SEQUENCES IN \textit{COMK} PROPHAGE JUNCTION FRAGMENTS
CLUSTER ISOLATES OF \textit{LISTERIA MONOCYTOGENES} INTO
SUBCLONES THAT ARE UNIQUE TO INDIVIDUAL MEAT AND
POULTRY PROCESSING PLANTS: A POSSIBLE MODEL FOR
RAPID NICHE-SPECIFIC ADAPTATION

A Thesis in
Food Science

by

Mei Lok

© 2010 Mei Lok

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Master of Science

August, 2010
The thesis of Mei Lok was reviewed and approved* by the following:

Stephen J. Knabel  
Professor of Food Science  
Thesis Advisor

Edward G. Dudley  
Assistant Professor of Food Science

Bhushan M. Jayarao  
Professor of Veterinary Science

Yi Chen  
Research Microbiologist  
U.S. Food and Drug Administration  
Special Signatory

John D. Floros  
Professor of Food Science  
Head of the Department of Food Science

* Signatures are on file in the Graduate School
ABSTRACT

*Listeria monocytogenes* is a foodborne pathogen that can cause listeriosis, a potentially fatal disease. It is also a saprophyte found to persist in different environments including food processing plants; however the mechanism(s) responsible for this persistence remain(s) unknown. This research first confirmed that the presumptive epidemic clone II (ECII) isolates from FSIS and Eifert et al. (2005) were ECII by ECII PCR and multi-virulence-locus sequence typing (MVLST). Based on sequences in both upstream and downstream *comK* prophage junction fragments, most ECII and ECIII isolates and the Canada outbreak strains were grouped into plant-specific subclones. This finding strongly suggested that specific subclones of *L. monocytogenes* had evolved due to natural selection to adapt to individual meat and poultry processing plants. Moreover, the results suggested that *comK* prophage junction fragment sequences may represent excellent molecular subtyping markers with high epidemiologic concordance. Lastly, based on the observations of spontaneous induction of the *comK* prophage, the apparent packing of defective phage DNA into phage particles, and extensive recombination between *comK* prophages, a model for rapid niche-specific adaptation in *L. monocytogenes* was proposed. This model may explain the rapid evolution and niche specific adaptation of specific subclones in food processing plants. Further research is needed to unveil the roles of putative “adapton” genes within the *comK* prophage in the persistence of *L. monocytogenes* in food processing environments, and to validate the proposed model in the present study. Such information may enhance our understanding of how *L. monocytogenes* rapidly adapts to different environments, and will allow
accurate tracking of specific subclones and subsequent implementation of more effective intervention strategies for their control.
TABLE OF CONTENTS

LIST OF FIGURES ........................................................................................................... ix

LIST OF TABLES ............................................................................................................. xi

LIST OF ABBREVIATIONS ............................................................................................... xii

ACKNOWLEDGEMENTS ................................................................................................. xv

Chapter 1 Statement of the problem .............................................................................. 1

Chapter 2 Literature review ............................................................................................ 5

2.1 Listeria monocytogenes ............................................................................................ 5

2.1.1 Characteristics of L. monocytogenes and listeriosis .............................................. 5

2.1.2 A model organism for the study of host-pathogen and environment-saprophyte interactions ........................................................................................................ 6

2.1.3 Population structure of L. monocytogenes ................................................................ 7

2.1.3.1 Serotypes ........................................................................................................... 7

2.1.3.2 Lineages ............................................................................................................. 7

2.1.3.3 Brief descriptions of lineages and some serotypes that deserve special attention ........................................................................................................ 7

2.1.3.4 Epidemic clones of L. monocytogenes .............................................................. 9

2.1.3.4.1. Epidemic clone I ..................................................................................... 10

2.1.3.4.2. Epidemic clone II .................................................................................. 10

2.1.3.4.3. Epidemic clone III ................................................................................ 11

2.1.3.4.4. Epidemic clone IV ................................................................................. 11
2.2 Epidemiology of \textit{L. monocytogenes} .................................................. 12

2.2.1 Importance of epidemiologic investigation of listeriosis outbreaks... 12

2.2.2 Epidemiologic methods for investigating outbreaks.......................... 13

2.2.2.1 Conventional epidemiology ...................................................... 13

2.2.2.2 Molecular epidemiology .......................................................... 14

2.2.2.2.1 Important criteria for molecular epidemiologic methods.... 14

2.2.2.2.1.1 Performance criteria important for molecular epidemiologic methods ........................................ 14

2.2.2.2.1.2 Practical criteria important for molecular epidemiologic methods .................................................. 16

2.2.2.2.2 Molecular subtyping methods ............................................ 17

2.2.2.2.2.1 Fragment-based methods................................................. 17

2.2.2.2.2.2 Sequence-based methods................................................. 20

2.3 Persistent strains of \textit{L. monocytogenes} ........................................... 23

2.3.1 Persistence of strains of \textit{L. monocytogenes} in food processing and
retail environments .................................................................................. 23

2.3.2 Possible mechanism(s) responsible for the persistence of specific
strains of \textit{L. monocytogenes} in food processing and retail
environments .......................................................................................... 26

2.4 Prophages ............................................................................................ 35

2.4.1 Possible roles of defective prophages in increasing the fitness of the
bacterial host ............................................................................................ 37

2.4.1.1 Virulence factors carried within defective prophages.............. 38
2.4.1.2 Homologous recombinations between defective prophages confer selective advantage to the lysogenic host

2.4.2 The \textit{comK} prophage of \textit{L. monocytogenes} ........................................... 39

2.5 Conclusions ........................................................................................................ 44

2.6 References ........................................................................................................ 45

Chapter 3 Sequences in \textit{comK} prophage junction fragments cluster isolates of \textit{Listeria monocytogenes} into subclones that are unique to individual meat and poultry processing plants: A possible model for rapid niche-specific adaptation ........................................................................................................ 72

3.1 Abstract ............................................................................................................. 74

3.2 Introduction ....................................................................................................... 75

3.3 Materials and Methods .................................................................................... 84

3.4 Results .............................................................................................................. 90

3.5 Discussion ....................................................................................................... 95

3.6 Acknowledgements ......................................................................................... 110

3.7 References ..................................................................................................... 111

3.8 List of Figure Legends ................................................................................... 126

3.9 Figures ........................................................................................................... 130

3.10 Tables ......................................................................................................... 140

Chapter 4 Conclusions and future research .......................................................... 156
4.1 Conclusions ................................................................. 156

4.2 Future research .......................................................... 157

APPENDIX ................................................................. 160

APPENDIX A Prophage induction of ECII (H7858) and III (N3-031) with
Mitomycin C (MMC) ......................................................... 161
LIST OF FIGURES

FIG. 2.1. Electron micrograph of negatively stained A118 phage particles ............... 42

FIG. 2.2. Schematic diagram of phage A118 integration into and excision out of the *L. monocytogenes* chromosome at *comK* ................................................. 43

FIG. 3.1. Schematic diagram of phage A118 integration into and excision out of the *L. monocytogenes* chromosome at *comK* .............................................. 130

FIG. 3.2. PCR amplification of the 1998 and 2002 outbreak clone markers in *L. monocytogenes* ECII isolates .......................................................... 131

FIG. 3.3. PCR amplification of ECII isolates with primer pair 1/4, which targets the downstream junction fragment in the *comK* prophage of *L. monocytogenes* ...... 132

FIG. 3.4. Cluster diagrams based on upstream and downstream *comK* prophage junction fragment sequences in *L. monocytogenes* isolates described in TABLE 3.1 ...... 133

FIG. 3.5. Recombination breakpoints map for the upstream (A) and downstream (B) junction fragments in the different subclones of *L. monocytogenes* as detected by different algorithms implemented in RDP program ........................... 136

FIG. 3.6. Results of PCR amplifications of *L. monocytogenes* strain H7858 indicating spontaneous induction .............................................................. 137

FIG. 3.7. Genes located in the *comK* prophage in *L. monocytogenes* strain H7858 .... 138

FIG. 3.8. Proposed model for rapid niche-specific adaptation in *L. monocytogenes* .... 139
SUPPLEMENTARY FIG. 1. Effect of MMC concentration on the viability of *L. monocytogenes* strains H7858 (A) and N3-031 (B) incubated at 35°C in TSBYE for different lengths of time.
LIST OF TABLES

TABLE 2. Summary of reports that certain strains of *L. monocytogenes* can become persistent in various types of food processing and retail environments ...... 31

TABLE 3.1. Origin, establishment, source and date of isolation, and PFGE profiles of the thirty-eight *L. monocytogenes* isolates analyzed in the present study and two Canada outbreak strains included in the cluster analysis.................. 140

TABLE 3.2. Primer pair name, primer sequences, target and allelic locations within *L. monocytogenes* H7858, amplicon size and annealing temperature for each fragment analyzed in the present study ........................................ 145

TABLE 3.3. Results of prophage PCR using the method developed by Chen and Knabel .................................................................................................................. 147

TABLE 3.4. SNP locations in the upstream (A) and downstream (B) junction fragments that differentiated subclones of *L. monocytogenes* in the present study ... 148

TABLE 3.5. Detection of spontaneous induction of the *comK* prophage using *attP* and *attB* PCR, *comK* phage using int PCR, and plaque formation by *L. monocytogenes* strains F2365, H7858, N3-031 and 1001::A118 on the indicator *L. monocytogenes* strain 1001. ......................................................... 153

TABLE 3.5. Differences among three types of transductions.............................. 154

TABLE 3.7. Differences among three types of phage cycles.............................. 155
**LIST OF ABBREVIATIONS**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFLP</td>
<td>Amplified Fragment Length Polymorphism</td>
</tr>
<tr>
<td>bp</td>
<td>Base Pair</td>
</tr>
<tr>
<td>°C</td>
<td>Degree Celsius</td>
</tr>
<tr>
<td>CDC</td>
<td>Centers for Disease Control and Prevention</td>
</tr>
<tr>
<td>CFU</td>
<td>Colony Forming Unit</td>
</tr>
<tr>
<td>D</td>
<td>Discriminatory Power</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic Acid</td>
</tr>
<tr>
<td>E</td>
<td>Epidemiologic Concordance</td>
</tr>
<tr>
<td>EC</td>
<td>Epidemic Clone</td>
</tr>
<tr>
<td>FSIS</td>
<td>Food Safety and Inspection Service</td>
</tr>
<tr>
<td>g</td>
<td>gram</td>
</tr>
<tr>
<td>h</td>
<td>hour</td>
</tr>
<tr>
<td>log</td>
<td>logarithm (when preceding numbers means log to the base 10)</td>
</tr>
<tr>
<td>MLEE</td>
<td>Multilocus Enzyme Electrophoresis</td>
</tr>
<tr>
<td>min</td>
<td>minute</td>
</tr>
<tr>
<td>ml</td>
<td>milliliter</td>
</tr>
<tr>
<td>mM</td>
<td>micromoles per liter</td>
</tr>
<tr>
<td>MLST</td>
<td>Multilocus Sequence Typing</td>
</tr>
<tr>
<td>MLVA</td>
<td>Multiple-Locus Variable number of tandem repeat Analysis</td>
</tr>
<tr>
<td>MVLST</td>
<td>Multi-Virulence-Locus Sequence Typing</td>
</tr>
<tr>
<td>MMC</td>
<td>Mitomycin C</td>
</tr>
<tr>
<td>NCBI</td>
<td>National Center for Biotechnology Information</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>ng</td>
<td>nanogram</td>
</tr>
<tr>
<td>OC</td>
<td>Outbreak Clone</td>
</tr>
<tr>
<td>ORF</td>
<td>Open Reading Frame</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase Chain Reaction</td>
</tr>
<tr>
<td>PFU</td>
<td>Plaque Forming Unit</td>
</tr>
<tr>
<td>RAI</td>
<td>Rapid Adaptation Island</td>
</tr>
<tr>
<td>REP-PCR</td>
<td>Repetitive Extragenic Palindromic element-based PCR</td>
</tr>
<tr>
<td>PFGE</td>
<td>Pulsed-field Gel Electrophoresis</td>
</tr>
<tr>
<td>RAPD</td>
<td>Randomly Amplified Polymorphic DNA</td>
</tr>
<tr>
<td>RFLP</td>
<td>Restriction Fragment Length Polymorphism</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic Acid</td>
</tr>
<tr>
<td>rpm</td>
<td>revolutions per minute</td>
</tr>
<tr>
<td>RTE</td>
<td>Ready-To-Eat</td>
</tr>
<tr>
<td>RDP</td>
<td>Recombination Detection Program</td>
</tr>
<tr>
<td>s</td>
<td>second</td>
</tr>
<tr>
<td>SNP</td>
<td>Single-Nucleotide Polymorphism</td>
</tr>
<tr>
<td>spp.</td>
<td>Species (plural)</td>
</tr>
<tr>
<td>ST</td>
<td>Sequence Type</td>
</tr>
<tr>
<td>TB</td>
<td>Tryptose Broth</td>
</tr>
<tr>
<td>TEM</td>
<td>Transmission Electron Microscopy</td>
</tr>
<tr>
<td>TR</td>
<td>Tandem Repeat</td>
</tr>
<tr>
<td>tRNA</td>
<td>transfer Ribonucleic Acid</td>
</tr>
<tr>
<td>TSAYE</td>
<td>Tryptic Soy Agar with Yeast Extract</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------------------</td>
</tr>
<tr>
<td>TSBYE</td>
<td>Tryptic Soy Broth with Yeast Extract</td>
</tr>
<tr>
<td>µl</td>
<td>microliter</td>
</tr>
<tr>
<td>µm</td>
<td>micrometer</td>
</tr>
<tr>
<td>USDA</td>
<td>United States Department of Agriculture</td>
</tr>
<tr>
<td>V</td>
<td>Volt</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENTS

I would like to thank my parents, Sio Peng Chang and Kaiman Lok, for their deep love for me and their wisdom to guide me through the life journey.

I would like to express sincere gratitude to my mentor, Dr. Knabel for showing me the beauty of science and nurturing me to become an independent thinker.

I also want to thank my committee members, Dr. Dudley and Dr. Jayarao, and special signatory, Dr. Chen for their guidance during my research.

I would like to convey special thanks to Dr. Bindhu Verghese for her contributions to my research and her encouragement for my growth into an independent researcher.

I would like to thank my lab mates Fenyun Liu, Jia Wen, Michelle Orner, Carrie Pyle, Jabari Hawkins and Bek Yilma for making me enjoy my stay here. Also I would like to thank all the faculty, staff and graduate students in Food Science Department for their support.

I would like to thank USDA for providing financial support through a special Milk Safety Grant.
Chapter 1

Statement of the problem

Listeria monocytogenes is both a foodborne pathogen in animals and a common saprophyte in diverse environments, including food processing plants. It can cause listeriosis, a potentially fatal disease, especially among immunocompromised individuals, and is responsible for significant outbreaks and numerous costly recalls. Though much is known about L. monocytogenes as a pathogen, its saprophytic lifestyle remains relatively little understood. Specific subclones of L. monocytogenes are found to persist in food processing and retail environments, and subsequently contaminate ready-to-eat (RTE) foods manufactured in these facilities. Multiple mechanisms have been proposed to explain the persistence of specific strains in different processing plants, including niche-specific adaptation to different environmental surfaces, enhanced adherence to surfaces and formation of complex biofilms, inhibition of other bacteria in the same environment, increased resistance to unfavorable environmental factors, etc. However, which mechanism(s) is/are responsible for the persistence of L. monocytogenes is still unknown. L. monocytogenes epidemic clone II (ECII) was associated with large multistate outbreaks in the U.S. and was found to be temporally and geographically widespread in recent years. Presumptive ECII isolates, based on pulsed-field gel electrophoresis (PFGE) and multilocus genotyping (MLGT) profiles, were isolated by the Food Safety and Inspection Service (FSIS) of the United States Department of Agriculture (USDA) as a result of its various RTE meat and poultry products monitoring
programs (4). Additionally, *L. monocytogenes* isolates with ECII-specific genetic markers were also recently isolated from two turkey processing plants in the U.S. by Eifert et al. (1). However, whether or not the above isolates were truly ECII was not known. Also, it was not known whether or not the presumptive ECII isolates differed in their *comK* prophage sequences. Most recently, a cluster of listeriosis cases in Belgium between 2006 and 2007 were shown to be ECII (Dr. Thierry De Baere, personal communication). Several studies have utilized whole genome sequencing to analyze ECIII isolates from a 1988 sporadic case and a 2000 outbreak (3), as well as isolates from the 2008 Canada outbreak (2). As the ECIII sporadic case and outbreak isolates were only from one plant in Texas and the 2008 Canada outbreak isolates came only from one plant in Ontario, Canada, sequence divergence among isolates from multiple plants could not be determined in these studies. Also, whole genome sequencing revealed significant sequence variations in the *comK* prophages, compared to backbone genomes among epidemiologically related isolates. Appropriate interpretation of sequence variations in both the backbone genome and the *comK* prophage in the context of molecular epidemiology is needed in order to detect reservoirs, sources and transmission pathways of *L. monocytogenes*. Due to the fact that epidemiologically unrelated isolates were not included in the same study and markers with high discriminatory power and epidemiologic concordance were not identified, the contributions of whole genome sequencing to epidemiologic investigations are limited. Additionally, defective prophages in different bacteria were suggested to facilitate the adaptation of the bacterial host to rapidly changing conditions.
environments in previous studies. However whether the comK prophage also facilitates the adaptation of *L. monocytogenes* in diverse environments is still unknown. Therefore, this research first confirmed that the *L. monocytogenes* isolates from FSIS and Eifert et al. (1) were ECII, and then showed comK prophage junction fragment sequences of most ECII and ECIII isolates, and the Canada outbreak strains were specific to different food processing plants. comK prophage junction fragment sequencing might allow tracking specific subclones of *L. monocytogenes* to identify reservoirs, sources and pathways of transmission. Further investigation of the comK prophage might help us better understand its roles in enabling the colonization and transmission of *L. monocytogenes* in food processing environments. This information may allow food safety experts to develop more effective intervention strategies to control this dangerous foodborne pathogen.


Chapter 2

Literature review

2.1  *Listeria monocytogenes*

2.1.1 Characteristics of *L. monocytogenes* and listeriosis

*L. monocytogenes* is a gram-positive, intracellular foodborne pathogen which causes listeriosis, a potentially severe disease (57). Clinical manifestations of listeriosis range from non-invasive gastroenteritis in immunocompetent individuals, to more severe invasive forms in immunocompromised individuals (3). Invasive listeriosis includes sepsis, meningitis, rhombencephalitis, perinatal infections, and abortions (3). Listerialiosis has the highest hospitalization rate (89.1%) and one of the highest case-fatality rates (~13% on average) of all foodborne pathogens (31). In 2008 Canada outbreak, the case-fatality rate was as high as 40% (64). It is responsible for an estimated 2,500 human cases and 500 deaths annually in the United States (114). This foodborne pathogen infects not only humans, but also many wild and domestic animals, for example, mammals such as cows, sheep, pigs, etc and birds including crows, chickens, turkeys, etc (57). *L. monocytogenes* also lives as a saprophyte in various environments like soil, water, silage, sewage and processing plants (57). It is common in food processing plants and can grow at refrigeration temperatures, and thus is very common in many ready-to-eat (RTE) foods and causes numerous costly recalls. In order to protect consumers against listeriosis, most
countries have implemented and currently enforce a zero-tolerance policy for *L. monocytogenes* in RTE foods (152).

### 2.1.2 A model organism for the study of host-pathogen and environment-saprophyte interactions

Since its discovery in 1926, *L. monocytogenes* has been recognized as an excellent model in cellular immunology and infection biology over the past twenty years (44). Analysis of *L. monocytogenes*’ host-pathogen interactions has provided considerable insight into the intricate interactions between this bacterium and the immune system, especially how bacteria invade, adapt to and manipulate hosts, move intracellularly, and disseminate in tissues (147).

Recent research has shed light on previously unknown regulatory mechanisms in *L. monocytogenes*, including the controlled expression of a single virulence regulator PrfA, which is linked with nutrient availability (106), transcriptional regulatory networks (146) and coordinated global transcriptional changes during *L. monocytogenes* infection (154). Therefore, this opportunistic pathogen has been recognized as a paradigm for infection biology, as well as a tool to address fundamental processes in cell biology (71).

In contrast to the abundance of knowledge on *L. monocytogenes* as a pathogen, its saprophytic lifestyle remains little understood. Considering its ubiquity in various environments, *L. monocytogenes* may also be an excellent model for environment-saprophyte interactions. Most previous research in this field has focused on various types of stress responses in *L. monocytogenes* (150). However,
we need to better understand how it transmits to and colonizes various food processing environments, so that we can develop more effective intervention strategies for its control. To implement more effective intervention strategies, it is critical to first identify mechanisms that allow it to colonize specific niches in food processing plants.

2.1.3 Population structure of *L. monocytogenes*

2.1.3.1 Serotypes

One phenotypic method for grouping *L. monocytogenes* is serotyping, which is based on somatic antigens (1/2, 3 and 4) and flagellar antigens (a, b, c, d and e) (144). Based on various combinations of somatic and flagellar antigens, *L. monocytogenes* can be grouped into 13 known serotypes (144).

