ATTACHMENT OF LISTERIA MONOCYTOGENES TO PLANT SURFACES AND HOST CELLS AND ITS SURVIVAL IN HOST AND NON-HOST ENVIRONMENTS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Sara Rose Milillo

May 2008

ATTACHMENT OF LISTERIA MONOCYTOGENES TO PLANT SURFACES AND HOST CELLS AND ITS SURVIVAL IN HOST AND NON-HOST ENVIRONMENTS

Sara Rose Milillo, Ph. D.

Cornell University 2008

Listeria monocytogenes was originally recognized as a pathogen of rabbits and guinea pigs following an outbreak of disease in animal care facilities in England in 1926. E.G.D. Murray subsequently performed the initial characterization of this gram-positive bacterium. The first widely recognized outbreak of human listeriosis (i.e. the illness arising from *L. monocytogenes* infection) due to consumption of contaminated food occurred in Canada and was reported by Schlech and colleagues in 1983. After this outbreak, research of *L. monocytogenes* transmission through the environment and research of its virulence mechanisms intensified. Scientists quickly identified several key *L. monocytogenes* virulence factors, one of which is encoded by the gene primarily responsible for host cell invasion (i.e. *inlA*). However, the factors that contribute to its environmental persistence, including possible reservoirs, as well as a complete understanding of its virulence mechanisms remain undefined. Research has continued in order to better define *L. monocytogenes* means of attachment to surfaces, which is a critical step in facilitating persistence and virulence. In addition, there is evidence that genetically, phenotypically, and epidemiologically distinct subsets of *L. monocytogenes* strains (i.e. lineages) exist and have different capabilities for surviving in natural and food processing environments and for causing disease. Several genes

with structures similar to *inlA*, designated internalin-like genes, are known to vary in presence among the lineages and these genes have been proposed to contribute to *L. monocytogenes* attachment abilities. On a larger scale, several genomic regions containing multiple genes have been identified that also vary in presence among the lineages. The variation in these internalin-like genes and large, unique genomic regions may represent the source(s) of the phenotypic and epidemiological differences observed among the lineages. The research presented here indicates that it is unlikely that the differing propensity to cause disease observed among *L. monocytogenes* strains results from the presence or absence in the *L. monocytogenes* genome of unique virulence factors, and rather is a consequence of another source, for example sequence divergence.

BIOGRAPHICAL SKETCH

Sara grew up in State College, Pennsylvania with her parents Steven and Roberta and brother Michael. There, she attended State College Area High School where she developed her interest in science, in particular microbiology, due largely to the influence of her AP molecular biology (Dr. Mary Knight) and AP chemistry (Mr. Patrick Gallagher) teachers who continually encouraged their students' questions as well as supported the use of logic to reach scientific conclusions. After high school, she attended The Pennsylvania State University (go lions!) and majored in Biotechnology with a minor in Microbiology. In May 2004, she graduated from college and moved to Ithaca, New York to attend graduate school at Cornell University. At Cornell, she studied the virulence of *Listeria monocytogenes* under Dr. Martin Wiedmann. While in graduate school, Sara volunteered frequently to coordinate and teach a variety of educational workshops and participate in other extension or outreach activities. Sara hopes to continue to contribute to both food safety research and science education in her future career.

ACKNOWLEDGMENTS

I would like to thank my major advisor, Dr. Martin Wiedmann and the Cornell University Department of Food Science for accepting me into their graduate school program and supporting my research for the last three and a half years. I have learned lessons in more than just microbiology and food safety research during my time in graduate school that I am sure I would not have been able to experience anywhere else. I will take the knowledge I have gained here with me throughout all my future endeavors.

I would also like to thank my committee members Dr. Barbara Crawford and Dr. Joseph Peters. Dr. Crawford has been a wonderful source of ideas and enthusiasm concerning the development of practical approaches to teaching science in the classroom as well as for developing new theories on the science of teaching itself. Dr. Peters, thank you for taking the time to talk with me about my research progress and career options. Thank you also to Janette Robbins, graduate student coordinator extraordinaire, for her cheerfulness and prompt answers to whatever question is at hand.

The support of my family and friends over the last few years has been invaluable. I cannot thank them enough for their patience (when I complain) and constant encouragement (when I'm struggling). Thank you specifically to my parents and my brother; I would not have the strength to pursue my goals without you by my side. To my friends, thank you for helping me have fun!

Lastly, I would like to thank God. I have been very blessed in my life and, while I am often not good at showing it, I know that God is the reason behind all my

achievements and for that I am grateful. I truly appreciate all the efforts of the Cornell Catholic Community staff to make God feel accessible and church/religion/spirituality relevant.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

APPENDIX

CHAPTER 1

MICROBIAL GENOMICS AND FOOD SAFETY

Introduction

The first usage of the term "genome" is attributed to Hans Winkler in his 1920 publication *Verbeitung und Ursache der Parthenogenesis im Pflanzen und Tierreiche* (Winkler, 1920). However, it was not until 1986 that the concept of genomics came together with the creation of a new journal by the same name (Anonymous, 1997). The study of genomics can be broadly defined as the use or application of "informatic tools" to study the features of a sequenced genome (Strauss, 1997). Today the field of genomics is typically considered to encompass efforts to wholly determine the nucleic acid DNA sequence of an organism as well as the expression of genetic information using high throughput, genome-wide methods, including transcriptome, proteome, and metabolome analyses. Functional genomics describes efforts focused on understanding the function of genes and other genomic elements. On the other hand, comparative genomics, a field that only has recently become feasible due to the advances in high throughout technologies, uses comparisons between genomic information for different organisms to gain insights into the organism's biology and evolution.

Foodborne disease continues to affect millions of people each year (Mead et al., 1999) and efforts are needed to develop new technologies and methods to maintain food safety and combat foodborne illness. Unlocking the information hidden in the genomes of foodborne pathogens can ameliorate food safety and help alleviate the burden of foodborne illness on society. The fields of functional and comparative

genomics have a significant impact on how researchers approach the study of foodborne pathogens and food safety, as discussed below.

Genome sequencing efforts

Over the last five years there has been a steep increase in the number of completely sequenced genomes. According to the Genomes Online database (Liolios, 2008), the number has risen dramatically in just the last three years from ~300 completed bacterial genomes in 2005 to nearly 700 completed bacterial genomes in 2008. This increase in available genetic information is expected to continue as sequencing methods become more automated, faster, and less cost restrictive. The genomes of many foodborne pathogens, including bacteria, protozoan, fungi, and viruses, have been sequenced. According to the Genome Database from the National Center for Biotechnology Information (NCBI), there are over 100 foodborne pathogen genomes sequenced to date representing 40 different species. These sequenced genomes include 84 bacterial, 5 protozoan or fungal, and 13 viral genomes. Among the bacterial foodborne pathogens, *Listeria monocytogenes* has the most genome sequences available (20), followed by *Staphylococcus aureus* subsp. *aureus* (14), then *Escherichia coli* (10), *Campylobacter* spp. (9), and *Salmonella enterica* subsp. *enterica* and *Shigella* spp. each with six genomes sequenced.

Advent of genome sequencing. The first organism to have its entire genome published was the bacteriophage MS2 (Fiers, 1976). The MS2 genome is RNA-based and, thus, it was not until 1977 that the sequence of the first DNA-based genome (bacteriophage Φ174) was published by Sanger and colleagues (Sanger et al., 1977A). Later that year, a group of researchers, again led by Sanger, published an article detailing the chain-termination sequencing technique (Sanger et al., 1977B). Briefly, Sanger sequencing reactions contain the desired DNA template (single-stranded),

labeled nucleotides, a DNA primer, and a DNA polymerase. One of four, labeled, chain-terminating dideoxynucleotides triphosphates (ddATP, ddTTP, ddCTP, or ddGTP), which once incorporated in the DNA strand prevent elongation as they lack the 3' –OH group required for synthesis of the phosphodiester bond between nucleotides, are added to each sequencing reaction. The newly synthesized, labeled DNA products from the four reactions are then run on a gel with a one-nucleotide resolution. The gel is imaged and each band represents a DNA fragment whose synthesis was terminated by incorporation of one of the four, labeled dideoxynucleotides. The DNA sequence is read from the top to bottom of the gel across the four lanes, beginning with the band of the smallest size.

For more than twenty years this method (commonly called Sanger sequencing) was the only sequencing method available (Schuster, 2008) and, surprisingly, its technology evolved little over the next three decades, despite some advances in automation (Smith, 1986). Using the Sanger method to devise the DNAsequence, genome sequencing was most often approached by so-called shot-gun sequencing (Anderson, 1981). In shot-gun sequencing, genomic DNA is randomly broken up and the pieces of a certain size are cloned into a carrier vector (generating a clone library), then the vector inserts are sequenced and the overlapping pieces are assembled to reflect contiguous portions of the genome ("contigs"), the small areas (gaps) remaining in the sequence can be directly sequenced using a variety of different methods, most often with PCR. The first bacterial genome sequenced using this method was published in 1995 (Fleischmann, 1995). The average cost of sequencing a genome using this method is ~$1 per kilobase (Shendure, 2005), so a 300Mb genome would cost at least $300,000. In addition to the high cost, the chain-termination method was limited by its low-throughput.

Current development of sequencing-by-synthesis technologies. The 1996 publication of an article entitled "Real-Time DNA Sequencing Using Detection of Pyrophosphate Release" signaled that a new course for sequencing technology was imminent (Ronaghi, 1996). In this manuscript, the researchers detailed a procedure they developed to sequence DNA without the need for gel electrophoresis, called pyrosequencing. The procedure worked by using a combination of three enzymes: a DNA polymerase, an ATP sulfurylase, and a luciferase. In this way, the researchers created an elegant system that took advantage of the pyrophosphate released every time the polymerase incorporated a nucleotide into the growing DNA strand to, via the sulfurylase, provide the ATP fuel needed for the luciferase to produce light, which could be easily detected and recorded using a luminometer. The luciferase produced light signals proportional to the amount of ATP available. The system requires that only one nucleotide at a time is available for the DNA polymerase, such that the light produced could only be due to incorporation of that particular nucleotide. Therefore, for the system to be adapted for large-scale use, sophisticated automation would be required.

In 2005, researchers from 454 Life Sciences (now partnered with Roche Applied Biosystems) and from the laboratory of George Church (Harvard Medical School) independently published two innovative, high-throughput sequencing-by-synthesis methods (Margulies, 2005; Shendure, 2005). In the 454 method, genomic DNA is sheared to create random fragment libraries that are then labeled with a common tag or adaptor. Then, the tagged fragments are diluted to isolate single strands that can be captured onto beads labeled with a complimentary tag. Beads are deposited into wells (one bead per well, i.e. one DNA fragment per well) where the synthesis reaction (as described above) proceeds. Reagents for the sequencing reactions are provided by "flows", with "wash" steps in between to ensure the accuracy (that light produced

from a well is due only to the currently available nucleotide). This method developed by 454 Life Sciences is now available commercially via Roche and called the GS-FLX Genome Analyzer. Using technology licensed from the Church lab, the company Illumina, Inc. also offers commercially available, post-Sanger generation, sequencing-by-synthesis (Schuster, 2008). The Illumina system, called Solexa sequencing technology, begins the same way as the method developed by 454 Life Sciences. Single stranded genomic DNA from the sample organism is prepared, randomly fragmented, and two different adaptors are added, one for each fragment end. The DNA fragments are then added to the flow cell, in which each channel is studded with complimentary adaptors to capture the fragmented DNA (one end of the DNA fragment will bind irreversibly to one flow channel adaptor and the other end will bind reversibly to a different adaptor, forming a horseshoe shape). Free, unlabeled DNA polymerase and nucleotides are added to the flow cell and amplification proceeds for each bound fragment. After one round of amplification two copies of the DNA fragment will result, each bound by one end to the flow cell channel. After several rounds of amplification there will be tightly grouped clusters of each DNA fragment. Next fluorescently labeled, reversibly chain-terminating nucleotides, DNA polymerase, and primers are added to the flow cell. The first nucleotide will be incorporated (after which amplification ceases) and then the first base in the sequence can be read using a laser to determine which labeled base was incorporated. The blocked 3' terminus of the nucleotide is removed and another round of amplification proceeds followed by another round of laser-excitation and determination of the second base.

While the techniques described above are still limited in read length (~100-250bp maximum), so too was the Sanger method when it was first detailed (~80bp) (Schuster, 2008). The read length of sequencing-by-synthesis methods is expected to

improve with time. Use of these technologies free researchers from relying on clone libraries and carrier vectors as well as greatly improve on the low degree of multiplicity in Sanger sequencing, already greatly reducing the cost of sequencing genomes. On average, a 300Mb genome can be sequenced using the GS-FLX Genome Analyzer by 454 Life Sciences for ~$24,000 and using the Solexa 1G Sequencer by Illumina a genome can be sequenced for ~$1,600. More recently Applied Biosystems has released a sequencer called the SOLiD system and, in the near future, third generation sequencing systems are expected to be released by VisiGen and Helicos (Schuster, 2008). The development of new technologies and even greater improvements in cost-reduction are anticipated to continue as the demand for genome sequences is enormous. A recent review reported by Genomes Online (Liolios, 2008) listed approximately 3,500 currently on-going sequencing projects.

Impact of genome sequencing studies on food safety

The availability and use of genome sequences for microbial foodborne pathogens has major impacts on food safety, including (i) annotating genomes and identifying markers for pathogenic organisms (which can be used for development of detection methods), (ii) developing subtyping methods based on full genome sequences, and (iii) identifying emerging pathogens.

Genome annotations and identification of markers for pathogenic organisms.

Following the completion of a genome sequence, the data must then be transformed from a string of A's, T's, C's, and G's into a more meaningful and useful collection of genes and other genetic elements. This process is called genome annotation and can be carried out in a variety of ways (Frishman, 2007). Predicted genes can be detected using software like GLIMMER (Delcher, 1999) and subsequently genes can be assigned putative functions based on their homology to other previously annotated

genes (e.g. using BLAST; Altschul, 1997), classified into a cluster of orthologous groups (e.g. with COG; Tatusov, 2000), by their association with a known pathway (e.g. via KEGG; Kanehisa, 2004), or by comparing their translated sequence to conserved families of proteins (e.g. using Pfam; Finn, 2006). Frequently, researchers use automated gene-prediction software to analyze the sequence and then manually revise the output as necessary (Stothard, 2006). Three common, recently developed genome annotation software packages are PUMA2, ASAP, and MaGe (Maltsev, 2006; Glasner, 2006; Vallenet, 2006). Annotation software is steadily improving, so for foodborne pathogens with older genome sequences several versions of the genome annotation may exist, as sequence data is often reanalyzed when new annotation software is released. Published genome sequences are publicly available via a number of web-based databases (e.g. GenBank [Benson, 2006], the European Molecular Biology Laboratory [EMBL; Cochrane, 2006], or the DNA Data Bank of Japan [Okubo, 2006]).

The power of genome annotation lies in its ability to facilitate genome comparisons. Annotation of a genome may reveal similarities of some/many genes to known virulence factors or genes involved in stress survival from other organisms, which could then be targeted for study (Raskin, 2006). For example, genome annotation of the extraintestinal, uropathogenic strain CFT073 revealed that it encoded many genes involved in fimbrial adhesion and phase-switching, two characteristics associated with colonization and persistence in the intestinal lumen (Welch, 2002). Furthermore, based on its genome annotation it was observed that *Mycobacterium tuberculosis* had two copies of the DNA polymerase *dnaE*. Researchers found that the second copy of dnaE (*dnaE2*) was critical for virulence *in vivo* and hypothesize that *dnaE2* may be an error prone DNA polymerase, not related to the Y family of error prone polymerases, that might contibute to inducible mutagenesis, as it is upregulated

following exposure to many DNA damaging treatments (Boshoff, 2003). In terms of food safety, these and other studies continue to build on the foodborne bacterial pathogen knowledge base and, through their contributions to the understanding of virulence mechanisms, will ultimately contribute to the development of treatment and prevention strategies.

Subtyping methods based on genome sequences. The investigation of foodborne illness outbreaks is critical to determining not only the origin of the outbreak but also the weaknesses in the food production industry. Increased pressure by consumers to deliver fresh and minimally processed foods (e.g. spinach) has subsequently increased the risk of outbreaks of foodborne illness on a national scale. Therefore, scientists use subtyping methods to discriminate between organisms isolated from food products and foodborne illness cases as well as to draw conclusions about contamination patterns, transmission routes, and source(s) of illness.

The original subtyping methods were developed based on organism phenotype, these include biotyping, serotyping, and phage typing (for review see Hyytia-Trees, 2007). However, researchers found these methods to be laborious and often expensive, requiring specific reagents (e.g. diagnostic antisera). Molecular subtyping methods were later introduced that provided improved discriminatory power (e.g. pulsed field gel electrophoresis [PFGE]) and portability (e.g. multi-locus sequence typing [MLST]). More recently, the availability of genome sequences stimulated development of two new sequence-based methods, multiple-locus variable-number tandem-repeat analysis (MLVA) and single nucleotide polymorphism (SNP)-based subtyping (Hoffmaster, 2002; Cebula, 2005).

MLVA is increasingly used to subtype foodborne pathogens and for foodborne disease surveillance. The MLVA technique relies on the presence of a variable number of tandem repeats (VNTR) in the genomic DNA. Short sequence DNA repeats

and specifically VNTR have been well documented in bacterial genomes (van Belkum, 1998). The strategy is simple: design a multiplex PCR system to simultaneously amplify multiple regions of tandem repeats known to vary in length and separate the PCR products by size using gel electrophoresis (with a resolution up to one base pair). The product sizes observed correspond the number of repeats and, thus, the allele present at that location. A recent study using a highly clonal organism (e.g., *Bacillus anthracis*) showed that MLVA was able to further discriminate isolates that were indistinguishable by two-enzyme PFGE (Hoffmaster, 2002). Recently, MLVA has been practically applied to foodborne disease surveillance and outbreak detection, for example with *Salmonella enterica subsp. enterica* serotype Typhimurium (*S.* Typhimurium) (Torpdahl, 2006) and with *E. coli* O157:H7 (Noller, 2003). In Denmark, MLVA identified an outbreak of foodborne illness caused by *S.* Typhimurium and linked the outbreak to contaminated, imported cured sausage (Nygard, 2007).

A SNP is the result of a single nucleotide change in an inter- or intragenic region of DNA of an one organism relative to another and detection of SNPs has become easier as more genome sequences have become available (Hyytia-Trees, 2007). SNPs in an organism's genome can be detected using a variety of methods, including PCR, sequencing of small SNP containing regions, and hybridization studies (Cebula, 2005). SNP-based subtyping can be particularly useful for discriminating between organisms that appear to evolve primarily clonally. Read and colleagues (2002) studying *B. anthracis* and Ducey et al. (2007) studying lineage I *L. monocytogenes* strains present two examples of researchers who demonstrated that SNP-based subtyping could distinguish between very similar bacterial isolates. In addition, the loci of some SNPs on the *S.* Typhimurium genome have been identified with potential for use in SNP-based subtyping strategies (Hu, 2006). The application of SNP-based

subtyping is expected to expand to use in more organisms as more genomes are sequenced and, thereby, more sequence data is available for SNP screening.

Subtyping of organisms using sequenced based techniques, like MLVA and SNP-based methods, is crucial to maintaining food safety. Sequence-based subtyping methods contribute to food safety by effectively and accurately tracking foodborne bacterial contaminants as well as quickly connect related cases to identify outbreaks of illness and facilitate rapid response.

Identification of emerging pathogens. At the turn of the century, *Salmonella* Typhi, *Brucella,* and *Mycobacterium tuberculosis* were considered major foodborne pathogens (Rosenau, 1926). Today, in industrialized nations foodborne illness caused by many of these pathogens has been significantly reduced (Tauxe, 2002). However, the burden of foodborne illness on society remains high; estimates in the United States have placed the incidence of foodborne illness as high as 76 million cases per year (Mead, 1999). This estimate implicates new or emerging pathogens as the likely cause of the high incidence of foodborne disease. Interestingly, researchers have approximated that once every two years a new (or newly characterized) foodborne pathogen becomes widely recognized (Tauxe, 2002). One of the major contributors to the identification of new or emerging pathogens is the dramatic increase in available genome sequences. In this way, genome sequencing impacts food safety directly by assisting in recognition and study of new foodborne pathogens of interest. The primary means for defining new bacterial taxonomic groups is based on the gene sequence for 16S rRNA and with the rise in sequence data available researchers can more easily compare organisms phylogenetically and more definitively determine their relationships (Jay, 2003). Some examples of new or emerging foodborne pathogens whose status was defined, at least in part, by genome sequence studies include, non-

jejuni/coli Campylobacter spp. (e.g. *concisus*), *Enterobacter sakazakii*, and enteroaggregative *Escherichia coli*.

The first example of a new or emerging foodborne pathogen defined using genome sequence studies is from the genus *Campylobacter*. *C. jejuni* and *C. coli* are two *Campylobacter* species commonly recognized as foodborne pathogens. Increasingly, epidemiological studies indicate that other *Campylobacter spp*. (e.g. *concisus*) should also be considered as possible causative agents of foodborne disease (Newell, 2005). Improved subtyping techniques based on genomic sequences will increase discrimination within and between the *Campylobacter* species (Fouts, 2005). A third example is *Enterobacter sakazakii*. In 1980 a new *Enterobacter* species was created separate from *E. cloacae*, called *E. sakazakii*, defined largely by its phenotypic properties (Farmer, 1980). *E. sakazakii* is now regarded as an emerging foodborne pathogen, most commonly linked to contamination of powdered infant formula (Drudy, 2006). Case mortality rates for *E. sakazakii* infections as high as 80% have been reported (Lai, 2001). A more recent study in 2004 based on gene sequences of 16S rRNA indicated that *E. sakazakii* is a genomically diverse species (Lehner, 2004) and its mechanisms of virulence are poorly understood (Drudy, 2004). Standardized subtyping methods are needed to improve epidemiological investigations and further *E. sakazakii* research (Drudy, 2004) and development in these areas will be greatly supported by the continual increases in sequenced genomes. Lastly, similar to *E. sakazakii*, in the 1980's a new subgroup of *Escherichia coli* called enteroaggregative *E. coli* (EAEC) was also described based on its phenotypic characteristic of aggregating on HEP-2 cells (Harrington, 2006). Despite research efforts, the virulence factors associated with EAEC pathogenicity are still unclear and strain heterogeneity remains an issue (Harrington, 2006). As more genome sequences become available,

EAEC strains can be better defined benefiting subtyping efforts and epidemiological studies.

Comparative genomic and comparative expression analyses of foodborne pathogens

Comparative genomic analyses include direct comparisons of organism's genomic DNA (as discussed above with sequencing and subtyping) as well as gene expression mapping. Gene expression mapping, or microarray, is a method extensively used by researchers to compare the level of gene expression in one cell versus another. A microarray is composed of a support (be it a glass slide, membrane, other material) spotted with oligonucleotide representations of part/most/all of a genome of interest (Schoolnik, 2002). Genomic DNA can then be hybridized to the array to examine sequence homology between the organism represented on the array and the organism represented by the genomic DNA sample. Alternatively, mRNA collected from an organism under two different environmental conditions can be converted to cDNA, differentially labeled, and then hybridized to the array measuring the relative gene transcription in the two conditions. This method is described further below.

Using a comparative microarry approach, researchers examined the change in gene transcription over time of *E. coli* O157:H7 grown in minimal media (Bergholz, 2007). They determined that there was a dramatic effect of growth phase on gene transcription, including for many genes thought to be involved in virulence. Specifically, they observed that transcription of the Shiga toxin genes increased significantly as the *E. coli* O157:H7 transitioned into the stationary phase of growth. In another study, gene expression mapping was recently used to monitor host immune response to another major foodborne pathogen, *Salmonella* Enteritidis (Lillehoj, 2007). In this study, researchers measured the changes in host cell gene transcription during *Salmonella* infection over 24 hours. They noted that many host cell genes

associated with cell adhesion and proliferation were down-regulated following infection. A similar approach was used to study the effects on the host cell of *Brucella abortus* infection (Eskra, 2003). It was noted that although *Brucella spp*. lipopolysacharide (LPS) induces a less dramatic immune response than *E. coli* LPS, it does not invade host cells silently as host cell defense related genes (e.g. encoding proinflammatory cytokines and chemokines) were up-regulated following infection (Eskra, 2003, Rajashekara, 2006).

Foodborne pathogens and functional genomic research

The efforts of researchers to interpret and organize the plethora of information that has become available as a result of the rise in sequenced genomes has given birth to a new field of study, functional genomics. The term functional genomics describes a collection of techniques that use genome sequences and derive from them likely functional attributes and interactions of the genes encoded therein (Suen, 2007). Interpreting the functions of genes and their possible contributions to virulence and stress response in foodborne pathogens is critical to improving overall food safety.

Micoarrays, protein expression analyses, and metabolic analyses are examples of experiment-based techniques that can be used for functional genomic applications. A series of sequence-based techniques designed to facilitate exploration of probable gene functions and protein interactions also exist often in the form of databases that compile gene homology and protein prediction services for a particular organism or subset of organisms (e.g. LEGER for *Listeria spp*. [Dieterich, 2006]) or for a wide range of organisms (e.g. Prolinks [Bowers, 2004], PLEX [Date and Marcotte, 2005], or STRING [von Mering, 2005]). More specifically, using the theory that homologous genes in homologous groups conserved across bacteria from different genuses are likely to be linked functionally (Suen, 2007), phylogenetic profiles for genes can be

developed that give clues as to their function. For example, the *Vibrio cholerae* genome was compared to 57 other bacteria and researchers were able to reconstruct in *V. cholerae* the SoxR oxidative stress response pathway using sequence homology and information known about the pathway in the other organisms (Date and Marcotte, 2003).

Functional Microarrays. In addition to use for comparative genomics, microarrays can be used for functional genomic analyses. The function of a gene of interest can be investigated when microarrays are designed to compare the gene expression profiles for a wild-type (WT) strain versus an isogenic null mutant. Researchers studied the role of extracytoplasmic function (ECF) alternative sigma factor E (σ^E) in *M. tuberculosis* using a microarray comparing gene transcription of a *sigE* null mutant with the WT *M. tuberculosis* (Manganelli, 2001). They found that 23 genes were up-regulated in the presence of σ^E (i.e. in WT cells) as compared to in the absence of σ^E (i.e. in the *sigE* null mutant) after exposure to the detergent sodium dodecyl sulfate. This observation indicated that in *M. tuberculosis,* σ^E has a functional role in regulating genes induced in response to cell wall damage. A study of the foodborne pathogen *Listeria monocytogenes* also used microarrays to interpret the function of a different alternative sigma factor, σ^B (Kazmierczak, 2003). They showed that in addition to regulating stress responsive systems, σ^B had a role in regulating the virulence factors InlA and InlB, which had not previously been demonstrated.

Protein Expression Analysis. Another means of studying gene function in an organism is to directly compare and contrast the proteins it expresses under different conditions. This approach is particularly desirable to researchers, as it has been noted repeatedly that cellular mRNA levels (like those measured with microarrays), while more convenient and simpler to measure, often do not correspond to protein levels (Gygi, 1999; Holt, 2000; Greenbaum, 2003). Studying gene function in terms of the

proteins produced is called proteomics and the word "proteome" was first used in 1995, refering to all proteins encoded by an organism's genome (Blackstock, 1999).

So called classical proteomic research mainly involves a two step process, step (1) protein separation using two-dimensional electrophoresis (2-DE) and step (2) peptide mass mapping for protein identification (Cash, 2003). Researchers investigating the proteome of *E. sakazakii* used a combination of 2-DE and matrix assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF-MS) (Riedel, 2007). First proteins isolated from *E. sakazakii* exposed to two different osmotic stresses were separated out using 2-DE. Protein spots of interest on the gel were excised and partially enzymatically digested. The digested proteins fragments separated and each fragments' mass was measured. Finally, the resulting mass profile was used to search fractionation databases for a protein match. It was shown that the two different osmotic stresses resulted in distinct but overlapping responses in terms of protein expression. Burns-Keliher and colleagues (1997) selectively labeled proteins produced by *S.* Typhimurium during infection of epithelial cells and, using a similar proteomic analysis, determined that 34 *S.* Typhimurium proteins were uniquely expressed in the intracellular environment. A subsequent study found that protein expression by *S.* Typhimurium changes in different host cell types (Burns-Keliher, 1998). They examined *S.* Typhimurium protein expression in five different host cell backgrounds and observed that 58 unique proteins were expressed during growth in Intestine-407 cells, 157 in J774.A cells, 40 in rat bone marrow-derived macrophages, 113 in mouse bone marrow-derived macrophages, and 91 in NMuLi cells. This study demonstrates that *S.* Typhimurium responds differently to different host environments, which could be an important consideration when designing new vaccines against *Salmonella*.

An alternative to 2-DE methods are protein arrays (also called protein chips or protein microarrays). Some of the earliest methods describing the development of protein arrays were reviewed by Nagayama in 1997 (Nagayama, 1997B). Currently, there are multiple versions of protein arrays (based on different support systems, e.g. glass slide or microwell-based) and each of the formats offer a high throughput means for analyzing protein-protein, protein-DNA, protein-RNA, protein-ligand, or protein-substrate interactions (Zhu, 2001). Together with protein mass-spectrophotometric identification methods, protein arrays have incredible potential for identifying and characterizing protein interactions and, subsequently, gene function (Walter, 2000). Using a nitrocellulose membrane base, researchers recently described a protein array designed to probe for better understanding of the immunoreactivity of 40 *Mycobacterium paratuberculosis* proteins (Bannantine, 2008). The researchers recombinantly expressed the 40 proteins using an expression vector in *E. coli*. After purifying and blotting the proteins onto the membrane, the researchers then compared sera collected from cattle experimentally infected with *M. paratuberculosis* and naturally infected cattle (i.e. those manifesting Johne's disease symptoms). They found that the antigenic profiles of cattle naturally and experimentally infected were strikingly similar, lending support to the theory that *M. paratuberculosis* is at least a contributing agent for Johne's disease. The main hurdle remaining to the advancement of protein array analyses and wide-spread usage is efficient, large-scale protein production and array development (Zhu, 2001; Bannantine, 2008).

