNP2	23
FSL R8-6176	B3	115	Persistent	CU-SNP3	28
FSL R8-6271	B4	116	Persistent	CU-SNP1	29
FSL R8-6317	B5	117	Persistent	CU-SNP1	2
FSL R8-6321	B6	118	Persistent	CU-57-267	10
FSL R8-6446	B7	119	Persistent	CU-SNP1	18
FSL R8-6480	B8	120	Persistent	CU-11-282	13
FSL R8-6649	B9	121	Persistent	CU-SNP1	2
FSL R8-6717	B10	122	Persistent	CU-57-267	10
FSL R8-6721	B11	123	Persistent	CU-SNP1	10
FSL R8-6765	B12	124	Persistent	CU-SNP2	23
FSL R8-6836	C1	125	Sporadic	CU-SNP3	24
FSL R8-7020	C2	126	Sporadic	CU-SNP1	12
FSL R8-7043	C3	127	Persistent	CU-57-267	10
FSL R8-7057	C4	128	Sporadic	CU-294-321	7
FSL R8-7121	C5	129	Persistent	CU-SNP1	29
FSL R8-7153	C6	130	Persistent	CU-11-282	13
FSL R8-7474	C7	131	Sporadic	CU-SNP1	11
FSL R8-7493	C8	132	Persistent	CU-SNP1	29
FSL R8-7534	C9	133	Persistent	CU-SNP1	16
FSL R8-7554	C10	134	Sporadic	CU-SNP1	16
FSL R8-7559	C11	135	Persistent	CU-262-318	21
FSL R8-7585	C12	136	Persistent	CU-SNP2	23
FSL R8-7599	D1	137	Persistent	CU-SNP2	23
FSL R8-7716	D2	138	Persistent	CU-SNP1	29
FSL R8-7722	D3	139	Persistent	CU-SNP3	29
FSL R8-7761	D4	140	Persistent	CU-294-321	16
FSL R8-7842	D5	141	Persistent	CU-SNP2	23
FSL R8-7926	D6	142	Persistent	CU-182-173	19
FSL R8-8106	D7	143	Persistent	CU-8-340	8
FSL R8-8110	D8	144	Persistent	CU-8-340	8
FSL R8-8439	D9	145	Persistent	CU-258-69	7
FSL R8-8453	D10	146	Persistent	CU-SNP1	16
FSL R8-8466	D11	147	Persistent	CU-SNP3	2
FSL R8-8477	D12	148	Sporadic	CU-SNP3	21
FSL R8-8481	E1	149	Sporadic	CU-SNP3	21
FSL R8-8505	E2	150	Persistent	CU-SNP1	24
FSL R8-8509	E3	151	Persistent	CU-SNP1	2
FSL R8-8514	E4	152	Persistent	CU-SNP3	2
FSL R8-8533	E6	154	Persistent	CU-SNP1	21
FSL R8-8665	E7	155	Persistent	CU-8-340	8
FSL R8-8728	E8	156	Persistent	CU-SNP3	2
FSL R8-8743	E9	157	Sporadic	CU-SNP1	7
FSL R8-8744	E10	158	Sporadic	CU-SNP1	7
FSL R8-8757	E11	159	Persistent	CU-SNP1	10
FSL R8-8765	E12	160	Persistent	CU-SNP1	10
FSL R8-8800	F1	161	Sporadic	CU-258-69	28

Table A.3 Cont

FSL R8-8850	F2	162	Persistent	CU-SNP1	24
FSL R8-8855	F3	163	Persistent	CU-SNP3	28
FSL R8-8867	F4	164	Persistent	CU-55-266	2
FSL R8-8876	F5	165	Persistent	CU-SNP3	2
FSL R8-8881	F6	166	Persistent	CU-55-266	2
FSL R8-8984	F7	167	Sporadic	CU-SNP1	29
FSL R8-9141	F8	168	Persistent	CU-8-340	8
FSL R8-9147	F9	169	Sporadic	CU-SNP1	29
FSL R8-9298	F10	170	Sporadic	CU-SNP1	7
FSL R8-9301	F11	171	Persistent	CU-258-69	7
FSL R8-9308	F12	172	Persistent	CU-SNP1	10
FSL R8-9314	G1	173	Sporadic	CU-SNP1	16
FSL R8-9353	G2	174	Persistent	CU-55-266	2
FSL R8-9354	G3	175	Persistent	CU-SNP1	2
FSL R8-9374	G4	176	Persistent	CU-SNP1	16
FSL R8-9383	G5	177	Persistent	CU-SNP1	21
FSL R8-9386	G6	178	Persistent	CU-SNP2	23
FSL R8-9404	G7	179	Persistent	CU-SNP1	16
FSL R9-0162	G8	180	Sporadic	CU-SNP1	7
FSL R8-5646	G9	181	Sporadic	CU-258-69	22
FSL R8-5726	G10	182	Sporadic	CU-258-69	27
FSL R8-6637	G11	183	Persistent	CU-55-266	2
FSL R8-6641	G12	184	Sporadic	CU-SNP1	2
FSL R8-7161	H1	185	Persistent	CU-SNP1	13
FSL R8-7348	H2	186	Sporadic	CU-294-321	2
FSL R8-7825	H3	187	Persistent	CU-SNP2	23
FSL R8-7833	H4	188	Persistent	CU-SNP3	23
FSL R8-8485	H5	189	Persistent	CU-SNP1	24
FSL R8-8495	H6	190	Persistent	CU-SNP3	28
FSL R8-8778	H7	191	Persistent	CU-SNP1	21
FSL R8-8789	H8	192	Persistent	CU-SNP2	23
FSL R8-9272	H9	193	Persistent	CU-SNP1	10
FSL R8-9285	H10	194	Persistent	CU-SNP1	2
FSL R9-0165	H11	195	Persistent	CU-SNP1	24
BLANK	H12	196	--	--	--
BLANK	--	197	--	--	--
BLANK	--	198	--	--	--
BLANK	--	199	--	--	--
EMPTY	--	200	--	--	--

[a] – Location number where the isolates were found