2.1.3.2 Lineages

The major lineages of *L. monocytogenes* were identified by multilocus enzyme electrophoresis (MLEE) typing (15). MLEE is a phenotypic method which differentiates isolates by the electrophoretic mobility of their major metabolic enzymes (15). Subsequent studies employing different genotypic methods, including ribotyping (165), high-density DNA array (52) and DNA sequence-based methods (134) have revealed similar clustering.

2.1.3.3 Brief descriptions of lineages and some serotypes that deserve special attention

Lineage I includes serotypes 1/2b, 3b, 4b 4d, and 4c. Serotype 4b is the most prevalent serotype in clinical, but not environmental isolates (152). Serotype
4b is associated with three of the four defined epidemic clones of *L. monocytogenes* (ECI, ECII and ECIV) (87, 152) and also most outbreaks not associated with epidemics (86). In contrast to serotype 4b, serotype 1/2b of Lineage I has not been linked to invasive listeriosis, though the two serotypes are closely related at the genomic level (48, 87, 139). Thus, the genomic comparison of Lineage I serotype 4b and 1/2b strains may help explain why serotype 4b was found to be predominant among invasive outbreak strains while no strains of serotype 1/2b have been associated with invasive outbreaks (41).

Lineage II consists of serotypes 1/2a, 3a, 1/2c, and 3c. Serotype 1/2a is most frequently identified among environmental and food isolates and exhibits the highest genetic diversity of all *L. monocytogenes* serotypes (87). The strains of this serotype were linked to ECIII (described below) and most recently a large outbreak in Canada in 2008 (4, 64). The Canada outbreak reportedly claimed 23 lives and was due to the consumption of contaminated deli meat products manufactured at a processing plant in Toronto, Ontario, Canada (4, 64). Additionally, a large proportion of isolates of serotype 1/2a and 1/2c of Lineage II expressed truncated internalin A (the gene product of *inlA*) (81). In contrast, all isolates of serotype 4b and 1/2b of Lineage I expressed full-length internalin (81). As truncated internalin would not be able to recognize the E-cadherin receptor, adequate uptake of the bacteria by the host cell may not be possible (81, 83). Such findings help explain the relatively low prevalence of serotype 1/2a and 1/2c in human disease.
Lineage III consists of serotypes 4a and 4c, as well as certain strains of serotype 4b (25, 52, 133-134, 140, 162, 165). However, the majority of serotype 4b isolates are in Lineage I (25, 52, 133-134, 140, 162, 165). Isolates of Lineage III are largely confined to animals (162, 165) and demonstrate variable virulence (100) and pronounced genetic diversity (41).

Disproportional distribution of serotypes among isolates of different sources has been observed. Serotype 4b is the most prevalent serotype in clinical isolates, and serotype 1/2a is the most prevalent serotype in environmental and food isolates. Though differences in virulence may account for the observed serotype distribution, attributes that enhance certain serotypes’ transmission, colonization and adaptation in food processing environments cannot be ignored (87). Additionally, higher sensitivity to selective enrichments might result in a serotype’s presence in foods and food processing environments being underestimated (87).

2.1.3.4 Epidemic clones of *L. monocytogenes*

An outbreak is defined as a cluster of cases of a disease caused by a source strain in excess of what is expected during a specified period of time (135). An epidemic can be defined as one or more outbreaks caused by an epidemic clone that survives and spreads over a long period of time (135). A clone has been defined as a group of isolates descended from a recent common ancestor which possesses similar genetic characteristics (131).

Many major outbreaks have been caused by the four epidemic clones (ECI, ECII, ECIII and ECIV) of *L. monocytogenes* (40). Two major epidemic clones
(ECII and ECIII) caused epidemics associated with post-lethality contamination of RTE foods, in which the growth of *L. monocytogenes* reached infectious level after refrigerated storage.

2.1.3.4.1. **Epidemic clone I**

Epidemic clone I (ECI) of serotype 4b, Lineage I was responsible for the 1981 Canada coleslaw outbreak, the 1985 California soft cheese outbreak, and the 1983-1987 Switzerland soft cheese outbreak caused by unpasteurized soft cheese (109). Isolates of the same serotype and multilocus enzyme electrophoretic subtype as ECI have been frequently isolated from animals, but only sporadically isolated in food and food processing environments (16). Additionally, the whole genome sequence of strain F2365 associated with the 1985 California soft cheese outbreak was published in 2004 (121). Comparative genomic analysis between the genome of this strain and the genomes of other *L. monocytogenes* strains may explain why ECI isolates are more associated with animals than food processing plants (16).

2.1.3.4.2. **Epidemic clone II**

In the U.S., epidemic clone II (ECII) of serotype 4b, Lineage I caused two large multistate outbreaks due to consumption of contaminated RTE hot dogs in 1998-1999 (87, 114) and turkey deli meats in 2002 (66). Isolates from these two outbreaks showed close relatedness based on the results of pulsed-field gel electrophoresis (PFGE) (67, 88), ribotyping (32), hybridization with genomic markers and DNA array (56, 88), and prophage PCR and sequencing (36). This was further confirmed by the finding of identical subtypes based on multi-virulence-
locus sequence typing (MVLST) (39). Additionally, open access to the genome sequence of the 1998 ECII outbreak strain H7858 made identification of genomic fragments specific to ECII possible. *L. monocytogenes* ECII is temporally and geographically widespread. In 2005, Eifert et al. reported isolates that harbored ECII-specific genetic markers were persistent in two turkey processing plant during the two-year study (55). Between 2006 and 2007, *L. monocytogenes* strains with similar AscI and ApaI patterns as the 1998 ECII outbreak clone were also isolated from a time-linked cluster of cases in Belgium (167).

### 2.1.3.4.3. Epidemic clone III

Unlike the other three epidemic clones, epidemic clone III (ECIII) is serotype 1/2a and belongs to Lineage II. ECIII caused a sporadic clinical case in 1988 (30) and an outbreak in 2000 due to contaminated RTE turkey products from the same meat processing plant in Texas, U.S.A. (128). Isolates from the sporadic case and the outbreak had identical ribotypes, AscI and ApaI PFGE subtypes (87). Therefore, these isolates were considered to be epidemiologically related.

### 2.1.3.4.4. Epidemic clone IV

Epidemic clone IV (ECIV) of serotype 4b, Lineage I, which includes isolates from the 1979 Boston vegetable outbreak, the 1983 Boston milk outbreak, the 1989 UK pate outbreak, and the 1997 northern Italian corn outbreak (8), was reclassified based on MVLST (40, Dr. Bindhu Verghese, personal communication). Other studies grouped these three outbreaks plus the 1983 Boston milk outbreak into ECIIa, based on multilocus enzyme electrophoresis and ribotyping (87) and multilocus
sequence typing (49). These subtyping methods have lower discriminatory power than MVLST. Multilocus genotyping (MLGT), a single-well DNA sequence-based SNP typing method utilizing 60 allele-specific probes, was also employed to subtype isolates from the 1979 Boston vegetable outbreak, the 1983 Boston milk outbreak, the 1989 UK pate outbreak (53). The results of MLGT suggested that strains from these outbreaks are sufficiently distinct from ECI (53), which agreed with the results of MVLST. The whole genome sequence of ECIV strain HPB2262 isolated from the 1997 northern Italian corn outbreak is available to the public. This will further facilitate investigations into the yet unclear phenotypic and genetic characteristics of ECIV.

2.2 Epidemiology of *L. monocytogenes*

2.2.1 Importance of epidemiologic investigation of listeriosis outbreaks

Public health officials place a high priority on epidemiologic investigations to identify outbreaks and stop them from incurring additional morbidity and mortality (152). However, a small proportion of listeriosis cases are associated with outbreaks; the rest are due to individual sporadic cases (152). Effective intervention strategies to prevent future outbreaks and sporadic cases can be developed if the reservoirs, sources, and pathways of transmission can be identified through epidemiologic investigation.
2.2.2 Epidemiologic methods for investigating outbreaks

2.2.2.1 Conventional epidemiology

Extensive data are collected during outbreak investigations by public health agencies using conventional and molecular epidemiology (148). An outbreak is likely to have occurred when the incidence of listeriosis observed is higher than what is expected during a specified period of time (135). In order to determine if the time-linked cluster of cases is really due to an outbreak, conventional epidemiology is conducted to determine if the outbreak can be traced back to a common food vehicle. This is done by reviewing the food history of case patients versus control people who do not show the symptoms of listeriosis. However, conventional epidemiologic investigations of listeriosis are particularly complex due to the intrinsic nature of listeriosis. Such features include: 1) the long incubation period (5-70 days) for invasive listeriosis, which makes it very difficult to obtain accurate food histories from case-patients and control people, and 2) the majority of listeriosis case-patients are immunocompromised, which makes it problematic to select appropriate immunocompromised control people for case-control studies (152). In order to overcome the limitations of conventional epidemiology, molecular subtyping of L. monocytogenes isolates from patients has been applied to rapidly detect and control outbreaks.
2.2.2.2 Molecular epidemiology

2.2.2.2.1 Important criteria for molecular epidemiologic methods

Molecular subtyping methods are evaluated based on two broad categories, performance criteria and practical criteria. Performance criteria are the most important and discussed first.

2.2.2.2.1.1 Performance criteria important for molecular epidemiologic methods

The performance criterion that has been most commonly used to evaluate different subtyping methods has been discriminatory power (D), which is defined as the probability that a typing system will assign a different type to two unrelated strains randomly sampled in the microbial population of a given taxon (151).

The equation for calculating discriminatory power is given below.

\[ D = 1 - \frac{1}{N(N-1)} \sum_{j=1}^{S} n_j (n_j - 1) \]

Where D is the index of discriminatory power; N is the number of unrelated strains tested; S is the number of different types; and \( n_j \) is the number of strains belonging to the jth type.

Another important performance criterion for molecular subtyping is epidemiologic concordance (E), which is defined as the probability that epidemiologically related strains derived from presumably single-clone outbreaks are determined to be similar enough to be classified into the same clone (151).

The equation for calculating epidemiologic concordance is given below.

\[ E = \frac{N_e}{N} \]
Where \( E \) is the index of epidemiologic concordance; \( N_e \) is the number of strains assigned to epidemic clones; and \( N \) is the number of strains tested from well-defined outbreaks.

Molecular subtyping methods with high epidemiologic concordance can allow outbreaks and epidemics to be quickly detected, terminated and prevented from happening again (152). It is thus important that epidemiologic subtyping methods target epidemiologically relevant markers and the speed at which the molecular markers undergo genetic changes should match the scope and purpose of the relevant study (38). However, in most studies epidemiologic concordance is often overlooked and more emphasis has been placed on discriminatory power. Therefore, epidemiologic concordance deserves more attention than it has previously received.

Besides discriminatory power and epidemiologic concordance, other performance criteria like typeability, reproducibility and portability are also important for evaluating a molecular subtyping method. Typeability is the proportion of strains that are assigned a type by a subtyping method; reproducibility is the ability of a subtyping method to assign the same type to a strain in independent assays; portability refers to the ease by which subtyping data can be exchanged electronically (38).

In conclusion, discriminatory power and epidemiologic concordance are the most important performance criteria for evaluating a subtyping method. High discriminatory power is often desired for a subtyping method to separate outbreak-
unrelated strains during epidemiologic investigations. However, high epidemiologic concordance is critical, because it enables a subtyping method to correctly cluster all outbreak-related strains and separate them from outbreak-unrelated strains. In the past, efforts were mainly focused on achieving high discriminatory power. Now it is generally accepted that a desirable subtyping method should yield both high discriminatory power and high epidemiologic concordance. Additionally, an ideal subtyping method should also have 100% typeability, 100% reproducibility and high portability. To be applicable to all strains of foodborne pathogens, subtyping results should be reproducible within the same laboratory and between different laboratories and should be highly portable.

2.2.2.1.2 Practical criteria important for molecular epidemiologic methods

In addition to the performance criteria discussed above, there are other criteria related to practical concerns, such as ease of use, high throughput and low cost (151). Simple and affordable molecular subtyping methods accessible to local health agencies would facilitate rapid identifications of local outbreaks and prevent them from developing into multistate outbreaks. High throughput methods that can analyze a large number of samples at the same time at low cost are also needed by state and federal health agencies in order to analyze large sample sets and thus speed up epidemiologic investigations (38).
**2.2.2.2 Molecular subtyping methods**

Various molecular subtyping methods have been widely employed to study the epidemiology of foodborne pathogens including *L. monocytogenes*, because these methods can differentiate strains more accurately at sub-species levels than phenotypic methods. Moreover, molecular subtyping methods can be applied to subtype cultures rendered non-viable due to prior administration of antimicrobial agents, which phenotypic methods fall short of (2). Fragment-based and sequence-based are two major categories of molecular subtyping methods.

**2.2.2.2.1 Fragment-based methods**

Fragment-based methods usually target certain genomic locations, utilize gel electrophoresis to separate digested or amplified DNA fragments, and then characterize strains by their unique banding patterns. Major fragment-based methods include the followings:

- **Restriction fragment length polymorphism (RFLP)**

  PCR-RFLP targeting virulence-associated genes can generate phylogenetically meaningful classifications of *L. monocytogenes* (158) and was subsequently confirmed by many other molecular subtyping methods. Phylogenetic information provides more useful insight into genomic divisions than the epidemiologic investigations of *L. monocytogenes* (158).

  Another method based on RFLP is pulsed-field gel electrophoresis (PFGE), in which the whole genome is targeted and digested by rare-cutting restriction enzymes (67). PFGE is currently the Centers for Disease Control and Prevention’s
(CDC’s) gold standard for subtyping *L. monocytogenes* and other foodborne pathogens (30). Additionally, PFGE is also the gold standard employed by PulseNet, the national molecular subtyping network of public health and food regulatory laboratories for foodborne disease surveillance (68). PFGE utilizes the entire genome, is universally applicable, and provides higher discriminatory power than other fragment-based methods (21, 68, 164). When compared with conventional epidemiology, PFGE sometimes was seen to produce too little or too much discriminatory power during listeriosis investigations (12, 40, 64). In some cases, isolates from the same outbreak may not have an identical PFGE subtype (141). In other cases, an identical PFGE subtype is sometimes detected in two samples that are not epidemiologically linked (61). Variations in PFGE patterns are sometimes due to dynamic chromosomal rearrangements in the homologous prophage regions as seen in subclutures of the same *Escherichia coli* O157:H7 isolate (80). When PFGE detects small genetic differences that are not epidemiologically relevant, the results are confounding, instead of epidemiologically meaningful (38). Moreover, PFGE sometimes suffers from poor reproducibility and the results are difficult to interpret and share via the internet (38). To sum up, though widely used, PFGE suffers from several drawbacks that limit its use for accurate epidemiologic investigation of listeriosis outbreaks.

- Repetitive element polymorphisms

  In repetitive extragenic palindromic element-based PCR or REP-PCR, the repetitive extragenic palindromic elements are amplified and the product-size
polymorphisms are visualized on a gel (42). Using a set of 128 *L. monocytogenes* isolates from human and various seafood sources, Chou et al. (42) showed that REP-PCR had similar discriminatory power to PFGE.

Another method based on repetitive element polymorphisms is multiple-locus variable number of tandem repeat analysis (MLVA). In this assay, tandem duplications of short DNA sequences within bacterial genomes are amplified with multiple primer pairs simultaneously (119). The sizes and numbers of tandem repeats (TRs) at each locus are then analyzed by computer (119). Since TRs are important in the adaptation of bacteria and targets of evolutionary events, MLVA has the potential to differentiate genetically unrelated strains with accuracy (38). MLVA has the additional advantages of ease of use, standardization, automation and interpretation when compared to PFGE (38). However, the typeability and stability of the assay do not seem to be satisfactory (38).

- **Multiplex PCR**

  Multiplex PCR allows simultaneous amplification of several targets of interest in one reaction by using more than one primer pair. Multiplex PCR has the advantages of low cost and simplicity (38). Therefore, it can be employed for screening large numbers of isolates during outbreak investigations (38). However, multiplex PCR lacks discriminatory power, which is needed for further epidemiologic investigation (38). A multiplex PCR assay developed by Doumith et al. was able to separate the four major *L. monocytogenes* serotypes (1/2a, 1/2b, 1/2c, and 4b) by utilizing various primer pairs (51). In order to provide a rapid method
for further subgrouping *L. monocytogenes*, Chen and Knabel (36) developed a multiplex PCR that allowed rapid screening for *Listeria, L. monocytogenes*, serotypes 1/2a and 4b, and EC I, II, and III of *L. monocytogenes* by including serotype markers identified by Doumith et al. (51).

- DNA probes

Some fragment-based methods utilize the hybridization of genomic DNA with a panel of subtype-specific DNA probes (38). Different strains may generate different hybridization banding patterns (38). Evans et al. (56) used Southern blotting to differentiate ECII from other serotype 4b strains. However, the genetic marker was negative in ECII isolates. This leads one to question the typeability of the method, or whether it can be applied to isolates with more diverse genetic backgrounds.

### 2.2.2.2.2 Sequence-based methods

Sequence-based methods have been developed based on the knowledge of whole genome sequences (38), in which single-nucleotide polymorphisms (SNPs) in the sequences of specific genes (37, 161, 168) or whole genomes (64, 129) are investigated. DNA sequence-based methods have gained popularity due to their high and unambiguous information content, high reproducibility and portability, increased time and cost efficiency, and ease of use. DNA sequences also contain phylogenetic information, thus they can be used for direct analysis of evolutionary relationships among different bacterial isolates (38). With the abovementioned advantages, sequence-based strategies are emerging as next generation subtyping.
methods with both high discriminatory power and high epidemiologic concordance for improving the molecular epidemiology of foodborne pathogens (63, 79).

Whole genome sequencing undoubtedly offers more detailed information than just sequencing certain genes. For example, DNA microarrays based on whole genome sequences identified many lineage- and serotype-specific genomic regions in more than 100 different test strains of *L. monocytogenes* (52). The identification of such regions provided a powerful tool for rapid tracing of listeriosis outbreaks (52).

Additionally, several studies have utilized whole genome sequencing to analyze ECIII isolates from a 1988 sporadic case and a 2000 outbreak (129), as well as isolates from the 2008 Canadian outbreak (64). As the ECIII sporadic case and outbreak isolates were only from one plant in Texas and the 2008 Canadian outbreak isolates came only from one plant in Ontario, Canada, sequence divergence among isolates from multiple plants could not be determined in these studies. Also, whole genome sequences can reveal SNPs among isolates from the same outbreak (64, 129). However, inappropriate interpretation of SNPs present in genes that do not affect epidemiology may confound epidemiologic investigations. Due to the fact that epidemiologically unrelated isolates were not included in the same study and markers with high discriminatory power and epidemiologic concordance were not identified, the contributions of whole genome sequencing to epidemiologic investigations are limited.
Therefore, in order to obtain high epidemiologic concordance in molecular subtyping, markers that play certain roles in causing outbreaks and epidemics should be selected. Such markers might include genes responsible for the transmission of the pathogen to the host and also virulence genes responsible for the growth and disease production in the host. Markers that fit this description can be sequenced to identify SNPs that help clarify the epidemiology of pathogens, including *L. monocytogenes*.

Multilocus sequence typing (MLST) based on housekeeping gene sequences was developed to study the population genetics of bacterial species (113). In order to improve the discriminatory power of MLST for subtyping *L. monocytogenes*, Zhang et al. (168) developed multi-virulence-locus sequence typing (MVLST), which was based on SNPs in virulence gene sequences. MVLST was able to accurately identify and differentiate the four epidemic clones of *L. monocytogenes* (ECI, ECII, ECIII and ECIV) and outbreak clones that were not involved in an epidemic, and sporadic case isolates, with both very high discriminatory power (D = 0.99) and epidemiologic concordance (E = 1.0) (40). This finding was confirmed by analyzing additional virulence genes and virulence gene regions in *L. monocytogenes* (107). In spite of the satisfactory performance in differentiating epidemic clones, MVLST was not able to discriminate very closely related outbreak clones within epidemic clones of *L. monocytogenes* (40, 107). To overcome this limitation, hypervariable regions within the *comK* and PSA prophages of *L. monocytogenes* were targeted (37). Sequence variations in the selected prophage
regions were able to accurately differentiate the known outbreak clones within ECII, ECIII and ECIV of this pathogen (37).

SNPs in virulence genes and prophage sequences appear to be excellent molecular markers for the investigation of epidemics and outbreaks caused by *L. monocytogenes* (36-37, 39-40, 168). Sequence conservation in virulence genes may be due to the critical functions virulence genes have in causing epidemics and the immune system response to *L. monocytogenes*’s invasion. Thus, virulence genes may be under strong negative selective pressure to remain unchanged (39-40, 168). On the other hand, certain prophage genes may be under positive selective pressure, as prophage genes are known to evolve rapidly via horizontal gene transfer and recombination (129). This may explain why prophage genes are more variable than both housekeeping genes and virulence genes in epidemic clones of *L. monocytogenes* (37).