Metabolomic Studies. Metabolomics is a third avenue with which to explore functional genomic research. Metabolomics is considered the analysis, including identification and measurement, of all metabolites within a biochemical system (Oldiges, 2007). The field of metabolomic research has been exploding in recent years, and it was shown that the rate of publication of metabolomic research projects

has been nearly exponential (Oldiges, 2007). Researchers used a metabolomic study of *L. monocytogenes* membranes to reveal that there is a change in membrane lipid composition to permit growth at low temperatures (Mastronicolis, 2006). For cells at 5ºC, the researchers reported a dramatic 10-fold increase in the total lipid anteiso-15:0/anteiso-17:0 fatty acid ratio as well as a 30% increase in neutral lipid membrane content. They believe that these two changes reflect the mechanistic response of *L. monocytogenes* to low (i.e. 5ºC) temperatures, which it may commonly experience during cold storage of fresh produce and other food products.

Impacts of comparative and functional genome analyses on food safety

While the use of genome sequence data has clearly already impacted food safety, including the development of new technologies and assays that are routinely used to detect and characterize foodborne pathogens, the impact of functional genomics (i.e. gene function analyses) on food safety has predominantly focused on the improved understanding of pathogen biology in both the host and the environment.

Impacts of comparative genomics – identification of virulence factors and targets for growth inhibition. For example, comparative genomics quickly made an impact on the study of *L. monocytogenes* virulence when a specific factor contributing to virulence was identified following whole genome comparisons between *L. monocytogenes* and the closely related, non-pathogen species *L. innocua* (Glaser, 2001; Vazquez-Boland, 2001, Dussurget, 2002). Dussurget and colleagues (2002) located a bile salt hydrolase gene specific to the *L. monocytogenes* genome and subsequently determined that this gene contributed to bile resistance, intestinal colonization, and systemic infection of the liver. The identification of new, previously uncharacterized virulence factors is a strength of comparative genomic analyses and has lasting implications for understanding foodborne pathogen virulence and

foodborne illness treatment. Another example of a recent study that used comparative genomics to further food safety research was by Chan et al. (2007). In this study, the researchers described a microarray-based analysis of genes expressed by *L. monocytogenes* during growth at refrigeration temperatures. They found that a number of genes required for cell proliferation (e.g. RNA helicases) were upregulated under these conditions while virulence factors were down-regulated. This study identified potential targets to inhibit *L. monocytogenes* growth at low temperatures, thereby leading to effective ways to control *L. monocytogenes* growth during food storage prior to consumption. In combination with similar studies of other foodborne pathogens, this research will establish an approach, using strategies to control pathogen growth, for improving food safety.

Impacts of functional genomics – development of new therapeutics and vaccines. Microarray technology can be applied in several ways that impact discovery of new therapeutics and vaccine development to treat and prevent foodborne illness. An array constructed of 198 small molecules with putative antibacterial effects was used recently to probe for novel antibacterial agents against *Staphylococcus aureus* (Bowman, 2007). While further studies are required to determine the safety and applicability of the antibacterial agents identified in this study, the significance of the results, in terms of future contributions to treatment of *Staphylococcal* infections, is clear. In addition, studies such as this smooth the way for similar methods to be developed for other pathogens, increasing the hope that novel therapeutics may soon be available, particularly for known multi-drug resistant foodborne pathogens such as non-typhoidal *Salmonella* (Alcaine, 2007). In a similar protein based study to that mentioned above for *M. paratuberculosis*, researchers used *in vitro* transcription and translation to study 197 *B. anthracis* proteins and their immunoreactivity (Gat, 2006). The researchers noted that currently available *B. anthracis* vaccines rely on the use of

B. anthracis culture supernatant, which is primarily composed of just one antigen, called the protective antigen. The researchers identified 30 novel, immunoreactive proteins related to virulence. These 30 proteins represent numerous targets for development of new, sensitive, and more powerful *B. anthracis* vaccines. As functional genomic methodologies continue to be refined, the emergence of novel therapeutics and vaccines will increase ultimately improving the treatment of foodborne illness.

Conclusions

The field of genomics has clearly had a considerable impact on microbial food safety research. While discussions on the impact of genomics on food safety often focus on advances in our understanding of pathogen genomes, insights into aspects of virulence and control of transmission also represent important contributions of the field of genomics to food safety. Future challenges in the field of genomics and its application to food safety, thus, not only include improved understanding of pathogen genomics and translation of this knowledge into practical applications that can reduce the incidence of foodborne illnesses, but also integration of pathogen genomics with human and intermediate host (e.g., food animals, plants) genomic information.

REFERENCES

Alcaine, S. D., Warnick, L. D. & Wiedmann, M. (2007). Antimicrobial resistance in nontyphoidal *Salmonella*. J Food Prot, 70, 780-790.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). Gapped blast and psi-blast: A new generation of protein database search programs. Nucleic Acids Res, 25, 3389-3402.

Anderson, S. (1981). Shotgun DNA sequencing using cloned DNAse i-generated fragments. Nucl. Acids Res., 9, 3015-3027.

Anonymous (1997). Genomics: Structural and functional studies of genomes. Genomics, 45, 244-249.

Bannantine, J. P., Waters, W. R., Stabel, J. R., Palmer, M. V., Li, L., Kapur, V. & Paustian, M. L. (2008). Development and use of a partial *Mycobacterium avium* subspecies *paratuberculosis* protein array. PROTEOMICS, 8, 463-474.

Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J. & Wheeler, D. L. (2006). Genbank. Nucleic Acids Res, 34, D16-20.

Bergholz, T., Wick, L., Qi, W., Riordan, J., Ouellette, L. & Whittam, T. (2007). Global transcriptional response of *Escherichia coli* O157:H7 to growth transitions in glucose minimal medium. BMC Microbiology, 7, 97.

Blackstock, W. P. & Weir, M. P. (1999). Proteomics: Quantitative and physical mapping of cellular proteins. Trends in Biotechnology, 17, 121-127.

Boshoff, H. I. M., Reed, M. B., Barry, C. E. & Mizrahi, V. (2003). Dnae2 polymerase contributes to *in vivo* survival and the emergence of drug resistance in *Mycobacterium tuberculosis*. Cell, 113, 183-193.

Bowers, P. M., Pellegrini, M., Thompson, M. J., Fierro, J., Yeates, T. O. & Eisenberg, D. (2004). Prolinks: A database of protein functional linkages derived from coevolution. Genome Biol, 5, R35.

Bowman, M. D., O'Neill, J. C., Stringer, J. R. & Blackwell, H. E. (2007). Rapid identification of antibacterial agents effective against staphylococcus aureus using small-molecule macroarrays. Chem Biol, 14, 351-357.

Burns-Keliher, L., Nickerson, C. A., Morrow, B. J. & Curtiss, R., 3rd (1998). Cell-specific proteins synthesized by *Salmonella* Typhimurium. Infect Immun, 66, 856-861.

Burns-Keliher, L. L., Portteus, A. & Curtiss, R., 3rd (1997). Specific detection of *Salmonella* Typhimurium proteins synthesized intracellularly. J Bacteriol, 179, 3604-3612.

Cash, P. (2003). Proteomics of bacterial pathogens. Adv Biochem Eng Biotechnol, 83, 93-115.

Cebula, T. A., Jackson, S. A., Brown, E. W., Goswami, B. & LeClerc, J. E. (2005). Chips and snps, bugs and thugs: A molecular sleuthing perspective. J. Food Prot., 68, 1271-1284.

Chan, Y. C., Raengpradub, S., Boor, K. J. & Wiedmann, M. (2007). Microarray-based characterization of the *Listeria monocytogenes* cold regulon in log- and stationary-phase cells. Appl Environ Microbiol, 73, 6484-6498.

Cochrane, G., Aldebert, P., Althorpe, N., Andersson, M., Baker, W., Baldwin, A., Bates, K., Bhattacharyya, S., Browne, P., van den Broek, A., Castro, M., Duggan, K., Eberhardt, R., Faruque, N., Gamble, J., Kanz, C., Kulikova, T., Lee, C., Leinonen, R., Lin, Q., Lombard, V., Lopez, R., McHale, M., McWilliam, H., Mukherjee, G., Nardone, F., Pastor, M. P., Sobhany, S., Stoehr, P., Tzouvara, K., Vaughan, R., Wu, D., Zhu, W. & Apweiler, R. (2006). EMBL nucleotide sequence database: Developments in 2005. Nucleic Acids Res, 34, D10-15.

Date, S. V. & Marcotte, E. M. (2003). Discovery of uncharacterized cellular systems by genome-wide analysis of functional linkages. Nat Biotech, 21, 1055-1062.

Date, S. V. & Marcotte, E. M. (2005). Protein function prediction using the protein link explorer (plex). Bioinformatics, 21, 2558-2559.

Delcher, A. L., Harmon, D., Kasif, S., White, O. & Salzberg, S. L. (1999). Improved microbial gene identification with glimmer. Nucleic Acids Res, 27, 4636-4641.

Dieterich, G., Karst, U., Fischer, E., Wehland, J. & Jansch, L. (2006). Leger: Knowledge database and visualization tool for comparative genomics of pathogenic and non-pathogenic *Listeria* species. Nucleic Acids Res, 34, D402-406.

Drudy, D., Mullane, N. R., Quinn, T., Wall, P. G. & Fanning, S. (2006). Food safety: *Enterobacter sakazakii*: An emerging pathogen in powdered infant formula. Clinical Infectious Diseases, 42, 996-1002.

Ducey, T. F., Page, B., Usgaard, T., Borucki, M. K., Pupedis, K. & Ward, T. J. (2007). A single-nucleotide-polymorphism-based multilocus genotyping assay for subtyping lineage I isolates of *Listeria monocytogenes*. Appl Environ Microbiol, 73, 133-147.

Dussurget, O., Cabanes, D., Dehoux, P., Lecuit, M., Buchrieser, C., Glaser, P. & Cossart, P. (2002). *Listeria monocytogenes* bile salt hydrolase is a PrfA-regulated virulence factor involved in the intestinal and hepatic phases of listeriosis. Mol Microbiol, 45, 1095-1106.

Eskra, L., Mathison, A. & Splitter, G. (2003). Microarray analysis of mRNA levels from raw264.7 macrophages infected with *Brucella abortus*. Infect. Immun., 71, 1125-1133.

Farmer, J. J. I., Asbury, M. A., Hickman, F. W. & Brenner, D. J. (1980). *Enterobacter sakazakii*: A new species of "*Enterobacteriaceae*" Isolated from clinical specimens. Int J Syst Bacteriol, 30, 569-584.

Fiers W, C. R., Duerinck F, Haegeman G, Iserentant D, Merregaert J, Min Jou & W, M. F., Raeymaekers A, Van den Berghe A, Volckaert G, Ysebaert M. (1976). Complete nucleotide sequence of bacteriophage ms2 RNA: Primary and secondary structure of the replicase gene. Nature, 260, 500-507.

Finn, R. D., Mistry, J., Schuster-Bockler, B., Griffiths-Jones, S., Hollich, V., Lassmann, T., Moxon, S., Marshall, M., Khanna, A., Durbin, R., Eddy, S. R., Sonnhammer, E. L. & Bateman, A. (2006). Pfam: Clans, web tools and services. Nucleic Acids Res, 34, D247-251.

Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., Kerlavage, A. R., Bult, C. J., Tomb, J. F., Dougherty, B. A., Merrick, J. M. & al, e. (1995). Whole-genome random sequencing and assembly of *Haemophilus influenzae* rd. Science, 269, 496-512.

Fouts, D. E., Mongodin, E. F., Mandrell, R. E., Miller, W. G., Rasko, D. A., Ravel, J., Brinkac, L. M., DeBoy, R. T., Parker, C. T., Daugherty, S. C., Dodson, R. J., Durkin, A. S., Madupu, R., Sullivan, S. A., Shetty, J. U., Ayodeji, M. A., Shvartsbeyn, A., Schatz, M. C.,

Badger, J. H., Fraser, C. M. & Nelson, K. E. (2005). Major structural differences and novel potential virulence mechanisms from the genomes of multiple *Campylobacter* species. PLoS Biology, 3, e15.

Frishman, D. (2007). Protein annotation at genomic scale: The current status. Chem. Rev., 107, 3448-3466.

Gat, O., Grosfeld, H., Ariel, N., Inbar, I., Zaide, G., Broder, Y., Zvi, A., Chitlaru, T., Altboum, Z., Stein, D., Cohen, S. & Shafferman, A. (2006). Search for *Bacillus anthracis* potential vaccine candidates by a functional genomic-serologic screen. Infect Immun, 74, 3987-4001.

Glaser, P., Frangeul, L., Buchrieser, C., Rusniok, C., Amend, A., Baquero, F., Berche, P., Bloecker, H., Brandt, P., Chakraborty, T., Charbit, A., Chetouani, F., Couve, E., de Daruvar, A., Dehoux, P., Domann, E., Dominguez-Bernal, G., Duchaud, E., Durant, L., Dussurget, O., Entian, K. D., Fsihi, H., Garcia-del Portillo, F., Garrido, P., Gautier, L., Goebel, W., Gomez-Lopez, N., Hain, T., Hauf, J., Jackson, D., Jones, L. M., Kaerst, U., Kreft, J., Kuhn, M., Kunst, F., Kurapkat, G., Madueno, E., Maitournam, A., Vicente, J. M., Ng, E., Nedjari, H., Nordsiek, G., Novella, S., de Pablos, B., Perez-Diaz, J. C., Purcell, R., Remmel, B., Rose, M., Schlueter, T., Simoes, N., Tierrez, A., Vazquez-Boland, J. A., Voss, H., Wehland, J. & Cossart, P. (2001). Comparative genomics of *Listeria* species. Science, 294, 849-852.

Glasner, J. D., Rusch, M., Liss, P., Plunkett, G., 3rd, Cabot, E. L., Darling, A., Anderson, B. D., Infield-Harm, P., Gilson, M. C. & Perna, N. T. (2006). Asap: A resource for annotating, curating, comparing, and disseminating genomic data. Nucleic Acids Res, 34, D41-45.

Greenbaum, D., Colangelo, C., Williams, K. & Gerstein, M. (2003). Comparing protein abundance and mRNA expression levels on a genomic scale. Genome Biology, 4, 117.

Gygi, S. P., Rochon, Y., Franza, B. R. & Aebersold, R. (1999). Correlation between protein and mRNA abundance in yeast. Mol. Cell. Biol., 19, 1720-1730.

Harrington, S. M., Dudley, E. G. & Nataro, J. P. (2006). Pathogenesis of enteroaggregative *Escherichia coli* infection. FEMS Microbiology Letters, 254, 12-18.

Hoffmaster, A. R., Fitzgerald, C. C., Ribot, E., Mayer, L. W. & Popovic, T. (2002). Molecular subtyping of *Bacillus anthracis* and the 2001 bioterrorism-associated anthrax outbreak, United States. Emerg Infect Dis, 8, 1111-1116.

Holt, L. J., Enever, C., de Wildt, R. M. & Tomlinson, I. M. (2000). The use of recombinant antibodies in proteomics. Curr Opin Biotechnol, 11, 445-449.

Hu, H., Lan, R. & Reeves, P. R. (2006). Adaptation of multilocus sequencing for studying variation within a major clone: Evolutionary relationships of *Salmonella enterica* serovar Typhimurium. Genetics, 172, 743-750.

Hyytia-Trees, E. K., Cooper, K., Ribot, E. M. & Gerner-Smidt, P. (2007). Recent developments and future prospects in subtyping of foodborne bacterial pathogens. Future Microbiology, 2, 175-185.

Jay, J. M. (2003). A review of recent taxonomic changes in seven genera of bacteria commonly found in foods. J. Food Prot., 66, 1304-1309.

Kanehisa, M., Goto, S., Kawashima, S., Okuno, Y. & Hattori, M. (2004). The KEGG resource for deciphering the genome. Nucleic Acids Res, 32, D277-280.

Kazmierczak, M. J., Mithoe, S. C., Boor, K. J. & Wiedmann, M. (2003). *Listeria monocytogenes* sigma B regulates stress response and virulence functions. J Bacteriol, 185, 5722-5734.

Lai, K. K. (2001). *Enterobacter sakazakii* infections among neonates, infants, children, and adults. Case reports and a review of the literature. Medicine (Baltimore), 80, 113-122.

Lehner, A., Tasara, T. & Stephan, R. (2004). 16s rRNA gene based analysis of *Enterobacter sakazakii* strains from different sources and development of a PCR assay for identification. BMC Microbiol, 4, 43.

Lillehoj, H. S., Kim, C. H., Keeler, C. L., Jr. & Zhang, S. (2007). Immunogenomic approaches to study host immunity to enteric pathogens. Poult Sci, 86, 1491-1500.

Liolios, K., Mavromatis, K., Tavernarakis, N. & Kyrpides, N. C. (2008). The genomes on line database (GOLD) in 2007: Status of genomic and metagenomic projects and their associated metadata. Nucl. Acids Res., 36, D475-479.

Maltsev, N., Glass, E., Sulakhe, D., Rodriguez, A., Syed, M. H., Bompada, T., Zhang, Y. & D'Souza, M. (2006). Puma2--grid-based high-throughput analysis of genomes and metabolic pathways. Nucleic Acids Res, 34, D369-372.

Manganelli, R., Voskuil, M. I., Schoolnik, G. K. & Smith, I. (2001). The *Mycobacterium tuberculosis* ecf sigma factor sigmae: Role in global gene expression and survival in macrophages. Mol Microbiol, 41, 423-437.

Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S., Bemben, L. A., Berka, J., Braverman, M. S., Chen, Y.-J., Chen, Z., Dewell, S. B., Du, L., Fierro, J. M., Gomes, X. V., Godwin, B. C., He, W., Helgesen, S., Ho, C. H., Irzyk, G. P., Jando, S. C., Alenquer, M. L. I., Jarvie, T. P., Jirage, K. B., Kim, J.-B., Knight, J. R., Lanza, J. R., Leamon, J. H., Lefkowitz, S. M., Lei, M., Li, J., Lohman, K. L., Lu, H., Makhijani, V. B., McDade, K. E., McKenna, M. P., Myers, E. W., Nickerson, E., Nobile, J. R., Plant, R., Puc, B. P., Ronan, M. T., Roth, G. T., Sarkis, G. J., Simons, J. F., Simpson, J. W., Srinivasan, M., Tartaro, K. R., Tomasz, A., Vogt, K. A., Volkmer, G. A., Wang, S. H., Wang, Y., Weiner, M. P., Yu, P., Begley, R. F. & Rothberg, J. M. (2005). Genome sequencing in microfabricated high-density picolitre reactors. Nature, 437, 376-380.

Mastronicolis, S. K., Boura, A., Karaliota, A., Magiatis, P., Arvanitis, N., Litos, C., Tsakirakis, A., Paraskevas, P., Moustaka, H. & Heropoulos, G. (2006). Effect of cold temperature on the composition of different lipid classes of the foodborne pathogen *Listeria monocytogenes*: Focus on neutral lipids. Food Microbiology, 23, 184-194.

Mead, P. S., Slutsker, L., Dietz, V., McCaig, L. F., Bresee, J. S., Shapiro, C., Griffin, P. M. & Tauxe, R. V. (1999). Food-related illness and death in the United States. Emerg Infect Dis, 5, 607-625.

Nagayama, K. (1997). Protein arrays: Concepts and subjects. Adv Biophys, 34, 3-23.

Newell, D. G. (2005). *Campylobacter concisus*: An emerging pathogen? European Journal of Gastroenterology & Hepatology, 17, 1013-1014.

Noller, A. C., McEllistrem, M. C., Pacheco, A. G., Boxrud, D. J. & Harrison, L. H. (2003). Multilocus variable-number tandem repeat analysis distinguishes outbreak and sporadic *Escherichia coli* O157:H7 isolates. J Clin Microbiol, 41, 5389-5397.

Nygard, K., Lindstedt, B. A., Wahl, W., Jensvoll, L., Kjelso, C., Molbak, K., Torpdahl, M. & Kapperud, G. (2007). Outbreak of *Salmonella* Typhimurium infection traced to imported cured sausage using mlva-subtyping. Euro Surveill, 12, E070315 070315.

Okubo, K., Sugawara, H., Gojobori, T. & Tateno, Y. (2006). Ddbj in preparation for overview

of research activities behind data submissions. Nucleic Acids Res, 34, D6-9.

Oldiges, M., Lutz, S., Pflug, S., Schroer, K., Stein, N. & Wiendahl, C. (2007). Metabolomics: Current state and evolving methodologies and tools. Appl Microbiol Biotechnol, 76, 495-511.

Rajashekara, G., Eskra, L., Mathison, A., Petersen, E., Yu, Q., Harms, J. & Splitter, G. (2006). *Brucella*: Functional genomics and host-pathogen interactions. Anim Health Res Rev, 7, 1-11.

Raskin, D. M., Seshadri, R., Pukatzki, S. U. & Mekalanos, J. J. (2006). Bacterial genomics and pathogen evolution. Cell, 124, 703-714.

Read, T. D., Salzberg, S. L., Pop, M., Shumway, M., Umayam, L., Jiang, L., Holtzapple, E., Busch, J. D., Smith, K. L., Schupp, J. M., Solomon, D., Keim, P. & Fraser, C. M. (2002). Comparative genome sequencing for discovery of novel polymorphisms in *Bacillus anthracis*. Science, 296, 2028-2033.

Riedel, K. & Lehner, A. (2007). Identification of proteins involved in osmotic stress response in *Enterobacter sakazakii* by proteomics. Proteomics, 7, 1217-1231.

Ronaghi, M., Karamohamed, S., Pettersson, B., Uhlen, M. & Nyren, P. (1996). Real-time DNA sequencing using detection of pyrophosphate release. Analytical Biochemistry, 242, 84-89.

Rosenau, M. J. (1926). Preventative medicine and hygiene. D. Appleton and Company, New York.

Sanger, F., Air, G. M., Barrell, B. G., Brown, N. L., Coulson, A. R., Fiddes, C. A., Hutchison, C. A., Slocombe, P. M. & Smith, M. (1977). Nucleotide sequence of bacteriophage phi x174 DNA. Nature, 265, 687-695.

Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A, 74, 5463-5467.

Schoolnik, G. K. (2002). Functional and comparative genomics of pathogenic bacteria. Curr Opin Microbiol, 5, 20-26.

Schuster, S. C. (2008). Next-generation sequencing transforms today's biology. Nat Methods, 5, 16-18.

Shendure, J., Porreca, G. J., Reppas, N. B., Lin, X., McCutcheon, J. P., Rosenbaum, A. M., Wang, M. D., Zhang, K., Mitra, R. D. & Church, G. M. (2005). Accurate multiplex polony sequencing of an evolved bacterial genome. Science, 309, 1728-1732.

Smith, L. M., Sanders, J. Z., Kaiser, R. J., Hughes, P., Dodd, C., Connell, C. R., Heiner, C., Kent, S. B. & Hood, L. E. (1986). Fluorescence detection in automated DNA sequence analysis. Nature, 321, 674-679.

Stothard, P. & Wishart, D. S. (2006). Automated bacterial genome analysis and annotation. Curr Opin Microbiol, 9, 505-510.

Strauss, E. J. & Falkow, S. (1997). Microbial pathogenesis: Genomics and beyond. Science, 276, 707-712.

Suen, G., Arshinoff, B. I., Taylor, R. G. & Welch, R. D. (2007). Practical applications of bacterial functional genomics. Biotechnol Genet Eng Rev, 24, 213-242.

Tatusov, R. L., Galperin, M. Y., Natale, D. A. & Koonin, E. V. (2000). The COG database: A tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res, 28, 33-36.

Tauxe, R. V. (2002). Emerging foodborne pathogens. Int J Food Microbiol, 78, 31-41.
Torpdahl, M., Sorensen, G., Ethelberg, S., Sando, G., Gammelgard, K. & Porsbo, L. J. (2006). A regional outbreak of *S*. Typhimurium in Denmark and identification of the source using MLVA typing. Euro Surveill, 11, 134-136.

Vallenet, D., Labarre, L., Rouy, Z., Barbe, V., Bocs, S., Cruveiller, S., Lajus, A., Pascal, G., Scarpelli, C. & Medigue, C. (2006). Mage: A microbial genome annotation system supported by synteny results. Nucleic Acids Res, 34, 53-65.

van Belkum, A., Scherer, S., van Alphen, L. & Verbrugh, H. (1998). Short-sequence DNA repeats in prokaryotic genomes. Microbiol Mol Biol Rev, 62, 275-293.

Vazquez-Boland, J. A., Dominguez-Bernal, G., Gonzalez-Zorn, B., Kreft, J. & Goebel, W. (2001). Pathogenicity islands and virulence evolution in listeria. Microbes Infect, 3, 571-584.

Walter, G., Bussow, K., Cahill, D., Lueking, A. & Lehrach, H. (2000). Protein arrays for gene expression and molecular interaction screening. Curr Opin Microbiol, 3, 298-302.

Welch, R. A., Burland, V., Plunkett, G., 3rd, Redford, P., Roesch, P., Rasko, D., Buckles, E. L., Liou, S. R., Boutin, A., Hackett, J., Stroud, D., Mayhew, G. F., Rose, D. J., Zhou, S., Schwartz, D. C., Perna, N. T., Mobley, H. L., Donnenberg, M. S. & Blattner, F. R. (2002). Extensive mosaic structure revealed by the complete genome sequence of uropathogenic *Escherichia coli*. Proc Natl Acad Sci U S A, 99, 17020-17024.

Winkler, H. (1920). Verbeitung und ursache der parthenogenesis im pflanzen und tierreiche. Springer Verlag, Jena.

Zhu, H. & Snyder, M. (2001). Protein arrays and microarrays. Curr Opin Chem Biol, 5, 40-45.

CHAPTER 2

GROWTH AND PERSISTENCE OF *LISTERIA MONOCYTOGENES* ISOLATES ON THE PLANT MODEL *ARABIDOPSIS THALIANA*

Abstract

While the majority of human listeriosis cases appear to be linked to consumption of processed ready-to-eat foods (e.g., deli meats), some listeriosis outbreaks have been linked to consumption of contaminated vegetables. In this study, we assessed four isolates representing the major *Listeria monocytogenes* lineages for their abilities to attach to and grow on *Arabidopsis thaliana*, a well-characterized plant model. When plants were dipped for 5 min into 3 ml of water containing 8.8 log CFU of *L. monocytogenes* and rinsed repeatedly, *L. monocytogenes* was recovered from the leaves at densities from 1.52 to 2.17 log CFU/cm^2. Ten days after exposure, bacterial numbers had increased over initial numbers by 2.60 to 2.95 log CFU/cm^2. Using *L. monocytogenes* expressing GFP, bacteria were visualized in the intercellular spaces of *A. thaliana* leaves, suggesting internalization through stomata. These data indicate that *L. monocytogenes* can rapidly attach to and multiply on plant surfaces and colonize intercellular spaces in *A. thaliana* leaves where it may be protected from sanitation treatments. When *A. thaliana* seeds were exposed to *L. monocytogenes,* between 4.23 and 4.57 log CFU/cm^2 were recovered from leaves 7 days post germination, suggesting that contaminated seeds can produce contaminated plants. Overall, our study demonstrates that prevention of *L. monocytogenes* contamination of plants throughout growing stages is critical, consistent with recommendations for other produce-transmitted foodborne pathogens.

Introduction

Recently, increasing numbers of foodborne illness outbreaks have been traced to consumption of plant-derived foods (Sivapalasingam et al., 2004). While, in particular, a number of outbreaks caused by *E. coli* O157:H7 and *Salmonella enterica* have been linked to consumption of contaminated fruit and vegetable produce (Soderstrom et al., 2005, Sandt et al., 2006, Gupta et al., 2007), many other foodborne pathogens can also be transmitted by plant-based foods. For example, a number of food- and waterborne parasites, including *Cryptosporidium* and *Giardia,* have been reported to be transmitted through plant-based foods (Slifko et al., 2000). Although contamination of fruits and vegetables, which are often consumed without prior heat treatment, can occur throughout the food distribution chain, pre-harvest contamination of fruits and vegetables represents a particular concern. Interestingly, a number of studies (Guo et al., 2001, Seo et al., 1999, Solomon et al., 2002, Cooley et al., 2003) have shown that, in addition to attaching to the outer surfaces of fruits and vegetables, various bacterial pathogens can enter plant tissues via different routes and, thus, be found in internal plant tissues. For example, *E. coli* O157:H7 and *Salmonella* serovar Typhimurium have been recovered from the interior of surface sterilized lettuce plants grown from seeds inoculated with the bacteria (Jablasone et al., 2005). Similarly, *E. coli* O157:H7 was found internalized in lettuce seedlings grown on manure-amended soil containing *E. coli* O157:H7 (Solomon et al., 2002). Both *E. coli* O157:H7 and *Salmonella* appear to be able to enter plants through root systems, with subsequent internal dispersal to edible tissues (Solomon et al., 2002, Jablasone et al., 2005). Guo et al. (2001) also found that inoculation of *Salmonella* onto tomato flowers resulted in *Salmonella* in the tomato pulp, suggesting internalization of bacterial contaminants during fruit development. Internalization of bacterial pathogens into the edible portions of plants

is of particular concern as internalized pathogens are unlikely to be removed by washing or surface sanitization methods (Jablasone et al., 2005).

While *Listeria monocytogenes* is found in diverse environments and has been isolated from environmental soil, water, and manure samples as well as from a variety of animals (Pell, 1997), the 2003 USDA-FDA risk assessment suggests that most human listeriosis cases are caused by contaminated ready-to-eat (RTE) meat and poultry, dairy, and seafood products (U.S. FDA and FSIS, 2003). However, a recent case control study in the US determined that human *L. monocytogenes* infections were significantly associated with consumption of melons or hummus prepared in commercial establishments (Varma et al., 2007). Although *L. monocytogenes* has been isolated from a variety of plant-derived foods, including salad vegetables, alfalfa sprouts, and mushrooms (Sagoo et al., 2003, Caggia et al., 2004, Samadpour et al., 2006) and some human listeriosis outbreaks have been linked to consumption of produce (Schlech et al., 1983, Ho et al., 1986), the mean estimated numbers of human listeriosis cases in the US resulting from consumption of contaminated fruits and vegetables are only 0.9 and 0.2 cases per annum, respectively (U.S. FDA and FSIS, 2003). Nonetheless, various lines of evidence indicate that *L. monocytogenes* contamination of fruits and vegetables may contribute to the burden of human listeriosis infections, particularly since *L. monocytogenes* is able to survive and multiply under various stress conditions, including those likely encountered on plant surfaces (Brandl, 2006, Garner et al., 2006, Vermeulen et al., 2007). The relatively rare linkage between consumption of fruits and vegetables and human listeriosis outbreaks and cases may partially be due to the difficulty of identifying food sources of *L. monocytogenes* infections, given the long incubation period of listeriosis (Linnan et al., 1988), and the inherent low power of case control studies in identifying associations of foodborne disease cases with commonly consumed foods, particularly

for diseases with long incubation periods. In addition, human disease typically requires exposure to high numbers of *L. monocytogenes* (U.S. FDA and FSIS, 2003) that may not be easily reached throughout the shelf life of many fruits and vegetables even with high starting levels of contamination (Li et al., 2002, Bari et al., 2005).

L. monocytogenes attachment to and growth on some plants, including spinach and alfalfa sprouts, has been documented (Gorski et al., 2004, Jablasone et al., 2005). However, studies on *L. monocytogenes* – plant interactions using genetically traceable, well-characterized plant species and pathogen strains are needed to further our understanding of interactions between foodborne pathogens and plants. Cooley et al. (2003) reported that *E. coli* O157:H7 and *Salmonella* Serovar Newport showed similar levels of growth on *Arabidopsis thaliana* as has been noted on vegetable plants (Charkowski et al., 2002), suggesting this plant as an appropriate model for studying interactions between plants and Gram-negative foodborne pathogens. To extend this line of inquiry to a Gram-positive bacterial pathogen, we used the plant model *A. thaliana*, which has a completed genome sequence and established tools for mutant generation (The Arabidopsis Genome Initiative, 2000), to create an experimental system to study the interactions between plants and different *L. monocytogenes* isolates. *L. monocytogenes* isolates were selected to represent the major genetic lineages of this pathogen, including lineage I (representing predominantly serotypes 1/2a and 1/2c), lineage II (representing predominantly serotypes 1/2b and 4b), as well as lineages IIIA and IIIB (both representing serotypes 4a, 4c, as well as atypical serotype 4b strains [Liu et al., 2006, Roberts et al., 2006]).

Methods and Materials

Plant maintenance and propagation. Wildtype *Arabidopsis thaliana* (Columbia ecotype) were used for all experiments. Seeds were stored dry at 4°C. Before

planting, seeds were surface sterilized essentially as described previously (Ketelaar et al., 2007) with some modifications as detailed below. Seeds were immersed for 1 min in 70% ethanol, rinsed three times in sterile distilled water, and then immersed for 30 s in 30% bleach (1595 ppm available chlorine) followed by another three rinses in sterile distilled water. Excess water was removed from the seeds before they were placed into magenta boxes containing sterile Murashige and Skoog (MS) media containing vitamins (obtained from Sigma-Aldrich); five seeds were planted per box. The planted seeds were cold-shocked at 4°C for three days and then incubated at 25°C under a growth light (16/8 h light/dark cycle) for germination and growth, as detailed in the guidelines from the Arabidopsis Biological Resource Center (http://www.biosci.ohio-state.edu/~plantbio/Facilities/abrc/index.html).

To study interactions between *L. monocytogenes* and *A. thaliana,* two different bacterial inoculation procedures were used as detailed below, including (i) inoculation of plants at 21 days post germination and (ii) inoculation of seeds.

***L. monocytogenes* isolates and growth.** Isolates representing the major *L. monocytogenes* lineages (i.e., lineages I, II, IIIA, and IIIB [Wiedmann et al., 1997, Roberts et al., 2006]) were used in this study (Table 2.1). Lineage II isolates have not only been reported to be isolated more frequently from contaminated foods than lineage I and III isolates, but are also, on average, present at higher numbers in foods as compared to lineage I isolates (Chen et al., 2006). Lineage I isolates are responsible for the majority of human listeriosis outbreaks and the majority of listeriosis cases in most countries (Jeffers et al., 2001, Gray et al., 2004). Lineage III isolates are typically linked to animal infections and are rarely associated with human

Table 2.1 *L. monocytogenes* isolates used in this study.

Isolate Designation	Description	Source	Reference
10403S	Lineage II, serotype 1/2a	environmental	Bishop and Hinrichs, 1987
F2365	Lineage I, serotype 4b	food, 1985 listeriosis outbreak in Los Angeles	Linnan et al., 1988
FSL J2-071	Lineage III, serotype 4c	animal	Roberts et al., 2006
FSL J1-208	Lineage III, serotype 4a	animal	Roberts et al., 2006
DH1039	10403S containing integrated pPL3+GFP	-	Shen et al., 2006

disease (Jeffers et al., 2001) or isolated from foods (Gray et al., 2004); recent evidence suggests that isolates classified into lineage III may represent at least two distinct genetic lineages, designated IIIA and IIIB (Roberts et al., 2006). For all four isolates used in these studies, genome sequences have either been completed (F2365; Nelson et al., 2004) or are in progress (10403S, FSL J2-071, FSL J1-208; Broad Institute of Harvard and MIT, http://www.broad.mit.edu).

All *L. monocytogenes* isolates were stored at -80ºC and plated onto Brain Heart Infusion (BHI) agar before use in the experiments detailed. To prepare bacterial broth cultures, one *L. monocytogenes* colony from a BHI plate was inoculated into 5 ml of sterile BHI broth (pH 7.4) followed by incubation with shaking (210 rpm) for 12 to 18 h at 37ºC. The resulting culture was diluted 1:100 into 5 ml of fresh BHI broth and grown with shaking at 37ºC to mid-log phase (OD_{600}=0.4). This mid-log phase culture was diluted 1:200 into 10 ml of fresh BHI broth and grown with shaking at 37ºC to stationary phase (i.e., to an OD_{600}=1 followed by an additional 3 h of incubation). The stationary phase cultures were centrifuged to pellet the cells, the supernatant was removed, and the cells were re-suspended in sterile phosphate buffered saline (PBS, pH 7.4) to an approximate concentration of either $2x10^9$ or $2x10^{10}$ CFU/ml. Stocks were flash frozen in liquid nitrogen and stored at -80°C until they were used for plant inoculations as described below. *L. monocytogenes* frozen with this protocol showed high survival rates (70-80%, data not shown) and pre-prepared frozen bacterial stocks have been used in other studies of *L. monocytogenes* (e.g., Raffelsbauer et al., 1998).

***L. monocytogenes* attachment and growth after inoculation of *A. thaliana* plants.** Inoculation of *L. monocytogenes* onto leaves was performed by dip inoculating *A. thaliana* in a bacterial suspension, similar to the procedure described by Prithiviraj et al. (2005). Briefly, *A. thaliana* plants at 21 days post germination were removed from the growth boxes and the roots and superfluous stem tissue were

removed with a sterile razor blade. Using sterile tweezers, the leaves of each plant were gently submersed into a 3 ml suspension containing about $2x10^8$ CFU *L. monocytogenes*/ml, with care taken to prevent the cut end of the stem from exposure to the suspension. This general approach was used to specifically test the attachment and persistence of *L. monocytogenes* on leaf surfaces, without allowing invasion from the plant roots. The bacterial suspension used for these experiments was prepared from the frozen *L. monocytogenes* stock, which was diluted in sterile PBS containing 0.02 % Silwet surfactant (Lehle Seeds, Round Rock, TX). Silwet was added to facilitate the inoculation process (as previously described by Prithiviraj et al., 2005); without the surfactant the leaves do not submerge into the aqueous suspension. In contrast to some studies which use forced contact by pressure inoculation of bacteria onto leaf surfaces (Rahme et al., 1995, Bais et al., 2004, Prithiviraj et al., 2005), use of a surfactant is more likely to mimic natural exposure routes. While the high level of the inoculum used may not reflect natural exposure, it was used to allow for quantification of *L. monocytogenes* numbers without the need for MPN procedures, consistent with other previous studies that used similarly high inoculum levels (e.g., Ukuku and Fett [2002]). After 5 min of exposure to *L. monocytogenes*, the plants were then dipped three times (approx. 3 sec per dip) in 3 ml of sterile water containing 0.02 % Silwet to remove unattached *L. monocytogenes*. A brief (5 min) exposure was chosen to simulate the short contact time pathogens may have with plant surfaces, e.g. through irrigation with contaminated water. Other studies on *L. monocytogenes* attachment to cantaloupe, cabbage, and lettuce have also used similarly brief exposure times (1 to 10 min) (Li et al., 2002, Ukuku and Fett, 2002, Ells and Hansen, 2006). After exposure, the plants were transferred to fresh, sterile MS media. All manipulations were performed in a laminar flow hood.

L. monocytogenes were enumerated either immediately after inoculation (i.e., < 5 min after inoculation, designated as T=0 days) or after the plants were incubated for 10 days after inoculation (T=10 days) at 25°C with a 16/8 h light/dark cycle. A total of three replicate experiments were performed for T=0 and T=10 days; in each replicate, each *A. thaliana* plant was inoculated with one of the four *L. monocytogenes* strains. Bacterial enumeration at T=0 and T=10 was performed by removing four leaves from each inoculated plant and individually macerating each leaf in sterile PBS. The solutions of macerated leaves were serially diluted and plated onto BHI plates using a Spiral Plater (Spiral Biotech, Norwood, MA). Plates were incubated at 37°C for 24 h before the *L. monocytogenes* colonies were counted. As expected, since all *A. thaliana* plants were grown from surface sterilized seeds, colonies recovered on BHI showed a single colony morphology typical for *Listeria* spp., suggesting no background flora or contamination of the plants. Data are reported as CFU/cm^2 as opposed to CFU/g due to the low weight of *A. thaliana* leaves. Leaves were grouped into three categories according to size (small, medium, and large) and the surface area of a representative leaf from each group was calculated using the equation "surface area = $r_1 * r_2 * \pi$" (the surface area value was then multiplied by 2 to account for potential attachment to both sides of the leaf).

Confocal microscopy to evaluate *L. monocytogenes* presence and location on leaves from inoculated *A. thaliana* plants. Confocal microscopy was used to visually evaluate *L. monocytogenes* presence and location on leaves from inoculated *A. thaliana* plants. For these experiments, *L. monocytogenes* DH1039 (isolate 10403S expressing green fluorescent protein [GFP]) was used (Table 1; Shen et al., 2006). *L. monocytogenes* DH1039 was grown to early stationary phase (OD_{600}=1+3 h) as described above. Inoculation of 21 day old *A. thaliana* plants was performed as described above; inoculated *A. thaliana* plants were incubated on MS media for 24 h

before leaves were collected for confocal microscopy. Leaves from two *A. thaliana* plants inoculated on different dates were examined for *L. monocytogenes* presence and distribution; leaves from an uninoculated *A. thaliana* plant were examined as a negative control.

Confocal microscopy was performed using a Leica TCS SP2 System. Images were processed with Leica Confocal Software version 2.61 (Leica Microsystems, Mannheim, Germany).

***L. monocytogenes* presence and persistence on *A. thaliana* plants grown from seeds inoculated with *L. monocytogenes*.** *A. thaliana* seeds were surface sterilized and 10 µl of a *L. monocytogenes* suspension (approx. 8.3 log CFU), prepared by thawing *L. monocytogenes* and adding Silwet surfactant to a final concentration of 0.02 %, was placed onto each seed. All five seeds in a given box were inoculated with the same *L. monocytogenes* isolate. Excess water from the bacterial suspension was evaporated off the seeds in a sterile laminar flow hood. As described above, the seeded boxes were cold-shocked at 4°C for three days and then transferred to germination and growth conditions (25°C with a 16/8 h light/dark cycle). Plants were allowed to grow for one week post-germination, then *L. monocytogenes* was enumerated on all available leaves from each germinated plant as described above. A total of 3 replicate experiments were performed.

Statistical Analysis. All data were analyzed with either One- or Two-way ANOVA or Chi-square test using Minitab (State College, PA, USA). P-values less than 0.05 were considered significant.

Results and Discussion

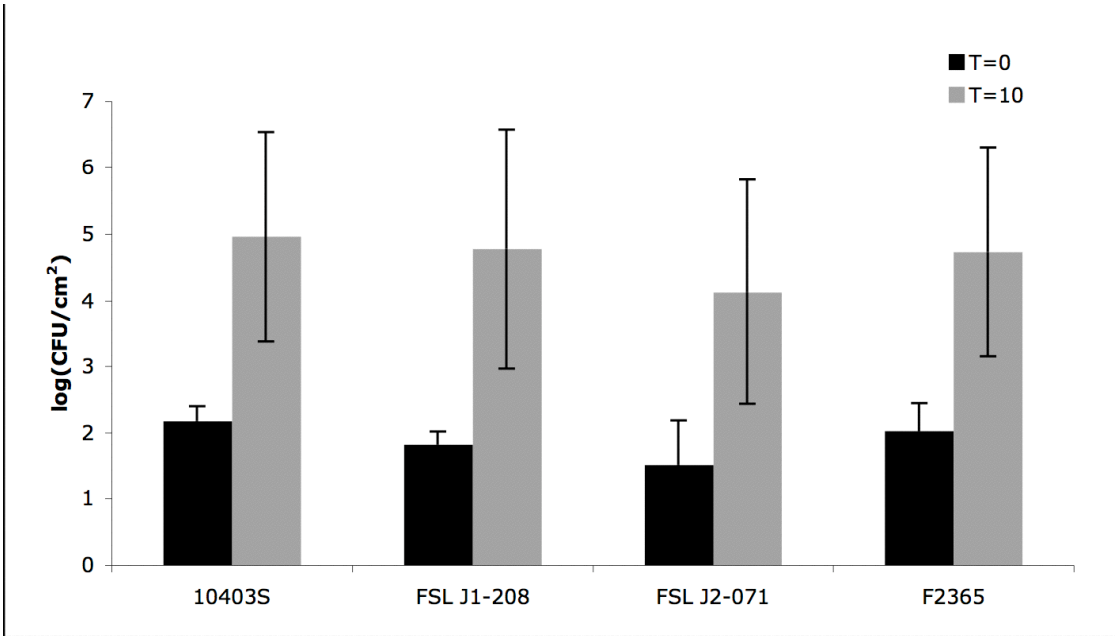
In this study, we used the plant model, *A. thaliana*, to develop an experimental system to gain a better understanding of pre-harvest interactions between *L. monocytogenes*

and plants. Even considering the typical limitations of laboratory studies (e.g., inoculation with high bacterial numbers, growth of plants in gnobiotic conditions that do not account for the effects of endogenous plant microflora), our data indicate that: (i) *L. monocytogenes* is able to rapidly attach to and multiply on *A. thaliana* after inoculation on leaves; (ii) *L. monocytogenes* can colonize intercellular spaces in leaves where it could be protected from decontamination and sanitation treatments; and (iii) contamination of seeds can generate *L. monocytogenes*-contaminated mature plants. These findings suggest that *L. monocytogenes* may require similar safety recommendations as for other produce-transmitted foodborne pathogens, such as enterohemorrhagic *E. coli* and *Salmonella* (National Advisory Committee on Microbiological Criteria for Foods, 1999).

Isolates representing the major *L. monocytogenes* lineages are able to attach to and grow on *A. thaliana* leaves. To develop a better understanding of the potential for *L. monocytogenes* to attach to and grow on plants after contamination of leaves, each *A. thaliana* plant was inoculated with one of the four *L. monocytogenes* isolates representing the main *L. monocytogenes* lineages. The goal of these experiments was to simulate pre-harvest exposure of produce plants to *L. monocytogenes*, e.g., through contaminated irrigation water. To accomplish this goal, we used an experimental system designed to mimic possible contamination events and examine the potential for *L .monocytogenes* attachment to and persistence on *A. thaliana*.

L. monocytogenes numbers recovered from *A. thaliana* leaves after short exposure (5 min) and repeated rinsing (Fig. 2.1; T=0 day) ranged from 1.52 (FSL J2-071) to 2.17 log CFU/cm^2 (10403S); numbers of recovered bacteria at T=0 did not differ significantly among the four isolates (ANOVA factor "isolate" was not significant; $p>0.05$). These data indicate that *L. monocytogenes* is capable of rapid

Figure 2.1 *L. monocytogenes* recovered from *A. thaliana* leaves immediately after inoculation (T=0) or following 10 days of incubation (T=10). *A. thaliana* plants (21 days after germination) were inoculated with the four *L. monocytogenes* strains and leaves were tested for *L. monocytogenes* (T=0; representing leaves collected after 5 min) or at 10 days after inoculation (T=10 days). Data represent the average *L. monocytogenes* numbers in log CFU/cm^2 from three independent replicate experiments (in each replicate 4 leaves from each of 1-2 plants were tested per isolate); error bars represent the standard deviations.


T=0
T=10
log(CFU/cm²)
0
1
2
3
4
5
6
7
10403S
FSL J1-208
FSL J2-071
F2365

attachment to leaf surfaces. Similar to our results, Ells and Hansen (2006) found that *L. monocytogenes* Scott A attached to cabbage after <5 min of exposure, yielding 4.3 log CFU *L. monocytogenes*/cm^2. Ukuku and Fett (2002) found that immersion of cantaloupe rinds (which represents a harsher environment as compared to cut leafy vegetables such as cabbage) in a mixture of four isolates at 10^8 CFU/ml for 10 min yielded 3.5 log CFU/cm^2 *L. monocytogenes* on the rinds. Other previous studies also revealed rapid (10 min) attachment of *L. monocytogenes* to abiotic surfaces (i.e., rubber and stainless steel) at densities of approximately 2.5 to 4 log CFU/cm^2 (Smoot and Pierson, 1998). Thus, *L. monocytogenes* appears to be able to rapidly attach to abiotic and biotic (plant) surfaces, suggesting that even brief exposure, e.g., during irrigation, could lead to surface contamination with *L. monocytogenes*.

To determine the ability of *L. monocytogenes* to survive and grow on *A. thaliana* leaves, each plant inoculated with one of the four *L. monocytogenes* isolates was also incubated for 10 days (Fig. 2.1; T=10 days). Average *L. monocytogenes* numbers recovered from the leaves at 10 days post-inoculation ranged from 4.12 (FSL J2-071) to 4.96 log CFU/cm^2 (10403S). *L. monocytogenes* growth on leaves (defined as the difference in average recovered bacterial numbers from leaves between T=10 and T=0 days) ranged from 2.60 (FSL J2-071) to 2.95 log CFU/cm^2 (FSL J1-208). Bacterial numbers increased significantly between T=0 and T=10 (ANOVA, factor "time" was significant; $p<0.001$). No significant statistical interaction was observed between the factors "time" and "isolate" ($p=0.996$), indicating that the four isolates did not differ in ability to grow on the leaves. While Gorski et al. (2004) found significant strain differences in the ability of *L. monocytogenes* to attach and grow on alfalfa sprouts, this group also did not report any significant differences in attachment among *L. monocytogenes* lineages (based on nine different isolates). Kalmokoff and collaborators (2001) reported that *L. monocytogenes* attachment to abiotic surfaces

also varied by isolate, but they, too, did not find any associations between lineage and attachment. On the other hand, others have previously reported evidence for differences in attachment and biofilm formation among *L. monocytogenes* lineages (e.g., Djordjevic et al., 2002). Taken together, these findings suggest that there are no consistent differences among *L. monocytogenes* lineages in their ability to attach to and grow on plants, even though individual isolates may have unique phenotypic characteristics that affect plant attachment and growth.

No previous studies have used *A. thaliana* to study *L. monocytogenes* attachment to and growth on plants, although previous studies have tested *L. monocytogenes* attachment and survival on a number of vegetable plants, including lettuce, spinach, radishes, and alfalfa sprouts (Ukuku and Fett, 2002, Gorski et al., 2004, Jablasone et al., 2005, Ells and Hansen, 2006). For example, Ells and Hansen (2006) showed that *L. monocytogenes* Scott A numbers increased on cabbage by approximately 1.5 log over 24 h after initial attachment of 4.3 log CFU *L. moncoytogenes*/cm^2. Li et al. (2002) also found that *L. monocytogenes* numbers generally increased on contaminated lettuce stored at 15°C for 7 days. On the other hand, Ukuku and Fett (2002) reported that *L. monocytogenes* numbers on cantaloupe decreased by 1 to 2 log over 15 days. Thus, while many leafy plants appear to allow growth of *L. monocytogenes* deposited on plant leaves, some fruit surfaces may not, possibly reflecting limited nutrient availability on fruit surfaces.

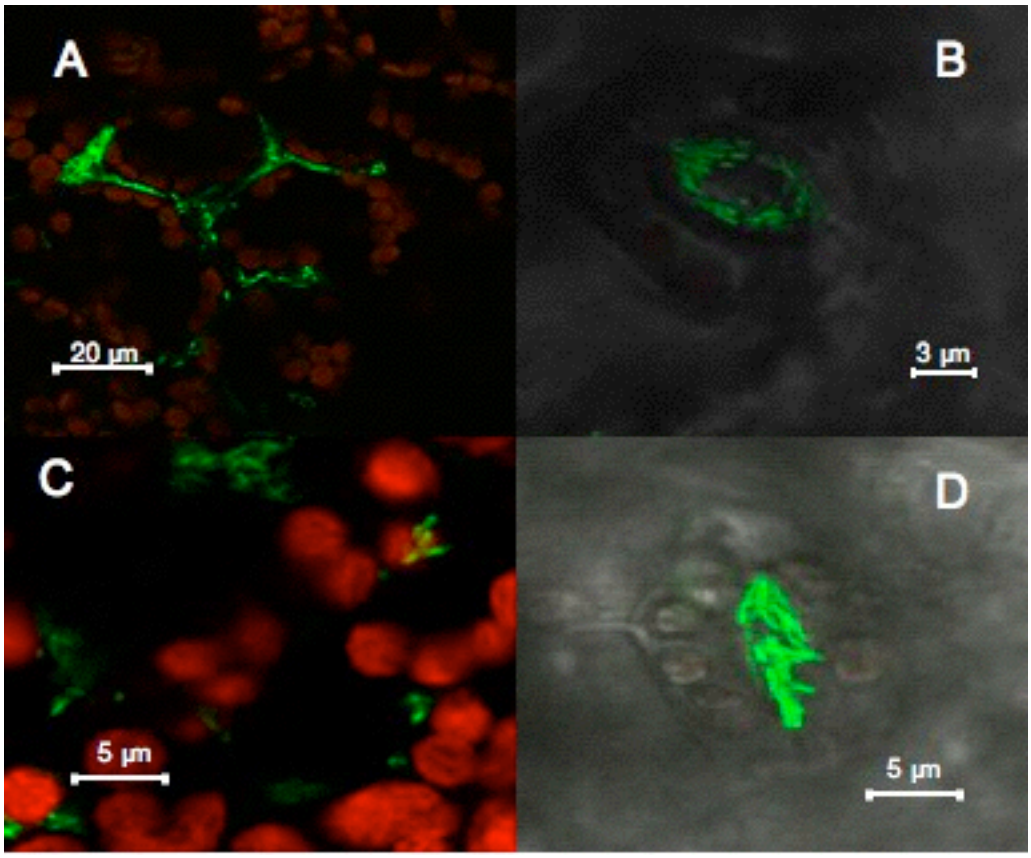
Studies of other foodborne pathogens have revealed that *E. coli* O157:H7 and *Salmonella* can attach to and grow on a variety of plants, including lettuce, spinach, carrots, and tomatoes (Solomon et al., 2002, Jablasone et al., 2005). *E. coli* O157:H7 and *Salmonella* Serovar Newport were noted to have similar growth on *A. thaliana* as on vegetable plants (Charkowski et al., 2002, Cooley et al., 2003). The present study

extends the use of *A. thaliana* as an appropriate model for studying interactions between leafy vegetables and *L. monocytogenes*.

***L. monocytogenes* deposited on *A. thaliana* leaves can internalize via the stomata, but does not invade the leaf cells.** Confocal microscopy of *A. thaliana* leaves inoculated with *L. monocytogenes* expressing GFP allowed visualization of *L. monocytogenes* inside the stomatal openings on the leaf surface after 24 h of incubation (Figure 2.2B & D). *L. monocytogenes* cells were also observed in intercellular spaces, deeper in the leaf tissue (Figure 2.2A & C). There was no evidence of *L. monocytogenes* inside leaf cells. Uninoculated control leaves showed no fluorescence patterns such as those observed for the inoculated leaves. Interestingly, Seo et al. (1999) showed that *E. coli* O157:H7 can accumulate in stomata and then penetrate deeper into leaf tissue in inoculated lettuce leaves. Localization of *L. monocytogenes* or other foodborne pathogens within leaf stomata could confer some protection against pathogen removal, consistent with a number of studies that have shown limited reductions in bacterial pathogen numbers following washing and surface sanitizing treatments of contaminated vegetables (e.g., Wei et al., 1995, Zhang and Farber, 1996, Beuchat and Ryu, 1997, Hellström et al., 2006). Future studies to determine the duration of *L. monocytogenes* persistence in stomata (beyond 24 h post inoculation) and to quantify *L. monocytogenes* removal from contaminated plants (when *L. monocytogenes* is located in the leaf stomata) will help us to further understand *L. monocytogenes* interactions with leafy plants and their relevance for pre- and post-harvest food safety.

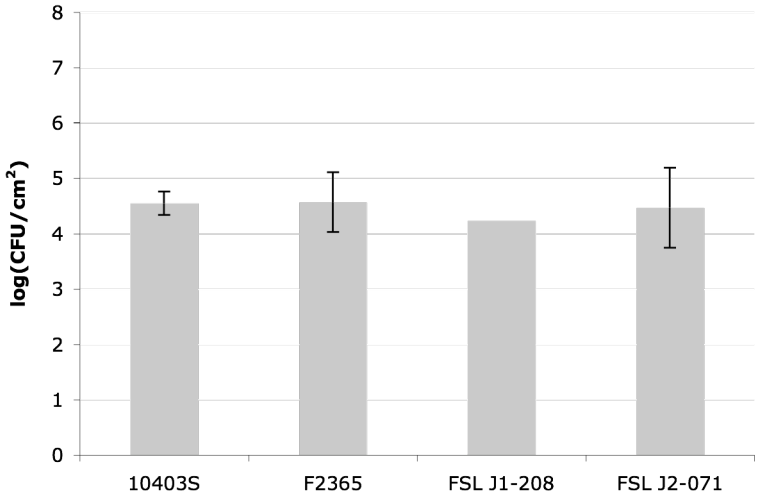
Seeds exposed to *L. monocytogenes* can germinate and produce contaminated plants. Use of pathogen-contaminated manure used as fruit and vegetable crop fertilizer may represent a risk factor for contamination of plants, with pathogens persisting to harvest time (Solomon et al., 2002, Hutchison et al., 2005,

Figure 2.2 Confocal images of *L. monocytogenes* in *A. thaliana* leaf intercellular spaces. Panels A and D show the GFP expressing *L. monocytogenes* cells interspersed between the *A. thaliana* leaf cells (the red circular structures are chloroplasts, which define the leaf cell boundaries). Panels B and C show *L. monocytogenes* cells inside *A. thaliana* stomata; the guard cells surrounding the stomata form a distinctive elliptical shape and faint red spots corresponding to deeper chloroplasts can be seen within the guard cells. Bars indicate length in microns (μm).


A
20 µm
B
3 µm
C
5 µm
D
5 µm

Johannessen et al., 2005). To mimic conditions in nature where *L. monocytogenes* could contaminate germinating seeds, *L. monocytogenes* was surface-inoculated onto *A. thaliana* seeds that were allowed to germinate and grow for one week before leaves were harvested for *L. monocytogenes* enumeration. While for three isolates (10403S, F2365, and FSL J2-071), between 24 and 36% of all inoculated seeds successfully germinated, only 4% (1/25) of seeds inoculated with isolate FSL J1-208 germinated; by comparison an average of 45% of uninoculated seeds geminated (based on 227 uninoculated planted seeds). Using a 4x2 contingency table and Pearson Chi-square test, there was an overall significant effect of isolate on germination ($p<0.05$). Statistical analysis of individual 2x2 contingency tables showed that the germination rate of both FSL J2-071 and F2365 were significantly higher than that of FSL J1-208 (Fisher's Exact test $p<0.05$; the germination rate of 10403S were borderline significantly higher than that for FSL J1-208, $p<0.10$; Fisher's exact test). The low germination rate for seeds contaminated with *L. monocytogenes* FSL J1-208 suggests that this isolate may, by an unknown mechanism, reduce or prevent *A. thaliana* seed germination. *L. monocytogenes* numbers recovered from the leaves of plants grown from inoculated seeds ranged from 4.23 log CFU/cm^2 (FSL J1-208) to 4.57 log CFU/cm^2 (F2365); bacterial numbers did not differ significantly between 10403S, F2365, and FSL J2-071 (Fig. 2.3; FSL J1-208 was not included in the statistical analyses as only a single data point was available for this isolate due to the low germination rate of seeds contaminated with this isolate). Our findings are consistent with previous studies that also recovered *L. monocytogenes* from plants grown from contaminated seeds, including alfalfa sprouts, garden peas, spinach, and tomatoes (Gorski et al., 2004, Jablasone et al., 2005, Saroj et al., 2006). Interestingly, Cooley et al. (2003) found that seeds produced by *A. thaliana* grown in *E. coli* O157:H7-

Figure 2.3 Recovery of *L. monocytogenes* from leaves from one-week *A. thaliana* plants grown from *L. monocytogenes*-inoculated seeds. *L. monocytogenes* isolates were inoculated onto surface sterilized *A. thaliana* seeds, which were allowed to germinate and grow seven days before enumeration of *L. monocytogenes* on each available leaf (2-4 leaves per plant). Data (in log CFU/cm^2 of *L. monocytogenes* recovered) represents the average from at least three independent experiments for isolates 10403S, F2365, and FSL J2-071; error bars represent standard deviations. The bar for isolate FSL J1-208 reflects only one replicate because only one seed of 25 germinated (thus, no error bars for FSL J1-208 are available).