2.3 Persistent strains of *L. monocytogenes*

2.3.1 Persistence of strains of *L. monocytogenes* in food processing and retail environments

Persistent strains of *L. monocytogenes* in food processing and retail environments can remain members of the established microbial flora for months or years, as seen in different countries (TABLE 2.1). Sample sites positive for persistent *L. monocytogenes* strains include both the product and the environment, particularly in the postprocessing environment (9-10, 12-14, 33, 46-47, 72, 76, 82,
significant example of persistent *L. monocytogenes* is ECIII of serotype 1/2a in a single meat processing plant in Texas, U.S.A. for over 12 years (30, 128).

Persistent strains of *L. monocytogenes* have been detected in plants that process different types of foods. It is not surprising to find persistent strains of *L. monocytogenes* in meat processing plants (9-10, 33, 110-111, 118, 122, 145), given that deli meats made from beef, pork and poultry are currently the RTE foods with the highest risk of causing listeriosis (5). Moreover, other types of food processing plants have shown the presence of persistent *L. monocytogenes* strains, including those that process poultry (9-10, 13-14, 33, 55, 108, 111, 116, 129, 136), dairy (9-10, 84, 156), seafood (9-10, 124, 137, 166) and vegetables (10) and also in retail environments (142-143). As a result of their dominance and persistence, these persistent (resident) strains of *L. monocytogenes* tend to contaminate a variety of foods processed and/or handled in these environments more frequently than sporadic (transient) strains (9-10, 13, 86, 110, 112).

Persistent strains may be introduced into the processing plants via raw products, where some persistent strains were detected (13-14). When raw products are processed and subjected to different conditions, some *L. monocytogenes* may be selected while others may be eliminated (159). Those selected *L. monocytogenes* persist in the plant, and may contaminate products during processing (13-14, 86, 110, 112), as the persistent strains are often isolated from surfaces of food.
equipment (machines, conveyor belts, cutting boards) or food environments (walls, floors, drains) (16, 82, 110, 136-137, 155, 159).

Despite the fact that most persistent strains of *L. monocytogenes* have not been implicated in illness (TABLE 2.1), the possibility of a more virulent strain becoming established in food processing or retail environments and contaminating RTE foods cannot be ignored. If such foods are eaten by a susceptible population, an outbreak can occur (155). Therefore, the food industry and governmental agencies must continue to treat all *L. monocytogenes* strains as potentially pathogenic (155).

Methods employed to subtype *L. monocytogenes* isolated from food processing and retail environments have evolved from phenotypic methods like serotyping and phage typing to more discriminating genotypic methods. These new methods have included pulsed-field gel electrophoresis (PFGE), ribotyping, restriction fragment length polymorphism (RFLP), amplified fragment length polymorphism (AFLP) and *actA* sequencing (TABLE 2.1). Ribotyping revealed two epidemic-associated ribotypes of serotype 4b Lineage I, which persisted in both dairy processing and retail environments (TABLE 2.1). One was ribotype DUP-1044A linked to the 1998 ECII outbreak (40). The other was ribotype DUP-1042B linked to ECIV, the 1983 Boston milk outbreak and the 2000 North Carolina soft cheese outbreak (40). These molecular subtyping methods provided much deeper insight into the microbial ecology of food operation environments and provided guidance for improving the control of *L. monocytogenes* (155).
The persistence of specific strains in different food processing plants is thought to be driven by niche-specific adaptation (54). Niche-specific adaptation within *L. monocytogenes* lineages was also suggested by Nightingale et al. (123) using cluster analysis to identify source-associated clades (human-, animal-, and food-associated). However, it was difficult to predict which genes were responsible for persistence. Therefore, the molecular mechanism(s) responsible for generating these persistent strains within a natural microbial community with myriad ecological interactions have remained obscure (89).

2.3.2 Possible mechanism(s) responsible for the persistence of specific strains of *L. monocytogenes* in food processing and retail environments

In order to eradicate *L. monocytogenes* in the food processing and retail environment, it is crucial to understand the mechanism(s) responsible for the persistence of specific strains. So far no specific evolutionary lineage seems to be associated with persistent strains, even though persistent strains have different subtypes than sporadic strains (9). Since different strains are established in different environments, different properties may be important for persistence in each particular niche (118). On the other hand, it is equally possible that some common properties critical for persistence are shared by various strains, which are being introduced to each environment and become niche-specific (118).

Likely reasons for persistence of *L. monocytogenes* include,

- Enhanced ability to adhere to food contact surfaces and form biofilms (17, 110, 112, 125). This mechanism is discussed in detail later.
• Continual re-introduction of the same strain over extended periods of time (9, 13-14, 33, 129). This might explain why some persistent strains were detectable only during processing operations, but not after cleaning (33).

• Resistance to disinfectants. This mechanism is also described in more detail later.

• The ability to compete for nutrients. Strains of *L. monocytogenes* unable to mount a stringent response and undergo physiological adaptation to nutrient deprivation were found to be impaired in surface-attachment and consequent colonization (153).

• Production of type E monocin and resistance to heavy metal. Production of type E monocin and resistance to heavy metal like cadmium were reported to occur more frequently in persistent strains than in sporadic strains of *L. monocytogenes* (72).

• The ability to grow and survive at low temperatures. This ability may be important for persistence of *L. monocytogenes* as strains show variability in cold stress tolerance phenotypes (6), but this remains to be determined.

• Long-term survival phase. The increased barotolerance and thermotolerance of *L. monocytogenes* upon extended incubation (163) might prevent *L. monocytogenes* in food from being completely destroyed by pressure or heat, generating persistent *L. monocytogenes* strains found in food.

• Resistance to antibiotics (72). Such characteristic might prevent *L. monocytogenes* in animals fed with antibiotics from being inactivated,
resulting in *L. monocytogenes* strains found in raw material.

- Resistance to drying (118).
- Combination(s) of the above (118) have also been proposed to explain persistence of specific strains of *L. monocytogenes* in specific environments.

Many studies speculated the most likely reason for the persistence of specific strains in processing plant environments is enhanced ability to adhere to food contact surfaces and form biofilms (17, 110, 112, 125), especially at short contact times (112). Bacteria adhere to food-contact-surface proteins within conditioning films (149). A common conditioning film may be produced by a common food with specific surface properties (20). Therefore, a common food might select for adherence and biofilm formation by a specific strain (20). Biofilms are formed when bacteria attach to surfaces and aggregate in a hydrated polymeric matrix of their own synthesis (45). From a mature biofilm, bacteria can become detached and contaminate food products (17). It is noteworthy that cells within biofilms are generally more difficult to remove mechanically and more resistant to hostile environment factors, like antimicrobial agents than planktonic cells (45, 118). In other words, the complex, stable, and yet undefined biofilm community (91) can affect survival and colonization of *L. monocytogenes* (13, 155).

The hypothesis that enhanced adherence and biofilm formation is critical for persistent strains is consistent with the suggested role of cell surface proteins in bacterial evolution. Cell surface genes are more associated with horizontal gene transfer (120) and cell surface proteins are under positive selection (130).
motor for this positive selection was suggested to be host-pathogen and environment-saprophyte interactions (132). Those selected cell surface proteins tend to enhance the ability \textit{L. monocytogenes} to survive in diverse environments (24).

However, some authors think differently about the impact of adherence and biofilm formation in bacterial persistence. Moreto and Langsrud (118) argued that the differences of adherence and biofilm formation observed between persistent and sporadic strains were of small magnitude and some sporadic strains even formed thicker biofilms than did persistent isolates. Tompkin (155) also stated that biofilm formation is less important for bacterial persistence than the fact that \textit{L. monocytogenes} survives in niches that are difficult to sanitize and where moisture and food debris are present. However, the hypotheses that niches or biofilms being important for persistence are not necessarily contradictory, since biofilm formation may lead to protection against disinfection and thus persistence (118). Given the above, other characteristics besides enhanced adherence and biofilm formation may also make persistent strains adapted to survive in food processing environments (118).

Resistance to disinfectants was also considered an attribute of some persistent strains as they are present before and after cleaning and disinfecting procedures (33). Resistance to disinfectants is probably due to changes in cell wall composition, which include cell surface proteins (115). However, opinions on the importance of resistance to disinfectants are mixed. On one hand, one study reported that resistance
to disinfectants was significantly higher among persistent strains of *L. monocytogenes* than sporadic strains (1). Additionally, *L. monocytogenes* developed resistance to disinfectants after regular exposure to sublethal concentrations (1). This might partially account for the persistence of *L. monocytogenes*. On the other hand, another study reported that persistence of *L. monocytogenes* was not necessarily related to resistance to disinfectants (110). In the latter study, the persistent strain showed significantly higher adherence to stainless steel surfaces than did the nonpersistent strains (110). Probably enhanced ability to adhere is a more important contributor to the persistence of *L. monocytogenes* than resistance to disinfectants. Nevertheless, it can be speculated that biofilm-forming *L. monocytogenes* strains that are resistant to disinfectants may have a competitive advantage for survival in food-processing environments (118).

In conclusion, all of the mechanisms mentioned above might explain the persistence of a common strain (genotype/subclone) in a processing plant that manufacturers the same food product over a long period of time, a phenomenon which has been observed in various types of food processing plants, including meat, poultry, dairy, seafood and vegetable processing plants (10, 118).
TABLE 2.1. Summary of reports that certain strains of *L. monocytogenes* can become persistent in various types of food processing and retail environments

<table>
<thead>
<tr>
<th>Environment and food type</th>
<th>Sample sites</th>
<th>Time of persistence</th>
<th>Country</th>
<th>Implicated in illness?</th>
<th>Subtyping method</th>
<th>Reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Processing environment</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Meat</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cooked meat</td>
<td>Environment</td>
<td>15 months</td>
<td>Finland</td>
<td>No</td>
<td>PFGE(^b)</td>
<td>(110)</td>
</tr>
<tr>
<td>Meat</td>
<td>Products and environment</td>
<td>2 years</td>
<td>Switzerland</td>
<td>No</td>
<td>Phenotyping, PFGE</td>
<td>(145)</td>
</tr>
<tr>
<td><strong>Poultry</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poultry and pork</td>
<td>Raw materials, products, environment, and equipment</td>
<td>Several months</td>
<td>France</td>
<td>No</td>
<td>PFGE</td>
<td>(33)</td>
</tr>
<tr>
<td>Poultry</td>
<td>Raw and cooked products, and environment</td>
<td>1 year</td>
<td>U.S.A</td>
<td>No</td>
<td><em>actA</em> sequencing</td>
<td>(14)</td>
</tr>
<tr>
<td>Poultry</td>
<td>Raw products, environment, and personnel</td>
<td>4 months</td>
<td>U.S.A.</td>
<td>No</td>
<td><em>actA</em> sequencing</td>
<td>(13), Dr. Mark Berrang, personal communication</td>
</tr>
<tr>
<td>Poultry</td>
<td>Raw materials, and</td>
<td>1 year</td>
<td>Spain</td>
<td>No</td>
<td>PCR-based</td>
<td>(108)</td>
</tr>
<tr>
<td>Category</td>
<td>Type</td>
<td>Duration</td>
<td>Location</td>
<td>Tested</td>
<td>Methods</td>
<td>References</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>-----------------------------</td>
<td>----------</td>
<td>-------------</td>
<td>--------</td>
<td>------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Poultry</td>
<td>Products and patients</td>
<td>16 months</td>
<td>Norway</td>
<td>No</td>
<td>MLEE, PFGE</td>
<td>(136)</td>
</tr>
<tr>
<td>Poultry</td>
<td>Products and environment</td>
<td>2 years</td>
<td>Finland</td>
<td>No</td>
<td>Serotyping, PFGE</td>
<td>(116)</td>
</tr>
<tr>
<td>Dairy</td>
<td>Environment</td>
<td>3 years</td>
<td>U.S.A.</td>
<td>No</td>
<td>Ribotyping</td>
<td>(76)</td>
</tr>
<tr>
<td>Cheese</td>
<td>Environment</td>
<td>10 weeks</td>
<td>U.S.A.</td>
<td>Yes</td>
<td>Ribotyping</td>
<td>(47)</td>
</tr>
<tr>
<td>Cheese</td>
<td>Environment</td>
<td>Over 2</td>
<td>U.S.A.</td>
<td>Yes</td>
<td>Ribotyping</td>
<td>(46)</td>
</tr>
<tr>
<td>Raw milk and nondairy foods</td>
<td>Products</td>
<td>Several months</td>
<td>Northern Ireland</td>
<td>No</td>
<td>MLEE, PFGE</td>
<td>(72)</td>
</tr>
<tr>
<td>Cheese</td>
<td>Products and environment</td>
<td>6 months</td>
<td>U.S.A.</td>
<td>Yes</td>
<td>Ribotyping, PCR-RFLP</td>
<td>(84)</td>
</tr>
<tr>
<td>Cheese</td>
<td>Products and environment</td>
<td>7 years</td>
<td>Scandinavian</td>
<td>No</td>
<td>PFGE</td>
<td>(156)</td>
</tr>
<tr>
<td>Category</td>
<td>Products and environmental</td>
<td>Duration</td>
<td>Country</td>
<td>Presence</td>
<td>Method(s)</td>
<td>Reference</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>---------------------------</td>
<td>------------</td>
<td>-----------</td>
<td>----------</td>
<td>-------------------</td>
<td>-----------</td>
</tr>
<tr>
<td>Seafood</td>
<td>Fish, smoked</td>
<td>14 months</td>
<td>Finland</td>
<td>No</td>
<td>PFGE</td>
<td>(82)</td>
</tr>
<tr>
<td></td>
<td>Raw materials, finished</td>
<td>4 months</td>
<td>U.S.A.</td>
<td>No</td>
<td>Ribotyping</td>
<td>(124)</td>
</tr>
<tr>
<td></td>
<td>Seafood</td>
<td>4 years</td>
<td>Norway</td>
<td>Yes</td>
<td>MLEE, PFGE</td>
<td>(137)</td>
</tr>
<tr>
<td></td>
<td>Fish</td>
<td>1.5 years</td>
<td>Denmark</td>
<td>No</td>
<td>RAPD</td>
<td>(166)</td>
</tr>
<tr>
<td></td>
<td>Fish, smoked</td>
<td>4 years</td>
<td>Denmark</td>
<td>No</td>
<td>RAPD, PFGE, AFLP</td>
<td>(159)</td>
</tr>
<tr>
<td>Retail environment</td>
<td>RTE and not RTE foods</td>
<td>5 years</td>
<td>U.S.A.</td>
<td>Yes</td>
<td>Ribotyping, PCR-RFLP</td>
<td>(142)</td>
</tr>
<tr>
<td></td>
<td>RTE deli meat and</td>
<td>8 – 19</td>
<td>U.S.A.</td>
<td>Yes</td>
<td>Ribotyping,</td>
<td>(143)</td>
</tr>
<tr>
<td></td>
<td>RTE deli salads</td>
<td>months</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Meat, poultry, dairy,</td>
<td>1 year</td>
<td>European</td>
<td>No</td>
<td>PFGE</td>
<td>(10)</td>
</tr>
<tr>
<td></td>
<td>fish, vegetables and</td>
<td></td>
<td>countries</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

33
<table>
<thead>
<tr>
<th>Products</th>
<th>1 year</th>
<th>Finland</th>
<th>No</th>
<th>PFGE, AFLP</th>
<th>(9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meat, poultry, dairy, fish</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **a**: Epidemic clone-associated ribotype is indicated in parentheses, following epidemic/outbreak clone information
- **b**: PFGE, pulsed-field gel electrophoresis
- **c**: MLEE, multilocus enzyme electrophoresis
- **d**: RFLP, restriction fragment length polymorphism
- **e**: AFLP, amplified fragment length polymorphism
2.4 Prophages

Lateral gene transfer is recognized as a major driving force in bacterial evolution, diversification and speciation (126). The significance of lateral gene transfer lies in that it can lead to formation of genomic “fitness islands” (blocks of DNA with signatures of mobile genetic elements whose functions increase bacterial fitness) and dramatic changes in bacterial behavior (69), as it allows bacteria to rapidly acquire a large number of genes (126). Additionally, lateral gene transfer has been demonstrated to account for the majority of intraspecies genome differences in some bacterial species (27).

Transduction is one type of bacterial lateral gene transfer, in which DNA is transferred from one cell to another via a bacterial virus, or bacteriophage (phage) (126). Transduction was shown to be a powerful source of “genetic novelties” in the evolution of cellular lineages based on informational protein phylogenies (59). For example, phages allow chromosomal rearrangements, insertions and deletions, probably via homologous recombination, within the same species like Escherichia coli (80, 90) and Xylella fastidiosa (157). Moreover, phages efficiently transfer large DNA fragments between distantly related bacteria (126). Thus phages are a likely motor for short-term bacterial evolution and diversification (11, 27-29).

Prophages are lysogenic phages that have been integrated (inserted) into bacterial host genomes (104). They are present in two-thirds of the sequenced low GC Gram-positive bacteria (28), including many strains of L. monocytogenes (70, 77). Significant sequence variation was observed in prophages compared to backbone genomes (bacterial host chromosomes not including prophages) in many Gram-positive and Gram-negative...
bacteria (6, 29, 97, 129). In other words, prophages are typically major contributors to
the genetic diversity between genomes of closely related strains of bacteria (27-28). One
possible reason for this observation is that homologous recombination between an
incoming phage and the prophage of the infected cell can potentially occur (27, 43, 117).
This phenomenon results in the capture of viral genes by cellular genomes (59) and the
rapid reshuffling of gene modules and/or gene sequences within the population (22, 27).
Therefore, prophages are often important players in the rapid evolution of their hosts.

When a phage infects a nonlysogenized cell, it can either kill the sensitive bacteria
(lytic cycle) or lysogenize the survivors (lysogenic cycle) (18). Lysogenic conversion
results in the survivors being resistant to superinfection by the same or related phages
(18). The reason for resistance to superinfection is that the lysogen synthesizes a
repressor which prevents transcription of most phage genes except the repressor gene
(98). Any transduced phage of the same type as the prophage will also be repressed. In
other words, its genes will not be expressed. Thus the transduced phage cannot enter the
lytic cycle (98). Due to resistance to superinfection, the fitness of the lysogen population
as a whole is enhanced as the outcome of the lysogenic conversion (18).

Any resulting recombinants and newly converted lysogens would then be
subjected to natural selection, resulting in a certain degree of co-evolution and mutualism
between the prophage and bacterium (50). Possible outcomes of this selection include
increased fitness of both the prophage and the lysogenic bacterium (50) and potentiated
horizontal transfer of virulence determinants (7, 73, 138). For example, increased
sensitivity to inducing signal(s) was suggested to have selected for the evolution of Shiga
toxin (Stx)-encoding prophages (101), as a greater fraction of lysogens with Stx-encoding
prophages were induced compared to lysogens with non-Stx-encoding prophages. Nevertheless, the fraction of induced lysogens with prophages carrying stx genes is still low. It is plausible that among lysogens with prophages carrying stx genes, a few have been sacrificed to give a growth advantage to the surviving lysogens with prophages carrying stx genes. The desired trait expressed by the dead lysogens is maintained in the population of survivors. Such phenomena are quite likely to result in increased fitness of both the host and prophage populations (101). However, much more evidence is needed to substantiate the role of transduction in bacterial evolution.

2.4.1 Possible roles of defective prophages in increasing the fitness of the bacterial host

Defective prophages can be found in different sequenced bacterial genomes, as a result of inactivating point mutations, inactivating DNA insertions or progressive DNA deletions (28, 50, 126). The presence of defective prophage can be explained over long-term evolutionary time scales, since fixing of entire prophages would not be expected as it increases the metabolic burden of extra DNA synthesis (27, 29). The genes that work optimally together to increase the fitness of the bacteria would remain in the genome (19, 74-75, 96). In fact, many phage-related genes that encode proteins that function as phage receptors or are important in the life cycle of phages were shown to be under positive selection, hypothesized to be driven by the dynamic co-evolutionary interaction between the prophage and host genome, similar to pathogen-host interactions (132). Therefore, the remaining phage-related genes are thought to enhance the fitness of the host; otherwise they would be rapidly lost (27-28, 93-94, 96, 103).
Besides phage-related genes, non-phage genes termed “morons” (for more DNA) or “lysogenic conversion genes” can be found within prophages (22). Morons may have been acquired by imprecise prophage excision from backbone genomes (bacterial host chromosomes not including the prophage) during specialized transduction as they usually locate near the prophage attachment site (23, 58). Morons have been postulated to change the phenotype or fitness of the lysogen (28).

Considering the possible roles of phage-related genes and morons, defective prophages could be a major source of novel cellular genes that enhance the fitness of the host bacterium in rapidly changing environments (7, 59, 70). Consequently, the ecological range of the bacterial hosts may extend to the body surfaces of new host species (i.e., domesticated animals and humans) and new food processing environments (22) in which bacteria survive and proliferate, or nutrient-depleted marine environment in which prophages enhance the fitness of the host by lowering its growth rate (35). To sum up, defective prophages are now deemed as active players in bacterial short-term evolution, rather than just passive cargo of bacterial chromosomes (28-29).