8
7
6
5
4
3
2
1
0
log(CFU/cm2)
10403S
F2365
FSL J1-208
FSL J2-071

contaminated soil were positive for *E. coli* O157:H7, indicating that for some bacterial foodborne pathogens contamination can even persist through seed production.

Conclusions

Using *A. thaliana* as a model system, we have shown that *L. monocytogenes* contamination of plants can occur by different mechanisms. The observation that contaminated seeds can grow into plants with *L. monocytogenes* leaf contamination suggests that exposure to *L. monocytogenes*-contaminated soil and manure, even early in the life cycle of plants, may present an important risk factor for contamination of vegetables, consistent with findings by other researchers (Guo et al., 2001, Solomon et al., 2002, Cooley et al., 2003). We also demonstrated that *L. monocytogenes* deposited on mature plant leaves can rapidly attach, grow and persist over time, suggesting that application of contaminated irrigation water represents another risk factor for contamination of vegetables. However, further studies are necessary to determine the effects of endogenous microflora on *L. monocytogenes* attachment and persistence on plants and to determine attachment and persistence after contamination with low *L. monocytogenes* levels. The observation that *L. monocytogenes* can be present in internal leaf spaces, presumably entering through the stomata, may provide a functional explanation for the limited efficacy of washing procedures in eliminating *Listeria* from produce (Kilonzo-Nthenge et al., 2006) and emphasizes the importance of intervention strategies that are capable of preventing pathogen invasion or eliminate pathogens found inside leaves.

Our findings on the ability of *L. monocytogenes* to attach to and grow on *A. thaliana* leaves and to yield contaminated plants after inoculation of seeds are in concordance with findings in different vegetable plants, including spinach and lettuce (Gorski et al., 2004, Jablasone et al., 2005). We thus propose that *A. thaliana* provides

a suitable model for studies of *L. monocytogenes*-plant interactions that may contribute to development of improved pre-harvest control strategies for foodborne pathogens. Use of a plant with a completed genome sequence (as well as tools for genetic modifications) in combination with pathogen strains with available genome sequences will enable future advanced studies on *L. monocytogenes*-plant interactions using a genomics approach. In addition, *A. thaliana* could be used as a model system to explore genetic modification in plants, including expression of anti-listerial compounds or antimicrobials that could ultimately be used to minimize contamination of fruits and vegetables.

REFERENCES

Arabidopsis Biological Resource Center. Plant and Seed Handling Protocols, Last accessed on 10-1-07. Available at http://www.biosci.ohio-state.edu/~plantbio/Facilities/abrc/abrchome.htm.

Bais, H.P., Fall, R., Vivanco, J.M., 2004. Biocontrol of *Bacillus subtilis* against infection of *Arabidopsis* roots by *Pseudomonas syringae* is facilitated by biofilm formation and surfactin production. Plant Physiol. 134, 307-319.

Bari, M.L., Nakauma, M., Todoriki, S., Juneja, V.K., Isshiki, K., and Kawamoto, S., 2005. Effectiveness of irradiation treatments in inactivating *Listeria monocytogenes* on fresh vegetables at refrigeration temperature. J. Food Prot. 68, 318-323.

Beuchat, L.R. and Ryu, J.H., 1997. Produce handling and processing practices. Emerg. Infect. Dis. 4, 459-465.

Bishop, D.K. and Hinrichs, D.J., 1987. Adoptive transfer of immunity to *Listeria monocytogenes*. The influence of in vitro stimulation on lymphocyte subset requirements. J. Immunol. 139, 2005-2009.

Brandl, M.T., 2006. Fitness of human enteric pathogens on plants and implications for food safety. Ann. Rev. Phytopathol. 44, 367-392.

Broad Institute of Harvard and MIT. *Listeria monocytogenes* sequencing project. Last accessed on 9-26-07. Available at http://www.broad.mit.edu.

Caggia, C., Randazzo, C. L., Di Salvo, M., Romeo, F., and Giudici, P., 2004. Occurrence of *Listeria monocytogenes* in green table olives. J. Food. Prot. 67, 2189-94.

Charkowski, A. O., Barak, J. D., Sarreal, C. Z., and Mandrell, R. E., 2002. Differences in growth of *Salmonella* Enterica and *Escherichia coli* O157:H7 on alfalfa sprouts. Appl. Environ. Microbiol. 68, 3114-3120.

Chen, Y., Ross, H., Gray, M.J., Wiedmann, M., Whiting, R.C., and Scott, V.N., 2006. Attributing risk to *Listeria monocytogenes* subgroups: Dose response in relation to

genetic lineages. J. Food. Prot. 69, 335-344.

Cooley, M.B., Miller, W.G., and Mandrell, R.E., 2003. Colonization of *Arabidopsis thaliana* with *Salmonella enterica* and Enterohemorrhagic *Escherichia coli O157:H7* and competition by *Enterobacter asburiae*. Appl. Environ. Microbiol. 69, 4915-4926.

Djordjevic, D., Wiedmann, M., and McLandsborough, L.A., 2002. Microtiter plate assay for assessment of *Listeria monocytogenes* biofilm formation. Appl. Environ. Micro. 68, 2950-2958.

Ells, T. C. and Truelstrup Hansen, L., 2006. Isolate and growth temperature influence *Listeria* spp. attachment to intact and cut cabbage. Int. J. Food Micro. 111, 34-42.

Garner, M.R., James, K.E., Callahan, M. C., Wiedmann, M., and Boor, K. J., 2006. Exposure to salt and organic acids increases the ability of *Listeria monocytogenes* to invade Caco-2 cells but decreases its ability to survive gastric stress. Appl. Environ. Microbiol. 72, 5384-5395.

Gray, M. J., Zadoks, R. N., Fortes, E. D., Dogan, B., Cai, S., Chen, Y., Scott, V. N., Gombas, D. E., Boor, K. J., and Wiedmann, M., 2004. Food and human isolates of *Listeria monocytogenes* form distinct but overlapping populations. Appl. Environ. Microbiol. 70, 5833-5841.

Gorski, L., Palumbo, J. D., and Nguyen, K. D., 2004. Strain-specific differences in the attachment of *Listeria monocytogenes* to alfalfa sprouts. J. Food Prot. 67, 2488-2495.

Guo, X., Chen, J., Brackett, R. E., and Beuchat, L. R., 2001. Survival of *Salmonellae* on and in tomato plants from the time of inoculation at flowering and early stages of fruit development through fruit ripening. Appl. Environ. Microbiol. 67, 4760-4764.

Gupta, S. K., Nalluswami, K., Snider, C., Perch, M., Balasegaram, M., Burmeister, D., Lockett, J., Sandt, C., Hoekstra, R. M., and Montgomery, S., 2007. Outbreak of *Salmonella* Braenderup infections associated with Roma tomatoes, northeastern United States, 2004: a useful method for subtyping exposures in field investigations. Epidemiol. Infect. Published online by Cambridge University Press 05 Feb 2007.

Hellström, S., Kervinen, R., Lyly, M., Ahvenainen-Rantala, R., and Korkeala, H., 2006. Efficacy of disinfectants to reduce *Listeria monocytogenes* on precut iceberg lettuce. J. Food Prot. 69, 1565-1570.

Ho, J. L., Shands, K. N., Friedland, G., Eckind, P., and Fraser, D.W., 1986. An outbreak of type 4b *Listeria monocytogenes* infection involving patients from eight Boston hospitals. Arch. Intern. Med. 146, 520-524.

Jablasone, J., Warriner, K., and Griffiths, M., 2005. Interactions of *Escherichia coli* O157:H7, *Salmonella* Typhimurium and *Listeria monocytogenes* plants cultivated in a gnotobiotic system. Int. J. Food Microbiol. 99, 7-18.

Jeffers, G. T., Bruce, J. L., McDonough, P.L., Scarlett, J., Boor, K.J., and Wiedmann, M., 2001. Comparative genetic characterization of *Listeria monocytogenes* isolates from human and animal listeriosis cases. Microbiol. 147, 1095–1104.

Johannessen, G. S., Bengtsson, G. B., Heier, B. T., Bredholt, S., Wasteson, Y., and Rorvik, L. M., 2005. Potential uptake of *Escherichia coli* O157:H7 from organic manure into Crisphead Lettuce. Appl. Environ. Microbiol. 71, 2221-2225.

Kalmokoff, M. L., Austin, J. W., Wan, X.D., Sanders, G., Banerjee, S., and Farber, J.M., 2001. Adsorption, attachment and biofilm formation among isolates of *Listeria monocytogenes* using model conditions. J. Appl. Microbiol. 91, 725-734.

Ketelaar, T., Allwood, E.G., Hussey, P.J., 2007. Actin organization and root hair development are disrupted by ethanol-induced overexpression of *Arabidopsis* actin interacting protein 1 (AIP1). New Phytol. 174, 57-62.

Kilonzo-Nthenge, A., Chen, F., Godwin, S.L., 2006. Efficacy of home washing methods in controlling surface microbial contamination on fresh produce. J. Food. Prot. 69, 330-334.

Li, Y., Brackett, R. E., Chen, J., and Beuchat, L.R., 2002. Mild heat treatment of lettuce enhances growth of *Listeria monocytogenes* during subsequent storage at 5ºC or 15ºC. J. Appl. Microbiol. 92, 269-275.

Linnan, M. J., Mascola, L., Lou, X. D., Goulet, V., May, S., Salminen, C., Hird, D.W., Yonekura, M. L., Hayes, P., and Weaver, R., 1988. Epidemic listeriosis associated with Mexican style cheese. N. Engl. J. Med. 319, 823-828.

Liu, D., Lawrence, M. L., Wiedmann, M., Gorski, L., Mandrell, R.E., Ainsworth, J.A.,

and Austin, F.W., 2006. *Listeria monocytogenes* Subgroups IIIA, IIIB, and IIIC delineate genetically distinct populations with varied pathogenic potential. J. Clin. Microbiol. 44, 4229-4233.

National Advisory Committee on Microbiological Criteria for Foods. 1999. Microbiological safety evaluations and recommendations on sprouted seeds. Int. J. Food Microbiol. 52, 123-153.

Nelson, K., Fouts, D., Mongodin, E.F., Ravel, J., DeBoy, R.T., Kolonay, J.F., Rasko, D.A., Angiuoli, S.V., Gill, S.R., Paulsen, I.T., Peterson, J., White, O., Nelson, W.C., Nierman, W., Beanan, M.J., Brinkac, L.M., Daugherty, S.C., Dodson, R.J., Durkin, A.S., Madupu, R., Haft, D.H., Selengut, J., Van Aken, S., Khouri, H., Fedorova, N., Forberger, H., Tran, B., Kathariou, S., Wonderling, L.D., Uhlich, G.A., Bayles, D.O., Luchansky, J.B., and Fraser, C.M., 2004. Whole genome comparisons of serotype 4b and 1/2a strains of the food-borne pathogen *Listeria monocytogenes* reveal new insights into the core genome components of this species. Nucleic Acids Res. 32, 2386-2395.

Pell, A.N., 1997. Manure and microbes: public and animal health problem? J. Dairy Sci. 80, 2673-2681.

Prithiviraj, B., Bais, H. P., Jha, A.K., Vivanco, J.M., 2005. *Staphylococcus aureus* pathogenicity on *Arabidopsis thaliana* is mediated either by a direct effect of salicylic acid on the pathogen or by SA-dependent, NPR1-independent host responses. The Plant Journal 42, 417-432.

Raffelsbauer, D., Bubert, A., Engelbrecht, F., Scheinpflug, J., Simm, A., Hess, J., Kaufmann, S.H.E., and Goebel, G., 1998. The gene cluster inlC2DE of *Listeria monocytogenes* contains additional new internalin genes and is important for virulence in mice. Mol. Gen. Genet. 260, 144-158.

Rahme, L.G., Stevens, E.J., Wolfort, S.F., Shao, J., Tompkins, R.G. and Ausubel, F.M., 1995. Common virulence factors for bacterial pathogenicity in plants and animals. Science 30, 1899-1902.

Roberts, A., Nightingale, K., Jeffers, G., Fortes, E., Kongo, J. M., and Wiedmann, M., 2006. Genetic and phenotypic characterization of *Listeria monocytogenes* lineage III. Microbiol. 152, 685-693.

Sagoo, S. K., Little, C. L., and Mitchell, R. T., 2003. Microbiological quality of open ready-to-eat salad vegetables: effectiveness of food hygiene training of management. J. Food Prot. 66, 1581-1586.

Saroj, S. D., Shashidhar, R., Pandey, M., Dhokane, V., Hajare, S., Sharma, A., and Bandekar, J. R., 2006. Effectiveness of radiation processing in elimination of *Salmonella* Typhimurium and *Listeria monocytogenes* from sprouts. J. Food Prot. 69, 1858-1864.

Samadpour, M., Barbour, M. W., Nguyen, T., Cao, T. M., Buck, F., Depavia, G. A., Mazengia, E., Yang, P., Alfi, D., Lopes, M., and Stopforth, J. D., 2006. Incidence of enterohemorrhagic *Escherichia coli*, *Escherichia coli* O157, *Salmonella*, and *Listeria monocytogenes* in retail fresh ground beef, sprouts, and mushrooms. J. Food Prot. 69, 441-443.

Sandt, C. H., Krouse, D. A., Cook, C. R., Hackman, A. L., Chmielecki, W. A., and Warren, N. G., 2006. The key role of pulsed-field gel electrophoresis in investigation of a large multiserotype and multistate food-borne outbreak of *Salmonella* infections centered in Pennsylvania. J. Clin. Microbiol. 44, 3208-3212.

Schlech, W. F., Lavigne, P. M., Bortolussi, R. A., Allen, A. C., Haldane, E. V., Wort, A. J., Hightower, A. W., Johnson, S. E., King, S. H., Nicholls, E. S., and Broome, C. V., 1983. Epidemic listeriosis--evidence for transmission by food. N. Engl. J. Med. 308, 203-206.

Seo, K. H. and Frank, J. F., 1999. Attachment of Escherichia coli O157:H7 to lettuce leaf surface and bacterial viability in response to chlorine treatment as demonstrated by using confocal scanning laser microscopy. J. Food Prot. 62, 3-9.

Shen, A. and Higgins, D. E., 2006. The MogR transcriptional repressor regulates nonhierarchal expression of flagellar motility genes and virulence in *Listeria monocytogenes*. PloS Pathog. 2, 282-295.

Slifko, T. R., Smith, H. V., and Rose, J.B., 2000. Emerging parasite zoonoses associated with water and food. Int. J. Parasitol. 30, 1379-1393.

Smoot, L. M. and Pierson, M. D., 1998. Influence of environmental stress on the kinetics and strength of attachment of *Listeria monocytogenes* Scott A to Buna-N Rubber and stainless steel. J. Food Prot. 61, 1286-1292.

Soderstrom, A., Lindberg, A., and Andersson, Y., 2005. EHEC O157 outbreak in Sweden from locally produced lettuce, August-September 2005. Euro. Surveill. 10, E050922.1.

Solomon, E. B., Yaron, S., and Matthews, K. R., 2002. Transmission of *Escherichia coli* O157:H7 from contaminated manure and irrigation water to lettuce plant tissue and its subsequent internalization. Appl. Environ. Microbiol. 68, 397-400.

Sivapalasingam, S., Friedman, C. R., Cohen, L., and Tauxe, R. V., 2004. Fresh produce: a growing cause of outbreaks of foodborne illness in the United States, 1973 through 1997. J. Food Prot. 67, 2342-2353.

The Arabidopsis Genome Initiative. 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408, 796-815.

Ukuku, D. O. and Fett, W., 2002. Behavior of *Listeria monocytogenes* inoculated on cantaloupe surfaces and efficacy of washing treatments to reduce transfer from rind to fresh-cut pieces. J. Food Prot. 65, 924-930.

U.S. Food and Drug Administration, Food Safety and Inspection Service. 2003. Quantitative assessment of the relative risk to public health from foodborne *Listeria monocytogenes* among selected categories of ready-to-eat foods, Washington, D.C. Available at: http://www.foodsafety.gov/~dms/lmr2-toc.html. Last accessed July 2007.

Varma, J. K., Samuel, M. C., Marcus, R., Hoekstra, R.M., Medus, C., Segler, S., Anderson, B.J., Jones, T.F., Shiferaw, B., Haubert, N., Megginson, M., McCarthy, P.V, Graves, L., Gilder, T.V., and Angulo, F.J., 2007. *Listeria monocytogenes* infection from foods prepared in a commercial establishment: A case-control study of potential sources of sporadic illness in the United States. Clin. Infect. Dis. 44, 521-528.

Vermeulen, A., Gysemans, K. P., Bernaerts, K., Geeraerd, A. H., Van Impe, J. F., Debevere, J., and Devlieghere, F., 2007. Influence of pH, water activity and acetic acid concentration on *Listeria monocytogenes* at 7 degrees C: Data collection for the development of a growth/no growth model. Int. J. Food Microbiol. 114, 332-341.

Wei, C. I., Huang, T. S., Kim, J.M., Lin, W.F., Tamplin, M.L., and Bartz, J.A., 1995.

Growth and survival of *Salmonella* Montevideo on tomatoes and disinfection with chlorinated water. J. Food Prot. 58, 829-836.

Wiedmann, M., Bruce, J. L., Keating, C., Johnson, A. E., McDonough, P. L., and Batt, C. A., 1997. Ribotypes and virulence gene polymorphisms suggest three distinct *Listeria monocytogenes* lineages with differences in pathogenic potential. Infect. Immun. 65, 2707-2716.

Zhang, S. and Farber, J. M., 1996. The effects of various disinfectants against *Listeria monocytogenes* on fresh-cut vegetables. Food Microbiol. 13, 311-321.

CHAPTER 3

CONTRIBUTIONS OF SIX LINEAGE-SPECIFIC INTERNALIN-LIKE GENES TO INVASION EFFICIENCY OF *LISTERIA MONOCYTOGENES*

Abstract

Listeria monocytogenes strains are divided into at least three lineages that seem to differ in virulence. Internalins are surface or secreted proteins that encode leucine rich repeats and *L. monocytogenes* encodes species- as well as lineage-specific internalin and internalin-like genes; internalin A and B have previously been shown to be critical for invasion of human host cells by *L. monocytogenes*. Transcription of selected internalins is regulated by the virulence gene regulator PrfA and/or the stress responsive alternative sigma factor σ^B. We hypothesized that lineage-specific internalin-like genes may contribute to differential virulence and niche adaptation of the *L. monocytogenes* lineages. Initial quantitative RT-PCR (qRT-PCR) showed that the six selected lineage-specific internalin-like genes were transcribed in cells grown in BHI at 16 and 37ºC. *lsiIIX* (*l*ineage-*s*pecific *i*nternalin-like gene, lineage *II*) showed significantly higher transcript levels in log-phase cells grown at 37 as compared to 16ºC. While none of the genes showed evidence for an upstream PrfA binding site, *lsiIA* was preceded by a putative σ^B-dependent promoter and showed σ^B-dependent transcription by qRT-PCR. None of the null mutants in lineage-specific internalin-like genes differed from their respective parent strain in the ability to invade either human intestinal epithelial or hepatocyte-like cell lines. In addition, all three mutants in lineage I-specific internalin-like genes exhibited the same growth condition-dependent invasion phenotype as their parent strain (approx. 1.5 log higher invasion efficiency when grown at 30 ºC without aeration versus 37ºC with aeration). Among six lineage-specific *L. monocytogenes* internalin-like genes, we identified one lineage I-specific,

σ^B-dependent gene and one lineage II- specific gene with growth condition-dependent transcription. Despite structural similarities to internalins with known roles in host cell attachment/invasion, none of the lineage-specific internalin-like genes appear to contribute to invasion. Combined with the observation that some non-pathogenic *Listeria* species also carry internalin genes, our findings suggest a broad role of *Listeria* internalins, not limited to attachment and invasion of human cells. Due to the broad host-range of *L. monocytogenes* and the fact that transcription of internalin-like genes can differ considerably depending on growth condition, elucidating the function of different internalins and internalin-like genes will remain a challenge.

Introduction

Among the six species in the genus *Listeria*, four species (*L. innocua, L. seeligeri, L. welshimeri,* and *L. grayi*) are considered non-pathogenic saprophytes. While the species *L. ivanovii* appears to cause disease only in ruminants, *L. monocytogenes* is an opportunistic, food-borne pathogen with a broad host range, including humans, a range of other mammals (e.g., goats, sheep, cattle) and birds (Pell, 1997). In humans, *L. monocytogenes* can cause a serious, often fatal, disease in vulnerable individuals, specifically the elderly, infants and young children, and pregnant women. Major manifestations of systemic listeriosis infection include septicaemia, meningitis, encephalitis, and abortion (Swaminathan and Gerner-Smidt, 2007). While human listeriosis is typically rare (0.1 to 11.3 cases/million population per year), it is characterized by a high case mortality rate of approx. 20% (Mead et al., 1999; Anonymous, 2004).

L. monocytogenes is a facultative intracellular pathogen and has been shown to invade a variety of mammalian host cells (Seveau et al., 2007), including non-phagocytic intestinal epithelial cells (Tilney and Portnoy, 1989) and hepatocytes (Dramsi et al., 1995). While a number of proteins appear to contribute to *L.*

monocytogenes host cell attachment and invasion (Seveau et al., 2007), internalin proteins (e.g., InlA and InlB) appear to be particularly important for adhesion and invasion of mammalian host cells (Galliard et al., 1991; Dramsi et al., 1995). InlA binds to the host cell surface molecule E-cadherin; this interaction is critical for invasion of intestinal epithelial cells (Lecuit et al., 2001) and may also contribute to crossing of the placental barrier (Lecuit et al., 2004). Interestingly, InlA binds to human and guinea pig E-cadherin, but not to mouse or rat E-cadherin (Lecuit et al., 2001). InlB has multiple identified host cell receptors and is thought to be important for invasion of heptocytes (Dramsi et al., 1995), and brain microvascular endothelial cells (Greiffenberg et al., 1998). The primary structural features of InlA and InlB are the presence of a cell wall binding motif (LPXTG, a common gram-positive cell wall anchor, in InlA and GW repeats in InlB) and a series of leucine-rich repeat domains (LRR) (Jonquieres et al., 1999; Glaser et al., 2001). LRR are common structural motifs among proteins known to facilitate protein-protein interactions (Kobe and Kajava, 2001). In addition to InlA and InlB, *L. monocytogenes* carries more than 25 other genes characterized by the presence of LRR (Bierne et al., 2007). While no other genes encoding proteins with LRR have been shown to contribute to invasion of tissue culture cells, some have been shown to contribute to virulence in animal models (e.g., *inlC*, *inlH*, and *inlJ*) (Engelbrecht et al., 1998; Schubert et al., 2001; Sabet et al., 2005). No clear roles in invasion or virulence have been identified for other genes encoding LRR (e.g. *inlE* and *inlI*) (Bergmann et al., 2002; Sabet et al., 2005) and for a number genes encoding LRR, null mutants have not yet been investigated. For the purposes of this study, previously characterized genes encoding LRR will be referred to as internalins (to be consistent with earlier published reports), while other genes encoding LRR will be referred to as internalin-like genes. While most of these LRR genes also encode a cell wall binding motif (typically an LPXTG motif), some encode secreted internalins (e.g., InlC) (Engelbrecht et al., 1998). A number of DNA

hybridization, genome sequence analyses and other studies have shown that *L. monocytogenes* strains also encode species-specific (e.g., InlA, InlB) as well as strain-specific internalins (e.g., LMOf6854_0338) (Bierne et al., 2007).

The alternative stress sigma factor σ^B and the main virulence gene regulator PrfA have each been shown to contribute to regulation of *L. monocytogenes* internalin genes (Dramsi et al., 1993; Lingnau et al., 1995; Engelbrecht et al., 1998; Kim et al., 2005; Luo et al., 2005). While *inlA* and *inlB* are regulated by both σ^B and PrfA, some internalin genes are regulated only by PrfA (e.g., *inlC*) or only by σ^B (e.g., *inlC2*, *inlD*); transcription of other internalin genes (e.g. *inlF*, *inlG*) appears to be independent of these two regulators (McGann et al., 2007a). Transcription of internalin genes by σ^B, which is activated in response to stress conditions typically encountered in the gastrointestinal system (e.g., acid and osmotic stress), might represent a mechanism that allows *L. monocytogenes* to induce transcription of internalins that may be critical for the gastrointestinal stages of infection (e.g., *inlA*) (McGann et al., 2007a,b). In addition to regulation by σ^B and PrfA, transcription of internalin genes has been shown to also be temperature-dependent (Dramsi et al., 1993; McGann et al., 2007b). For example, transcript levels of the σ^B-dependent internalins *inlC2* and *inlD* were found to be highest at 16 and lowest at 37ºC, while transcript levels of *inlG* and *inlJ* were highest at 37 and lowest at 16ºC (McGann et al., 2007b), suggesting roles for these internalins in different environments.

Population genetics and phylogenetic studies have shown that *L. monocytogenes* isolates represent at least three genetically distinct lineages, including two common lineages (I and II) and one rare lineage (III), which appears to be predominantly associated with animal listeriosis (Piffaretti et al., 1989; McLauchlin, 1990; Boerlin et al., 1997; Wiedmann et al., 1997; Jeffers et al., 2001). *L. monocytogenes* lineage I strains, which primarily represent serotypes 4b and 1/2b, are overrepresented among isolates from sporadic and outbreak-associated human

listeriosis cases. Lineage II strains, which primarily represent serotypes 1/2a and 1/2c, are overrepresented among food isolates and are, on average, found in higher number in foods as compared to lineage I strains (Chen et al., 2006). While strains in lineage II do cause human listeriosis (McLauchlin, 1990; Hayes et al., 1999; Gray et al., 2004), lineage II strains are on average less virulent than lineage I strains as supported by (i) tissue culture assays, which showed lower plaque size, on average, for lineage II strains (Wiedmann et al., 1997; Gray et al., 2004) (ii) modelling data which support a higher infectious dose for lineage II strains (Chen et al., 2006), and (iii) epidemiological data, which show that serotype 1/2a and 1/2c strains are rarely responsible for human listeriosis outbreaks. To date the majority of laboratory studies have focused on characterization of serotype 1/2a and 1/2c lineage II strains (e.g., EGD (Murray et al., 1926), 10403S (Bishop and Hinrichs, 1987), and LO28 (Vincente et al., 1985). Clearly, studies characterizing virulence associated characteristics in lineage I strains are needed, including studies focusing on lineage-specific genes. To further explore the genetic basic of the apparent virulence differences as well as the evolution of the two main *L. monocytogenes* lineages, we investigated the hypotheses that (i) lineage-specific internalin-like genes contribute to the ability of *L. monocytogenes* to invade human cells, and (ii) temperature-dependent transcription patterns of lineage-specific internalin-like genes may differ between lineages.

Methods and Materials

Bacterial strains and mutant construction. The *L. monocytogenes* lineage I, serotype 4b strain F2365, a food isolate from the 1985 listeriosis outbreak in Los Angeles (Linnan et al., 1988), and the lineage II serotype 1/2a strain 10403S (Bishop and Hinrichs, 1987) were used as parent strains for mutant construction (Table 3.1).

Table 3.1 *L. monocytogenes* strains used in this study

Strain Designation (FSL no.)	Strain Background	Description[a]	Lineage	Reference
F2365 (R2-574)	-	Lineage I type strain	I	[55]
Δ*sigB* (O1-002)	F2365	708 bp in-frame deletion in *sig*B	I	Oliver, Unpublished Data
Δ*lsiIA* (M4-009)	F2365	1734 bp in-frame deletion in *F-lmo_0374*	I	This Study
Δ*lsiIB* (M4-010)	F2365	3228 bp in-frame deletion in *F-lmo_1254*	I	This Study
Δ*lsiIC* (M4-031)	F2365	1035 bp in-frame deletion in *F-lmo_2416*	I	This Study
10403S (X1-001)	-	Lineage II type strain	II	[37]
Δ*sigB* (A1-254)	10403S	297 bp in-frame deletion in *sig*B	II	[63]
Δ*lsiIIX* (M4-019)	10403S	2421 bp in-frame deletion in *lmo_0171*	II	This Study
Δ*lsiIIY* (M4-021)	10403S	1806 bp in-frame deletion in *lmo_0801*	II	This Study
Δ*lsiIIZ* (M4-020)	10403S	1881 bp in-frame deletion in *lmo_2026*	II	This Study

[a] Locus tags for *L. monocytogenes* EGD-e are given for lineage II-specific genes, while locus tags for F2365 are given for lineage I-specific genes.

Primers for mutant construction (Table A3.1) were designed using the completed genome sequences for F2365 (Nelson et al., 2004) and EGD-e, another lineage II strain (Glaser et al., 2001). Construction of deletion mutants was performed using SOE-PCR (Horton et al., 1990) to construct internal deletion alleles in a temperature-sensitive suicide plasmid (pKSV-7) (Smith and Youngman, 1992) and subsequent allelic exchange mutagenesis (as described by Camilli et al. [1990]) in the appropriate parent strain (i.e., 10403S for lineage II-specific internalin-like genes and F2365 for lineage I-specific internalin-like genes). Plasmid inserts were sequenced to ensure the proper, in-frame deletions for each gene of interest.

Growth conditions. For all experiments, single colonies of the appropriate strains were initially inoculated into BHI (Brain Heart Infusion) broth and incubated for 12-18 h at 37ºC with aeration (shaking at 210 rpm). This culture was diluted 1:100 into fresh BHI broth and grown to log phase (defined as OD_{600}=0.4), followed by another 1:100 dilution into fresh BHI broth, which was subsequently incubated under different defined conditions, including (i) growth to log phase (OD_{600}=0.4) in BHI at 16°C without aeration, (ii) growth to log phase (OD_{600}=0.4) at 37°C in BHI without aeration, and (iii) growth to early-stationary phase (OD_{600}=1.0 followed by an additional 3 h of incubation) at 37°C in BHI with aeration (shaking at 210 rpm). For invasion assays with F2365 and its isogenic mutants, bacteria were also grown to early-stationary phase (OD_{600}=1.0 + 3 h) at 30°C in BHI without aeration.

RNA isolation and quantitative RT-PCR (qRT-PCR). RNA isolation was performed essentially as described by Kazmierczak et al. (2003). Briefly, *L. monocytogenes* RNA was stabilized using RNA Protect (Qiagen) and total RNA was extracted with the RNeasy Midi kit (Qiagen). Cells were lysed by sonication on ice at 18 watts, performed three times for 30 s each with 30 s rest periods in between. RNA was eluted from the column with RNase-free water and treated with Turbo DNase (Ambion) following the manufacturers protocol. The RNA was then ethanol

precipitated, centrifuged, washed in 70% ethanol, and finally re-suspended in RNase-free water. RNA preparations that required further purification were phenol/chloroform extracted, then ethanol precipitated and re-suspended in RNase-free water.

Transcript levels for the six lineage-specific internalin-like genes (Table 3.2) and the housekeeping gene *rpoB* were determined using TaqMan qRT-PCR performed on the ABI Prism 7000 Sequence Detection System (Applied Biosystems) as described previously (Sue et al., 2004). Primer Express software was used to design all TaqMan primers and probes (Table A3.1); TaqMan probes were synthesized with 6-carboxyfluorescein (6-FAM) reporter dye at the 5' end and QSY7 dark-quencher at the 3' end. The TaqMan assays were performed as described previously [46] using the TaqMan One-Step RT-PCR Master Mix Reagents kit (Applied Biosystems). Briefly, duplicate TaqMan reactions for each gene were run in 96-well MicroAmp optical reaction plates using cycling conditions that consisted of 1 cycle at 48ºC for 30 min, 1 cycle at 95ºC for 3 min, followed by 40 cycles at 95ºC for 15 s and 60ºC for 1 min. Genomic DNA contamination in each RNA isolation was quantified using duplicate reactions lacking the reverse transcriptase enzyme (Multiscribe, Applied Biosystems). Genomic DNA standard curves for each of the seven target genes were included in each assay to calculate absolute cDNA levels as described by Sue et al. (2004). Phenol/chloroform-purified genomic DNA for the standard curve was prepared from 16-18 h cultures of *L. monocytogenes* strains 10403S and F2365 as described by Flamm et al. (1984). Three independent qRT-PCR replicates using three independent RNA isolations were performed.

Hidden markov model (HMM) and Protein family (Pfam) searches. HMM searches to determine putative σ^B-dependent promoters were performed using HMMER (available at http://selab.janelia.org/) as described by Kazmierczak et al. (2003) and Raengpradub et al. (2007). The training alignments for the σ^B models

Table 3.2 *L. monocytogenes* internalin-like genes of interest

Gene	Lineage Specificity	Homolog in *L. innocua* CLIP11262[c]	Possible Regulatory Site[d]	Reference
lsiIA	I[a]	yes, *lin_0372*	σ^B (24)	[40]
lsiIB	I	yes, *lin_1204*	σ^B(214)	[40]
lsiIC	I	yes, *lin_2537*	none	[40]
lsiIIX	II[b]	no	σ^B (884)	[39]
lsiIIY	II	no	none	[39]
lsiIIZ	II	no	none	[39]

[a] Lineage I-specific genes were identified based on the comparative genomics analyses reported by Nelson et al. (see Supplemental Tables 2 & 5) [40]. *lsiIA*, *lsiB*, and *lsiC* were present in lineage I strains F2365 and H7858 but not the lineage II strains EGD-e or F6854.

[b] Lineage II-specific genes were selected based on the genome hybridization analysis reported by Doumith et al. [39]. *lsiIIX* was present in all 39 lineage II strains tested and absent in the 47 lineage I and the five lineage III strains tested. *lsiIIZ* was present in all 12 lineage II serotype 1/2c and 3c strains and in more than two-thirds of the 27 lineage II serotype 1/2a and 3a strains tested; none of the lineage I or III strains tested contained *lsiIIZ*. *lsiIIY* was present in all 12 lineage II serotype 1/2c and 3c strains and in more than two-thirds of the 27 lineage II serotype 1/2a and 3a strains tested; *lsiIIY* was absent in more than two-thirds of the lineage I strains, and present in only 1/5 lineage III strains tested.

[c] BLAST searches were used to identify homologs, for a given gene, in the *L. innocua* CLIP11262 genome sequence.

[d] Numbers in parenthesis reflect the distances of the putative regulatory site from the start codon of the target gene.

included 33 *L. monocytogenes* σ^B-dependent promoter sequences confirmed by microarray experiments (Kazmierczak et al. 2003; Raengpradub et al. 2007). The models were searched against the complete *L. monocytogenes* strain EGD-e genome sequence (Glaser et al., 2001) (which was used as a model for the 10403S genome that has not yet been closed) and the complete *L. monocytogenes* strain F2365 genome sequence (Nelson et al., 2004). Any hits within ~1kb of the start codon in any of the six genes of interest were considered (Table 3.2). The Pfam 22.0 database (http://pfam.janelia.org/) (Finn et al., 2006) was used to search the translated coding region sequence for each gene of interest. SignalP-HMM 2.0 was used to search for probable N-terminal signal sequences (http://www.cbs.dtu.dk/services/SignalP/) (Nielsen et al., 1998). Protein motifs identified by Pfam and SignalP-HMM in the six lineage-specific internalin-like genes are shown in Figure 3.1.

Cell culture and invasion assays. The human colorectal epithelial cell line Caco2 (ATCC HTB-37) and the hepatocellular carcinoma cell line HepG2 (ATCC HB-8065) were cultivated at 37ºC under a 5% CO_2 atmosphere. Caco2 cells were grown in Dulbecco's Minimum Essential Medium (DMEM) with Earle's Salts and 1 % sodium pyruvate containing 20% FBS (Fetal Bovine Serum), 1% non-essential amino acids, 0.15% sodium bicarbonate, penicillin G (100 Units/ml) and streptomycin (100 μg/ml). HepG2 cells were cultured in the same medium, except for a final FBS concentration of 10%. For invasion assays, Caco2 and HepG2 cells were seeded, using media without antibiotics, into 24 well plates at a density of 5.0 x10^4 cells per well for Caco2 cells and 7.5x10^4 cells per well for HepG2 cells, at 48 h before infection. At the time of infection, tissue culture cells were 85-95% confluent. Invasion assays were performed as described by Garner et al. (2007). Briefly, approximately 2x10^7 CFU of *L. monocytogenes* were inoculated into a given well; the exact inoculum numbers for each experiment were quantified on BHI agar plates using a spiral plater (Spiral Biotech). At 30 min post-infection, host cells were washed three times with

sterile PBS and fresh pre-warmed media without antibiotics was added back to each well. At 45 min post-infection, fresh media containing 150 μg/ml gentamicin was added to each well. At 90 min post-infection, the host cells were washed three more times with sterile PBS and the number of internalized bacteria per well was determined by lysing infected cells in ice-cold, sterile water. Subsequent numeration of bacteria was performed on BHI agar plates. Invasion efficiency for all experiments was defined as log (CFU recovered/CFU infected).

In addition to standard invasion assays as described above, we also performed competitive invasion assays in which tissue culture cells were infected with a 1:1 mixture of the parent strain and an isogenic mutant strain (containing approx. $1x10^6$ CFU of each strain). Invasion assays were otherwise performed as described above. A total of 20 colonies recovered from each competitive invasion assay were tested with a PCR assay using primers external to the deletion (primers F-ex and R-ex; Table A3.1) to determine the parent/mutant ratio. Three independent trials of all invasion assays were performed.

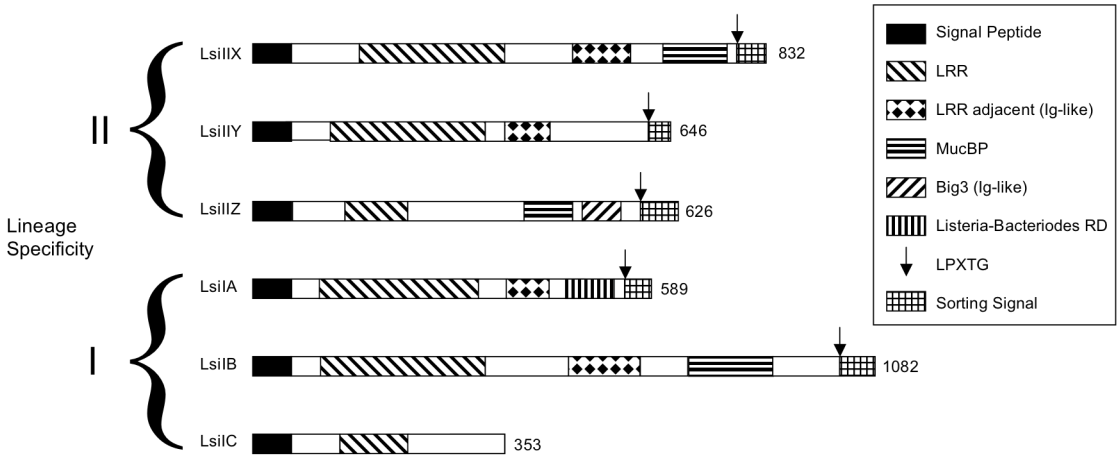
Statistical analysis. All statistical analyses were conducted using Minitab 14 (Minitab, Inc., State College, PA). General linear model ANOVA and post hoc Tukey method for multiple comparisons were used to analyse all TaqMan data. For the individual comparison of *lsiIA* transcript levels in the parent strain and *sigB* null mutant background a two-sided t-test was used. Invasion assay data were analysed with one-way ANOVA with post hoc Tukey multiple comparisons correction. Competitive invasion assay data were analysed using goodness-of-fit chi-square tests. P-values <0.05 were considered statistically significant.

Results

Construction and growth characterization of null mutants in selected lineage-specific internalin-like genes. A total of six genes encoding internalin-like proteins,

including three genes unique to lineage I and three genes unique to lineage II (Table 3.2), were selected for construction of null mutants to explore the role of lineage-specific internalin-like genes in *L. monocytogenes*. Lineage-specific internalin-like genes were selected based on (i) lineage specificity as previously determined by genome hybridization studies (Doumith et al., 2004) and full genome sequence analyses (Bierne et al., 2007; Nelson et al., 2004) and (ii) presence of sequences encoding leucine-rich repeat domains (LRR; Fig. 3.1). Lineage I-specific genes were identified based on the comparative genomics analysis performed by Nelson et al. (see their Supplemental Tables 2 & 5) (Nelson et al., 2004); all three lineage I-specific genes were present in the F2365 and H7858 lineage I genomes and absent in the genomes for the lineage II strains EGD-e and F6854 (Bierne et al., 2007; Nelson et al., 2004). BLAST analysis of the lineage I-specific genes indicated they are also present in the *L. innocua* CLIP11262 genome (Glaser et al., 2001). All three lineage II-specific genes were present in the sequenced genome of the lineage II type strain EGD-e and two of these genes (i.e., *lsiIIX* and *lsiIIY*) were also present in the genome sequence of the lineage II strain F6854. All three lineage II- specific genes were absent from the genomes of the lineage I strains F2365 and H7858 as well as the genome of *L. innocua* strain CLIP11262 (Bierne et al., 2007). The genome hybridization analyses reported by Doumith et al. (2004) showed that *lsiIIX* was present in all 39 lineage II strains tested and absent from the 47 lineage I and the five lineage III strains. *lsiIIZ* was present in all 12 lineage II serotype 1/2c and 3c strains and in more than two-thirds of the 27 serotype 1/2a and 3a strains tested; none of the lineage I or III strains tested contained *lsiIIZ. lsiIIY* was present in all 12 lineage II serotype 1/2c and 3c strains and in more than two-thirds of the 27 serotype 1/2a and 3a strains tested, while it was absent from more than two-thirds of the lineage I strains, and present only in 1/5 lineage III strains tested. While *lsiIIY* is, thus, not exclusively found in lineage II, it is more commonly found in lineage II strains as compared to

Figure 3.1 Gene sequences for lineage-specific internalin-like genes obtained from the genomes for *L. monocytogenes* EGD-e (Glaser et al., 2001) and F2365 (Nelson et al., 2004) were used to predict protein motifs using Pfam analysis. All lineage II (Bierne et al., 2007) and lineage I genes have a signal peptide at the N-terminus as well as leucine rich repeats (LRRs). Five of the six proteins have LPXTG motifs at the carboxy-terminus indicating that they can be covalently bound to the peptidoglycan layer on the bacterial cell surface. The numbers directly following the carboxy-terminus indicate the predicted amino acid length for each protein.


Lineage Specificity
II
I
LsilIX
LsilIY
LsilIZ
LsilA
LsilB
LsilC
832
646
626
589
1082
353
Signal Peptide
LRR
LRR adjacent (Ig-like)
MucBP
Big3 (Ig-like)
Listeria-Bacteriodes RD
LPXTG
Sorting Signal

lineage I strains and therefore, for the purposes of this study, is included with the lineage II-specific genes.

Five of six selected lineage-specific internalin-like genes were characterized by the presence of a sequence encoding the gram-positive cell surface anchor LPXTG (Fig. 3.1). Two lineage I- (*lsiIA* and *lsiIB*) and two lineage II- (*lsiIIX* and *lsiIIY*) specific genes also encode for a domain called LRR adjacent, a region found on the carboxy-terminal end of the LRR; this motif is thought to stabilize but not directly participate in protein binding (Freiberg et al., 2004). *lsiIIX*, *lsiIIZ*, and *lsiIB* also encode domains named for their similarity to mucin binding proteins (MucBP). *lsiIIZ* in addition encodes for a bacterial immunoglobulin-like fold (Fig. 3.1); bacterial proteins containing immunoglobulin-like folds found have been shown to play a role in host cell adherence in other pathogens, e.g. *Escherichia coli* and *Yersinia pseudotuberculosis* (Hamburger, et al., 1999; Luo et al., 2000). Pfam determined that *lsiIA* contains a *Listeria-Bacteroides* repeat domain (RD), which is found in many *Listeria* species (e.g., *L. ivanovii*, *L. monocytogenes* and *L. welshimeri*) as well as *Bacteroides* and *Geobacter* species.

Non-polar null mutants for the three lineage I-specific internalin-like genes (constructed in *L. monocytogenes* F2365) and for the three lineage II-specific internalin-like genes (constructed in *L. monocytogenes* 10403S) did not differ from their parent strains in growth to early-stationary phase (defined as OD_{600}=1.0 followed by an additional 3 h of incubation) in aerated BHI at 37°C (Table A3.2).

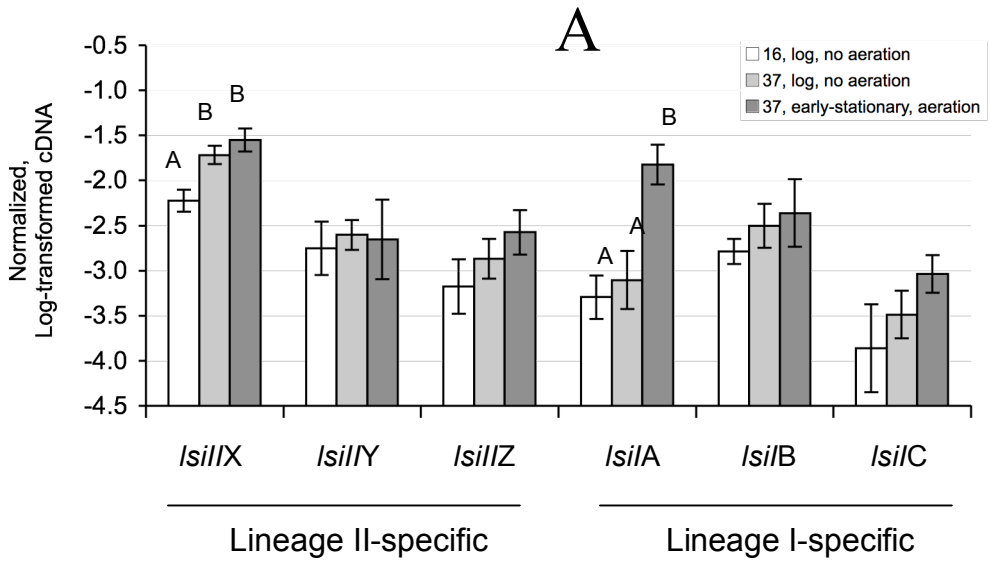
Transcript levels of lineage-specific internalin-like genes differ under different environmental conditions. In order to initially assess whether the selected lineage-specific internalin-like genes are likely to be functional and have the potential to play a role in invasion of mammalian cells, we evaluated transcript levels of these internalin-like genes under three different conditions, including (i) growth to log phase (OD_{600}=0.4) at 16°C in non-aerated BHI, (ii) growth to log phase (OD_{600}=0.4) at 37°C

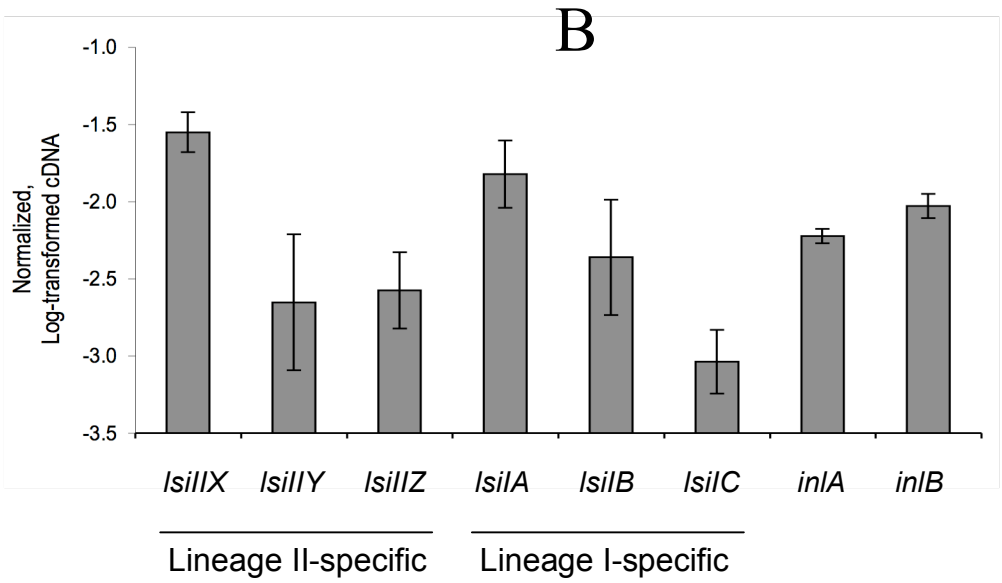
in non-aerated BHI, and (iii) growth to early-stationary phase (OD_{600}=1.0+3 h) at 37°C in aerated BHI. In addition to transcript levels for the six internalin-like genes, we also determined the transcript levels for the housekeeping gene *rpoB* for all strains grown under the conditions described above; *rpoB* levels were not significantly affected by either strain background or condition ($p>0.05$; ANOVA). Initial analyses on the raw non-normalized transcript levels showed detectable transcript levels for all six internalin-like genes ranging from 2.62 (*lsiIC*; bacteria grown to log phase at 16°C) to 5.20 (*lsiIIX*; bacteria grown to early-stationary phase at 37°C) log copy numbers/25 ng of total RNA. All subsequent analyses used internalin-like gene transcript levels normalized to the levels of *rpoB* (Fig. 3.2A).

Analysis of the numerical normalized internalin-like gene transcript levels showed two consistent trends across all genes, including (i) higher transcript levels in bacteria grown to log phase at 37 as compared to 16°C (possibly supporting a role for these genes in warm-blooded animals) and (ii) higher transcript levels in bacteria grown at 37ºC to early-stationary phase with aeration as compared to those grown to log phase without aeration at 37ºC (observed in all genes except for *lsiIIY*; Fig. 3.2A). Despite these numerical trends, a significant effect of growth condition on transcript levels was only observed for two genes (*lsiIIX* and *lsiIA*; ANOVA, $p<0.05$). *lsiIIX* transcript levels were significantly ($p<0.05$) higher in bacteria grown at 37ºC as compared to 16ºC, regardless of growth phase or aeration. *lsiIA* transcript levels were significantly ($p<0.05$) higher for bacteria grown at 37ºC to early-stationary phase with aeration as compared to bacteria grown to log-phase without aeration at either 16 or 37ºC.

When comparing transcript levels between genes, *lsiIIX* showed the highest transcript levels under all three conditions, while *lsiIC* showed the lowest transcript levels under all conditions (differences in transcript levels between these genes were

Figure 3.2 Panel A displays transcript levels for three lineage I- and three lineage II-specific internalin-like genes determined, using qRT-PCR, for *L. monocytogenes* grown to (i) log phase (OD_{600}=0.4) at 16°C in BHI without aeration, (ii) log phase (OD_{600}=0.4) at 37°C in BHI without aeration, and (iii) early-stationary phase (OD_{600}=1.0 followed by an additional 3 h of incubation) at 37°C in BHI with aeration. Transcript levels for each gene are reported as $\log_{10}$ (*gene of interest*/*rpoB*). Each bar represents the average of three independent replicates and error bars represent standard deviations. For genes with significant differences in transcript levels between conditions, letters signifying statistical groupings are shown (transcript levels with different letters are significantly different). Panel B depicts the transcript levels for the internalin-like genes investigated in this study as well as transcript levels for *inlA* and *inlB* (as previously reported by McGann et al. (2007b) for bacteria grown to early-stationary phase (OD_{600}=1.0 followed by an additional 3 h of incubation) at 37°C in BHI with aeration.


A
16, log, no aeration
37, log, no aeration
37, early-stationary, aeration
Normalized, Log-transformed cDNA
-0.5
-1.0
-1.5
-2.0
-2.5
-3.0
-3.5
-4.0
-4.5
A
B
B
A
A
B
lsiIIX
lsiIIY
lsiIIZ
lsiIA
lsiIB
lsiIC
Lineage II-specific
Lineage I-specific


B
Normalized, Log-transformed cDNA
-1.0
-1.5
-2.0
-2.5
-3.0
-3.5
lsiIIX
lsiIIY
lsiIIZ
lsiIA
lsiIB
lsiIC
inlA
inlB
Lineage II-specific
Lineage I-specific

significant for all conditions; $p<0.05$). Transcript levels for all six genes in bacteria grown to early-stationary phase with aeration were also compared to previously reported transcript levels, in bacteria grown under the same conditions, for the known virulence genes *inlA* and *inlB* (McGann et al., 2007b). *inlA* and *inlB* showed transcript levels of -2.22 and -2.03 log copy number/log copy number of *rpoB*, respectively (negative log copy numbers indicate that the copy number for these two genes were 2.22 and 2.03 log lower than the corresponding *rpoB* numbers); log copy numbers for the internalin-like genes evaluated here were considerably higher for two genes (i.e., *lsiIA* and *lsiIIX*), in the same range but slightly lower for three genes (i.e., *lsiIIZ, lsiIIY* and *lsiIB*), and considerably lower for one gene (i.e., *lsiIC*; Fig. 3.2B)

One lineage-specific internalin shows σ^B-dependent transcription. HMM analyses did not identify PrfA binding sites upstream of any of the six lineage-specific internalin-like genes, but identified putative σ^B-dependent promoters upstream of three of the lineage-specific internalin-like genes (*lsiIA*, *lsiIB*, and *lsiIIX*; Table 2). We thus used qRT-PCR to evaluate that transcript levels for these three genes in the appropriate parent strains as well as corresponding isogenic Δ*sigB* strains (i.e., *L. monocytogenes* 10403SΔ*sigB* for *lsiIIX* and F2365Δ*sigB* for *lsiIA* and *lsiIB*), grown to early-stationary phase with aeration (Fig. 3.3). While transcript levels for *lsiIIX* and *lsiIB* were not significantly affected by the *sigB* deletion (Fig. 3.3), *lsiIA* transcript levels were significantly ($p<0.05$) lower in the Δ*sigB* strain as compared to its isogenic parent strain (F2365), confirming σ^B-dependent transcription of this lineage-specific internalin-like gene. The σ^B independence of *lsiIIX* and *lsiIB* transcript levels observed here may indicate that transcription of these genes is either σ^B-independent or only σ^B-dependent under specific conditions.

Invasion efficiencies of lineage-specific internalin-like gene null mutants in Caco2 and HepG2 cells. The invasion efficiencies of the six strains carrying null mutations

in lineage-specific internalin-like genes and their parent strains were tested using a human intestinal epithelial-like cell line (Caco2) and a human hepatocellular carcinoma cell line (HepG2). Initial invasion assays were conducted with bacteria grown, with aeration, to early-stationary phase at 37°C, as this growth condition was determined by qRT-PCR (see above) to yield the highest transcript levels for all but one of the six internalin-like genes. None of the deletion mutants in lineage II-specific internalin-like genes showed reduced invasion efficiency, as compared to the parent strain 10403S, in either Caco2 or HepG2 cells (Fig. 3.4A). Similarly, none of the null mutants in lineage I-specific internalin-like genes differed significantly from their parent strain (F2365) in their invasion efficiency in either Caco2 or HepG2 cells ($p>0.05$; Fig. 3.4B). To more sensitively measure the relative invasion efficiency of the lineage-specific internalin-like gene null mutants as compared to their parent strains, we also performed competitive index invasion assays in Caco2 cells. In these assays, Caco2 cells were inoculated with a 1:1 mixture of a given null mutant and its isogenic parent strain and the ratios of intracellular mutant to intracellular parent strain recovered at 1.5 h after infection were determined by plating and subsequent allele-specific PCR to differentiate the parent and mutant strains. The PCR data were then used to calculate a parent to mutant ratio. Using this competitive index assay, none of the six null mutants showed significantly different recovery as compared to the parent strain ($p>0.05$, chi-square test; Fig. 3.5), further suggesting that these six internalin-like genes have limited or no importance for invasion of human intestinal epithelial cells.

Nightingale et al. (2007) previously reported that the lineage I parent strain F2365 showed considerably lower invasion of Caco2 cells when grown at 37°C to early-stationary with aeration as compared to growth at 30°C to early-stationary phase without aeration. Therefore, we also tested the invasion efficiency in Caco2 cells *of L. monocytogenes* F2365 and the three lineage I-specific internalin-like null mutants

Figure 3.3 qRT-PCR was used to determine *lsiIA* and *lisIB* transcript levels in *L. monocytogenes* F2365 and F2365 Δ*sigB* and *lsiIIX* transcript levels in strains 10403S and 10403S Δ*sigB* using RNA isolated from bacteria grown to early-stationary phase (OD_{600}=1.0 followed by an additional 3 h of incubation) at 37°C in BHI with aeration. Transcript levels for each gene are reported as $\log_{10}$ (*gene of interest*/ *rpoB*). Each bar represents the average of three independent replicates and error bars represent standard deviations; bars marked with an asterisk represent transcript levels with significant differences between the Δ*sigB* strain and the parent strain (t-test; $p<0.05$).

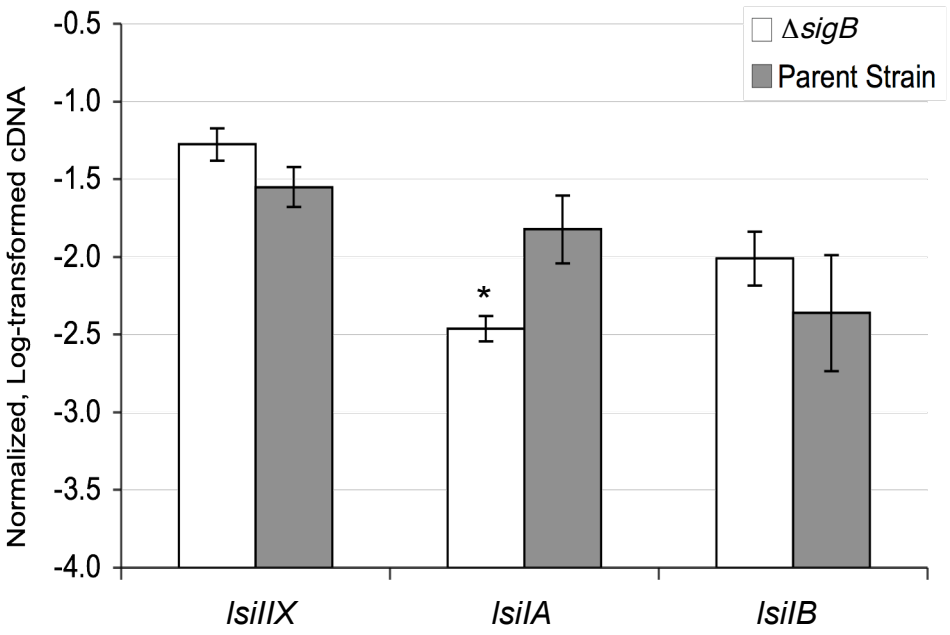
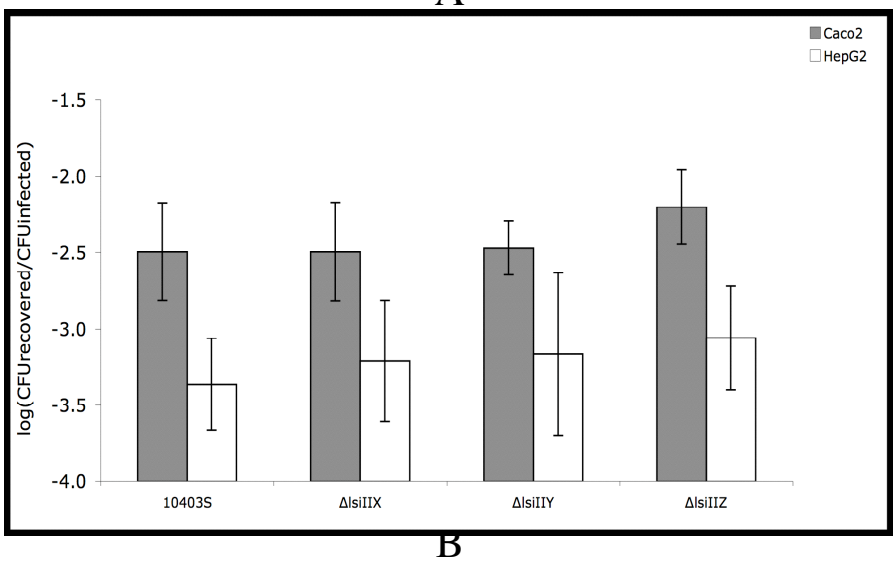

ΔsigB
Parent Strain
Normalized, Log-transformed cDNA
-0.5
-1.0
-1.5
-2.0
-2.5
-3.0
-3.5
-4.0
*
IsiIIX
IsiIA
IsiIB

Figure 3.4 Invasion efficiencies in Caco2 and HepG2 cells were determined for (A) the lineage II parent strain and the three null mutants in lineage II-specific, internalin-like genes and (B) the lineage I parent strain and the three null mutants in lineage I-specific, internalin-like genes. In preparation for invasion assays, all strains were grown to early-stationary phase (OD_{600}=1.0 followed by an additional 3 h of incubation) at 37ºC in BHI with aeration. Invasion efficiency was calculated as log (CFU recovered/CFU infected). Each bar represents the average of three independent replicates and error bars represent standard deviations. There was no significant effect of strain on invasion efficiency (ANOVA; $p>0.05$; separate ANOVA tests were performed for lineage I and II mutants within each of the two cell lines). While Δ*lsiIB* may appear to have a higher invasion efficiency in HepG2 cells as compared to the parent strain, the overall ANOVA for invasion efficiency in HepG2 cells of the lineage II strains was not significant ($p=0.090$).