2.4.1.1 Virulence factors carried within defective prophages

Prophages play an important and, in some cases, a decisive role in the emergence of new bacterial pathogens and epidemic clones (22). In some cases, a single prophage encodes virulence factors related to the cause of a specific disease, as seen in *E. coli* O157, *Corynebacterium diphtheriae* and *Clostridium botulinum* (27, 60). In other cases, a multitude of prophages within the same genome encode various virulence factors, each making an incremental contribution to the virulence of the lysogen, as seen in *Vibrio cholerae, Staphylococcus aureus, Streptococcus pyogenes, and Salmonella enterica*
serovar Typhimurium (26-27, 50, 62, 78, 85, 127, 160). Prophages not only harbor virulence factors, but also potentiate their horizontal transfer (28-29, 50, 95, 104). Based on their recognized roles in pathogenicity, prophages have emerged as prime suspects in the adaptation of pathogens to new hosts (22).

2.4.1.2 Homologous recombinations between defective prophages confer selective advantage to the lysogenic host

As mentioned earlier, homologous recombination is possible between defective prophages or other genetic components in the same or in a different genome. In *E. coli* O157:H7, homologous recombination was seen between the Stx1 and Stx2 defective prophages, which encode potent cytotoxins (7). Homologous recombination was also seen in the nonpathogenic bacterium *Rhodobacter capsulatus*, after the defective-prophage-like particle transferred random segments of the donor cell’s genome to recipient cells (92). The genetic exchange process was suggested to be in response to growth-limiting environmental conditions (92). Homologous recombination was even possible between the prophages of the human pathogen *Streptococcus pyogenes* and those bacteriophages infecting dairy starters, despite the fact that their hosts are distantly related. The prophages of *S. pyogenes* were predicted to increase the ecological fitness of their lysogenic host. Therefore, homologous recombination between defective prophages in bacteria occupying different ecological niches in nature (69, 92) can confer a selective advantage to these lysogenic hosts confronting hostile environments (92).

2.4.2 The comK prophage of *L. monocytogenes*

The *comK* prophage of *L. monocytogenes* is formed by the integration of phage A118, a Sfi11-like *Siphoviridae* phage, into the 785-bp *comK* gene located in the host
chromosome (105). Phage A118 has a capsid diameter of approximately 60 nm, a long and rather flexible tail of roughly 300 nm (FIG. 2.1), and a circular genome of an average length of 43.4 kb (102). ComK is a putative transcriptional activator for various factors involved in competence for DNA uptake in Bacillus subtilis (105). However, its function in L. monocytogenes is unknown. In order to prove that phage A118 had indeed inserted within comK and was present in the L. monocytogenes chromosome as a prophage, Loessner et al. (105) developed primer sets that could specifically amplify the attP site within phage A118, the attB attachment site within comK in L. monocytogenes strain EGDe, and the upstream and downstream comK prophage junction fragments in L. monocytogenes lysogenic for A118, EGDe::A118, as illustrated in FIG. 2.2. Sequence comparison of the attP sequence in phage A118 and attB sequence in EGDe with the upstream and downstream junction fragment sequences from EGDe::A118 confirmed that phage A118 had indeed inserted within comK (FIG. 2.2).

The comK prophage is rendered defective (inactive) in L. monocytogenes by gene deletions and insertions after phage A118 integrates within comK (34). These deletions and insertions have created various comK prophages in different lineages and strains of L. monocytogenes. The comK prophage of EGDe differs from phage A118 substantially in nonstructural genes which include the lysogeny module (29). This might explain why A118 can be propagated on EGDe containing the comK prophage (65, 105). Prophages in L. monocytogenes ECII (121) and ECIII strains (121, 129) showed extensive diversifications from phage A118, in terms of both gene content and gene sequence. Additionally, strong expression of certain phage-related genes observed during
intracellular growth indicated some unknown role for the comK prophage (34). In short, many L. monocytogenes strains harbor defective comK prophage with unclear roles.

Spontaneous induction of the comK prophage has been seen in many strains of L. monocytogenes (104). After phage particles are formed, they are likely to transduce into other strains that also carry the comK prophage (129). Superinfection by the same or related phages would not lead to the lytic cycle (18, 98, 103), but probably homologous recombination described earlier (27). The resulting modular exchange may explain the relatedness of Listeria phages to viruses of other closely related Gram-positive bacteria (103, 105). Therefore, spontaneous induction, transduction of defective prophages, and recombination may be mechanisms by which various fitness factors are mobilized into L. monocytogenes, as seen in other pathogenic bacteria (21, 27-28, 50, 99, 105).

Extensive recombination was speculated to have contributed to the short-term genome evolution of L. monocytogenes (129). Approximately 10,000-fold more SNPs on a per nucleotide basis were seen in the comK prophage than the backbone genome in isolates evolved from L. monocytogenes ECIII over a 12-year period of time in a food processing plant (129). The extensive recombination throughout a large section of the comK prophage appeared to have occurred between ECIII serotype 1/2a isolates in Lineage II and a serotype 1/2b strain in Lineage I (129). Therefore extensive recombination may be possible between strains in different lineages of L. monocytogenes that carry the comK prophage (129). Given the above, it is reasonable to explain the numerous SNPs found within the comK prophage in ECII and ECIII as due to extensive recombination (37).
FIG. 2.1. Electron micrograph of negatively stained A118 phage particles (102).
FIG. 2.2. Schematic diagram of phage A118 integration into and excision out of the *L. monocytogenes* chromosome at comK (modified from Loessner et al. (105)). A. Phage A118 with *attP* attachment site. Horizontal arrows indicate PCR priming sites for amplifying the fragment containing *attP*. B. *L. monocytogenes* comK gene containing *attB* attachment site. Horizontal arrows indicate priming sites for amplifying comK with *attB* attachment site. C. Lysogenized *L. monocytogenes* showing the locations of the forward and reverse primers (horizontal arrows) for PCR amplification of upstream and downstream comK prophage junction fragments. The locations of the prophage PCR targets (LMOh7858_2426 and LMOh7858_2422) in the lysogenized strain (36) are shown with upward-pointing arrows.
2.5 Conclusions

*L. monocytogenes* is both a foodborne pathogen and a saprophyte in various environments, including food processing plants. Though much is known about *L. monocytogenes* as a pathogen, its saprophytic lifestyle remains little understood. Specific strains of *L. monocytogenes* are found to persist in food processing and retail environments, and subsequently contaminate RTE foods manufactured in these facilities. Multiple mechanisms have been proposed to explain the persistence of specific strains in different processing plants. However, which mechanism(s) is/are responsible for the persistence of *L. monocytogenes* is still unknown. Whole genome sequencing has been utilized to analyze isolates from epidemics and outbreaks. It revealed significant sequence variations in the *comK* prophages, compared to backbone genomes among epidemiologically related isolates. Due to the fact that epidemiologically unrelated isolates were not included in the same study and markers with high discriminatory power and epidemiologic concordance were not identified, the contributions of whole genome sequencing to epidemiologic investigations are limited. Additionally, defective prophages in different bacteria were suggested to facilitate the adaptation of the bacterial host to rapidly changing environments in previous studies. However whether the *comK* prophage also facilitates the adaptation of *L. monocytogenes* in diverse environment is still unknown. Therefore, this research investigated *comK* prophage junction fragment sequences of ECII and ECIII isolates, and the Canada outbreak strains to see whether they were specific to different food processing plants, and proposed a model based on the *comK* prophage for rapid niche-specific adaptation of *L. monocytogenes* in food processing environments.
2.6 References


Chapter 3

Sequences in *comK* prophage junction fragments cluster isolates of *Listeria monocytogenes* into subclones that are unique to individual meat and poultry processing plants:

A possible model for rapid niche-specific adaptation

Mei Lok¹, Bindhu Verghese¹, Yi Chen²,

Peter Evans³, Sophia Kathariou⁴ and Stephen J. Knabel¹*

1. Department of Food Science, The Pennsylvania State University, University Park, PA, 16802
2. Microbial Methods Development Branch, Division of Microbiology, Office of Regulatory Science, Center for Food Safety and Applied Nutrition, US Food and Drug Administration, College Park, MD, 20740
4. Department of Food Science, North Carolina State University, NC, 27695

Submitted to Applied and Environmental Microbiology on June 4, 2010

* Corresponding author. Mailing address: Department of Food Science, The Pennsylvania State University, University Park, PA 16802. Phone: (814) 863-1372. Fax: (814) 863-6132. E-mail: sjk9@psu.edu.
Unless stated otherwise, all research in this section was conducted by the author of the thesis.
3.1 Abstract

Numerous reports have shown that specific strains of *L. monocytogenes* persist in food processing plants and contaminate foods for many years; however, the mechanism(s) responsible for this persistence remain largely unknown. Presumptive epidemic clone II (ECII) isolates of *L. monocytogenes* from the Food Safety and Inspection Service of USDA and from Eifert et al. (31) were first shown to be ECII by ECII PCR and multi-virulence-locus sequence typing (MVLST). Based on sequences in both upstream and downstream *comK* prophage junction fragments, specific subclones of ECII, ECIII and the 2008 Canada outbreak strain were identified and shown to be specific to individual meat and poultry processing plants. Recombination analysis of upstream and downstream junction fragment sequences suggested that these subclones had evolved due to recombination within the junction fragments in the *comK* prophage. Comparison of *L. monocytogenes* sequenced genomes revealed five ORFs that are present in all *comK* prophage-containing genomes and four ORFs that are unique to serotype 1/2a strains. Based on the results of the present study, a model for rapid niche-specific adaptation was proposed. This model may explain how *L. monocytogenes* rapidly adapts to different food processing and retail environments to yield persistent subclones that frequently contaminate ready-to-eat (RTE) food products. Given their congruence with individual food processing plants, *comK* prophage junction fragment sequences may be excellent molecular markers for subtyping and tracking dangerous subclones of *L. monocytogenes* both between and within food processing plants and retail
establishments. This information can then be used to design more effective intervention strategies to reduce contamination and thus enhance food safety.

3.2 Introduction

Listeria monocytogenes is a unique foodborne pathogen which causes listeriosis, which ranges from febrile gastroenteritis to more severe life-threatening invasive diseases, especially for immunocompromised individuals (89). It is widely distributed in many wild and domestic animals and various natural environments and is resistant to a wide variety of environmental stresses (33). L. monocytogenes is considered a model organism for the study of host-pathogen interactions, especially as a model for intracellular pathogens of humans (44). However, it may also be an excellent model for environment-pathogen interactions, because it is well-known to cycle between being a pathogen in many wild and domestic animals and a saprophyte in diverse environments, including various types of food processing environments (41). As a saprophyte, L. monocytogenes lives on dead organic matter. The evolution of unique gene-regulation strategies is thought to allow this pathogen to maintain a balance between the above two different lifestyles (24, 41, 90). However, while much is known about the pathogenic lifestyle of L. monocytogenes, much less is known about its saprophytic lifestyle; including the genetic and phenotypic factors affecting its colonization and persistence in food processing and retail environments and subsequent transmission to foods.

Ready-to-eat (RTE) foods often become contaminated with L. monocytogenes after pasteurization due to transfer from contaminated food-contact surfaces (7, 56,
This has resulted in numerous costly recalls and occasional cases and outbreaks, which are often associated with high mortality (89). *L. monocytogenes* is known to colonize food processing plant environments, especially in harborage sites that are difficult to clean and sanitize (91), and form biofilms that protect them against various environmental stresses (9, 53, 67). Deli meats made from beef, pork and poultry are currently the RTE foods with the highest risk of causing listeriosis (2). Of those listeriosis cases and deaths attributed to deli meats, most are now thought to be associated with deli meats sliced at retail (1). Once contaminated, RTE meat and poultry products are known to support the relatively rapid growth of this psychrotrophic pathogen (55), which further enhances the risk of foodborne illness (1). As a result, a zero-tolerance policy for *L. monocytogenes* currently exists for RTE meat and poultry products and other RTE foods manufactured in the U.S.

Persistent (resident) strains of *L. monocytogenes* have been found repeatedly in various types of food processing plants, and are genetically distinct from those transient strains that are isolated sporadically (4-5). Persistent strains have been observed in plants that process meat (4-5, 37, 39, 62, 85), poultry (5-6, 31, 62, 76, 81), dairy (4-5, 36, 45, 49, 92), seafood (4-5, 72, 82), vegetables (5), sandwiches (7) and also in retail environments (84). As a result of their predominance and persistence, these resident strains of *L. monocytogenes* tend to contaminate foods processed and/or handled in these environments more frequently than sporadic strains (5). While many reports have clarified the role of numerous virulence genes and their coordinated expression in causing listeriosis, the mechanism(s) responsible for colonization and persistence of specific subclones of *L. monocytogenes* in food
processing and retail environments and subsequent transmission to foods and humans remain obscure. Lunden et al. (63) and Norwood and Gilmour (73) reported that persistent strains of *L. monocytogenes* showed enhanced adherence to food contact surfaces compared to non-persistent strains. Autio et al. (5) found that some pulsotypes of *L. monocytogenes* were recurrently found in the same product from the same producer, while other pulsotypes were recurrently found in products from multiple producers. Autio et al. (4) subsequently concluded that even though persistent strains differed from sporadic strains there did not seem to be any specific evolutionary lineage associated with persistent strains. Many authors have suggested that the most likely reason for persistent strains is that these strains have become better adapted to attach and form biofilms in specific processing plant environments (67). However, continual re-introduction of the same strain over extended periods of time, ability to grow and survive at low temperatures, competitiveness for nutrients, ability to mount a stringent response and undergo physiological adaptation to nutrient deprivation, resistance to sanitizers and/or heavy metals and antibiotics, formation of a stable ecosystem within biofilms or some combination of the above have also been hypothesized to explain the persistence of specific strains of *L. monocytogenes* in specific environments (67).

The population structure of *L. monocytogenes* is highly clonal with three distinct lineages (23, 70, 79-80) that appear to differ in their histories of horizontal gene transfer (66, 71, 77). Most clinical isolates fall into seven well-demarcated clonal complexes and three serotypes 1/2a, 1/2b and 4b, with each clonal complex containing a dominant serotype (80). Although rare among isolates from food
processing facilities, serotype 4b strains cause most outbreaks of listeriosis (89, 94). 
*L. monocytogenes* serotype 4b is also of most concern because strains of this serotype are associated with more severe clinical presentation and high mortality rates, and include three of the four defined epidemic clones of *L. monocytogenes* (ECI, ECII and ECIV), which have been responsible for multiple outbreaks worldwide (20-21, 30, 50, 89). The other epidemic clone, ECIII in serotype 1/2a, caused a clinical case in 1988 (17) and an outbreak in 2000 due to contaminated RTE turkey products from the same meat processing plant in Texas, U.S.A. (75). A 1/2a strain of *L. monocytogenes* was recently associated with a large outbreak in Canada due to consumption of contaminated deli meat products manufactured at a particular processing plant in Toronto, Ontario, Canada (37). Most recently, 1/2a strains of *L. monocytogenes* were associated with a listeriosis outbreak in Austria and Germany, due to consumption of ‘Quargel’, an acid curd cheese (36). Both of these 1/2a outbreaks had case-fatality-rates that were similar to the above 4b outbreaks, indicating outbreak strains of these two serotypes may be equally virulent.

In order to implement more effective intervention strategies to control *L. monocytogenes* we must first understand the epidemiology of this unique foodborne pathogen. Selecting the right molecular subtyping method for molecular epidemiology is pivotal for increasing our understanding of how *L. monocytogenes* is transmitted to foods (35, 89). Since molecular epidemiology involves the use of various types of molecular markers, selecting the right molecular marker and applying it to a relevant collection of strains are the keys to generating an accurate molecular epidemiology (21). Molecular markers for molecular subtyping of *L.
L. monocytogenes have included fragment-based markers such as restriction sites (40), repetitive elements (68) and various primer and probe annealing sites (19, 29, 32), and sequence-based markers such as sequences of genes or whole genomes and the subsequent identification of single nucleotide polymorphisms (SNPs) between different strains (20, 22, 76, 93). Discriminatory power (D) has traditionally been the main criterion used to evaluate different subtyping methods. The Centers for Disease Control and Prevention (CDC) has relied on PFGE for subtyping L. monocytogenes and other foodborne pathogens; however, in the case of L. monocytogenes this has sometimes produced too much discriminatory power (21-22, 76), which is often due to dynamic chromosomal rearrangements in homologous prophage regions (20, 48, 76-77). Based on both PFGE analysis (31, 39, 51) and whole genome sequencing (37, 76) genomic differentiation of L. monocytogenes appears to be taking place in established populations in meat and poultry processing facilities, yielding isolates with closely related, but distinct, whole genome profiles. Therefore, perhaps a more important criterion than discriminatory power for accurate molecular epidemiology is epidemiologic concordance (E) (88). It is critical that molecular subtyping methods have high epidemiologic concordance so that outbreaks and epidemics can be quickly identified, terminated and prevented from happening again (89). Genes that would be concordant with epidemiology would include those responsible for transmission of the infectious agent to the host and the virulence genes responsible for growth and disease production in the host. Chen et al. (22) demonstrated that multi-virulence-locus sequence typing (MVLST) of L. monocytogenes (97) possessed both very high discriminatory power (D = 0.99) and very high epidemiologic concordance (E = 1.0)
when applied to a large collection of outbreak and non-outbreak strains of *L. monocytogenes* (22). Both MVLST (22) and multiplex PCR (19) very accurately identified and differentiated the four epidemic clones of *L. monocytogenes* (ECI, ECII, ECIII and ECIV), outbreak clones that did not belong to a specific epidemic clone and non-outbreak isolates. The excellent epidemiologic concordance of MVLST was subsequently confirmed by Lomonaco et al. (61) using six additional virulence genes in *L. monocytogenes*. However, MVLST was not able to differentiate outbreak clones within epidemic clones of *L. monocytogenes* (22, 61). To overcome this limitation Chen and Knabel (20) targeted regions within the *comK* and PSA prophages of *L. monocytogenes* and designed primers that amplified sequences which differentiated the known outbreak clones within the epidemic clones of this pathogen. Orsi et al. (76) subsequently performed whole genome sequencing on Lineage II serotype 1/2a ECIII strains of *L. monocytogenes* from the 1988 sporadic case and the 2000 outbreak linked to RTE poultry products from the same processing plant in Texas. Their findings confirmed those of Chen and Knabel (20) that prophages in *L. monocytogenes* are excellent markers for differentiating outbreak clones within epidemic clones of this pathogen. The results of Orsi et al. (76-77) also revealed extensive recombination had occurred throughout the *comK* prophage between the 2000 serotype 1/2a ECIII strain in Lineage II and a serotype 1/2b strain in Lineage I. However, SNPs in the genome backbone of the 1/2a ECIII isolates did not accurately differentiate the 1988 sporadic case and the 2002 outbreak (76). Therefore, whole genome sequencing actually identified markers (SNPs) in the genome backbone that reduced epidemiologic concordance.
Lateral gene transfer is increasingly seen as a major driving force in bacterial evolution, because it allows bacteria to very rapidly acquire existing successful mutations, genes and gene clusters (74), which can lead to formation of genomic “fitness islands” and evolution in “quantum leaps” (42). Tailed phages probably constitute an absolute majority of “organisms” on the planet in sheer numbers (12) and allow efficient transfer of large DNA fragments between different bacteria (74). Most bacterial cells, including *L. monocytogenes* (58), contain prophage DNA which can interact with infecting phage (46). As a result, phage-prophage interactions are thought to be a dynamic force acting on bacterial populations (46) and account for a major portion of bacterial evolution that is occurring via lateral gene transfer (11, 15-16). Two-thirds of the sequenced low GC Gram-positive bacteria (15), including many strains of *L. monocytogenes* (15, 43, 47) contain identifiable prophages. Lysogenic phages which integrate their genomes into host genomes, or that exist in a carrier state for a long evolutionary period, could be a major source of novel cellular genes that enhance the fitness of the host (34, 43). Although most if not all *Listeria* strains carry functional or defective prophages, the potential influence of lysogeny on the host phenotype, including selective benefits, remains unknown (16, 28, 58, 76). Prophages typically constitute the major differences between genomes of closely related strains of bacteria (14-15), especially in the case of epidemic clones and outbreak clones of *L. monocytogenes*, which are highly clonal in terms of their backbone genome (genome excluding prophage) sequences (23, 70, 76, 79-80). Although *Listeria* phage genomes feature a conserved organization, they also show extensive mosaicism within the genome building blocks (28). The *comK* prophage of
*L. monocytogenes* can spontaneously induce and form phage, including phage A118, which is capable of forming plaques on host strains of *L. monocytogenes* (59). Loessner et al. (60) subsequently sequenced the entire phage A118 genome and developed primer sets that could specifically amplify the upstream and downstream *comK* prophage junction fragments, proving that phage A118 had indeed inserted within *comK* and was present in the *L. monocytogenes* chromosome as a prophage (FIG. 3.1). Chatterjee et al. (18) reported that after phage A118 inserts within *comK* it is rendered inactive (defective) by gene deletions. This is consistent with the case of ECII strains and ECIII strains (70), where the *comK* prophage showed extensive diversification from the original A118 insert in terms of both gene content and gene sequence (70, 76).