A


Caco2
HepG2
log(CFUrecoverec/CFUinfected)
-1.5
-2.0
-2.5
-3.0
-3.5
-4.0
10403S
ΔlsiIIX
ΔlsiIIY
ΔlsiIIZ

B

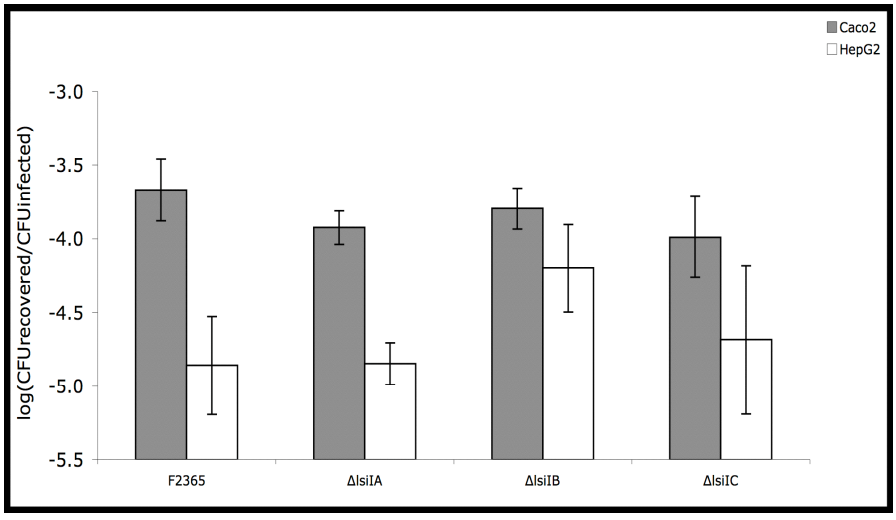
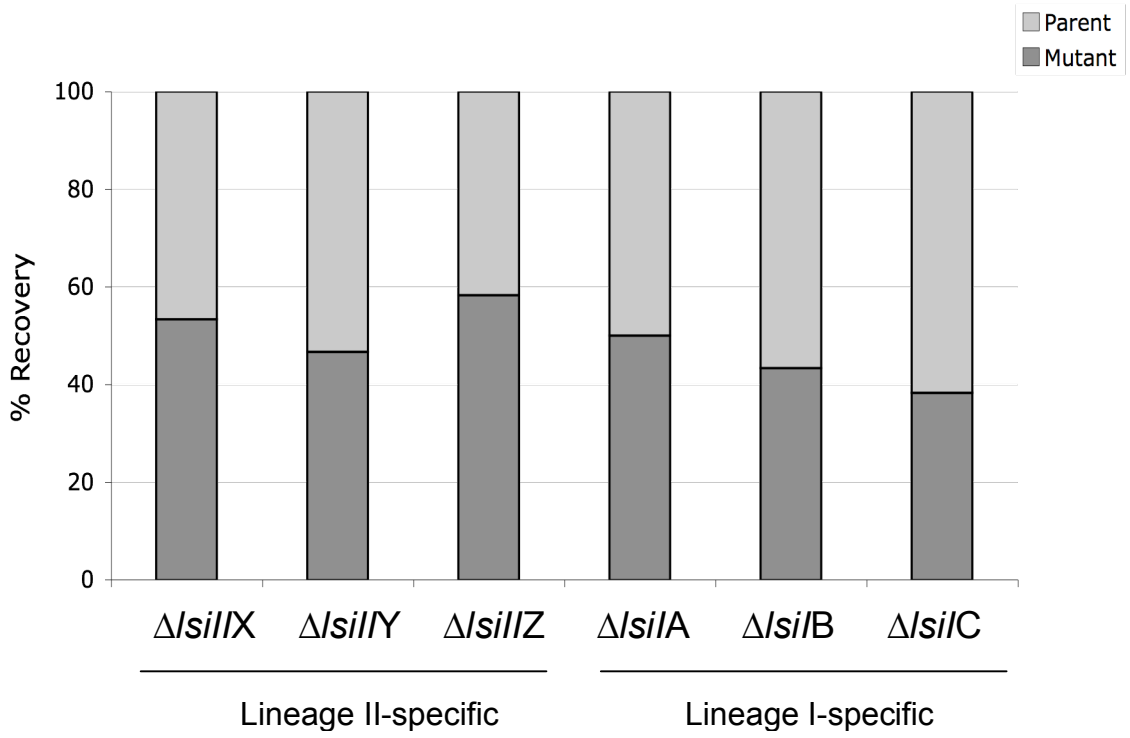

Caco2
HepG2
log(CFUrecovered/CFUinfected)
-3.0
-3.5
-4.0
-4.5
-5.0
-5.5
F2365
ΔlsiIA
ΔlsiIB
ΔlsiIC

Figure 3.5 Competitive invasion assays were conducted by infecting Caco2 cells with a 1:1 mixture of each null mutant and its respective parent strain. Recovered intracellular parent and null mutant bacteria were differentiated using allele-specific colony PCR (primers are listed in Table A3.1). Analyses using a goodness-of-fit chi-square tests showed no statistically significant differences in the recovery of any null mutants as compared to the parent strains; the difference between recovery of the parent strain F2365 and Δ*lsiIC* was borderline statistically significant (p=0.071). All assays were performed in triplicate and the numbers represent the relative recovery based on 60 colonies (20 colonies per assay).


Parent
Mutant
% Recovery
100
80
60
40
20
0
ΔlsilIX
ΔlsilIY
ΔlsilIZ
ΔlsilA
ΔlsilB
ΔlsilC
Lineage II-specific
Lineage I-specific

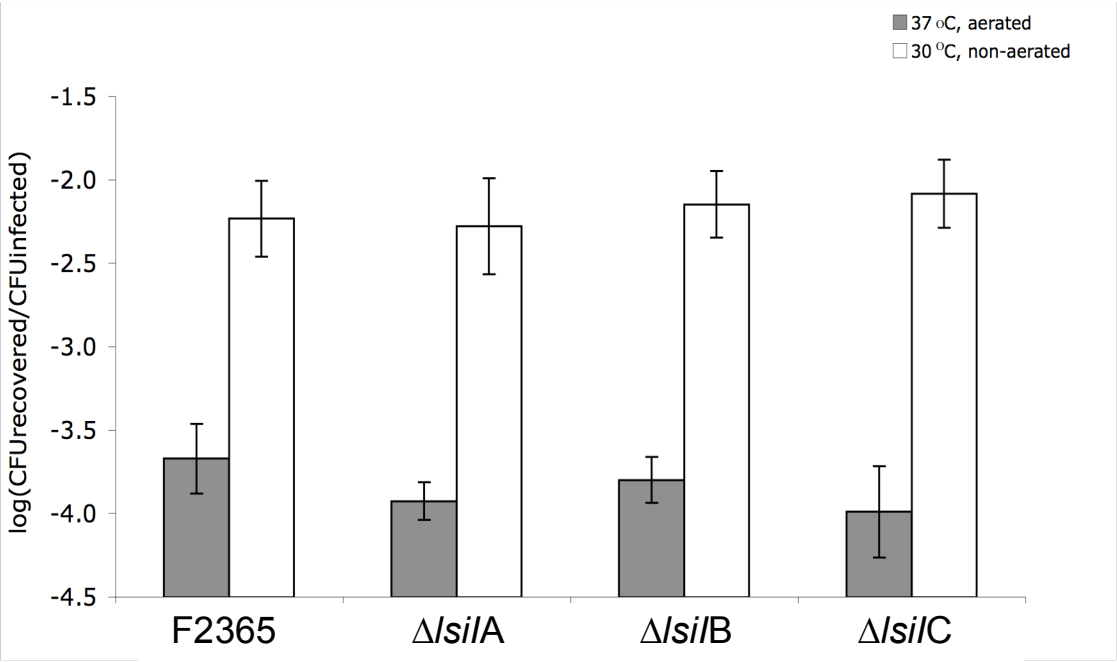
grown under these two conditions (Fig. 3.6). Even when grown to early-stationary phase (without aeration) at 30ºC, none of the three null mutants in lineage I-specific internalin-like genes differed in their invasion efficiencies from their parent strain. Our data confirmed that there is a significant effect of temperature and growth condition on the invasiveness for Caco2 cells for *L. monocytogenes* F2365 and derived null mutants ($p<0.001$, One-way ANOVA; Fig. 3.6). Strains grown at 30ºC to early-stationary phase without aeration showed 1.44-1.84 log higher invasion efficiency as compared to strains grown at 37ºC to early-stationary phase with aeration.

Discussion

While a number of genes common in all *L. monocytogenes* strains and lineages have clearly assigned roles in virulence, our understanding of the contributions of lineage-specific internalin-like genes to *L. monocytogenes* virulence are still limited. We investigated transcription of six lineage-specific internalin-like genes as well as their roles in invasion of selected human cells. Overall, our data show that (i) selected lineage-specific internalin-like genes show temperature- and σ^B-dependent transcription, and (ii) the selected lineage-specific internalin-like genes make limited or no contributions to invasion of human cell lines. Lineage-specific internalin-like genes may contribute to interactions between *L. monocytogenes* and select warm-blooded hosts other than humans, perhaps indicating unique host ranges for the different *L. monocytogenes* lineages.

Selected lineage-specific internalin-like genes show temperature- and σ^B-dependent transcription. Overall, our data showed that all lineage-specific internalin-like genes tested here had higher transcript levels at 37 as compared to 16°C, even through this difference was only significant for one lineage II-specific internalin-like gene (*lsiIIX*). Interestingly, McGann et al. (2007b) also observed

Figure 3.6 Invasion efficiencies are shown for *L. monocytogenes* F2365 and its isogenic null mutants in lineage I-specific internalin like genes; strains were grown to early-stationary phase (OD_{600}=1.0 followed by an additional 3 h of incubation) in BHI at 30°C (without aeration) or at 37ºC (with aeration); data shown for 37ºC are the same data shown in Fig. 3.4. Data are presented as log (CFU recovered/CFU infected). Each bar represents the average of three independent replicates and error bars represent standard deviations. Growth condition had a significant effect on invasion efficiency (ANOVA; $p<0.001$).


37 oC, aerated
30 oC, non-aerated
log(CFUrecovered/CFUinfected)
-1.5
-2.0
-2.5
-3.0
-3.5
-4.0
-4.5
F2365
ΔlsilA
ΔlsilB
ΔlsilC

temperature-dependent regulation of *inlA* and *inlB*, with both genes showing significantly higher transcript levels in bacteria grown at 37 as compared to 16°C. These temperature-dependent transcription patterns are consistent with the fact that *inlA* and *inlB* encode two internalins with known roles in invasion of human intestinal epithelial (Galliard et al., 1991) and hepatic (Dramsi et al., 1995) cell lines. We also found that all but one (*lsiIC*) of the lineage-specific internalin-like genes studied here showed higher (i.e., *lsiIIX* and *lsiIA*) or comparable but slightly lower (*lsiIIY*, *lsiIIZ,* and *lsiIB*) transcript levels at 37°C as compared to *inlA* and *inlB*, providing further support for the hypothesis that at least most of these lineage specific internalin-like genes contribute to interaction with warm-blooded hosts.

The observed σ^B-dependent transcription of *lsiA* is consistent with observations that a number of other internalin genes (e.g., *inlA, inlB, inlC2,* and *inlD*) also show σ^B-dependent transcription; σ^B-dependent transcriptional regulation of internalin-like genes may assure expression of these genes in a specific, possibly host-related, environment. The observation that a number of genes encoding proteins with roles in the intestinal stage of *L. monocytogenes* infection (e.g., *inlA*, *bsh*, *opuCA*) (Kazmierczak et al., 2003; Sue et al., 2004) are regulated by σ^B, suggests the specific hypothesis that *lsiA* encodes a protein that may be important for intestinal stage of infection of a warm-blooded non-human host. Interestingly, the *lsiA* homologue in *L. innocua* (*lin0372*), which is considered a non-pathogenic *Listeria* spp., also showed σ^B-dependent transcription (Raengpradub et al., 2007), further supporting that the product of this gene is not critical for pathogenic interactions with a human host, as *L. innocua* has not been identified as a human pathogen. Another gene that encodes a cell-wall associated protein (with an LPXTG motif and a collagen binding domain) in both *L. monocytogenes* and *L. innocua* (*lmo0880* and *lin0879*) also showed σ^B-dependent transcription in both species (Raengpradub et al., 2007). One could thus

hypothesize that surface proteins encoded by σ^B-dependent genes may generally be important for interactions, by different *Listeria* spp., with warm-blooded hosts, including possibly commensal relationships (Lammerding et al., 1992).

The selected lineage-specific internalin-like genes make limited or no contributions to invasion of human cell lines. Overall, we did not find evidence that any of the six selected lineage-specific internalin-like genes contributed to invasion of either human intestinal epithelial cells (i.e., Caco2 cells) or human hepatocellular carcinoma cells (i.e., HepG2 cells). While an EGD transposon mutant in *lsiIIZ* (previously designated ORF626 [Autret et al., 2001]), showed reduced presence in the brain of infected mice, this transposon mutant also showed reduced invasion of HepG2 cells (Autret et al., 2001), while our non-polar mutant showed no reduced invasion in HepG2 cells. This suggests a polar effect of the transposon or possibly strain differences in the contributions of *lisIIZ* to invasion (as our strain was generated in a 10403S background while the transposon mutant was generated in a EGD background). Caco2 cells have previously been used to show that *L. monocytogenes* InlA is required for invasion of human intestinal epithelial cells (Galliard et al., 1991). Similarly, HepG2 cells have previously been used to show that InlB contributes to invasion of liver cells (Dramsi et al., 1995). Therefore, these cell lines are appropriate models for evaluating the contributions of *L. monocytogenes* surface molecules to invasion of human cells. While our initial hypothesis was that lineage-specific internalin-like genes may contribute to differences in virulence among *L. monocytogenes* lineages (and might have helped to explain why lineage I strains are overrepresented among human listeriosis cases relative to lineage II strains) (Gray et al., 2004), our tissue culture and transcriptional data along with the observed presence of selected internalin-like genes in *L. innocua*, suggest that lineage-specific internalin-like genes may be important for interactions, possibly of a non-pathogenic nature, with non-human warm-blooded animals. This conclusion is consistent with the broad

environmental distribution of *Listeria* spp. and the extensive host range of *L. monocytogenes*. Alternatively, some internalin-like genes, particularly those found in *L. monocytogenes* but absent from the non-pathogenic *L. innocua*, may contribute to invasion of specific tissue and cell types not evaluated in our experiments.

While some previous studies using null mutants in internalin genes also did not find a reduced invasion phenotype in initial tissue culture assays, some of these studies found evidence for reduced virulence of their mutants in animal models. For example, a *L. monocytogenes* EGD-e *inlGHE* null mutant showed reduced ability to cause systemic disease in mice, even though it did not show reduced invasion of a number of tissue culture cell lines, including Caco2, TIB-37, HBMEC, and J774 cells (Raffelsbauer et al., 1998; Bergmann et al., 2002). Sabet et al. (2005) also found no evidence that either InlI nor InlJ were important for *in vitro* host cell invasion (using nine different cell lines including Caco2, HepG2, the human embryonic kidney cell line HEK293, and the mouse macrophage-like cell line J774), even though their *inlJ* null mutant exhibited reduced virulence in a transgenic mouse expressing human E-cadherin. Due to the species-specific interactions between InlA and its receptor, E-cadherin, *inlA* null mutants show no evidence for reduced virulence in a non-transgenic mouse models (Lecuit et al., 2004), despite showing reduced invasion in human intestinal epithelial cells (Galliard et al., 1991) and reduced virulence in oral infections of guinea pig and transgenic mice expressing human E-cadherin (Lecuit et al., 2001). Use of animal models to evaluate the null mutants constructed here would thus have likely been of limited value considering (i) the apparently broad host range of *L. monocytogenes* and the species-specific importance of some of its virulence genes as well as (ii) the observation that a number of the lineage-specific internalin-like genes studied here are also present in *L. innocua*, which does not invade human intestinal cells (Lecuit et al., 1999), does not cause disease in typically used animal models (e.g., guinea pigs (Lecuit et al., 2001), mice (Hof et al., 1988; Nishibori et al.,

1995), and has not been associated with disease in humans or other animals. Rather, further studies on the ecology and distribution of the different *Listeria* species and different *L. monocytogenes* lineages will hopefully identify animal species that may be naturally be infected by or carry *Listeria*, thus providing a lead on appropriate animal models to select for future studies on the mutants described here.

Conclusions

Our study identified six lineage-specific *L. monocytogenes* internalin-like genes that are actively transcribed. Despite their structural similarities to internalins with known roles in host cell attachment and invasion, these lineage-specific internalin-like genes do not appear to contribute to invasion of human intestinal cells or hepatocytes. Our findings, combined with the observation that non-pathogenic *Listeria* species (i.e., *L. innocua*, *L. seeligeri* (Schmid et al., 2005) also carry internalin genes, suggest that the functions of *Listeria* internalins are not limited to attachment and invasion of human cells.

REFERENCES

Anonymous: Microbiological Risk Assessment Series. Technical Report, WHO, Rome 2004, No. 5.

Autret N, Dubail I, Trieu-Cuot P, Berche P, Charbit A: Identification of new genes involved in the virulence of *Listeria monocytogenes* by signature-tagged transposon mutagenesis. *Infect Immun* 2001, 69:2054-2065.

Bergmann B, Raffelsbauer D, Kuhn M, Goetz M, Hom S, Goebel W: InlA- but not InlB mediated internalization of *Listeria monocytogenes* by non-phagocytic mammalian cells needs the support of other internalins. *Mol Microbiol* 2002, 43:557-570.

Bierne H, Sabet C, Personnic N, Cossart P: Internalins: a complex family of leucine-rich repeat-containing proteins in *Listeria monocytogenes*. *Microbes Infect* 2007, 9:1156-1166.

Bishop DK, Hinrichs DJ: Adoptive transfer of immunity to *Listeria monocytogenes*. The influence of in vitro stimulation on lymphocyte subset requirements. *J Immunol* 1987, 139:2005-2009.

Boerlin P, Boerlin-Petzold F, Bannerman E, Bille J, Jemmi T: Typing *Listeria monocytogenes* Isolates from Fish Products and Human Listeriosis Cases. *Appl Environ Microbiol* 1997, 63:1338-1343.

Camilli A, Portnoy A, Youngman P: Insertional mutagenesis of *Listeria monocytogenes* with a novel Tn917 derivative that allows direct cloning of DNA flanking transposon insertions. *J Bacteriol* 1990, 172:3738-3744.

Chen Y, Ross WH, Gray MJ, Wiedmann M, Whiting RC, Scott VN: Attributing Risk to *Listeria monocytogenes* Subgroups: Dose Response in Relation to Genetic Lineages. *J Food Prot* 2006, 69.

Dramsi S, Kocks C, Forestier C, Cossart P: Internalin-mediated invasion of epithelial cells by *Listeria monocytogenes* is regulated by the bacterial growth state, temperature and the pleiotropic activator prfA. *Mol Microbiol* 1993, 9:931-94.

Dramsi S, Biswas I, Maguin E, Braun L, Mastroeni P, Cossart P: Entry of *Listeria monocytogenes* into hepatocytes requires expression of inIB, a surface protein of the internalin multigene family. *Mol Microbiol* 1995, 16:251-261.

Doumith M, Cazalet C, Simoes N, Frangeul L, Jacquet C, Kunst F, Martin P, Cossart P, Glaser P, Buchrieser C: New aspects regarding evolution and virulence of *Listeria monocytogenes* revealed by comparative genomic and DNA arrays. *Infect Immun* 2004, 72:1072-1083.

Engelbrecht F, Chun S, Ochs C, Hess J, Lottspeich F, Goebel W, Sokolovic Z: A new PrfA regulated gene of *Listeria monocytogenes* encoding a small, secreted protein which belongs to the family of internalins. *Mol Microbiol* 1996, 21:823-837.

Finn RD, Mistry J, Schuster-Bockler B, Griffiths-Jones S, Hollich V, Lassmann T, Moxon S, Marshall M, Khanna A, Durbin R, et al: Pfam: clans, web tools and services. *Nucl Acids Res* 2006, 34:D247-251.

Flamm R, Hinrichs D, Thomashow M: Introduction of pAM beta 1 into *Listeria monocytogenes* by conjugation and homology between native *L. monocytogenes* plasmids. *Infect Immun* 1984, 44:157-161.

Freiberg A, Machner MP, Pfeil W, Schubert WD, Heinz DW, Seckler R: Folding and stability of the leucine-rich repeat domain of internalin B from *Listeria monocytogenes*. *J Mol Biol* 2004, 337:453-461.

Gaillard J, Berche P, Frehel C, Gouin E, Cossart P: Entry of *L. monocytogenes* into cells is mediated by internalin, a repeat protein reminiscent of surface antigens from gram-positive cocci. *Cell* 1991, 65:1127-1141.

Garner M, Njaa B, Wiedmann M, Boor K: Sigma B contributes to *Listeria monocytogenes* gastrointestinal infection but not to systemic spread in the guinea pig infection model. *Infect Immun* 2006, 74:876-886.

Glaser P, Frangeul L, Buchrieser C, Rusniok C, Amend A, Baquero F, Berche P, Bloecker H, Brandt P, Chakraborty T, et al: Comparative Genomics of *Listeria* Species. *Science* 2001, 294:849-852.

Gray M, Zadoks R, Fortes E, Dogan B, Cai S, Chen Y, Scott V, Gombas D, Boor K, Wiedmann M: *Listeria monocytogenes* isolates from foods and humans form distinct but overlapping populations. *Appl Environ Microbiol* 2004, 70:5833-5841.

Greiffenberg L, Goebel W, Kim K, Weiglein I, Bubert A, Engelbrecht F, Stins M, Kuhn M: Interaction of *Listeria monocytogenes* with human brain microvascular endothelial cells:

InlB-dependent invasion, long-term intracellular growth, and spread from macrophages to endothelial cells. *Infect Immun* 1998, 66:5260-5267.

Hamburger ZA, Brown MS, Isberg RR, Bjorkman PJ: Crystal Structure of Invasin: A Bacterial Integrin-Binding Protein. *Science* 1999, 286:291-295.

Hayes PS, Graves LM, Ajello GW, Swaminathan B, Weaver RE, Wenger JD, Schuchat A, Broome CV: Comparison of Cold Enrichment and U.S. Department of Agriculture Methods for Isolating *Listeria monocytogenes* from Naturally Contaminated Foods. *Appl Environ Microbiol* 1991, 8:2109-2113.

Hof H, Hefner P: Pathogenicity of *Listeria monocytogenes* in comparison to other *Listeria* species. *Infection* 1988, 16:S141-144.

Horton R, Cai Z, Ho S, Pease L: Gene splicing by overlap extension: tailor-made genes using the polymerase chain reaction. *Biotechniques* 1990, 8:528-535.

Jeffers GT, Bruce JL, McDonough PL, Scarlett J, Boor KJ, Wiedmann M: Comparative genetic characterization of *Listeria monocytogenes* isolates from human and animal listeriosis cases. *Microbiol* 2001, 147:1095–1104.

Jonquieres R, Bierne H, Fiedler F, Gounon P, Cossart P: Interaction between the protein InlB of *Listeria monocytogenes* and lipoteichoic acid: a novel mechanism of protein association at the surface of Gram-positive bacteria. *Mol Microbiol* 1999, 34:902-914.

Kazmierczak MJ, Mithoe SC, Boor KJ, Wiedmann M: *Listeria monocytogenes* sigma B regulates stress response and virulence functions. *J Bacteriol* 2003, 185:5722-5734.

Kim H, Boor K, Marquis H: *Listeria monocytogenes* sigmaB contributes to invasion of human intestinal epithelial cells. *Infect Immun* 2004, 72:7374-7378.

Kim H, Marquis H, Boor KJ: SigmaB contributes to *Listeria monocytogenes* invasion by controlling expression of inlA and inlB. *Microbiol* 2005, 151:3215-3222.

Kobe B, Kajava AV: The leucine-rich repeat as a protein recognition motif. *Curr Opin Struct Biol* 2001, 11:725-732.

Lammerding AM, Glass KA, Gendron-Fitzpatrick A, Doyle MP: Determination of virulence of different strains of *Listeria monocytogenes* and *Listeria innocua* by oral inoculation of pregnant mice. *Appl Environ Microbiol* 1992, 58:3991-4000.

Lecuit M, Dramsi S, Gottardi C, Fedor-Chaiken M, Gumbiner B, Cossart P: A single amino acid in E-cadherin responsible for host specificity towards the human pathogen *Listeria monocytogenes*. *EMBO* 1999, 18:3956-3963.

Lecuit M, Vandormael-Pournin S, Lefort J, Huerre M, Gounon P, Dupuy C, Babinet C, Cossart P: A transgenic model for listeriosis: role of internalin in crossing the intestinal barrier. *Science* 2001, 292:1722-1725.

Lecuit M, Nelson D, Smith S, Khun H, Huerre M, Vacher-Lavenu M, Gordon J, Cossart P: Targeting and crossing of the human maternofetal barrier by *Listeria monocytogenes*: role of internalin interaction with trophoblast E-cadherin. *PNAS* 2004, 101:6152-6157.

Lingnau A, Domann E, Hudel M, Bock M, Nichterlein T, Wehland J, Chakraborty T: Expression of the *Listeria monocytogenes* EGD inlA and inlB genes, whose products mediate bacterial entry into tissue culture cell lines, by PrfA-dependent and -independent mechanisms. *Infect Immun* 1995, 63:3896-3903.

Linnan MJ, Mascola L, Lou XD, Goulet V, May S, Salminen C, Hird DW, Yonekura ML, Hayes P, Weaver R: Epidemic listeriosis associated with Mexican-style cheese. *N Engl J Med* 1988, 319:823-828.

Luo Y, Frey EA, Pfuetzner RA, Creagh AL, Knoechel DG, Haynes CA, Finlay BB, Strynadka NCJ: Crystal structure of enteropathogenic *Escherichia coli* intimin-receptor complex. *Nature* 2000, 405:1073-1077.

Luo Q, Rauch M, Marr AK, Muller-Altrock S, Goebel W: *In vitro* transcription of the *Listeria monocytogenes* virulence genes inlC and mpl reveals overlapping PrfA-dependent and independent promoters that are differentially activated by GTP. *Mol Microbiol* 2004, 52:39-52.

Mc Gann P, Wiedmann M, Boor KJ: The alternative sigma factor {sigma}B and the virulence gene regulator PrfA both regulate transcription of *Listeria monocytogenes* internalins. *Appl Environ Microbiol* 2007a, 73:2664-2606.

Mc Gann P, Ivanek R, Wiedmann M, Boor KJ: Temperature-dependent expression of *Listeria monocytogenes* internalin and internalin-like genes suggests functional diversity for these proteins among the listeriae. *Appl Environ Microbiol* 2007b, 73:2923-2906.

McLauchlin J: Distribution of serovars of *Listeria monocytogenes* isolated from different

categories of patients with listeriosis. *Eur J Clin Microbiol Infect Dis* 1990, 9:210-213.

Mead PS, Slutsker L, Dietz V, McCaig LF, Bresee JS, Shapiro C, Griffin PM, Tauxe RV: Food-Related Illness and Death in the United States. *Emerg Infect Dis* 1999, 5:607-625.

Murray EGD, Webb RE, Swann MBR: A disease of rabbits characterized by a large mononuclear leucocytosis, caused by a hitherto undescribed bacillus *Bacterium monocytogenes*. (n. sp.). *J Pathol Bacteriol* 1926, 29:407-439.

Nelson K, Fouts D, Mongodin E, Ravel J, DeBoy R, Kolonay J, Rasko D, Angiuoli S, Gill S, Paulsen I, et al: Whole genome comparisons of serotype 4b and 1/2a strains of the food-borne pathogen *Listeria monocytogenes* reveal new insights into the core genome components of this species. *Nucleic Acids Res* 2004, 32:2386-2395.

Nielsen H, Krogh A: Prediction of signal peptides and signal anchors by a hidden Markov model. *Proc Int Conf Intell Syst Mol Biol* 1998, 6:122-130.

Nightingale KK, Milillo SR, Ivy RA, Ho AJ, Oliver HF, Wiedmann M: *Listeria monocytogenes* F2365 carries several authentic mutations potentially leading to truncated gene products, including inlB, and demonstrates atypical phenotypic characteristics. *J Food Prot* 2007 70:482-488.

Nishibori T, Cooray K, Xiong H, Kawamura I, Fujita M, Mitsuyama M: Correlation between the presence of virulence-associated genes as determined by PCR and actual virulence to mice in various strains of *Listeria* spp. *Microbiol Immunol* 1995, 39:343-349.

Pell AN: Manure and Microbes: Public and Animal Health Problem? *J Dairy Sci* 1997, 80:2673-2681.

Piffaretti J-C, Kressebuch H, Aeschbacher M, Bille J, Bannerman E, Musser JM, Selander RK, Rocourt J: Genetic Characterization of Clones of the Bacterium *Listeria monocytogenes* Causing Epidemic Disease. *PNAS* 1989, 86:3818-3822.

Raengpradub S, Wiedmann M, Boor KJ: Comparative Analysis of the {sigma}B-Dependent Stress Response in *Listeria monocytogenes* and *Listeria innocua* Exposed to Selected Stress Conditions. *Appl Environ Microbiol* 2007, 74:951-907.

Raffelsbauer D, Bubert A, Engelbrecht F, Scheinpflug J, Simm A, Hess J, Kaufmann S: The

gene cluster inlC2DE of *Listeria monocytogenes* contains additional new internalin genes and is important for virulence in mice. *Mol Gen Genet* 1998, 260:144-158.

Rasmussen OF, Skouboe P, Dons L, Rossen L, Olsen JE: *Listeria monocytogenes* exists in at least three evolutionary lines: evidence from flagellin, invasive associated protein and listeriolysin O genes. *Microbiol* 1995, 141:2053-2061.

Sabet C, Lecuit M, Cabanes D, Cossart P, Bierne H: LPXTG Protein InlJ, a Newly Identified Internalin Involved in *Listeria monocytogenes* Virulence. *Infect Immun* 2005, 73:6912-692217.

Schmid MW, Ng EY, Lampidis R, Emmerth M, Walcher M, Kreft J, Goebel W, Wagner M, Schleifer KH: Evolutionary history of the genus *Listeria* and its virulence genes. *Syst Appl Microbiol* 2005, 28:1-18.

Schubert W-D, Gobel G, Diepholz M, Darji A, Kloer D, Hain T, Chakraborty T, Wehland J, Domann E, Heinz DW: Internalins from the human pathogen *Listeria monocytogenes* combine three distinct folds into a contiguous internalin domain. *J Molecular Biology* 2001, 312:783-794.

Seveau S, Pizarro-Cerda J, Cossart P: Molecular mechanisms exploited by *Listeria monocytogenes* during host cell invasion. *Microbes Infect* 2007, 9:1167-1175.

Smith K, Youngman P: Use of a new integrational vector to investigate compartment-specific expression of the *Bacillus subtilis* spoIIM gene. *Biochemie* 1992, 74:705-711.

Sue D, Fink D, Wiedmann M, Boor KJ: sigmaB-dependent gene induction and expression in *Listeria monocytogenes* during osmotic and acid stress conditions simulating the intestinal environment. *Microbiol* 2004, 150:3843-3855.

Swaminathan B, Gerner-Smidt P: The epidemiology of human listeriosis. *Microbes Infect* 2007, 9:1236-1243.

Tilney LG, Portnoy DA: Actin filaments and the growth, movement, and spread of the intracellular bacterial parasite, *Listeria monocytogenes*. *J Cell Biol* 1989, 109:1597-1608.

Vincente MF, Baquero F, Perez-Diaz JC: Cloning and expression of the *Listeria monocytogenes* haemolysin in E. coli. *FEMS Microbiol Lett* 1985, 30:77-79.

Wiedmann M, Bruce JL, Keating C, Johnson AE, McDonough PL, Batt CA: Ribotypes and

Virulence Gene Polymorphisms Suggest Three Distinct *Listeria monocytogenes* Lineages with Differences in Pathogenic Potential. *Infect Immun* 1997, 65:2707-2716.

Wiedmann M, Arvik T, Hurley R, Boor K: General stress transcription factor sigmaB and its role in acid tolerance and virulence of *Listeria monocytogenes*. *J Bacteriol* 1998, 180:3650-3656.

APPENDIX

Table A3.1 Primers and Probes used for mutant construction, TaqMan analysis, and allele-specific PCR.

Gene Target	Primer/Probe Name[a]	Sequence (5' – 3')
*lsiII*X	SRM14 SOE-A	[b]GGGGTACCCTACCAAGTAAGGCTGGG
	SRM15 SOE-B	[c]*GAGAGGGTTAATCCAGCTAAT* AGCAAGAGAAACAACAATGG
	SRM16 SOE-C	ATTAGCTGGATTAACCCTCTC
	SRM17 SOE-D	CGGGATCCCACACTCTTATGAATATTATCAGTG
	SRM41 Fex	CAGATGTAGATCCAGGG
	SRM42 Rex	ATAGCGATCAAGGTTCGG
	SRM63 TaqMan F	CAACTCCGGGAGATTATACAGTCA
	SRM64 TaqMan R	CCGGATCTGCCTTTAAGTCAAA
	TaqMan probe	CTACATGCTAAAAATG
*lsiII*Y	SRM20 SOE-A	GGGGTACCTCAATCAAACATTGGTCCAG
	SRM21 SOE-B	*TCGAAGATAAGAAATACCCGCC* AAAAATAATATTACGGTGCA
	SRM22 SOE-C	GCGGGTATTTCTTATCTTCGA
	SRM23 SOE-D	CGGGATCCTTTCTAATTTCGCCCCAATAC
	SRM43 Fex	TAGAAGAGAAAACAGCGG
	SRM44 Rex	TCAGTGCATTTTCTGGGG
	SRM65 TaqMan F	TGGCCTTGTTCCGAATGC
	SRM66 TaqMan R	GCTGGTGGTGCTACTTGTAAATTATTA
	TaqMan probe	CTGAATCACCATCAGCTG
*lsiII*Z	SRM26 SOE-A	GGGGTACCGGTTGCGTATATGATTGGGT
	SRM27 SOE-B	*GGTTGCCCATCTTTTTTCTCA*TTCAATCCATGAAGTGGAGGC
	SRM28 SOE-C	TGAGAAAAAAGATGGGCAACC
	SRM29 SOE-D	CGGGATCCTTTACCACATTTACTTTGGACAC
	SRM45 Fex	AATTCAATGATGCATACCC
	SRM46 Rex	TGTGCCCTTGTATGACC
	SRM67 TaqMan F	TCCAAAAATAGGCAAGACATTACTTTAG
	SRM68 TaqMan R	GATTGAATCTGCACCAAATACATGTA
	TaqMan probe	CTGATTTGGAAACCATCC
*lsiI*A	SOE-A	AACTGCAGATGCTGGTCTTAGTTTCCG
	SOE-B	AAAAACAGTCTCCCTTTGGA
	SOE-C	*TCCAAAGGGAGACTGTTT*TTAAGAAAATAAAGGAAAAAGCAG
	SOE-D	GCTCTAGAATGCTTTCCAAAGGTCGCG
	F	AACGGCCGCTGGTGATAAGAAAATGGGAA
	R	CGGGATCCAGCAGCTGCGTTATGCAGC
	SRM01 Fex	GATTTACCAAATGCGGAGG
	SRM02 Rex	TGGTTTAACAGAAGTTCGG

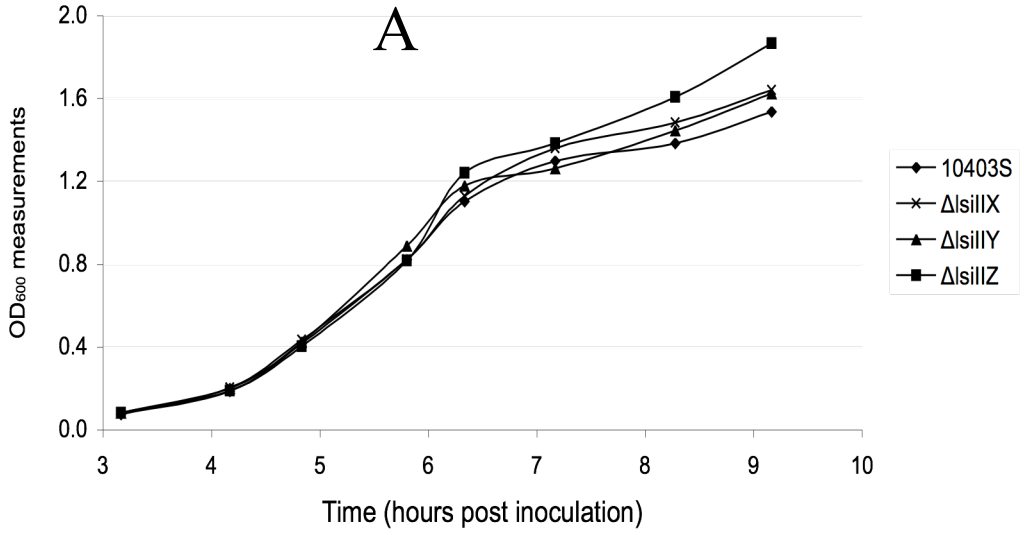
	SRM69 TaqMan F	AGATATTAACGACGCGCAAGTTACT
	SRM70 TaqMan R	GCCTAAACCCTTTAGATTTGTTAAACC
	TaqMan probe	TCACGCCACTGAGC
*lsiI*B	SOE-A	AACTGCAGGATTATTGAGATGATTGATCC
	SOE-B	TCTTTTCATGATTAGTCTCCT
	SOE-C	*AGGAGACTAATCATGAAAAGA*AAAGATAAATCTCATAACTGC
	SOE-D	GCTCTAGAAGTCAATACATTTAAATCAGGC
	SRM03 Fex	GAATCTCATTGAAAACGAAG
	SRM04 Rex	CCATTCTCGGTGTTTCCC
	SRM71 TaqMan F	TGATGATTTCCCACAAATGACTACTT
	SRM72 TaqMan R	TAGTTTTGGCATCTTTGTTTTAAGGTTAC
	TaqMan probe	AACTCCGTCAGTGTTTTA
*lsiI*C	SOE-A	AACTGCAGTCAGGTGTCAATCCTCCAG
	SOE-B	TAGATTTTTCATCCTATCTCC
	SOE-C	*GGAGATAGGATGAAAAATCTA* AATTAATGGAAAACTTGTGAC
	SOE-D	GCTCTAGACACCGTATTAGCATTTCGC
	F	AACGGCCGGGGGAATTTCCTCGTCCT
	R	CGGGATCCGCAACTATTATTTAGCATGTC
	SRM05 Fex	CAACGACCAATTGGGTAG
	SRM06 Rex	AAAATATGGAGATGGTGGG
	SRM73 TaqMan F	TTTAGCAGAAGATATGGTGGAGAACTT
	SRM74 TaqMan R	TGTCACCGCATCCACATCAT
	TaqMan probe	CCGATGTCATTACCC

[a] Primers containing "SOE" in the name were used for for construction of mutant alleles uisng SOEing PCR; primers ending in "ex" were designed to be external to the SOEing primers and were used in allele-specific PCR used to confirm the presence of the null mutant allele; oligonucleotides with TaqMan F, R, or probe represent the forward and reverse primers used in qRT-PCR analyses and the qRT-PCR probe, respectively.

[b] Underlined sequences designate restriction enzyme recognition sites.

[c] Italicized sequences indicate complimentary, overlapping regions.

Figure A3.2 OD_{600} measurements were used to monitor growth in BHI at 37^{o}C until early-stationary phase (OD_{600}=1+3h) of lineage II-specific null mutants and their parent strain 10403S (A) and lineage I-specific null mutants and their parent strain F2365 (B).


A
2.0
1.6
1.2
0.8
0.4
0.0
OD600 measurements
3
4
5
6
7
8
9
10
Time (hours post inoculation)
10403S
ΔlsiIIX
ΔlsiIIY
ΔlsiIIZ

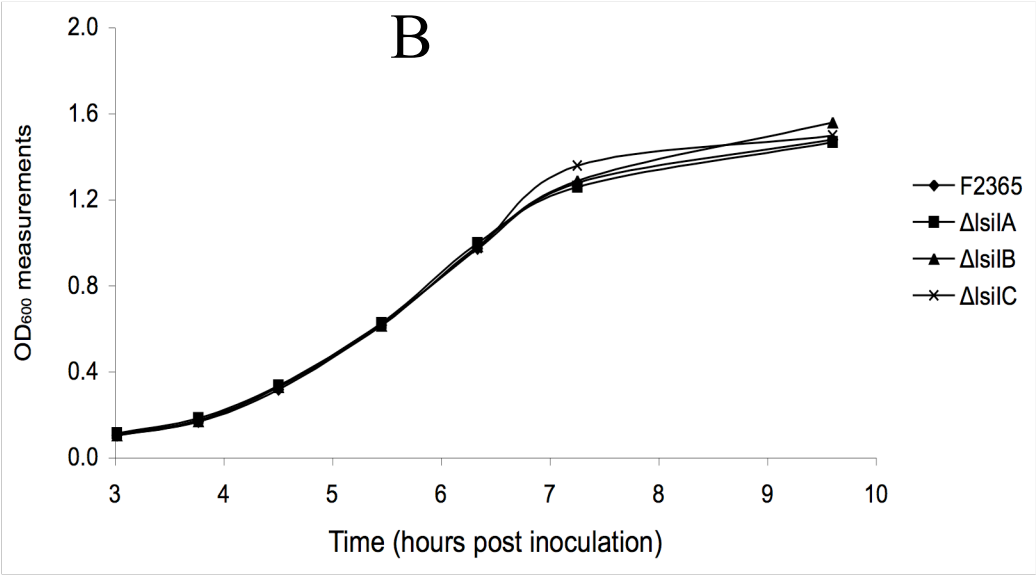
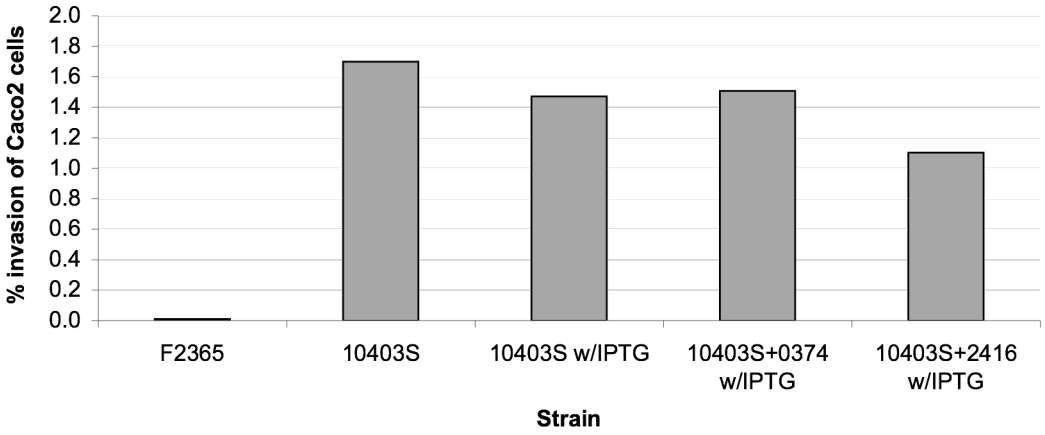

B
2.0
1.6
1.2
0.8
0.4
0.0
OD600 measurements
3
4
5
6
7
8
9
10
Time (hours post inoculation)
F2365
ΔlsiIA
ΔlsiIB
ΔlsiIC

Figure A3.3 This figure depicts the Caco2 cell invasion results (n=1) of two *L. monocytogenes* complementation mutants. The mutants were constructed in a 10403S (lineage II) background using the integration vector pLIV2 (containing an IPTG inducible promoter) with either *F-lmo_0374* (*lsiIA*) or *F-lmo_2416* (*lsiIC*) cloned into the MCS (both genes are lineage I specific). Inducteion of gene transcription was achieved by adding IPTG (final concentration of 0.5mM) to the growing cultures such that at OD_{600}1+3h the cells had been incubated with IPTG for 5 h. Data is reported as CFUrecovered/CFUinfected*100.


% invasion of Caco2 cells
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
F2365
10403S
10403S w/IPTG
10403S+0374 w/IPTG
10403S+2416 w/IPTG
Strain

CHAPTER 4

CONTRIBUTIONS TO SELECTED PHENOTYPIC CHARACTERISTICS OF LARGE SPECIES- AND LINEAGE-SPECIFIC GENOMIC REGIONS IN *LISTERIA MONOCYTOGENES*

Abstract

Listeria monocytogenes is a human foodborne pathogen that represents at least three phylogenetic lineages that appear to differ in their ability to cause human disease as well as their ability to grow and multiply in natural environments. We hypothesized that genomic regions specific to *L. monocytogenes* or selected *L. monocytogenes* lineages or strains may contribute to virulence and phenotypic differences among *L. monocytogenes* strains. A whole genome alignment of two completed *L. monocytogenes* genomes and the one completed *L. innocua* genome initially identified 28 genomic regions of difference (RD) > 4 kb that were found in one or both *L. monocytogenes* genomes, but absent from the non-pathogenic *L. innocua* as well as 14 RDs found only in *L. innocua*. A total of 3 and 8 RDs were only found in the lineage I strain F2365 or the lineage II strain EGD-e, respectively. Strain and lineage specific distribution of alls RDs was further characterized using BLAST searches against 18 unfinished *L. monocytogenes* genomes; 15 RDs were found in all or most of the unfinished *L. monocytogenes* genomes (i.e., >15/18); three RDs were found in >6 lineage I genomes, but no lineage II genomes and four RDs were found in all lineage II genomes, but no lineage I genomes. Two *L. monocytogenes*-specific RDs (found in all or most *L. monocytogenes*) and one lineage II-specific RD were chosen for further analyses. Null mutants with deletions of the full RD 16 (6.8 kb), 25 (6.5 kb), and 30 (6.5 kb) were generated and showed no apparent growth defects. None of these three

mutants was impaired in their ability to invade human intestinal epithelial cells and or their ability to grown in activated mouse macrophage cells. In a competition experiment in pH 5.5 minimal media, the ΔRD30 null mutant showed reduced ability to grow as compared to its parent strain though, indicating that RD30 may have a role in *L. monocytogenes* growth under limited nutrient conditions at acidic pH. Our data show that *L. monocytogenes* contain a number of species- and lineage-specific genomic regions. Deletion of two regions investigated (one widely distributed among *L. monocytogenes* lineages and one lineage II specific) did not reveal any contributions to *L. monocytogenes* survival and growth under specific environmental conditions. We determined that one widely distributed region, RD30, contributed to *L. monocytogenes* growth and survival in nutrient restricted environments with acidic pH. RD30 could, therefore, have a role in supporting *L. monocytogenes* survival in acidic food and/or host environments.

Introduction

The genus *Listeria* contains two pathogenic species (*L. monocytogenes* and *L. ivanovii*) as well as four non-pathogenic species (*L. innocua*, *L. welshimeri*, *L. seeligeri*, and *L. grayi*). The most well known of these species, *L. monocytogenes*, is a human foodborne pathogen that can also cause infections in a broad range of other animal species (Pell, 1997). *L. monocytogenes* can be divided into at least three lineages (Wiedmann et al., 1997, Rasmussen et al., 1995, Piffaretti et al., 1989; Roberts et al., 2006, which may have adapted to different ecological niches. *L. monocytogenes* lineage I strains have not only been responsible for the majority of human listeriosis outbreaks (Swaminathan & Gerner-Smidt, 2007) and are overrepresented among human clinical isolates (Gray et al., 2004), but also appear to show a significantly higher probability of causing human disease as compared to

lineage II strains (Chen et al., 2006). In addition, lineage I strains seem to consistently show a significantly higher ability to spread from cell-to-cell *in vitro* (Wiedmann et al., 1997, Gray et al., 2004) as compared to lineage II strains, further supporting that lineage I may be a human host adapted lineage. While also common among isolates from sporadic human listeriosis cases, lineage II strains are overrepresented among food isolates (Gray et al., 2004) and are, on average, found at higher levels in foods as compared to lineage I strains (Chen et al., 2006), suggesting that these strains are better adapted to non-host associated environments. Lineage III strains are most frequently isolated from animal listeriosis cases (Jeffers et al., 2001), are sometimes found among isolates from human clinical cases.

Initial genome comparisons between *L. monocytogenes* EGD-e and *L. innocua* (Glaser et al., 2001) not only confirmed that the major, approx. 9 kb, *L. monocytogenes* virulence gene cluster, which contains the virulence genes *prfA*, *hly*, *plcA*, *plcB*, *actA*, and *mpl*) is absent from *L. innocua*, but also identified a number of larger genomic regions and specific genes unique to either *L. monocytogenes* EGD-e or *L. innocua* CLIP11262. A number of the *L. monocytogenes* specific genes identified through these genome comparisons were subsequently shown to play a role in *L. monocytogenes* virulence. For example, *aut*, which was present in *L. monocytogenes* EGD-e and absent in *L. innocua,* was found to encode a surface associated autolysin that contributes to *L. monocytogenes* virulence (Cabanes, 2004). Subsequent genome comparisons between two *L. monocytogenes* serotype 1/2a (lineage II) strains and two serotype 4b (lineage I) strains (Nelson et al., 2004) as well as genomic DNA array studies on larger set of *L. monocytogenes* isolates (Doumith et al., 2004; Zhang et al., 2003) also identified a number of lineage specific genes and regions of differences, including 83 serotype 1/2a specific and 51 serotype 4b specific genes identified by Nelson et al. (2004). Interestingly, one region specific to *L.*

monocytogenes lineage II was found to represent an operon encoding an extracytoplasmic function (ECF) sigma factor (σ^C) as well as two other proteins with unclear functions (Zhang, 2003). Deletions of *sigC* and one of the other genes in this region (i.e., lmo0422) lead to reduced heat resistance, suggesting that this operon encodes a thermal resistance system (Zhang, 2005), which may contribute to increased environmental stress resistance of lineage II strains.

As previous data clearly suggest, the *L. monocytogenes* species- and lineage–specific genes and genomic regions can contribute to relevant phenotypic characteristics of this foodborne pathogen. We hypothesized that additional *L. monocytogenes* species- and lineage–specific genomic regions may contribute to virulence or other phenotypic characteristics associated with foodborne transmission. To investigate this hypothesis, we performed initial *in silico* genome analyses, utilizing available finished genomes for one *L. monocytogenes* lineage I and one lineage II strain as well as unfinished genome sequences for 18 additional *L. monocytogenes* strains, to identify lineage- and species-specific genomic regions. Lineage and species specificity of selected regions was subsequently confirmed on a larger diversity strain set and null mutants were constructed in selected regions to characterize the phenotypic contributions of these regions.

Methods and Materials

Genome sequences. The sequence files for the four closed *Listeria* genomes were downloaded using the NCBI GenBank FTP service (ftp://ftp.ncbi.nih.gov/genbank/genomes/; accessed on 7/25/07); the files used were (i) AL592022.gbk (*L. innocua* CLIP11262; Glaser et al., 2001); (ii) AE017262.gbk (*L. monocytogenes* F2365, a serotype 4b, lineage I strain; Nelson et al., 2004); (iii) AL591824.gbk (*L. monocytogenes* EGD-e, a serotype 1/2a, lineage II strain; Glaser et al., 2001); and (iv)

AM263198 (*L. welshimeri* SLCC5334, Hain et al., 2006). The unclosed genomes for *L. monocytogenes* strains H.7858 and F.6854 (CMR version 15.2 for both genomes; Table 4.1; Nelson et al., 2004) were obtained from TIGR's Comprehensive Microbial Resource website (http://cmr.tigr.org/tigr-scripts/CMR/CmrHomePage.cgi). Supercontig assemblies of draft genome sequences for an additional 16 strains (Table 4.1) were downloaded (on 10/4/07) from the Broad Institute of Harvard and MIT website (*Listeria monocytogenes* Sequencing Project, http://www.broad.mit.edu).

Genome alignments. The program Multiple Alignment of Conserved Regions in Genome Sequences (MAUVE; available at http://gel.ahabs.wisc.edu/mauve/index.php; Darling et al., 2004) was used to generate an initial alignment of the genomes for *L. innocua* CLIP11262, *L. monocytogenes* F2365, and *L. monocytogenes* EGD-e using the following program settings: Default seed weight=15, Minimum LCB weight=default, Aligner=Muscle3.6, Minimum island size=50, Minimum backbone gap size=50, and Minimum backbone size=50. The backbone.xls file generated by MAUVE upon alignment completion was used to identify all regions > 4kb that were absent in at least one of three genomes; these regions were complied into a list of regions of difference (RD; Table 4.2).

Verification of distribution of regions of differences in draft genome sequences and *L. welshimeri* genome sequence. To determine the distribution among all available *Listeria* genome sequences for each region of differences identified by the initial MAUVE analyses described above, two in silico approaches utilizing additional unfinished *L. monocytogenes* genome sequences and the *L. welshimeri* genome sequence were used including (i) BLAST searches of the full sequences for the 50 RDs against these genome sequences, and (ii)

Table 4.1 *Listeria* genomes used in MAUVE alignments and Blast searches.