In the U.S., ECII serotype 4b caused large multistate outbreaks due to consumption of contaminated RTE hot dogs in 1998-1999 (65) and turkey deli meats in 2002 (38). *L. monocytogenes* isolates from the above two outbreaks were assigned to two distinct clonal clusters by AscI PFGE (39, 51), unique chromosomal markers (32, 51) and prophage PCR and sequencing (20). Eifert et al. (31) also isolated serotype 4b strains of *L. monocytogenes* with positive genetic markers typical of ECII from two turkey processing facilities in the U.S., and isolates with these markers were repeatedly isolated from one of these facilities over a two-year period. Most recently, Yde et al. (96) reported a cluster of cases in Belgium between 2006 and 2007 due to a strain of *L. monocytogenes* with AscI and ApaI patterns that were very similar to the 1998 outbreak clone of ECII. Further analysis using ECII PCR (19) confirmed that these strains were ECII (Thierry De Baere, Scientific Institute of Public Health,
Brussels, Belgium, personal communication). The above findings are consistent with other reports that epidemic clones are temporally and geographically widespread contaminants of RTE foods and processing facilities (30) and can persist in both food processing facilities (31, 94-95) and retail facilities (84) for extended periods of time. In the U.S. this scenario is especially problematic in the case of ECII outbreaks clones, because these clones have been associated with recent major multistate outbreaks and continue to be isolated by the Food Safety and Inspection Service (FSIS) of the United States Department of Agriculture (USDA) from establishments that manufacture RTE meat and poultry products (94).

FSIS is responsible for monitoring the safety of RTE meat and poultry products manufactured in the U.S. As part of this responsibility, FSIS conducts risk-based L. monocytogenes testing programs in establishments producing post-lethality exposed RTE meat and poultry products. FSIS gives special attention to deli meats and hot dogs, because they are known to become contaminated with L. monocytogenes after thermal processing and can allow the growth of L. monocytogenes to unsafe levels during the extended refrigerated storage (60-90 days) typical of these products. Thus, these products are in the very highest risk category for listeriosis among selected categories of RTE foods (2). As a result of its L. monocytogenes sampling program, FSIS has generated more than 500 L. monocytogenes isolates from different RTE meat processing facilities throughout the U.S. (94). A small fraction (~4%) of these isolates were ECII based on PFGE profiles and multilocus genotyping (MLGT) (94); however, these methods could not determine whether or not they represented different subclones of ECII. Eifert et al. (31) also isolated L. monocytogenes strains from two
different turkey processing plants from non-adjoining states in the U.S. that had ECII markers and PFGE patterns similar to the 1998 and 2002 *L. monocytogenes* outbreak clones (31). Therefore, the purposes of the present study were: 1) To test the hypothesis that the *L. monocytogenes* 4b isolates from FSIS and Eifert et al. (31) were not ECII using ECII PCR and MVLST; 2) To determine which subclones these serotype 4b ECII isolates, as well as the 1/2a ECIII and 2008 Canada outbreak isolates, belonged to, using both *comK* prophage PCR and upstream and downstream *comK* prophage junction fragment sequencing; 3) To analyze the sequences of the *comK* prophage junction fragments in different 4b and 1/2a strains of *L. monocytogenes* to determine if recombination is occurring in these fragments; and 4) To develop a possible model of how sequence variations in the *comK* prophage might be driving rapid niche-specific adaptation and persistence of *L. monocytogenes* in different meat and poultry processing plants.

### 3.3 Materials and Methods

**Bacterial isolates and DNA extraction.** Isolate identification numbers, sources and dates of isolation are given in TABLE 3.1. A total of ten isolates were obtained from the *Listeria* collection at CDC and eighteen isolates were obtained from FSIS of the USDA. For the ten isolates obtained from CDC, six were associated with two outbreaks of listeriosis involving ECII and four isolates were ECIII isolates from the same plant in Texas (two isolates associated with one sporadic case in 1988 and two isolates associated with one outbreak of listeriosis in 2000). The eighteen isolates obtained from FSIS were isolated by FSIS at different times from different meat
processing facilities as part of their *L. monocytogenes* testing programs. FSIS had determined that these isolates had similar PFGE and MLGT subtypes as previously characterized ECII isolates. Bacterial isolates were streaked on tryptic soy agar with 0.6% yeast extract (TSAYE) (BD, Franklin Lakes, NJ) with incubation at 35°C for 24 h. For each strain, one colony on the plate was inoculated into 10 ml of tryptic soy broth with yeast extract (TSBYE) and then incubated at 35°C overnight. Cultures grown overnight were adjusted to an optical density of 0.2 at 650 nm, which is equivalent to approximately 10^7 CFU/ml. For all isolates, bacterial genomic DNA was extracted using an UltraClean microbial DNA extraction kit (Mo Bio Laboratories, Solana Beach, CA) and stored at −20°C before use. Genomic DNA of ten isolates described by Eifert et al. (31) from two different turkey processing plants was provided by Sophia Kathariou at North Carolina State University.

**ECII PCR.** ECII PCR was performed with the following PCR program: hot lid, 106°C; initial denaturation at 94°C for 15 min followed by 31 cycles of 94°C for 30 s, 59°C for 1.5 min, and 72°C for 1.5 min, final extension at 72°C for 10 min. A total of 7.5 ng of purified genomic DNA was mixed with 3 µl of each primer (1 µM) (TABLE 3.2) and PCR master mix (Qiagen Inc, Valencia, CA) to achieve a final volume of 50 µl for each PCR amplification.

**MVLST.** Intragenic regions of the six virulence genes (*prfA, inlB, inlC, dal, clpP*, and *lisR*) were amplified as previously described by Zhang et al. (97).

**comK prophage PCR.** PCR amplifications of the hypervariable regions (LMOh7858_2426 and LMOh7858_2422) within the *comK* prophage were performed as previously described by Chen and Knabel (20).
Amplifications of the upstream and downstream junction fragments. Amplifications of the upstream and downstream junction fragments within comK were performed using methods modified from Loessner et al. (60). These modifications included redesigned primers (TABLE 3.2) and different cycling conditions. Briefly, a single PCR program (initial denaturation at 95°C for 15 min followed by 40 cycles of 94°C for 30 s, 63°C for 1.5 min, and 72°C for 1.5 min, final extension at 72°C for 10 min) was used. A total of 50 ng of purified genomic DNA was mixed with 1 µl of each primer (10 µM) (TABLE 3.2) and PCR master mix (Qiagen) to achieve a final volume of 50 µl for each PCR amplification.

Detection of PCR products. After amplification, PCR products were detected by gel electrophoresis at 110 V in a 1.5% agarose gel. Following ethidium bromide staining PCR products were visualized and photographed using a Model EC3 310 Imaging System (UVP, Upland, CA).

DNA sequencing. After detection by gel electrophoresis, all PCR products were purified using ExoSAP-IT® (USB Corp., Cleveland, OH) prior to sequencing. DNA cycle-sequencing reactions were performed at the Pennsylvania State University Shared Nucleic Acid Facility using an MJ Research Tetrad thermal cycler, 3′BigDye-labeled dideoxyribonucleotide triphosphates (v 3.1 dye terminators), and protocol 43032337 (Applied Biosystems, Foster City, CA). Cycle-sequencing reaction products were separated and analyzed on an ABI 3730XL DNA analyzer using the ABI Data Collection program (v 2.0). Data were analyzed with ABI Sequencing Analysis software (v 5.1.1). Both forward and reverse PCR primers were used as sequencing primers in separate runs. The sequences of two clinical isolates from the
2008 Canada outbreak were downloaded from NCBI Genome (http://www.ncbi.nlm.nih.gov/sitesgenome/).

**Sequence analysis.** Sequence analysis was performed on a total of 40 isolates in the present study (TABLE 3.1) using Lasergene (v8.1). Multiple sequence alignments were performed using molecular evolutionary genetic analysis software (MEGA version 4.0). Different sequence types were assigned to those isolates with at least a 1-nucleotide difference. MEGA 4.0 was used to construct cluster diagrams based on the upstream and downstream junction fragment sequences of *L. monocytogenes* isolates.

**Recombination analysis** (This analysis was conducted by Dr. Bindhu Verghese based on the sequences obtained by the author of the thesis). The extent of recombination within the upstream and downstream junction fragments of *L. monocytogenes* was determined using the Recombination Detection Program (RDP), GENECONV, Maximum Chi-square, Chimera, and Sister Scan recombination detection methods implemented in RDP v 3.38 (64). Since no single program provided optimal performance under all conditions, a stringent criterion was applied, such that any event for detection of positive recombination breakpoints had to be supported by five or more methods with $p \leq 10^{-5}$.

**Detection of spontaneous induction of the *comK* prophage.** Spontaneous induction of the *comK* prophage was detected by *attP* and *attB* PCR in twelve isolates from various outbreaks and processing plants in the present study (TABLE 3.1). Amplifications of the *attP* site within *comK* phage and the *attB* attachment site within *comK* were performed using methods modified from Loessner et al. (60). The
modifications included redesigned primers and cycling conditions as described above for amplification of the upstream and downstream junction fragments. A total of 50 ng of purified genomic DNA was mixed with 1 µl of each primer (10 µM) (TABLE 3.2) and PCR master mix (Qiagen) to achieve a final volume of 50 µl for each PCR amplification.

Detection of int in L. monocytogenes genomic DNA and comK phage DNA by PCR (The author of the thesis prepared the phage particles. Dr. Bindhu Verghese conducted subsequent phage DNA extraction and PCR). The phage integrase (int) gene in L. monocytogenes genomic DNA and comK phage DNA was targeted by PCR in strains F2365 (ECI), H7858 (ECII), N3-031 (ECIII) and 1001::A118. The ECI strain F2365 was isolated from Mexican-style soft cheese during the 1985 California outbreak. The ECII and ECIII strains are described in TABLE 3.1. Strain 1001::A118 is an artificially lysogenized serotype 1/2a strain. The cells were grown to early log phase (10^7 CFU/ml) at 35°C in 500 ml of tryptose broth (TB) supplemented with CaCl_2 to a final concentration of 10 mM. Cells were centrifuged at 12,000 rpm for 1 min and resuspended in 5 ml of fresh TB. Phage particles were isolated from the supernatant after overnight incubation at 35°C. Genomic DNA extraction was performed as described above. Phage DNA extraction was conducted as described by Asadulghani et al. (3). Briefly, phage DNA that was present in the phage head was detected by PCR amplification of int after DNase (Rockland, Gilbertsville, PA), RNase (EMD, Gibbstown, NJ) and proteinase K (Rockland, Gilbertsville, PA) treatment. PCR amplification conditions were 94°C for 15 min followed by 30 cycles of 94°C for 30 s, 62°C for 1.5 min, and 72°C for 1.5 min, final
extension at 72°C for 10 min. Phage DNA amplification was repeated at least three times for each isolate. The int primer sequences for phage detection are shown in TABLE 3.2. In order to see if there was genomic DNA contamination in phage DNA extraction samples, amplification of the virulence gene, prfA using the method described above was conducted on samples before proteinase K treatment and the final product of phage DNA extraction.

**Plaque assay.** To determine whether the prophages in strains H7858 (ECII) and N3-031 (ECIII) were defective, a plaque assay was performed on these strains. All media used in the plaque assay were supplemented with CaCl₂ to a final concentration of 10 mM. Cells were grown at 35°C in 180 ml of TB to early log phase (10⁷ CFU/ml). Cells were centrifuged at 12,000 rpm for 1 min and then resuspended in 1.8 ml of fresh TB. After overnight incubation in TB at 35°C, samples were centrifuged at 14,000 rpm for 10 min and the supernatant was filtered through a 0.2-μm sterile membrane filter (Thermo Scientific, Waltham, MA). 100 μl of filtrate was mixed with 100 μl of exponential-phase cells of the serotype 1/2a nonlysogenic indicator strain 1001 (provided by Martin Loessner, ETH, Zurich); the mixture was added to 5 ml of molten TB containing 0.75% agar at 50°C and then poured onto solidified TB containing 1.5% agar. Plaque formation was observed after overnight incubation at 30°C. The plaque assay was also performed on strains F2365 (ECI) and 1001::A118. Each plaque assay was replicated once.
3.4 Results

**ECII PCR.** Positive amplifications of the ECII marker (19) was observed, therefore the null hypothesis was rejected. All isolates from FSIS and Eifert et al. (31) were shown to be ECII (data not shown).

**MVLST.** Isolates from FSIS and Eifert et al. (31) had the same MVLST sequence type as ECII (22). *In silico* MVLST analysis of the two fully sequenced 2008 Canada outbreak isolates revealed that they had a MVLST sequence type that was different from ECIII, and thus were classified into their own unique sequence type in serotype 1/2a in Lineage II.

**comK prophage PCR.** The results of *comK* prophage PCR are summarized in TABLE 3.3. Two primer pairs (primer pair 2426 and primer pair 2664 in TABLE 3.2) targeting the prophage region LMOh7858_2426 were able to differentiate the 1998 and 2002 outbreak clones based on the sizes of the amplified products (20). Isolates from plants F, G, J, and O only had the 1998 ECII outbreak clone profile, the isolate from plant I had the 2002 ECII outbreak clone profile, and isolates from plant N had both the 1998 and 2002 ECII outbreak clone profiles. However, isolates from plants H, K, L and M showed no amplification with either of the above primer pairs (the amplification result of isolate from plant H is shown in FIG. 3.2). SNPs in sequences amplified with primer pair 2422 differentiated the 1998 and 2002 outbreak clones (20). Analysis of the sequences amplified with primer pair 2422 showed similar results as the prophage PCR targeting LMOh7858_2426 using primer pairs 2426 and 2664. Sequence analysis of the amplicons from PCR reaction using primer pair 2422 showed that isolates from plants F, G, M and O had the same sequence type as the
1998 ECII outbreak isolate H7858, isolates from plant N had the identical sequence type as another 1998 ECII outbreak isolate H7557, isolates from plant I had the identical sequence type as the 2002 outbreak isolates, the isolate from plant J did not share the same sequence type with any outbreak isolates in our collection. Isolates from plants H, K and L showed no amplification with primer pair 2422.

**Amplification of upstream and downstream junction fragments.** Positive amplifications of the ECII marker in isolates from plants H, K, L and M indicated the presence of genomic DNA (described above). PCR experiments targeting the upstream and downstream junction fragments were conducted to determine whether the negative amplification using the abovementioned prophage PCR primer pairs in these isolates was due to mismatching at priming sites or the absence of comK. Positive amplifications of the downstream junction fragment (FIG. 3.3) were seen in all isolates tested, including isolates from plant H, which had previously shown no amplification with prophage primer pairs 2426, 2664 and 2422 designed by Chen and Knabel (20). Nonspecific amplifications were observed using the upstream junction fragment and attP primer pairs designed by Loessner et al. (60). This was not surprising, because the primer pairs described by Loessner et al. were designed based on *Listeria* phage U153 of a serovar 1/2 strain. We next asked the question, how do the sequences in the upstream and downstream junction fragments vary among isolates? In order to obtain more specific amplification of these junction fragments, primer pairs were designed based on the published whole genome sequences of ECII and ECIII outbreak strains. These strains were H7858 (1998 ECII outbreak strain, accession number AADR00000000), F6900 (1989 ECIII sporadic case strain,
accession number NZ_AARU00000000), F6854 (1989 ECIII sporadic case strain, accession number NZ_AADQ00000000), J2818 (2000 ECIII outbreak strain, accession number NZ_AARX00000000) and J0161 (2000 ECIII outbreak strain, accession number NZ_AARW00000000). These primers for PCR amplification of the upstream and downstream junction fragments were defined as Upstream JF and Downstream JF in the present study (TABLE 3.2). Using these primer pairs amplification of the upstream and downstream junction fragments was observed among all isolates analyzed in the present study. Sequencing results for ECII isolates from plants K and M (both from NY) and plant L (from NJ) suggested that the comK prophage was present. However, more than one product was amplified using these junction fragment primer pairs (data not shown), even when cycling conditions were modified and alternative primer pairs targeting upstream and downstream junction fragments were employed. Therefore, these isolates could not be included in the cluster analysis in the present study.

Isolates with the same sequence type were assigned to the same subclone. SNPs within the junction fragments that differentiated the different subclones of L. monocytogenes are identified in TABLE 3.4. Sequence analysis was conducted to see what genes were present in the junction fragments. The upstream junction fragment contained partial comK from the backbone genome (genome minus prophage) and hypothetical protein 1 (HP1) and partial HP2 in the comK prophage. The downstream junction fragment contained partial integrase (int) of the comK prophage and partial comK from the backbone genome. The first 89 bp in the upstream junction fragment prior to comK and the last 103 bp of comK in the downstream
junction fragments contained no SNPs (TABLE 3.4).

Cluster diagrams were constructed based on the sequences of the upstream junction fragments (FIG. 3.4A), downstream junction fragments (FIG. 3.4B) and combined upstream and downstream junction fragment sequences (FIG. 3.4C). Isolates with an identical sequence type were assigned to the same subclone. Subclones 1 and 7 in the upstream junction fragment cluster diagram (FIG. 3.4A) were each split into two subclones in the downstream junction fragment cluster diagram to yield Subclones 1a and 1b, and 7a and 7b, respectively (FIG. 4B). In general, sequences in the downstream junction fragments were identical among isolates from individual processing plants, but different among isolates from different plants. The only two exceptions were Subclone 2 which contained isolates from plants G and O and Subclone 5 which contained isolates from plants B, C and I (FIG. 3.4). Cluster diagrams based on the downstream junction fragment (FIG. 3.4B) and the combined upstream and downstream junction fragment sequences (FIG. 3.4C) presented the same grouping but a different topology.

Recombination Analysis (This analysis was conducted by Dr. Bindhu Verghese based on the sequences obtained by the author of the thesis). Recombination analysis was conducted with the RDP v 3.38 program to evaluate the possibility of recombination events in the upstream and downstream fragments among isolates in current study. Three recombination events were found in the downstream junction fragment. Among them, only two recombination events (Subclones 1b/2 and 7a) were supported by at least four methods with high statistical significance \( p \leq 10^{-3} \) (FIG. 3.5). Four recombination events were detected in the upstream junction
fragment, and two of these recombination events (Subclones 2 and 3) were supported by five methods implemented in RDP program with very high statistical significance ($p \leq 10^{-5}$).

**Detection of spontaneous induction of the comK prophage.** Positive amplifications with $attB$ primer pairs were seen in all twelve isolates tested (FIG. 3.6 and data not shown). Positive amplifications with $attP$ primer pairs were seen in nine of the isolates tested, but not in isolates OB070181, OB080398 and OB080567 (FIG. 3.6 and data not shown). Comparison of the $attP$ and $attB$ sequences from strains H7858 and N3-031 with the corresponding upstream and downstream junction fragment sequences from these same strains confirmed that the fragments amplified were the $attP$ site within phage A118 and the $attB$ attachment site within comK.

**Detection of int in L. monocytogenes genomic DNA and comK phage DNA by PCR** (The author of the thesis prepared the phage particles. Dr. Bindhu Verghese conducted subsequent phage DNA extraction and PCR). Amplification of $int$ in genomic DNA was observed in *L. monocytogenes* strains H7858 (ECII), N3-031 (ECIII) and 1001::A118, but not in F2365 (ECI). Amplification of $int$ in comK phage DNA was observed in *L. monocytogenes* strains N3-031 and 1001::A118, but not in F2365 or H7858. Amplification of $int$ was confirmed by sequencing the amplicons. Negative amplification of *prfA* before proteinase K treatment and the final product of phage DNA extraction indicated the absence of genomic DNA contamination in the sample.
**Plaque assays.** Phage plaques were observed on plates containing filtrate from *L. monocytogenes* strain 1001::A118, but were not observed with ECI strain F2365, ECII strain H7858 or ECIII strain N3-031 (TABLE 3.5).

### 3.5 Discussion

The backbone genomes of *L. monocytogenes* strains are highly clonal, especially those in Lineage I (23, 70, 79-80). Within individual epidemic clones the backbone genomes are even more conserved. For example, Chen et al. (22) and Lomonaco et al. (61) could not find any sequence differences in 12 virulence genes within individual epidemic clones (ECI, ECII, ECIII and ECIV) of *L. monocytogenes*. Using an MLST scheme based on four housekeeping genes (*gap*, *prs*, *purM* and *ribC*), two virulence genes (*inlA* and *actA*) and one stress response gene (*sigB*) den Bakker et al. (26) also could not find any sequence differences within ECII or ECIII strains. Orsi et al. (76) used whole genome sequence comparison to analyze two 1988 sporadic case isolates and two 2000 outbreak isolates of ECIII that were associated with the same processing plant over a 12-year period. Only 11 total SNPs, six of which were non-coding and one which was synonymous, were found in the backbone genome between different strains of ECIII. Only a single SNP could differentiate the two 1988 ECIII isolates from the two 2000 ECIII isolates, and that SNP was in the tRNA Thr-4 prophage. Whole genome sequence analysis revealed sequence differences between two 2008 Canada outbreak isolates, including 28 SNPs and 3 indels (37). However, it is unlikely that these genetic differences would affect either transmission to the host or virulence within the host, because both isolates were involved in the
same outbreak. Therefore, these genetic differences would also not likely to be epidemiologically relevant. Therefore, similar to the results of Orsi et al. (76), the study by Gilmour et al. (37) also demonstrated that whole genome sequencing actually identified genetic markers that lacked epidemiologic concordance, and thus inappropriate interpretation of such markers may confound rather than clarify the molecular epidemiology of listeriosis outbreaks. In contrast, the ~40 kb comK prophage alignment in ECIII had 1,274 polymorphic sites that differentiated the two 1988 isolates from the two 2000 isolates (76). When comparing the backbone genome with the comK prophage, it can be seen that the comK prophage contained approximately 10,000-fold more SNPs on a per nucleotide basis than the backbone genome. This very large and disproportionate heterogeneity in the comK prophage within apparently so short a period of time (12 years or less) is very likely due to extensive recombination (76-77). This disproportionality between backbone genome and prophage sequence variation is also seen with many Gram-positive and Gram-negative bacteria (3, 15, 54). Most meat and poultry processing plants in the U.S. have been in existence for less than 100 years. Based on the observation that comK prophage junction fragment sequences differentiated subclones of the same epidemic clone, which had identical virulence gene sequence type, it is speculated that only the comK prophage and not the backbone genome of L. monocytogenes would appear to provide enough sequence variation to account for the rapid evolution (FIG. 3.4). Thus, rapid evolution was likely due to niche-specific adaptation in these plants within a short period of time.

In the present study, no amplifications were observed with primer pairs targeting
two internal prophage regions LMOh7858_2426 and LMOh7858_2422 in isolates from plant H. Therefore, cycling conditions were modified and primers targeting more conserved sites in LMOh7858_2426 and LMOh7858_2422 were employed. Despite these efforts, no amplifications were observed in the isolates from plant H. This observation raised the question, is the prophage absent in these isolates? Positive amplifications of the upstream and downstream junction fragments seen among all isolates indicated that the prophage was indeed present in all isolates in the present study (FIG 3.3). Therefore, failure to amplify the above internal prophage markers was likely due to extensive recombination having occurred within the comK prophage at these priming sites. Other possibilities, like SNPs at the priming sites and the integration of a different prophage cannot be eliminated. Full genome sequencing of these and other isolates described in the present study is currently underway in our laboratory (Knabel) to better characterize the true gene content and sequence divergence within both the backbone genomes and comK prophages among different L. monocytogenes isolates. This information will be used to design primers that can specifically amplify comK prophage junction fragments and other informative markers within the comK prophages of L. monocytogenes.

The comK prophage in L. monocytogenes may represent an excellent marker for differentiating outbreak clones of this pathogen (20). The junction fragments of the comK prophage in the 1988 and 2002 ECIII serotype 1/2a isolates from the same establishment in Texas had identical nucleotide sequences based on published complete genomes (76). This agrees with the results of the present study, where sequencing the upstream and downstream junction fragments also revealed sequence
identity between the 1988 and 2000 isolates from this poultry processing plant and thus assigned them to the same subclone, Subclone 6. *In silico* analysis of whole genome sequences of the two isolates from the recent 2008 serotype 1/2a outbreak due to consumption of RTE meats manufactured in Ontario, Canada revealed identical comK prophage junction fragment sequence, which were not shared by other subclones (37) and thus these isolates were also identified as a unique subclone (Subclone 3) in the present study. Analysis of the comK prophage junction fragments in ECII strains isolated from meat and poultry processing plants in the U.S. also revealed different subclones that were unique to individual processing plants, except in the case of Subclone 2, which was found in two processing plants, and Subclone 5, which was found in three processing plants (FIG. 3.4). Upstream junction fragment sequences assigned isolates from Est. N (state information not available) and Est. J in KS to Subclone 1 (FIG. 3.4A); however the downstream junction fragment separated Subclone 1 into Subclones 1a and 1b, and showed these two subclones were unique to individual establishments (FIG. 3.4B). Both the downstream and upstream junction fragments indicated that Subclone 2 contained isolates from establishment G in PA and one isolate from establishment O (state information not available). Interestingly, both establishment O (Sophia Kathariou, unpublished data) and establish G (FSIS, unpublished data) processed raw chicken. Transmission of raw poultry from establish O to establishment G might explain why Subclone 2 was found in both of these establishments. The 1998 ECII hot dog outbreak isolates were assigned to Subclone 4 by both the downstream and upstream junction fragments. Subclone 5 contained the 2002 ECII turkey deli outbreak isolates from establishment B in PA and
establishment C in NJ, and an environmental isolate from establishment I in IN. CDC had previous implicated both establishment B and establishment C in the 2002 turkey deli outbreak (38). Interestingly, establishments B and C were only approximately 30 miles apart and manufactured the same type of product, turkey deli meat, during the period of this outbreak. The isolation of Subclone 5 from establishment I in IN may have been due to some unknown mode of transmission of this subclone between establishments B and/or C and establishment I in IN. Such transmission has been hypothesized to occur due to meat or poultry, trucks, pallets, people, machinery, etc (20, 63, 91). Upstream junction fragment sequences assigned Subclone 7 to isolates from both establishment F in PA and establishment H in NC; however, downstream junction fragment sequences separated Subclone 7 isolates and assigned them to separate establishments; Subclone 7a isolates were assigned to establishment F in PA and Subclone 7b isolates were assigned to establishment H in NC. Looking just at the results in FIG. 3.4B, it can be seen that downstream junction fragment sequencing identified seven subclones that were unique to individual establishments and identified two additional subclones that were associated with two or three establishments. The strong congruence between the upstream and downstream junction fragment sequences and their correlation with specific establishments strongly suggests that these subclones were not evolving randomly, but may be the result of niche-specific adaptation mediated by genes located in both the upstream and downstream junction fragments and a “black block” region in Orsi et al. (76) that is immediately upstream of R3 (FIG. 3.7). The additional discrimination provided by the downstream junction fragment may be due to further niche-specific adaptation
being mediated by the downstream junction fragment. Interestingly, sequence analysis of the comK prophage revealed 5 open reading frames (ORFs) (LMOh2410, 2411, 2442, 2452 and 2475) located in the “black block” regions of Orsi et al. (76). These ORFs are present in all fully sequenced and publically available strains of L. monocytogenes that contain the comK prophage and showed sequence identity between the 1988 ECIII sporadic case isolates and the 2000 ECIII outbreak isolates (76). In addition, 4 other ORFs (LMOf2338, 2372, 2375 and 2377), which are unique to publically available and fully sequenced 1/2a strains and not present in the phage A118 genome or ECII can also be found in “black block” regions of Orsi et al. (76). Analysis of their dn/ds ratios in the present study (data not shown) indicated many of these ORFs were under positive selection.

Taken together, the above findings suggest that the subclones identified in the current study are the result of recombination followed by both negative (purifying) selection within processing plants and positive (diversifying) selection between different processing plants. The fact that some subclones are found in multiple establishments may be due to a common food source in these establishments. This would have been the case in the 1988 ECIII sporadic case and the 2000 ECIII outbreak, where RTE turkey products were manufactured in establishment D in Texas over a twelve year period of time, and also the case in the 2002 ECII outbreak, where both establishment B and establishment C were manufacturing the same RTE turkey deli meat product, cooked turkey deli meat, during this outbreak. A common food in different establishments would likely produce a common conditioning film (10) that might select for attachment and biofilm formation by a common subclone. This is
consistent with Nakamura et al. (69) who concluded that the biological functions of horizontally transferred genes are biased toward cell surface-related functions and with Cabanes et al. (13) who stated that the abundance of surface proteins in \textit{L. monocytogenes} is probably related to the ability of this bacterium to survive in diverse environments. It is also consistent with Petersen et al. (78) who reported that positive selection targets cell surface proteins, and with the results of Smoot and Pierson (87) who reported that addition of trypsin to an attachment medium resulted in a 99.9\% reduction in the adhered cell population of \textit{L. monocytogenes} when compared to controls. Different bacteria are known to adhere to different surface proteins within conditioning films and then go on to form biofilms that protect them against hostile environments (25). All of the above would explain the persistence of a common subclone in a processing plant that manufacturers the same food product over a long period of time, a phenomenon which has been observed in various types of food processing plants, including meat, poultry, dairy, seafoods and vegetable processing plants (5, 67, 82). Given the above, we speculate that the extensive genetic heterogeneity seen within the \textit{comK} prophage may result in surface protein structures that allow different subclones of \textit{L. monocytogenes} to attach to different types of food-conditioning films in different processing plants. This is analogous to the situation with another Gram-positive pathogen, where lysogeny of \textit{Streptococcus pneumoniae} with MMI phage led to improved adherence of this Gram-positive pathogen to inert surfaces and pharyngeal cells (57). Further research is needed to determine if the \textit{comK} prophage is playing a similar role in adherence and persistence of \textit{L. monocytogenes} on various surfaces in food processing plants.
One common theme that may help explain the results of the present study is the processing and shipment of raw poultry. Finding Subclone 1a and Subclone 2 in raw poultry processing plants, finding Subclone 2 in a RTE meat and poultry plant, and finding Subclone 4 and Subclone 5 isolates in RTE poultry processing plants indicates that raw poultry may be the original source for many of the subclones detected in the present study. Many of the ECII subclones identified were from processing plants from states in the eastern U.S. (NC, PA and NJ), many of which process raw poultry and/or RTE meat and poultry products. Similarly, an outbreak in Texas due to ECIII associated with the 2000 outbreak was linked to cooked turkey deli meat (75). Also, ECII was isolated from cooked turkey deli meat in the plant in MI that was associated with the 1998 hot dog outbreak (39). Based on the above, we speculate that ECII and ECIII of *L. monocytogenes* may have first evolved in raw poultry processing environments and then underwent further natural selection and niche-specific adaptation once they were transmitted to RTE meat and poultry processing plant environments. Further research is needed to test this hypothesis.

The finding of different subclones in different food processing plants is consistent with the ecotype concept of microbial evolution (52), which is thought to be driven by niche-specific selection (30). An ecotype has been defined as an ecologically distinct group whose diversity is limited by a force of cohesion, usually the genome-wide purging of diversity (52). Unfortunately it is difficult to predict the genes responsible for this ecological divergence. However, a DNA-sequenced based approach which targets ecologically relevant markers can overcome these challenges. In the present study, ecologically distinct subclones could be identified by sequencing
the comK prophage junction fragments which contained many of the genes that are common to all fully sequenced strains of L. monocytogenes that contain the comK prophage (FIG. 3.7). Interestingly, these genes are located at the terminal ends of the comK prophage (FIG. 3.7). Perhaps the terminal ends of the prophage provide a balance between variation due to recombination with other comK phage genomes and also conservation due to their close proximity to the highly conserved genome backbone. This scenario might produce novel gene sequences that are stable enough not to be rapidly eliminated by the extensive recombination seen within the comK prophage. This hypothesis is consistent with the fact that no SNPs were present in comK at the terminal ends of the upstream (nucleotide (nt) 0-89) and downstream (nt 581-673) comK prophage junction fragments nearest the genome backbone (TABLES 3.4A and 3.4B). Boerlin and Piffaretti (8) concluded that ECI strains of L. monocytogenes appeared to be host-adapted to animals, which might explain why outbreaks due to this epidemic clone were more closely associated with raw animal products than with food processing plant environments (50). This is consistent with the fact that ECI strains lack the comK prophage (70, TABLE 3.5), and is also consistent with the proposed role for the comK prophage in niche-specific adaptation of different subclones (ecotypes) in different processing plant environments proposed in the present study.

The above scenario begs the question as to how rapid recombination within the comK prophage might be occurring. Given the evidence for extensive recombination within prophages in the present and other studies (3, 20, 76-77), we speculate that transduction of the entire comK prophage might be mediating this rapid and extensive
recombination. In the present study the comK prophages in *L. monocytogenes* ECII and ECIII were shown to spontaneously induce, as positive amplifications of attP and attB were seen in all lysogenized isolates analyzed (FIG. 3.6 and data not shown). The results in TABLE 3.5 demonstrate that the comK phage spontaneously induced from ECIII strain N3-031 and 1001::A118 and the comK phage gene int was amplified after DNase and proteinase K treatments, supporting the conclusion that the comK phage genome was packaged within phage particles by these strains. These findings are consistent with those of Loessner et al. (59), which also revealed spontaneous induction of the comK prophage and subsequent phage formation.

Spontaneous induction of defective comK phage would likely result in transduction of the entire comK prophage. The transduction of defective comK phage has three principal differences compared to generalized and specialized transduction, 1) Type of DNA transduced; 2) Frequency of packaging and 3) Extent of homologous recombination (TABLE 3.6). Regarding type of DNA transduced, in transduction of defective comK phage, the entire comK prophage is transduced. In contrast, in generalized transduction fragments of chromosomal DNA are transduced, and in specialized transduction, only the ends of chromosome and viral DNA are transduced. The frequency of packaging refers to the percentage of phages that contain a specific type of DNA among all phages formed. The frequency of packaging for transduction of defective comK phage is close to 100%. This number is much higher than the frequency of packaging for generalized or specialized transduction (~10^-4) (58). The extent of recombination refers to the extent of transduced DNA which recombines with the chromosomal DNA of the recipient cell. Due to the presence of homologous
regions between transduced phage DNA and prophage DNA in the transduced host, the extent of homologous recombination during transduction of defective comK phage and specialized transduction would tend to be very high. In contrast, the extent of homologous recombination during generalized transduction tends to be low due to the absence of homologous regions between fragments of chromosomal DNA packed within phage particles and prophage DNA in the transduced host. The transduced defective comK phage could undergo frequent and large-scale homologous recombination with the comK prophage in recipient cells of L. monocytogenes, which would result in the rapid reshuffling of gene modules and/or gene sequences within the population (3, 11, 14, 76-77). The proposed mechanism of induction, transduction and recombination of defective comK phage would also be different than the lytic or lysogenic cycles of infective phage replication (TABLE 3.7), because induction, transduction and recombination of defective comK phage might be occurring for the purpose of rapid evolution of L. monocytogenes, not phage propagation. Therefore, induction, transduction and recombination of defective comK phage might be under the control of the bacterial host (11, 46). Many prophage genes from other sequenced bacterial genomes show inactivating point mutations, inactivating DNA insertions or progressive DNA deletions leading to defective prophages which cannot form infective phage, thus genes in defective prophages tend to be rapidly lost if they do not increase the fitness of the host bacterium (11, 15, 74). The above scenario is consistent with what is currently known about defective prophages, which are now seen not as passive cargo of bacterial chromosomes, but instead as important targets for selection working on bacterial genomes (16).
Defective prophages are also known to be especially important for rapid short-term evolution and adaptation of bacteria to rapidly changing environments (11, 15-16). Such environments may include the body surfaces of new host species that are rapidly proliferating in the ecosphere (i.e., domesticated animals and humans) and new food processing environments (11). On the basis of simple Darwinian reasoning a certain degree of co-evolution and mutualism between prophage and bacterium was predicted, where the expression of lysogenic conversion genes from prophages were postulated to increase the fitness of both the prophage and the lysogenic bacterium in various ecological niches (46). This has been observed in many pathogenic bacteria, where various virulence factors are now known to be carried within defective prophages inserted into bacterial chromosomes (3, 11, 15-16). Many of the phage-encoded virulence factors are located on the bacterial cell surface to help pathogens adapt to rapidly changing surfaces inside the human host (11). Defective prophages are known to transduce virulence factors between pathogens where they can either insert into or recombine with the host chromosome. For example, Ruzin et al. (83) described how Φ80α mobilized a pathogenicity island containing the gene for toxic shock toxin in Staphylococcus aureus and transduced it into other strains of S. aureus. Recently, Asadulghani et al. (3) reported that defective prophages in E. coli O157:H7 released as particulate phages, which subsequently transduced to other strains of E. coli O157:H7 where they recombined to generate new stx1 genes. This phenomenon of defective prophage induction, transduction and recombination has also been observed in non-pathogenic bacteria, which occupy different ecological niches in nature (42). Defective prophages in pathogenic and non-pathogenic bacteria are also
known to recombine with each other and confer selective advantages to the lysogenic hosts (27).

Extensive homologous recombination with a large and diverse phage pool in food processing plants may be driving rapid niche-specific adaptation in *L. monocytogenes*. We speculate that *L. monocytogenes* recombinants are generated via this mechanism and are subsequently subjected to negative and positive selection in food processing plants, which subsequently generated the unique subclones observed in the present study (FIGS. 3.4 and 3.8). Spontaneous induction and formation of defective transducing phages has been seen in various other microorganisms, and is a mechanism by which various virulence genes (3) and genomic islands, including pathogenicity islands (83) are mobilized into other strains. The results of Orsi et al. (76-77) indicated extensive recombination had occurred throughout the *comK* prophage between an ECIII serotype 1/2a strain in Lineage II and a serotype 1/2b strain in Lineage I (FIG. 3.7), which supports the above model for rapid niche-specific adaptation (FIG. 3.8). Those remaining genes within a prophage that are non-phage genes have been termed “morons” for “more DNA”, which have been defined as all extra genes present in prophage genomes which do not have a phage function, but may act as fitness factors for the lysogen (11). However, the term moron does not indicate what specific role such genes might play in the biology of the host bacterium. Similar to Pathogenicity Islands (PAIs) which carry virulence genes, defective *comK* prophages may carry “adaptons” (FIG. 3.7), which we here define as genes contained within defective prophages that allow a host bacterium to undergo rapid niche-specific adaptation in different environments. If the above
scenario proves to be accurate, we believe that defective \textit{comK} prophage would be more accurately described as a Rapid Adaptation Island (RAI) (FIG. 3.8). This RAI concept is consistent with the recent formulation of a novel concept of inter-phage interactions in defective prophage communities in \textit{E. coli} O157:H7 (3). The finding that no \textit{comK} phage DNA of ECII (H7858) was detected after PCR in the present study suggests that strain H7858 may be a donor, but not a recipient, in the proposed model. However, further work is needed to validate this and the various steps in the Rapid Niche-specific Adaptation model proposed in the present study (FIG. 3.8). For example, further research is needed to determine if the putative RAI is transduced into and recombines with the \textit{comK} prophage in recipient cells of \textit{L. monocytogenes} and generates recombinants, and if the putative “adaptons” within the \textit{comK} prophage that were identified in the present study might allow rapid niche-specific adaptation. If the function of such “adaptons” is to promote attachment and biofilm formation on various surfaces, then “adapton” mutants may be defective in their ability to attach to these surfaces, similar to those strains of \textit{L. monocytogenes} isolated from RTE meat processing plants that were found to be “weakly” adherent and achieved lower numbers in biofilms (53, 67). Additional research to test these various hypotheses is needed.

The combination of MVLST and \textit{comK} prophage junction fragment sequencing may now yield a more complete and accurate picture of both long-term and short-term evolution and epidemiology of \textit{L. monocytogenes}. Sequencing of virulence genes allowed very accurate identification of all four epidemic clones and numerous outbreak clones, but not outbreak clones within epidemic clones (22, 61). In contrast,
sequencing \textit{comK} prophage junction fragments in the present study allowed very accurate identification of different subclones (putative ecotypes) within ECII and ECIII and also allowed identification of a new 1/2a Canada outbreak clone of \textit{L. monocytogenes}. As can be seen in FIG. 3.4B, seven subclones were unique to individual processing plants, while two other subclones were present in multiple plants. A similar finding had been reported previously (4). Finding the same subclone in multiple plants may be due to cross contamination of a common subclone between plants and/or convergent evolution due to production of a common type of food in those plants. Cross contamination between plants could be occurring due to movement of contaminated trucks, pallets, meat, equipment and/or people between processing plants. Further research is needed to test these hypotheses.

In conclusion, the putative ECII isolates from FSIS and Eifert et al. (31) were shown to be ECII by ECII PCR and MVLST. Sequence typing of the \textit{comK} prophage junction fragments allowed very accurate identification of subclones within ECII and ECIII and also the 2008 1/2a Canadian outbreak clone of \textit{L. monocytogenes} and revealed these subclones were associated with individual food processing plants, or in the case of two ECII subclones (Subclones 2 and 5) multiple plants producing similar RTE meat and poultry products (FIG 3.4A and FIG. 3.4B). The integration of \textit{comK} prophage junction fragment sequencing and MVLST may yield a molecular subtyping method with both high discriminatory power and high epidemiologic concordance. Significant recombination was detected in the \textit{comK} prophage junction fragments of some subclones (FIG. 3.5), which was consistent with previous publications that demonstrated extensive recombination within the \textit{comK} prophage of
*L. monocytogenes.* A model for rapid niche-specific adaptation was developed, in which the defective *comK* prophage in *L. monocytogenes* was recharacterized as a “Rapid Adaptation Island” and some “morons” within the RAI were identified (FIG. 3.7) and recharacterized as “adaptons”, or genes that mediate rapid niche-specific adaptation in different food processing plant environments (FIG. 3.8). Further research is needed to test the model proposed in the present study, especially RAI transduction and recombination, and the possible role(s) of putative adaptons in generating persistent subclones of *L. monocytogenes*. If the model is confirmed to be accurate, it could lead to a more fundamental understanding of how specific subclones evolve, adapt to and persist in individual establishments. It could also help identify the routes by which specific subclones of *L. monocytogenes* are transmitted between and within processing plants and retail operations. This information would help food companies implement more effective intervention strategies by targeting harborage sites and eliminating transmission to prevent these dangerous subclones from contaminating various types of high-risk RTE foods, and thus better ensure food safety in the future.

### 3.6 Acknowledgements

The authors thank Martin Wiedmann at Cornell University for providing ECI and ECIII strains, Bala Swaminathan at CDC for providing ECII strains, and Martin Loessner at ETH, Zurich for providing strains 1001 and 1001::A118.

This research was funded by a United States Department of Agriculture Special Grant on Milk Safety to the Pennsylvania State University.
3.7 References


typing identifies single nucleotide polymorphisms which differentiate epidemic
clones and outbreak strains of *Listeria monocytogenes*. J. Clin. Microbiol. 45:835-
846.

epidemic clones, and population structure. *In* D. Liu (ed.), Handbook of *Listeria


sequence typing of outbreak-associated *Listeria monocytogenes* isolates to

2001. Comparative genomics reveals close genetic relationships between phages
from dairy bacteria and pathogenic *Streptococci*: evolutionary implications for

R. Calendar, and M. J. Loessner. 2009. Comparative genome analysis of
*Listeria* bacteriophages reveals extensive mosaicism, programmed translational


41. **Gray, M. J., N. E. Freitag, and K. J. Boor.** 2006. How the bacterial pathogen *Listeria monocytogenes* mediates the switch from environmental Dr. Jekyll to pathogenic Mr. Hyde. *Infect. Immun.* **74:**2505-2512.


   Persistent and nonpersistent *Listeria monocytogenes* contamination in meat and poultry processing plants. J. Food Prot. 66:2062-2069.


   Persistent *Listeria monocytogenes* strains show enhanced adherence to food contact surface after short contact times. J. Food Prot. 63:1204-1207.


Genes under positive selection in *Escherichia coli*. Genome Res. 17:1336-1343.

69. Piffaretti, J. C., H. Kressebuch, M. Aeschbacher, J. Bille, E. Bannerman, J.
M. Musser, R. K. Selander, and J. Rocourt. 1989. Genetic characterization of 
clones of the bacterium *Listeria monocytogenes* causing epidemic disease. Proc.

70. Ragon, M., T. Wirth, F. Hollandt, R. Lavenir, M. Lecuit, A. Le Monnier, and 
Pathog. 4:e1000146.

epidemiological survey of *Listeria monocytogenes* in broilers and poultry 

epidemiological survey of *Listeria monocytogenes* in seafoods and seafood-

73. Ruzin, A., J. Lindsay, and R. P. Novick. 2001. Molecular genetics of SaPI1--a 
mobile pathogenicity island in *Staphylococcus aureus*. Mol. Microbiol. 41:365-
377.


### 3.8 List of Figure Legends

FIG. 3.1. Schematic diagram of phage A118 integration into and excision out of the *L. monocytogenes* chromosome at *comK* (modified from Reference 60). A. Phage A118 with *attP* attachment site. Horizontal arrows indicate PCR priming sites for amplifying the fragment containing *attP*. B. *L. monocytogenes comK* gene containing *attB* attachment site. Horizontal arrows indicate priming sites for amplifying *comK* with *attB* attachment site. C. Lysogenized *L. monocytogenes* showing the locations of the forward and reverse primers (horizontal arrows) for PCR amplification of upstream and downstream *comK* prophage junction fragments. The locations of the prophage PCR targets (LMOh7858_2426 and LMOh7858_2422) in the lysogenized strain (20) are shown with upward-pointing arrows.

FIG. 3.2. PCR amplification of the 1998 and 2002 outbreak clone markers (20) in *L. monocytogenes* ECII isolates. Lanes 1, 8 and 15, Promega 100 bp DNA ladder (Promega, Madison, WI); lanes 2 to 7, amplification of the 1998 outbreak clone marker by primer pair 2426; lanes 9 to 14, amplification of the 2002 outbreak clone marker by primer pair 2664; lanes 2 and 9, negative controls of PCR reactions for
primer pairs 2426 and 2664; lanes 3 and 7, positive amplification of the 1998 outbreak clone and isolate OB070181 from plant J by primer pair 2426; lanes 4 to 6, negative amplification of the 2002 outbreak clone, isolate OB020621 from plant I and isolate OB030029 from plant J by primer pair 2426; lanes 11 and 13, positive amplification of the 2002 outbreak clone and isolate OB030029 from plant I by primer pair 2664; lanes 10, 12 and 14, negative amplification of the 1998 outbreak clone, isolate OB020621 from plant H and isolate OB070181 from plant J by primer pair 2664. Note, isolate OB020621 from plant H showed negative amplification with both primer pairs. Molecular sizes are given (in base pairs) on the right.

FIG. 3.3. PCR amplification of ECII isolates with primer pair 1/4, which targets the downstream junction fragment in the comK prophage of L. monocytogenes (60). Lanes 1, 7, 13, 14, 20 and 26, Promega 100 bp DNA ladder (Promega, Madison, WI); lane 2, negative control of PCR reaction; lanes 3 to 6, L. monocytogenes isolates H7858, H7557, J1816 and R2-765; lanes 8 to 12, L. monocytogenes isolates R2-764, J1703, OB040119, OB050272 and OB050273; lanes 15 to 19, L. monocytogenes isolates OB050226, OB050347, OB050350, OB050351 and OB050355; lanes 21 to 25, L. monocytogenes isolates OB070122, OB020621, OB020790, OB030029 and OB070181. Molecular sizes are given (in base pairs) on the right.

FIG. 3.4. Cluster diagrams based on upstream and downstream junction fragment sequences in L. monocytogenes isolates described in TABLE 3.1. A. Upstream junction fragment cluster diagram. B. Downstream junction fragment cluster diagram. NA: State information is not available.
FIG. 3.5. Recombination breakpoints map for the upstream (A) and downstream (B) junction fragments in the different subclones of *L. monocytogenes* as detected by different algorithms implemented in RDP program (64). Statistically significant breakpoints and their positions are indicated by lines and numbers. Regions that correspond to different genotypes are shaded.

FIG. 3.6. Results of PCR amplifications of *L. monocytogenes* strain H7858 indicating spontaneous induction. Lane 3, Novagen Perfect DNA™ 100 bp Ladder (Novagen, San Diego, CA); lanes 1, 2, 4 and 5, PCR products amplified by Upstream JF, Downstream JF, *attB* and *attP* primer pairs, respectively. All negative controls of the PCR amplifications were negative, but were not shown on the gel. Molecular sizes are given (in base pairs) on the right.

FIG. 3.7. Genes located in the *comK* prophage in *L. monocytogenes* strain H7858 (bottom of figure). LMOh designations 2410, 2411, 2442, 2452 and 2475 indicate those *comK* prophage genes that are present in all sequenced genomes of *L. monocytogenes* that contain the *comK* prophage and LMOf designations 2338, 2372, 2375 and 2377 (in bold) indicate those *comK* prophage genes that are unique to serotype 1/2a. From left to right: *comK’*, N-terminal *comK* fragment; *HP1*, hypothetical protein 1; *HP2*, hypothetical protein 2; *ssb*, single-stranded DNA binding protein; *gp45*, phage gp45 protein; *int*, integrase; *comK*, C-terminal *comK* fragment. Arrows point to the corresponding positions of these loci in the *comK* prophage in ECIII strains at the top of the figure, as described by Orsi et al. (76).

FIG. 3.8. Proposed model for rapid niche-specific adaptation in *L. monocytogenes*. The cycle starts at the top with spontaneous induction of a Rapid
Adaptation Island (RAI) in the donor cell, followed by RAI phage formation, transduction of the donor RAI into the recipient cell, which also contains an RAI integrated into its chromosome. Recombination between donor and recipient RAIs generates numerous RAI recombinants. Natural selection then acts on RAI recombinants to yield subclones that are unique and adapted to individual processing plants or multiple plants manufacturing the same type of food product.
3.9 Figures

FIG. 3.1. Schematic diagram of phage A118 integration into and excision out of the *L. monocytogenes* chromosome at *comK* (modified from Reference 60). A. Phage A118 with *attP* attachment site. Horizontal arrows indicate PCR priming sites for amplifying the fragment containing *attP*. B. *L. monocytogenes comK* gene containing *attB* attachment site. Horizontal arrows indicate priming sites for amplifying *comK* with *attB* attachment site. C. Lysogenized *L. monocytogenes* showing the locations of the forward and reverse primers (horizontal arrows) for PCR amplification of upstream and downstream *comK* prophage junction fragments. The locations of the prophage PCR targets (LMOh7858_2426 and LMOh7858_2422) in the lysogenized strain (20) are shown with upward-pointing arrows.
FIG. 3.2. PCR amplification of the 1998 and 2002 outbreak clone markers (20) in *L. monocytogenes* ECII isolates. Lanes 1, 8
and 15, Promega 100 bp DNA ladder (Promega, Madison, WI); lanes 2 to 7, amplification of the 1998 outbreak clone marker by
primer pair 2426; lanes 9 to 14, amplification of the 2002 outbreak clone marker by primer pair 2664; lanes 2 and 9, negative controls
of PCR reactions for primer pairs 2426 and 2664; lanes 3 and 7, positive amplification of the 1998 outbreak clone and isolate
OB070181 from plant J by primer pair 2426; lanes 4 to 6, negative amplification of the 2002 outbreak clone, isolate OB020621 from
plant I and isolate OB030029 from plant J by primer pair 2426; lanes 11 and 13, positive amplification of the 2002 outbreak clone and
isolate OB030029 from plant I by primer pair 2664; lanes 10, 12 and 14, negative amplification of the 1998 outbreak clone, isolate
OB020621 from plant H and isolate OB070181 from plant J by primer pair 2664. Note, isolate OB020621 from plant H showed
negative amplification with both primer pairs. Molecular sizes are given (in base pairs) on the right.
FIG. 3.3. PCR amplification of ECII isolates with primer pair I/4, which targets the downstream junction fragment in the comK prophage of *L. monocytogenes* (60). Lanes 1, 7, 13, 14, 20 and 26, Promega 100 bp DNA ladder (Promega, Madison, WI); lane 2, negative control of PCR reaction; lanes 3 to 6, *L. monocytogenes* isolates H7858, H7557, J1816 and R2-765; lanes 8 to 12, *L. monocytogenes* isolates R2-764, J1703, OB040119, OB050272 and OB050273; lanes 15 to 19, *L. monocytogenes* isolates OB050226, OB050347, OB050350, OB050351 and OB050355; lanes 21 to 25, *L. monocytogenes* isolates OB070122, OB020621, OB020790, OB030029 and OB070181. Molecular sizes are given (in base pairs) on the right.
FIG. 3.4. Cluster diagrams based on upstream and downstream *comK* prophage junction fragment sequences in *L. monocytogenes* isolates described in TABLE 3.1.

A. Upstream junction fragment cluster diagram. B. Downstream junction fragment cluster diagram. C. Cluster diagrams based on combined upstream and downstream junction fragment sequences. NA: State information is not available.
A Stringent criteria with five or more methods with $p \leq 10^{-5}$

b Two or more methods with $p \leq 10^{-3}$

* The actual breakpoint position is undetermined

FIG. 3.5. Recombination breakpoints map for the upstream (A) and downstream (B) junction fragments in the different subclones of L. monocytogenes as detected by different algorithms implemented in RDP program (64). Statistically significant breakpoints and their positions are indicated by lines and numbers. Regions that correspond to different genotypes are shaded.
FIG. 3.6. Results of PCR amplifications of *L. monocytogenes* strain H7858 indicating spontaneous induction. Lane 3, Novagen Perfect DNA™ 100 bp Ladder (Novagen, San Diego, CA); lanes 1, 2, 4 and 5, PCR products amplified by Upstream JF, Downstream JF, *attB* and *attP* primer pairs, respectively. All negative controls of the PCR amplifications were negative, but were not shown on the gel. Molecular sizes are given (in base pairs) on the right.
FIG. 3.7. Genes located in the comK prophage in L. monocytogenes strain H7858 (bottom of figure). LMOh designations 2410, 2411, 2442, 2452 and 2475 indicate those comK prophage genes that are present in all sequenced genomes of L. monocytogenes that contain the comK prophage and LMOf designations 2338, 2372, 2375 and 2377 (in bold) indicate those comK prophage genes that are unique to serotype 1/2a. From left to right: comK’, N-terminal comK fragment; HP1, hypothetical protein 1; HP2, hypothetical protein 2; ssb, single-stranded DNA binding protein; gp45, phage gp45 protein; int, integrase; comK, C-terminal comK fragment. Arrows point to the corresponding positions of these loci in the comK prophage in ECIII strains at the top of the figure, as described by Orsi et al. (76).
FIG. 3.8. Proposed model for rapid niche-specific adaptation in *L. monocytogenes*. The cycle starts at the top with spontaneous induction of a Rapid Adaptation Island (RAI) in the donor cell, followed by RAI phage formation, transduction of the donor RAI into the recipient cell, which also contains an RAI integrated into its chromosome. Recombination between donor and recipient RAIs generates numerous RAI recombinants. Natural selection then acts on RAI recombinants to yield subclones that are unique and adapted to individual processing plants or multiple plants manufacturing the same type of food product.
3.10 Tables

TABLE 3.1. Origin, establishment, source and date of isolation, and PFGE profiles of the thirty-eight *L. monocytogenes* isolates analyzed in the present study and two Canada outbreak strains included in the cluster analysis.

<table>
<thead>
<tr>
<th>Group and ID</th>
<th>State / Country</th>
<th>Establishment</th>
<th>Source</th>
<th>Date of isolation</th>
<th>AscI pattern / ApaI pattern</th>
<th>PFGE profile</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Epidemic clone II isolates from CDC</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1998 Sara Lee hot dog outbreak</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H7858</td>
<td>MI</td>
<td>A</td>
<td>Food</td>
<td>1998</td>
<td>GX6A16.0002 / GX6A12.0002</td>
<td>1</td>
</tr>
<tr>
<td><strong>2002 U.S. turkey deli outbreak</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>J1815</td>
<td>PA</td>
<td>B</td>
<td>Environmental</td>
<td>2002</td>
<td>GX6A16.0235 / GX6A12.0003</td>
<td>4</td>
</tr>
<tr>
<td>J1703</td>
<td>NA</td>
<td>NA</td>
<td>Human</td>
<td>2002</td>
<td>GX6A16.0235 /</td>
<td>3</td>
</tr>
</tbody>
</table>

*Na — Not available.*
<table>
<thead>
<tr>
<th>Source of Isolates</th>
<th>Source Details</th>
<th>Year</th>
<th>Accession Numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>N3-031</td>
<td>TX</td>
<td>D</td>
<td>Food</td>
</tr>
<tr>
<td>J1-101</td>
<td>TX</td>
<td>D</td>
<td>Human</td>
</tr>
<tr>
<td>R2-603</td>
<td>TX</td>
<td>D</td>
<td>Food</td>
</tr>
<tr>
<td>R2-499</td>
<td>TX</td>
<td>D</td>
<td>Human</td>
</tr>
<tr>
<td>2008 Canada RTE meat product outbreak</td>
<td>2008</td>
<td>2008</td>
<td>LMAAI.0001</td>
</tr>
<tr>
<td>08-5578</td>
<td>Canada</td>
<td>E</td>
<td>Human</td>
</tr>
<tr>
<td>08-5923</td>
<td>Canada</td>
<td>E</td>
<td>Human</td>
</tr>
<tr>
<td>Isolates from FSIS <em>L. monocytogenes</em> Testing Programs</td>
<td></td>
<td>06/2004</td>
<td>GX6A16.0026 / GX6A12.1645</td>
</tr>
<tr>
<td>OB040119</td>
<td>PA</td>
<td>F</td>
<td>Food</td>
</tr>
<tr>
<td>OB050272</td>
<td>PA</td>
<td>F</td>
<td>Food</td>
</tr>
<tr>
<td>OB050273</td>
<td>PA</td>
<td>F</td>
<td>Food</td>
</tr>
<tr>
<td>ID</td>
<td>Category</td>
<td>Grade</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>----------</td>
<td>-------</td>
<td>--------------------------------------</td>
</tr>
<tr>
<td>OB050226</td>
<td>PA</td>
<td>G</td>
<td>Food</td>
</tr>
<tr>
<td>OB050347</td>
<td>PA</td>
<td>G</td>
<td>Environmental, pallet</td>
</tr>
<tr>
<td>OB050350</td>
<td>PA</td>
<td>G</td>
<td>Environmental, packaging machine (bottom)</td>
</tr>
<tr>
<td>OB050351</td>
<td>PA</td>
<td>G</td>
<td>Environmental, side of handwash sink</td>
</tr>
<tr>
<td>OB070122</td>
<td>PA</td>
<td>G</td>
<td>Environmental, drain</td>
</tr>
<tr>
<td>OB020621</td>
<td>NC</td>
<td>H</td>
<td>Food</td>
</tr>
<tr>
<td>OB020790</td>
<td>NC</td>
<td>H</td>
<td>Food</td>
</tr>
<tr>
<td>OB030029</td>
<td>IN</td>
<td>I</td>
<td>Food</td>
</tr>
<tr>
<td>OB070181</td>
<td>KS</td>
<td>J</td>
<td>Food</td>
</tr>
<tr>
<td>Code</td>
<td>State</td>
<td>Level</td>
<td>Location</td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
<td>-------</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td>OB080183</td>
<td>NY</td>
<td>K</td>
<td>Food</td>
</tr>
<tr>
<td>OB080398</td>
<td>NJ</td>
<td>L</td>
<td>Environmental, floor</td>
</tr>
<tr>
<td>OB080487</td>
<td>NJ</td>
<td>L</td>
<td>Environmental, floor</td>
</tr>
<tr>
<td>OB080567</td>
<td>NJ</td>
<td>L</td>
<td>Environmental, floor</td>
</tr>
<tr>
<td>OB090318</td>
<td>NY</td>
<td>M</td>
<td>Environmental, drain</td>
</tr>
</tbody>
</table>

Isolates described in Eifert et al. (31)

<table>
<thead>
<tr>
<th>Code</th>
<th>State</th>
<th>Level</th>
<th>Location</th>
<th>Date</th>
<th>Code Type 1</th>
<th>Code Type 2</th>
<th>ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>1493</td>
<td>NA</td>
<td>N</td>
<td>Environmental, overhead pipe</td>
<td>08/2003</td>
<td>S3 / P3c</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>1495</td>
<td>NA</td>
<td>N</td>
<td>Environmental, floor</td>
<td>10/2003</td>
<td>S3 / P3</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>1496</td>
<td>NA</td>
<td>N</td>
<td>Environmental, carcass rinse</td>
<td>10/2003</td>
<td>S3 / P3-a</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>1498</td>
<td>NA</td>
<td>N</td>
<td>Environmental, dip tank</td>
<td>02/2004</td>
<td>S3 / P3</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>1503</td>
<td>NA</td>
<td>N</td>
<td>Environmental, ladder</td>
<td>02/2004</td>
<td>S3 / P3</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Sample Code</td>
<td>Sample</td>
<td>Type</td>
<td>Description</td>
<td>Date</td>
<td>Lab Code</td>
<td>Code</td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>--------</td>
<td>------</td>
<td>-------------------------</td>
<td>-------</td>
<td>-----------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>1506 (L0704)</td>
<td>NA</td>
<td>N</td>
<td>Environmental, sink and hose</td>
<td>05/2004</td>
<td>S3 / P3</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>1513 (L0719)</td>
<td>NA</td>
<td>N</td>
<td>Environmental, floor</td>
<td>05/2004</td>
<td>S3 / P3</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>1514 (L0720)</td>
<td>NA</td>
<td>N</td>
<td>Environmental, floor drain</td>
<td>05/2004</td>
<td>S3 / P3-b</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>1516 (L0724)</td>
<td>NA</td>
<td>N</td>
<td>Environmental, bin</td>
<td>05/2004</td>
<td>S3 / P3</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>1117 (34-6a)</td>
<td>NA</td>
<td>O</td>
<td>Environmental, floor drain</td>
<td>09/2003</td>
<td>S4 / P4</td>
<td>12</td>
<td></td>
</tr>
</tbody>
</table>

*a* NA, not available  
*b* Original numbers in Eifert et al. (31) are indicated within parentheses below lab code numbers  
*c* AscI pattern/ApaI pattern is described in Eifert et al. (31)
TABLE 3.2. Primer pair name, primer sequences, target and allelic locations within *L. monocytogenes* H7858, amplicon size and annealing temperature for each fragment analyzed in the present study.

<table>
<thead>
<tr>
<th>Primer pair name</th>
<th>PCR and sequencing primers&lt;sup&gt;a&lt;/sup&gt; (5’-3’)</th>
<th>Target</th>
<th>Amplicon size (bp)</th>
<th>Annealing temp (°C)</th>
<th>Reference or source</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECII</td>
<td>ATTATGCCAAGTGGTTACGGA (F) ATCTGGTTTGCGAGACCGTTGTC (R)</td>
<td>LMOh7858_0487.8 to <em>inlA</em> in strain H7858</td>
<td>889</td>
<td>59</td>
<td>(19)</td>
</tr>
<tr>
<td>2426</td>
<td>CAACCCGTTGATGGAGTATT (F) AAACGTTCATTTTTTAACCGATG (R)</td>
<td>An internal region of LMOh7858_2426 in the ECII 1998 outbreak clone</td>
<td>611</td>
<td>Touchdown PCR (56-52°C)</td>
<td>(20)</td>
</tr>
<tr>
<td>2664</td>
<td>CACCTGTACCCCGCTAT (F) AGTTTCCGGGAGGGTCTAAAT (R)</td>
<td>An LMOh7858_2426-homologous region in the ECII 2002 outbreak clone</td>
<td>313</td>
<td>Touchdown PCR (56-52°C)</td>
<td>(20)</td>
</tr>
<tr>
<td>2422</td>
<td>GGCGAGATAGACATGGGT (F) AAAAAACGAAAATCGAACGATA (R)</td>
<td>LMOh7858_2422 in strain H7858</td>
<td>678</td>
<td>53</td>
<td>(20)</td>
</tr>
<tr>
<td>1/4</td>
<td>TGTAACATGGAGTTCTGGCAATC (1) CTCATGAACCTGAAAATGCGG (4)</td>
<td>Downstream junction fragment <em>int-comK</em> in lysogenized strain EGDe::A118</td>
<td>743</td>
<td>55</td>
<td>(60)</td>
</tr>
<tr>
<td>Upstream JF&lt;sup&gt;b&lt;/sup&gt;</td>
<td>AACGCTTTTGACGAGCGAATA (F) GACTTGCAAAATTACCATGTGCT (R)</td>
<td>Upstream junction fragment in strain H7858, J2818, J0161, F6900 and F6854</td>
<td>684</td>
<td>63</td>
<td>This study, modified from (60)</td>
</tr>
<tr>
<td>Downstream JF</td>
<td>TCTTCGTTATTTGGAAGCAT (F) TTGCTAAATAATTTGGTAGTGTGT (R)</td>
<td>Downstream junction fragment int-comK in strain H7858</td>
<td>996</td>
<td>63</td>
<td>This study, modified from (60)</td>
</tr>
<tr>
<td>---------------</td>
<td>---------------------------------------------------------</td>
<td>--------------------------------------------------</td>
<td>-----</td>
<td>---</td>
<td>---------------------------</td>
</tr>
<tr>
<td>attP</td>
<td>AACGCTTTTGACGACGAATA (F) TTGCTAAATAATTTGGTAGTGTGT (R)</td>
<td>attP site within phage A118</td>
<td>956</td>
<td>63</td>
<td>This study, modified from (60)</td>
</tr>
<tr>
<td>attB</td>
<td>TCTTCGTTATTTGGAAGCAT (F) GACTTGCAAATTACCATGTGCT (R)</td>
<td>attB attachment site within comK in strain F2365</td>
<td>724</td>
<td>63</td>
<td>This study, modified from (60)</td>
</tr>
<tr>
<td>int (H7858)</td>
<td>TTGGAGAGCTGACAAACTCG (F) AAACCTCGTAGAAACCGCTGA (R)</td>
<td>Phage integrase gene</td>
<td>215</td>
<td>62</td>
<td>This study</td>
</tr>
<tr>
<td>int (N3-031)</td>
<td>TTGGAGAGCTGACAAACTCG (F) CAGCCATCATAGCATCAAGTTC (R)</td>
<td>Phage integrase gene</td>
<td>215</td>
<td>62</td>
<td>This study</td>
</tr>
</tbody>
</table>

\(^{a}\) PCR primers were synthesized at the Pennsylvania State University Shared Nucleic Acid Facility

\(^{b}\) JF = Junction Fragment
TABLE 3.3. Results of prophage PCR using the method developed by Chen and Knabel (20)

<table>
<thead>
<tr>
<th>Establishment</th>
<th>Amplification results of outbreak clone markers</th>
<th>Sequence type of 2422 amplicon</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1998 OC marker</td>
<td>2002 OC marker</td>
</tr>
<tr>
<td>F</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>G</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>H</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>I</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>J</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>K</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>L</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>M</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>N</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>O</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>
TABLE 3.4. SNP locations in the upstream (A) and downstream (B) junction fragments that differentiated subclones of *L. monocytogenes* in the present study

**A. Upstream Junction Fragment**

<table>
<thead>
<tr>
<th>Subclone</th>
<th>Isolate</th>
<th>Nucleotide sites where SNPs were detected</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>2 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3</td>
</tr>
<tr>
<td>1</td>
<td>OB070181</td>
<td>T C A G G G G G T G A A A A G A C A T C T G T</td>
</tr>
<tr>
<td>2</td>
<td>OB050226</td>
<td>. . . . . . . . . . . . . . . . . . . . . . .</td>
</tr>
<tr>
<td>3</td>
<td>08-5578</td>
<td>C T G A A A T C A . C - G A . . . . . . . .</td>
</tr>
<tr>
<td>6</td>
<td>N3-031</td>
<td>T G A A A T C A . C G . A C A T C T C T C T C</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subclone</th>
<th>Isolate</th>
<th>Nucleotide sites where SNPs were detected</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3</td>
</tr>
<tr>
<td>1</td>
<td>OB070181</td>
<td>C A A G G G A T G T C T A T A C T A - - G A G A A</td>
</tr>
<tr>
<td>2</td>
<td>OB050226</td>
<td>. . . . . . . . . . . . . . . . . . . . . . .</td>
</tr>
<tr>
<td>3</td>
<td>08-5578</td>
<td>. . G . . . . . . . . . . . . . . . . . . . . A</td>
</tr>
<tr>
<td>4</td>
<td>H7858</td>
<td>G G G A T G T C T A T A C T A - - G A G A A</td>
</tr>
<tr>
<td>5</td>
<td>J1816</td>
<td>G G G A T G T C T A T A C T A - - G A G A A</td>
</tr>
<tr>
<td>6</td>
<td>N3-031</td>
<td>A G G A T G T C T A T A C T A T A G A G A A</td>
</tr>
<tr>
<td>7</td>
<td>OB040119</td>
<td>. . G . . . . . . . . . . . . A C T A T A G A G A A</td>
</tr>
</tbody>
</table>
## Nucleotide sites where SNPs were detected

<table>
<thead>
<tr>
<th>Subclone</th>
<th>Isolate</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>OB070181</td>
<td>A</td>
<td>C</td>
<td>A</td>
<td>T</td>
<td>C</td>
<td>T</td>
<td>C</td>
<td>T</td>
<td>A</td>
<td>T</td>
<td>G</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>2</td>
<td>OB050226</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>4</td>
<td>H7858</td>
<td>C</td>
<td>G</td>
<td>T</td>
<td>A</td>
<td>T</td>
<td>C</td>
<td>T</td>
<td>G</td>
<td>G</td>
<td>A</td>
<td>-</td>
<td>C</td>
<td>A</td>
<td>.</td>
<td>G</td>
<td>T</td>
<td>G</td>
<td>G</td>
<td>A</td>
<td>.</td>
</tr>
<tr>
<td>7</td>
<td>OB040119</td>
<td>C</td>
<td>T</td>
<td>T</td>
<td>A</td>
<td>T</td>
<td>C</td>
<td>T</td>
<td>G</td>
<td>G</td>
<td>A</td>
<td>-</td>
<td>C</td>
<td>A</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>G</td>
<td>.</td>
<td>A</td>
<td>.</td>
</tr>
</tbody>
</table>

## Nucleotide sites where SNPs were detected

<table>
<thead>
<tr>
<th>Subclone</th>
<th>Isolate</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
<th>21</th>
<th>22</th>
<th>23</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>OB070181</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>A</td>
<td>T</td>
<td>C</td>
<td>T</td>
<td>C</td>
<td>T</td>
<td>C</td>
<td>G</td>
<td>A</td>
<td>T</td>
<td>A</td>
<td>C</td>
<td>T</td>
<td>A</td>
<td>T</td>
<td>T</td>
<td>G</td>
<td>C</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>OB050226</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>G</td>
<td>.</td>
<td>T</td>
<td>C</td>
<td>T</td>
<td>C</td>
<td>T</td>
<td>A</td>
<td>T</td>
<td>.</td>
<td>G</td>
<td>C</td>
<td>C</td>
<td>G</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>08-5578</td>
<td>.</td>
<td>A</td>
<td>T</td>
<td>G</td>
<td>T</td>
<td>C</td>
<td>T</td>
<td>C</td>
<td>T</td>
<td>C</td>
<td>.</td>
<td>A</td>
<td>T</td>
<td>.</td>
<td>G</td>
<td>C</td>
<td>C</td>
<td>G</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>OB040119</td>
<td>.</td>
<td>A</td>
<td>T</td>
<td>C</td>
<td>G</td>
<td>.</td>
<td>T</td>
<td>C</td>
<td>T</td>
<td>C</td>
<td>T</td>
<td>A</td>
<td>T</td>
<td>.</td>
<td>G</td>
<td>C</td>
<td>C</td>
<td>G</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td></td>
</tr>
<tr>
<td>Subclone</td>
<td>Isolate</td>
<td>Nucleotide sites where SNPs were detected</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>---------</td>
<td>-------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1 OB070181 T G T C G T C T T A C A G A T G T T G A A T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subclone</th>
<th>Isolate</th>
<th>Nucleotide sites where SNPs were detected</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 OB070181 G G A A C C T T C T T - G C -</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2 OB050226 . A . . T T C . T C T A T G</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3 08-5578 . . . T T C . T C C A T .</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4 H7858 A A C T T T C C T C A A T A</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5 J1816 A A C T T T C C T C A A T A</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6 N3-031 . A . . T T C . T C T A T G</td>
<td></td>
</tr>
<tr>
<td></td>
<td>7 OB040119 . A . . T T C . T C T A T G</td>
<td></td>
</tr>
</tbody>
</table>
### B. Downstream Junction Fragment

<table>
<thead>
<tr>
<th>Subclone</th>
<th>Isolate</th>
<th>Nucleotide sites where SNPs were detected</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1</td>
</tr>
<tr>
<td>1a</td>
<td>1493</td>
<td>C A C C A T C A T T G A C A G T C T T G C</td>
</tr>
<tr>
<td>3</td>
<td>08-5578</td>
<td>T T T G T G C . . G . . . C T C C A . T</td>
</tr>
<tr>
<td>5</td>
<td>J1816</td>
<td>G T T A C T G C C A G T T A C T . . .</td>
</tr>
<tr>
<td>7a</td>
<td>OB040119</td>
<td>G T T A C T G C C A G T T A C T . . .</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subclone</th>
<th>Isolate</th>
<th>Nucleotide sites where SNPs were detected</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1</td>
</tr>
<tr>
<td>1a</td>
<td>1493</td>
<td>A G G G C G T A A T G G T T T C T A A C T G</td>
</tr>
<tr>
<td>7a</td>
<td>OB040119</td>
<td>G A A A T A C . . . C . . . G . . G C T C T</td>
</tr>
<tr>
<td>Subclone</td>
<td>Isolate</td>
<td>1</td>
</tr>
<tr>
<td>----------</td>
<td>-----------</td>
<td>---</td>
</tr>
<tr>
<td>1a</td>
<td>1493</td>
<td>C</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subclone</th>
<th>Isolate</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
<th>21</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a</td>
<td>1493</td>
<td>A</td>
<td>C</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>C</td>
<td>A</td>
<td>T</td>
<td>C</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>1b</td>
<td>OB070181</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>2</td>
<td>OB050226</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>4</td>
<td>H7858</td>
<td>G</td>
<td>A</td>
<td>T</td>
<td>A</td>
<td>A</td>
<td>.</td>
<td>G</td>
<td>A</td>
<td>T</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>5</td>
<td>J1816</td>
<td>G</td>
<td>A</td>
<td>T</td>
<td>.</td>
<td>A</td>
<td>T</td>
<td>G</td>
<td>A</td>
<td>T</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>7a</td>
<td>OB040119</td>
<td>.</td>
<td>A</td>
<td>T</td>
<td>.</td>
<td>T</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>7b</td>
<td>OB020621</td>
<td>G</td>
<td>.</td>
<td>T</td>
<td>A</td>
<td>A</td>
<td>.</td>
<td>G</td>
<td>.</td>
<td>T</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
</tbody>
</table>
TABLE 3.5. Detection of spontaneous induction of the *comK* prophage using *attP* and *attB* PCR, *comK* phage using *int* PCR, and plaque formation by *L. monocytogenes* strains F2365, H7858, N3-031 and 1001::A118 on the indicator *L. monocytogenes* strain 1001

<table>
<thead>
<tr>
<th>Strain tested</th>
<th>Detection of spontaneous excision</th>
<th>Amplification of <em>int</em> in</th>
<th>Plaque formation</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>attP</em></td>
<td><em>attB</em></td>
<td>Lysogen</td>
<td>Phage</td>
</tr>
<tr>
<td>F2365 (ECI)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>H7858 (ECII)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>N3-031 (ECIII)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>1001::A118</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup> Epidemic clone designation is indicated within parentheses below strain tested

<sup>b</sup> Approximately 3.00 log<sub>10</sub> plaque-forming units/ml of filtrate as described in materials and method
TABLE 3.6. Differences among three types of phage transduction

<table>
<thead>
<tr>
<th>Type of phage transduction</th>
<th>Type of DNA transduced</th>
<th>Percentage of DNA packaged</th>
<th>Extent of recombination with the host chromosome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generalized transduction</td>
<td>Fragments of chromosomal DNA</td>
<td>0.01%</td>
<td>High</td>
</tr>
<tr>
<td>Specialized transduction</td>
<td>Ends of chromosomal and viral DNA</td>
<td>0.01%</td>
<td>Low</td>
</tr>
<tr>
<td>Transduction of defective comK phage</td>
<td>Entire defective comK prophage</td>
<td>~100%</td>
<td>High</td>
</tr>
</tbody>
</table>
TABLE 3.7. Differences among three types of phage cycle

<table>
<thead>
<tr>
<th>Type of phage cycle</th>
<th>Is the phage defective?</th>
<th>Does the phage integrate into host chromosome?</th>
<th>Does the phage recombine with host chromosome?</th>
<th>Biological function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lytic cycle</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Phage propagation</td>
</tr>
<tr>
<td>Lysogenic cycle</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Phage propagation</td>
</tr>
<tr>
<td>Induction, transduction and recombination of defective $comK$ phage</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Rapid evolution of $L. monocytogenes$</td>
</tr>
</tbody>
</table>
Chapter 4

Conclusions and future research

4.1 Conclusions

Based on the results of ECII PCR, multi-virulence-locus sequence typing (MVLST), this study confirmed that the *Listeria monocytogenes* isolates from FSIS’s *L. monocytogenes* testing programs and the isolates from Eifert et al. (2005) were indeed ECII.

Both the upstream and downstream *comK* prophage junction fragments of isolates from individual meat and poultry processing plants or epidemiologically related isolates had identical DNA sequences, but different sequences among isolates from different processing plants or epidemiologically un-related isolates. This strongly suggested that the specific subclones of *L. monocytogenes* ECII were persistent in individual meat and poultry processing plants. Additionally, ECIII and the 2008 Canadian outbreak clone, both of serotype 1/2a in Lineage II, were differentiated based on *in silico* MVLST analysis, which agreed with *comK* prophage junction fragment sequencing in the present study. Therefore, *comK* prophage junction fragment sequences may represent excellent molecular subtyping markers with high epidemiologic concordance.

Additionally, the *comK* prophages in *L. monocytogenes* ECII and ECIII were shown to spontaneously induce as positive amplifications of *attP* and *attB* were detected in all lysogenized isolates analyzed. The phage particles induced from *L. monocytogenes* ECIII contained the *int* gene of the *comK* phage, which indicated that the phage particles contained the *comK* phage. Lastly, the phages induced from *L. monocytogenes* ECII and
ECIII were defective in plaque formation, indicating prophages that spontaneously induce from *L. monocytogenes* ECII and ECIII could not lyse the indicator cells if transduced, and thus were defective.

Integrating *comK* prophage junction fragment sequencing, and MVLST, may help public health agencies trace subclones of *L. monocytogenes* both between and within individual meat and poultry processing facilities or retail stores to identify sources of contamination and routes of transmission. This information may enhance our understanding of how *L. monocytogenes* rapidly adapts to different environments and also allow implementation of more effective intervention strategies for controlling *L. monocytogenes*.

### 4.2 Future research

Further research is needed to compare sequence variations in the five *comK* prophage open reading frames (ORFs) present in all published *L. monocytogenes* genomes that contain the *comK* prophages. Three of these five ORFs were located in the upstream and downstream junction fragments of the *comK* prophage analyzed in the present study. The other two ORFs located internally within the *comK* prophage were not investigated in the present study. It would be interesting to construct cluster diagrams based on the sequences of these two ORFs in all isolates in the present study to see if they yield similar clusterings as the ones based on the upstream and downstream *comK* prophage junction fragments.

In order to better evaluate the sequence variations within both the backbone genomes and *comK* prophages among the isolates from different establishments, full
genome sequencing of these isolates is currently underway. This information should help clarify why no amplifications were observed in isolates using internal prophage primer pairs in the present study. This information will also reveal whether the \textit{comK} prophage junction fragment sequences in isolates from Est. K, L and M are specific to individual plants. Most importantly, the full genome sequence information will be used to develop \textit{L. monocytogenes} sequence-based subtyping methods with both high discriminatory power and high epidemiologic concordance by targeting appropriate markers in the \textit{comK} prophage.

Further research should be conducted to unveil the roles of the putative adaptons in the \textit{comK} prophage in the persistence of \textit{L. monocytogenes} in food processing environments. It is important to determine whether these putative adaptons are transcribed and translated in the isolates in the current study. If the putative adaptons are found to be transcribed and/or translated, mutant strains with the putative adaptons knocked out can be created. Then phenotypic comparison can be conducted among isogenic strain pairs that differ only by the presence of putative adaptons. As these putative adapton genes might facilitate the transmission, colonization and adaptation of bacterial hosts in various ways, phenotypes like the abilities to adhere to surfaces and form biofilms, inhibit other bacteria in the same environments and resistance to unfavorable environmental factors should be investigated.

More evidence is needed to test the ‘Adaptive Cycle of Lysogen Evolution’ model proposed in the present study, which involved spontaneous induction, transduction, recombination and niche-specific adaptation. Spontaneous induction and packaging of phage DNA in defective \textit{comK} phage particles were detected in the current study.
However, transductions of the entire *comK* phage, homologous recombination and niche-specific adaptation have not yet been demonstrated. A potential approach to address the question would be first to construct donor cells with three different antibiotic markers located in both ends and the middle of the *comK* prophage, and recipient cells with a different antibiotic marker in the backbone genome. If transduction of the entire *comK* phage and homologous recombination are occurring between the donor and recipient cells, then the new recombinant can be selected based on resistance to four different antibiotics. The *comK* prophage in the donor, recipient, and recombinant cells can then be sequenced and subjected to recombination analysis. Furthermore, niche-specific adaptation can be examined by comparing the abilities of different subclones to adhere to different food surfaces and form biofilms. Deeper insights into the adaptive cycle of lysogen evolution of *L. monocytogenes* would help us better understand how the *L. monocytogenes* transmits to and colonizes different food processing environments. With such knowledge, food safety personnel can develop more effective intervention strategies for its control.
APPENDIX

(Additional Data Not Presented in the Main Body of Thesis)
APPENDIX A

Prophage induction of ECII (H7858) and III (N3-031) with Mitomycin C (MMC)

**Purpose:** To determine the effect of MMC concentration on the viability of *Listeria monocytogenes* strains H7858 of ECII and N3-031 of ECIII incubated at 35°C in TSBYE for different lengths of time.

**H₀:** There is no effect of MMC concentration on the viability of *L. monocytogenes* strains H7858 of ECII and N3-031 of ECIII incubated at 35°C in TSBYE for different lengths of time.

**Methods:** Cells were grown at 35°C in 100 ml of TSBYE to early log phase (10⁷ CFU/ml), and MMC was added to the culture to a final concentration of 0.1 µg/ml or 10 µg/ml. At 2-h intervals, samples were plated on TSAYE with subsequent incubation at 35°C for 48 h before enumeration of CFUs. The limit of detection was 10 CFU/ml.

**Results:** A more rapid drop in viability and turbidity (data not shown) was observed when cells were treated with MMC at a concentration of 10 µg/ml compared with those treated with MMC at a concentration of 0.1 µg/ml or those left untreated with MMC during the first two hours of incubation (SUPPLEMENTARY FIG. 1). The viability of strain H7858 treated with MMC at a concentration of 10 µg/ml stayed below the detection limit (1 CFU/ml) after the first 2-h incubation with MMC. The viability of strain N3-031 treated with MMC at a concentration of 10 µg/ml was below the detection limit (1 CFU/ml) after the first 2-h incubation with MMC and then began to increase 4 hours later. The viabilities of cells treated with MMC at a concentration of 0.1 µg/ml or those left untreated with MMC were similar after 18-h incubation with or without MMC.
Conclusions: There is an effect of MMC concentration on the viability of *L. monocytogenes* strains H7858 of ECII and N3-031 of ECIII incubated at 35°C in TSBYE for different lengths of time. However, it is still unclear what the main reason is for the drastic decrease in bacterial viability after cells were treated with MMC at a concentration of 10 µg/ml. It is probably either cell lysis due to prophage induction or damage to bacteria host chromosome by MMC.
SUPPLEMENTARY FIG. 1. Effect of MMC concentration on the viability of *L. monocytogenes* strains H7858 (A) and N3-031 (B) incubated at 35°C in TSBYE for different lengths of time. Open triangles represent indicate no CFUs were observed at the limit of detection.