Genome ID[a]	Lineage	Serotype	Ribotype	Description
Sequenced and Closed *Listeria* Genomes				
F2365	I	4b	DUP 1038B	isolated from cheese and associated with a human listeriosis outbreak in CA
EGD-e	II	1/2a	DUP 1039C	lab strain originally associated with a listeriosis outbreak in rabbits
L. innocua CLIP2262	-	6a	-	nonpathogenic, closely related to *L. monocytogenes*
L. welshimeri SLCC5334	-	6b	-	Nonpathogenic
Draft *L. monocytogenes* Genomes				
F.6854	II	1/2a	DUP 1053A	RTE meat associated w/ human outbreak
H.7858	I	4b	-	RTE meat associated w/ human outbreak
J0161/FSL R2-499	II	1/2a	DUP 1053A	human outbreak
10403S	II	1/2a	DUP 1030A	human sporadic
FSL J1-194	I	1/2b	DUP 1042B	human sporadic
FSL R2-503	I	1/2b	DUP 1051B	human outbreak
J2818	II	1/2a	-	human outbreak
HPB2262	I	4b	-	human outbreak
FSL N1-017	I	4b	DUP 1042C	RTE fish product
FSL N3-165	II	1/2a	DUP 1045A	animal clinical
F6900	II	1/2a	DUP 1053A	Soil
LO28	II	1/2c	-	lab strain
FSL J2-003	II	1/2a	DUP 1039C	animal clinical
FSL J1-175	I	1/2b	DUP 1042A	Water
FSL F2-515	II	1/2a	DUP 1062A	RTE meat product
FSL J2-064	I	1/2b	DUP 1052	animal clinical
FSL J2-071	III	4c	DUP 1061A	animal clinical
FSL J1-208	III	4a	DUP 10142	animal clinical

Table 4.2 Characteristics and distribution of regions of difference (>4kb) identified in initial *L. innocua* and *L. monocytogenes* genome comparisons

Region Designation[a]	General Category	Genome 1[c]	Locus Tags	Size (bp)	Genome 2	Locus Tags	Size (bp)	Location (Mb)	Average %GC	Distribution[d,e]				
										Lineage I (8)	Lineage II (10)	Lineage III (2)	*L. innocua* (1)	*L. welshimeri* (1)
1	Transposon related	LI CLIP11262	lin_1618-1624	4863	-	-	-	~1.62	40.18	0	1	0	1	0
2		LI CLIP11262	lin_0864-0869	4815	-	-	-	~0.89	28.48	0	0	0	1	0
3		LI CLIP11262	lin_0384-0386	3925	LM EGDe	lmo_0365-0367	4115	~0.39	40.44	1	9	1	1	0
4	Metabolism (carbohydrate)	LI CLIP11262	lin_2454-2459	6010	-	-	-	~2.47	34.62	0	0	0	1	0
5		LM EGDe	lmo_0444-0448	8765	-	-		~0.47	36.24	3	4	1	0	0
6		LI CLIP11262	lin_0197-0201	5264	-	-	-	~0.18	39.17	1	0	0	1	1*
7	Cell wall associated	LM F2365	F-lmo_1974	6012	-	-	-	~2	35.23	7	0	0	0	1
8	Cell wall associated	LI CLIP11262	lin_2281	4868	LM F2365	F-lmo_2210	4589	~2.25	34.73	7	2	0	1	0
9	Cell wall associated	LI CLIP11262	lin_0665	2495	LM F2365	F-lmo_0693	5131	~0.69	32.14	7	1	0	1	0
10	Cell wall associated	LI CLIP11262	lin_1204	4992	LM F2365	F-lmo_1251-1254	5530	~1.2	36.63	7	0	0	1	0
11	Cell wall associated	LM EGDe	lmo_0842	6083	LM F2365	F-lmo_0859	6134	~0.875	35.95	8	10	0	0	0
12	Cell wall associated	LM EGDe	lmo_2576	4592	-	-	-	~2.65	37.67	0	10	1	0	0
13	Metabolism and cell wall associated	LM EGDe	lmo_1666	5136	LM F2365	F-lmo_1690	5136	~1.7	38.62	6	10	0	0	0
14	Metabolism and cell wall associated (internalins)	LM EGDe	lmo_0262-0264	4963	LM F2365	F-lmo_0278-0283	9991	~0.28	34.29	7	10*	1	0	1*
15	Cell wall associated (internalin-like)	LI CLIP11262	lin_0557-0561	4993	-	-	-	~0.58	35.47	0	0	0	1	0
16	***Cell wall associated (internalin-like)***	**LM EGDe**	***lmo_0332-0334***	***6860***	**LM F2365**	***F-lmo_0348-0351***	***9744***	***~0.36***	***35.02***	***8***	***10***	***1****	***0***	***0***
17		LI CLIP11262	lin_0801-0806	9146	-	-	-	~0.82	32.94	1*	0	0	1	0
18		LM F2365	F-lmo_2701-2707	5653	-	-	-	~2.74	28.94	1	0	0	0	0
19		LI CLIP11262	lin_0476-0478	5661	-	-	-	~0.495	28.15	0	0	0	1	0
20		LM EGDe	lmo_0141-0151	5275	LM F2365	F-lmo_0159-0166	2989	~0.14	29.95	5	10	1	0	1*
21	Metabolism (carbohydrate)	LI CLIP11262	lin_0759-0764	6530	LM F2365	lmo_0765-0770	6530	~0.78	35.02	6	10	1	1	1
22	Metabolism (carbohydrate)	LM EGDe	lmo_2732-2736	7481	LM F2365	F-lmo_2719-2723	7481	~2.76	39.34	8	10	1	0	0
23	Metabolism (carbohydrate)	LM EGDe	lmo_2780-2784	6289	LM F2365	F-lmo_2771-2775	6289	~2.83	37.95	8	10	1	0	0
24	Metabolism (carbohydrate)	LM EGDe	lmo_0630-0634	5164	LM F2365	F-lmo_0659-0663	5164	~0.67	36.47	8	10	0	0	1
25	***Metabolism (carbohydrate)***	**LM EGDe**	***lmo_0734-0739***	***6547***	**-**	**-**	**-**	***~0.76***	***36.54***	***0***	***10***	***0***	***0***	***1****
26	Metabolism (carbohydrate)	LM EGDe	F-lmo_0266-0272	8055	-	-	-	~0.27	35.35	8	5	0	0	1
27		LM EGDe	lmo_0746-0754	4765	LM F2365	F-lmo_0776-0778	2334	~0.77	36.41	8	10	0	0	1
28	Metabolism	LM EGDe	lmo_1030-1036	9087	LM F2365	F-lmo_1051-1057	9087	~1.05	37.23	8	10	1	0	0
29	Metabolism (carbohydrate)	LM EGDe	lmo_2771-2773	4230	LM F2365	F-lmo_2761-2763	4598	~2.81	37.61	8	10*	2*	0	0
30	***Metabolism (amino acid)***	**LM EGDe**	***lmo_0036-0041***	***6526***	**LM F2365**	***F-lmo_0045-0050***	***6526***	***~0.04***	***38.00***	***8***	***10***	***1***	***0***	***0***
31		LM EGDe	lmo_0459-0476	16887	LM F2365	F-lmo_0493-0504	10054	~0.495	34.12	7	9*	0	0	0
32		LM EGDe	F-lmo_2051-2059	11103	-	-	-	~2.08	35.99	4*	0	0	0	0
33	Transposon related	LM EGDe	lmo_1097-1122	26928	LM F2365	F-lmo_1111-1128	11663	~1.12	31.83	3	2	0	0	0
34	Transposon related and cell wall associated (internalin-like)	LM EGDe	lmo_1060-1063	4271	-	-	-	~1.08	33.14	1	10	0	0	1
35		LM EGDe	lmo_0171-0174	3698	-	-	-	~0.169	35.50	0	5	0	0	0
36	Metabolism and cell wall associated	LI CLIP11262	lin_0415-0419	5378	LM F2365	F-lmo_0413-0417	5448	~0.42	34.21	8	0	1*	1	1
37	Phage related	LI CLIP11262	lin_1697-1765	48131	-	-	-	~1.71	36.05	1	0	0	1	0
38	Phage related	LI CLIP11262	lin_2372-2426	38989	LM EGDe	lmo_2271-2332	41458	~2.36	35.99	5*	8	0	1	0
39	Phage related	LI CLIP11262	lin_1231-1302	49853	-	-	-	~1.24	34.71	0	0	1*	1	0
40	Phage related	LI CLIP11262	lin_0071-130	39863	-	-	-	~0.077	37.30	0	0	0	1	0
41	Phage related	LI CLIP11262	lin_2561-2610	37927	-	-	-	~2.58	35.10	1	0	0	1	0
42		LI CLIP11262	lin_0521-0525	8077	-	-	-	~0.54	33.83	1	0	0	1	1
LIPI1	Pathogenicity island	LM EGDe	lmo_0200-0206	9137	LM F2365	F-lmo_0211-0217	9037	~0.2	36.08	8	10	2	0	0
43	Metabolism	LM EGDe	lmo_1968-1974	5099	-	-	-	~2.04	37.43	0	10	0	0	1
44	Metabolism	LI CLIP11262	lin_1064-1069	11140	LM F2365	F-lmo_1093-1098	11140	~1.08	32.77	3	0	1	1	1*
45		LI CLIP11262	lin_2742-2744	4897	-	-	-	~2.76	33.69	0	1	0	1	0
46	Metabolism	LI CLIP11262	lin_1811-1814	4199	-	-	-	~1.81	33.26	0	0	0	1	0
47	Metabolism	LM F2365	F-lmo_2626-2631	3829	-	-	-	~2.67	32.04	4	1	0	0	0
48	Metabolism (carbohydrate)	LM EGDe	lmo_2788-2789	3793	-	-	-	~2.87	39.33	0	10	1	0	0
49	Metabolism (carbohydrate)	LM EGDe	lmo_1079-1085	10253	-	-	-	~1.11	33.19	4	9	0	0	0

[a] Regions designated in bold, italicized lettering were further investigated.

[b] LIPI1 refers to the main virulence gene cluster in *L. monocytogenes* (Vazquez-Boland, 2001)

[c] LI refers to *L. innocua* and LM refers to *L. monocytogenes*

[d] Genome counts include *L. monocytogenes* F2365, *L. monocytogenes* EGDe, and *L. innocua* CLIP11262 and reflect results of BLAST searches and with MAUVE alignment confirmation

[e] Entires marked with an asterisk indicate that in one or more genomes searched <50% of the region was present

construction of a progressive MAUVE alignment of the three genomes included in the initial alignment (F2365, EGD-e, and CLIP11262) and the additional genome sequences. BLAST searches were performed using the blastn program. For unfinished *L. monocytogenes* genomes available from the Broad Institute, BLAST searches allowing for gapped alignments were conducted through the Broad website. BLAST searches against the *L. monocytogenes* H.7858 and F.6854 genomes were performed via the CMR website with the contig database for each genome. BLAST searches against the *L. welshimeri* SLCC5334 genome (Hain, 2006) were performed via NCBI. The program ProgressiveMAUVE (also available at http://gel.ahabs.wisc.edu/mauve/index.php) can use contig assemblies of genome sequences, and was, thus, used to generate a whole genome alignment for all *Listeria* genome sequences available (finished and draft). The resulting alignment was visually assessed to confirm BLAST results for each RD.

Protein BLAST (blastp) searches of the predicted protein sequences corresponding to each ORF contained in selected RDs. Any RD identified by the initial whole genome alignment as specific to either or both *L. monocytogenes* strain (i.e. not found in *L. innocua*) and that contained either metabolism or cell-wall related ORFs was chosen for further investigation. Predicted protein sequences for each ORF contained in the selected RDs were searched using blastp (available at http://www.ncbi.nlm.nih.gov/blast/Blast.cgi) against the NCBI database for all non-redundant protein sequences. RDs with blastp hits of greater than 35% amino acid similarity corresponding to one or more contiguous ORFs found in another organism are listed in Table 3.

PCR screens to confirm presence/absence of three selected RDs in a diversity set of unsequenced *Listeria* isolates. A set of 48 diverse *L. monocytogenes*

isolates belonging to all three lineages (lineage I, n= 18; lineage II n=20; lineage III, n=10 and five diverse *L. innocua* isolates (Table A4.1; additional isolate information is available at www.pathogentracker.net.) was assembled to confirm the presence/absence patterns of three selected RDs. The thirty-eight lineage I and II isolates were selected as a subset of 132 *L. monocytogenes* isolates collected in NY state from the years 2000 to 2002 that included food, human clinical, and animal clinical isolates (Nightingale, 2005; Sauders, 2005). All ten lineage III isolates were previously published as part of a larger study of *L. monocytogenes* lineage III (Roberts, 2006). The lineage III isolates were selected to represent isolates from foods and human and animal clinical cases. The five *L. innocua* isolates were selected to represent five unique allelic types for *sigB*, which encodes the alternative sigma factor σ^B; sigB sequencing has previously been shown to allow for sensitive subtype discrimination of different *Listeria* spp. (Sauders, 2005).

Whole genomic DNA for use as a PCR template was prepared from all isolates using the QiaAMP DNA Mini kit (Qiagen). For each of the three selected RDs, a set of region specific PCR primers (Table A4.2) that flanked a given RD were designed. Primers were designed based on genome sequences for *L. monocytogenes* EGD-e and F2365 and *L. innocua* CLIP11262 to assure that primers would successfully amplify in both species; if necessary separate primers were designed for amplification of RDs in lineages I and II and *L. innocua* (Table A4.2). PCR was performed using the iProof DNA polymerase (BioRad) and cycling conditions according to the manufacturer's instructions. Amplification of a PCR fragment of the size expected for presence of a given RD was interpreted as indicating the presence of a given RD.

Mutant construction. Null mutants in three selected RDs (RD16, RD25, and RD30) were constructed in the lineage II *L. monocytogenes* strain EGD-e (Table 4.3); null mutants were constructed by deleting the complete RD and not just a specific

gene in a given RD. Construction of deletion mutants was performed using SOE-PCR (Horton *et al.*, 1990; see Table A4.2 for primers) to construct internal deletion alleles in a temperature-sensitive suicide plasmid (pKSV-7) (Smith and Youngman, 1992) and subsequent allelic exchange mutagenesis (as described by Camilli et al., 1990). Allelic exchange mutagenesis was confirmed by PCR and sequencing in the final mutant strains constructed.

Bacterial growth conditions. Standard growth conditions were employed to grow the parent strain and mutant strains for all phenotypic characterization experiments; cultures were always grown at 37^{o}C with aeration (shaking at 210 rpm). Briefly, one *L. monocytogenes* colony from a BHI agar plate was transferred into 5ml of sterile BHI broth and grown for 12-18h. This culture was diluted 1:100 into either 5ml fresh BHI or 5 ml minimal media (MM) broth and grown to an optical density of 0.4 (determined at 600 nm for BHI or at 480nm for MM). The OD=0.4 culture was diluted 1:100 into the final growth medium (either BHI or MM). Basic MM was made as described by Phan-Thanh and Gormon (1997) with one modification; glucose was added to a final concentration of 25mM (as opposed to 55mM). MM was prepared to a final pH of either 5.5 or 7.1 using a 1M solution of 2-(N-morpholino)ethanesulfonic acid (MES) buffer. For some growth experiments, MM was formulated with different concentration of selected nutrients (i.e., ribose, arginine, and citrulline). Specifically MM was prepared with (i) ribose concentrations of 0, 5, or 13mM (approx. 0, 50, or 100% of the ribose concentration in MM reported by Premaratne [1991]), (ii) arginine concentration of 0, 0.29, or 0.52 mM (approx. 0, 50, or 100% of the arginine concentration in MM reported by Premaratne [1991]), or (iii) citrulline concentrations of 0, 0.26, or 0.52 mM. Bacterial growth rate constants during logarithmic growth were calculated using the following equation: $\log_{10}N - \log_{10}N_0 = (\mu/2.303)(t - t_0)$ and $\mu=0.693/g$, where $N= OD_{480}$, t=time, μ=growth rate constant, and g=generation time.

Table 4.3 *Listeria* strains used for mutant construction, growth curves, invasion and intracellular growth assays.

Strain	Previous ID	Details	Reference
EGDe	-	*L. monocytogenes,* serotype 1/2a	Glaser et al, 2001
ΔRD16	-	*L. monocytogenes* EDGeΔRD16	this study
ΔRD25	-	*L. monocytogenes* EDGeΔRD25	this study
ΔRD30	-	*L. monocytogenes* EDGeΔRD30	this study
Δ*inlA*	DP-L4405	*L. monocytogenes* 10403SΔ*inlA*	Bakardjiev et al, 2004
Δ*hly*	DP-L2166	*L. monocytogenes* 10403SΔ*hly*	Jones & Portnoy, 1994

Competitive growth experiments. As initial growth curves performed separately with mutant strains and the parent strain showed no effects of the mutations on growth, we performed competitive growth experiments under selected conditions. For these experiments, cultures were grown in MM as described above to an OD_{480}=0.4; each mutant was then mixed 1:1 with the parent strain, followed by 1:100 dilution into the final growth media. These competition cultures were serially diluted in phosphate-buffered saline (PBS) and spiral plated (Spiral Biotech, Norwood MA) on BHI directly after mixing and at three other pre-determined time points corresponding to mid-log, early- and late-stationary phases (i.e. 15, 29, and 35h after inoculation for cultures grown in 25mM glucose at pH 5.5; 12, 18 and 35h for cultures grown in 25mM glucose at pH 7.1; and 12, 20, and 23h for cultures grown in 2.5mM glucose at pH 7.1). BHI plates were incubated for 24 h at 37^{o}C before the *L. monocytogenes* colonies were enumerated. For each time point, 12 colonies were subjected to PCR analysis to identify colonies as the mutant or parent strains. PCR was performed using GoTaq Green (Promega) with allele specific primers (Table A4.2). Cycle conditions were: 4 m at 95^{o}C; followed by 20 cycles of 95^{o}C for 30s, 62^{o}C for 30 s, and 72^{o}C for 1 m with a 0.5^{o}C decrease per cycle; then 20 cycles of 95^{o}C for 30s, 52^{o}C for 30 s, and 72^{o}C for 1 m; and lastly, 72^{o}C for 7 m. All PCR products were held at 4^{o}C until gel electrophoresis. All competitive growth experiments were performed in triplicate.

Caco2 invasion and intracellular growth assays. Bacterial strains used in all tissue culture experiments are listed in Table 4.3. Invasion and intracellular growth assays in Caco2 cells were performed to evaluate the ability of the different mutant strains to invade and grow in this intestinal epithelial cell line. Caco2 (ATCC HTB-37) cells were grown in Dulbecco's Minimum Essential Medium (DMEM) containing 20% FBS, 1% non-essential amino acids, 1% sodium pyruvate, 0.15% sodium

bicarbonate, and 100 Units/ml each of Penicillin G and Streptomycin (all obtained from Gibo). For all assays, Caco2 cells were seeded, at two days prior to infection, into 24 well plates at a density of 5.0 x10^4 cells per well. *L. monocytogenes* for both invasion and intracellular growth assays were grown in BHI as described above with the final BHI culture grown to early stationary phase (defined as OD_{600}=1.0 plus an additional 3h of incubation) at 37ºC with aeration.

Invasion assays were performed as described previously with a multiplicity of infection (MOI) of approx. 200 (Garner et al., 2006A). Intracellular growth assays were performed with an MOI of approx. 3 essentially as described by Kim et al. (2004). For either assay, the monolayer in each well was washed three times with sterile PBS at 30 min post-infection to remove any unassociated *L. monocytogenes*, followed by addition of fresh medium without antibiotics. At 45 min post-infection, fresh medium containing either 150 μg/ml (for invasion assays) or 50 μg/ml (for intracellular growth) gentamycin was added to kill any extracellular *L. monocytogenes*. For invasion assays, Caco2 cells were lysed with ice-cold dH_2O at 1.5 h after infection and intracellular bacteria were enumerated on BHI plates. For intracellular growth assays, cells were lysed at 1.5, 3, 5, and 7 h after infection, followed by enumeration of intracellular bacteria at each time point. Enumeration of intracellular bacteria was performed using a Spiral Plater (Spiral Biotech; Norwood, MA) to plate lysates on BHI agar plates; plates were incubated for 24 h at 37ºC before enumeration of colonies. Both invasion and intracellular growth assays were performed three times.

Intracellular growth assays in activated J774 cells. Intracellular growth assays in activated J774 cells (a murine macrophage-like cell line; ATCC TIB-67) were performed to evaluate the ability of the different mutant strains to multiply in activated macrophage like cells. J774 were cultured in the same medium that was

used for Caco2 cells (see above), except that the media contained 10% FBS and that no non-essential amino acids were added. For intracellular growth assays, J774 cells were seeded into 24 well plates at a density of $2x10^5$ cells per well. J774 cells were stimulated by the addition of 0.1 μg/ml LPS (*E.coli* O55:B5) 24 h prior to infection (Amano & Akamatsu, 1991). *L. monocytogenes* for J774 intracellular growth assays were grown to stationary phase as described above. Intracellular growth assays were conducted essentially as described previously (Conte et al., 2000), using an MOI of approx. 1. At 30 min post infection, monolayers were washed with PBS to remove unassociated *L. monocytogenes*, and fresh medium with 50 μg/ml gentamycin was added to kill any extracellular *L. monocytogenes*. At 1.5, 3, 5, and 7h post infection, infected J774 cells were lysed with ice-cold dH_2O and intracellular *L. monocytogenes* were enumerated using a spiral plater as detailed above. J774 intracellular growth assays were performed as three independent replicates.

Statistical Analyses. Invasion and intracellular growth assay data were analyzed in Minitab 14 (State College, PA) using one-way ANOVA with post-hoc Tukey "Honestly Significantly Different" method for multiple comparisons. Competitive growth curve data were analyzed with JMP (SAS) 7.0 using Pearson's Chi-square test. P-values of <0.05 were considered significant.

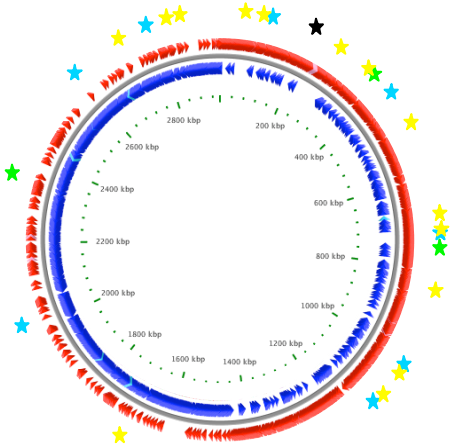
Results and Discussion

Whole genome alignment reveals that *L. monocytogenes* contains numerous large species- and lineage-specific genomic regions, which could contribute to niche adaptation. While previous genome hybridization and sequence analysis studies (Nelson et al., 2004, Doumith et al., 2004) have identified a number of genes and larger "regions of difference" that were found to be unique to *L. monocytogenes* or to specific *L. monocytogenes* lineages, no full genome alignments using genomes for the

two main *L. monocytogenes* lineages and *L. innocua* have previously been used to identify *L. monocytogenes* species and lineage specific regions of differences. We thus assembled, using MAUVE, a whole genome alignment of the previously sequenced and closed genomes for *L. monocytogenes* EGD-e (Lineage I), *L. monocytogenes* F2365 (lineage II), and *L. innocua* CLIP11262 (Table 1). This alignment revealed 50 large genomic regions of difference (RD; using a size cut-off >4 kb) that were identified as present in one or two but not all three of the genomes aligned. The identified RDs are distributed throughout the genomes (Fig. 4.1), ranged in size from 4 kb to approx. 50 kb, and showed a GC content from 28 to 40% (Table 4.2) as compared to a genome average GC content of 37.98 for *L. monocytogenes* EGD-e and 37.38 for *L. innocua* CLIP11262. Five RDs (2,18, 19, 20, and 33; Table 4.2) and two RDs (1 and 3) show GC contests clearly lower (<32%) and higher (>40%), respectively, than average genome wide GC contents. The GC content for only one RD (19) was a statistical outlier (outliers were defined as data points lying more than 1.5*[Interquartile Range] from the median), suggesting acquisition by horizontal gene transfer. A total of 14 and 17 RDs were specific to *L. innocua* or *L. monocytogenes* (both strains), respectively, including the well known *Listeria* virulence gene Island that is absent in *L. innocua* (Vazquez-Boland, 2001). In addition, three and eight RDs were specific to *L. monocytogenes* lineage I and II, respectively (and absent in *L. innocua*). Two RDs were found in *L. innocua* and the lineage II strain genome, while five RDs were found in *L. innocua* and the lineage I strain genome. The functional categories (as determined by the sequence annotations from Glaser et al. [2001] and Nelson et al. [2004]) of genes in RDs were used to classify regions based on the predominant function among the genes in a given RD; the RDs primarily fell into the following categories “carbohydrate metabolism” (10

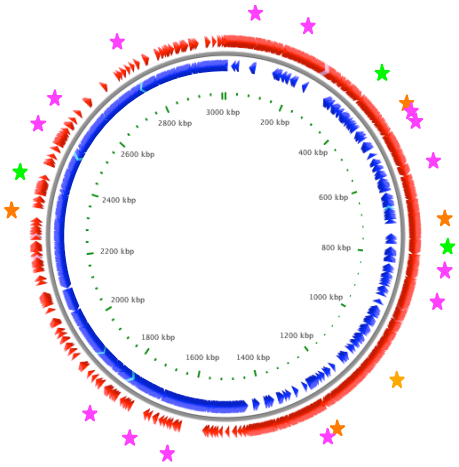
Figure 4.1 Genomes sequences for *L. innocua* CLIP11262, *L. monocytogenes* EGD-e (lineage II), and *L. monocytogenes* F2365 (lineage I) were obtained from GenBank and circular genomes were produced using CGview (Stothard et al., 2005). Stars surrounding each circular genome represent the approximate location of each genomic region of differences detailed in Table 4.2; color-coding of stars was used to indicate the distribution of the genomic regions of difference among the three genomes.

Listeria monocytogenes EGD-e



Accession: NC_003210 Length: 2,944,528 bp; Genes: 2,931

Listeria innocua CLIP11262

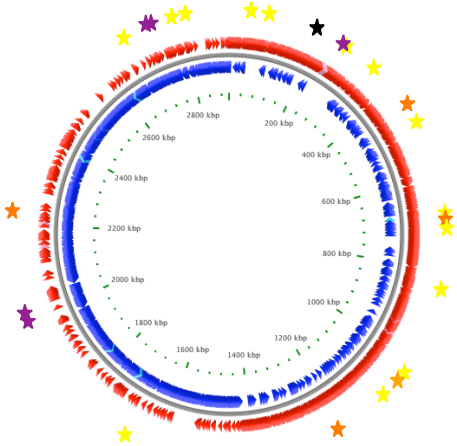


Accession: NC_003212 Length: 3,011,208 bp; Genes: 3,052

Legend:

- region in EGD-e only
- region in EGD-e & *L. innocua*
- region in EGD-e & F2365
- region in F2365 & *L. innocua*
- region in F2365 only
- region in *L. innocua* only
- LIPI

Listeria monocytogenes F2365



Accession: NC_002973 Length: 2,905,310 bp; Genes: 2,995

RDs), "cell wall associated" (8 RDs), and "phage related" (5 RDs) (Table 4.2). All RDs categorized as phage related were present in *L. innocua*, consistent with previous genome comparisons between *L. innocua* CLIP11262 and *L. monocytogenes* EGD-e (Glaser et al., 2001).

To further characterize species- and lineage-specific distribution of the RDs identified through the genome alignment detailed above, we utilized an additional 18 unfinished *L. monocytogenes* genomes (H7858, F6854 [Nelson et al., 2004], and 16 other genome sequences available from Broad [2006]), and a completed *L. welshimeri* genome sequence (Hain et al., 2006) (Table 4.1). These genome sequences were used to (i) perform BLAST searches using the full sequences for the 50 RDs against these genome sequences and to (ii) construct a progressive MAUVE alignment of the three genomes included in the initial alignment (F2365, EGD-e, and CLIP11262) and these additional 19 genome sequences; the total of 22 genomes analysed thus included 20 *L. monocytogenes* genomes (8 lineage I, 10 lineage II, and 2 lineage III genomes), as well as one genome each for *L. innocua* and *L. welshimeri*. These BLAST searches and alignments (Table 4.2) showed that 15 of the initially identified RDs were found in all or most of the *L. monocytogenes* genomes (i.e., >15/18) but absent from *L. innocua*, and 12 RDs were found in all or most of the *L. monocytogenes* genomes but absent from both *L. innocua* and *L. welshimeri*. Three RDs (7, 10, 36) were found in >6 of the 8 lineage I genomes, but no lineage II genomes. Four RDs were found in >8 lineage II genomes (12, 25, 43, and 48), but no lineage I genomes. The remaining RDs are found in low frequency and/or with no clear patterns of lineage specific distribution (e.g., RD 33, 47).

Species- and lineage-specific distribution of three selected RDs is confirmed by PCR-based screens on a larger diversity set of *L. monocytogenes* and *L. innocua.*

Three of the 50 RDs (i.e., RD16, RD25, and RD30; Table 5, Fig. 2) were chosen for follow-up investigation to (i) confirm their species- and lineage-specific distribution and to (ii) generate null mutants with deletions of the complete RDs for phenotypic characterization. The three regions for further investigation were chosen to represent (i) a range of putative functions and (ii) regions found to have a corresponding region present another organism. RD25 contains ORFs related to carbohydrate metabolism, RD16 encodes a cell-wall related ORF, and RD30 contains amino acid metabolism related ORFs (Table 2). In addition, most (RD25) or all (RD30) of the ORFs contained in RD25 and RD30 were identified using blastp searches of predicted protein sequences to also be present as a single region in sequenced *Lactococcus lactis* and *Pediococcus pentosaceus* strains respectively (Table A4.3).

Based on the whole genome alignment analysis, RD16 and RD30 were found in all lineage I and II genomes, and not found in the *L. innocua* or *L. welshimeri* genomes. RD16 encodes a large putative cell-wall associated protein containing leucine rich repeats (LRR), which is similar to the known virulence factor internalin A (previously designated as InlI; Sabet et al., 2005), as well as two hypothetical proteins. RD30 includes five ORFs putatively involved in arginine metabolism, amino acid transport, and also encodes an RpiR family transcriptional regulator (Table 4.4; Fig. 4.3); this RD was previously noted by Doumith et al. (2004) and Glaser et al. (2001). RD25 is a lineage II specific RD; it was found in all 10 lineage II genomes, but absent from lineage I and III genomes as well as absent from *L. innocua* and *L. welshimeri* genomes (except for a 2.3kb fragment of RD 25 with 82% identity found in the *L. welsihmeri* genome using BLAST). RD 25 encodes five ORFs putatively involved in ribose-related metabolism as well as one LacI-family transcriptional regulator.

PCR assays using primers external to the first and last genes in a given RD were used to probe for the presence of these regions in a diverse set of 48 *L. monocytogenes* and five *L. innocua* isolates; amplification of fragment size corresponding the presence of an RD in one of the sequenced reference strains was considered evidence for the presence of this RD in a given isolate. None of the three RDs were found in any the five *L. innocua* isolates, confirming the absence of these three RDs in the *Listeria* spp. most closely related to *L. monocytogenes* (Vazquez-Boland, 2001). The PCR assays further confirmed that RD16 and RD30 are commonly found in both lineage I and II strains (e.g., RD16 was identified in 13/18 lineage I and 15/20 lineage II isolates), even though both island were not detected in some isolates; by comparison genomic microarray data (Doumith et al., 2004) showed that *lmo0333* (located in RD16) was found in all lineage I and II isolates characterized in that study and also indicate that genes found in RD30 were found in all lineage I and II strains. Interestingly, RD16 was absent from all lineage III strains and RD30 was found 2/10 lineage III isolates; by comparison genomic microarray data (Doumith et al., 2004) showed that *lmo0333* (located in RD16) was absent from all (3) serotype 4a lineage III strains and present in all (2) serotype 4c lineage III strains. These finding are consistent with our full genome comparisons where we found RD16 in a serotype 4c genome (J2-071, see Table 4.1). Genomic microarray data (Doumith et al., 2004) also showed that *lmo0036* through *lmo0041* was absent from all lineage III strains. Overall, these data suggest variable presence of these RDs in lineage III strains, consistent with the observation that lineage III represents a highly diverse group of *L. monocytogenes* strains (Roberts, 2006). PCR assays also confirmed that RD25 is lineage II-specific; RD25 was detected in 19/20 lineage II strains and none of the lineage I and III strains (Table 4.4).

Table 4.4 Characteristics and distribution of three RDs (RD16, RD25, and RD30) selected for further characterization[a]

Region	Size (Kb)[b]	# of ORFs	ORF Annotations	PCR Results			
				Lineage I (18)	Lineage II (20)	Lineage III (10)[c]	*L. innocua* (5)
RD16	~6.8	3	cell wall attachment related, contains LRR, similar to internalins	13	16	0	0
RD25	~6.5	6	ribose metabolism, LacI-family transcriptional regulator	0	19	0	0
RD30	~6.5	6	arginine metabolism, amino acid permease, RpiR-family transcriptional regulator	18	20	2	0

[a] Genomes used in Blast searches are different than those used for PCR screen.

[b] Approximate region sizes reflect variation which may occur between strains.

[c] The two lineage III isolates found to contain RD30 were FSL M1-001 (ribotype10142) and FSL R2-128 (ribotype 1061A)

Figure 4.2 This schematic represents the order and orientation of the ORFs comprising each of the three regions of difference that were chosen for further characterization (i.e., RD16, RD25, and RD30). ORFs are shown they appear in the *L. monocytogenes* F2365 (green) and EGD (purple) genomes. Arrows indicate ORF direction and the locus tags for each ORF appear above each arrow; locus tags with "lmo" prefix refer to EGD-e, while locus tags with the "F-lmo" prefix refer to F2365. The corresponding areas in *L. innocua* CLIP11262 (blue), which lacks these regions, are also shown (locus tags with "lin" prefix refer to *L. innocua* CLIP11262). ORFs in the *L. monocytogenes* F2365 or EGD genomes homologous to ORFs in *L. innocua* CLIP11262 are indicated with the dashed arrows.

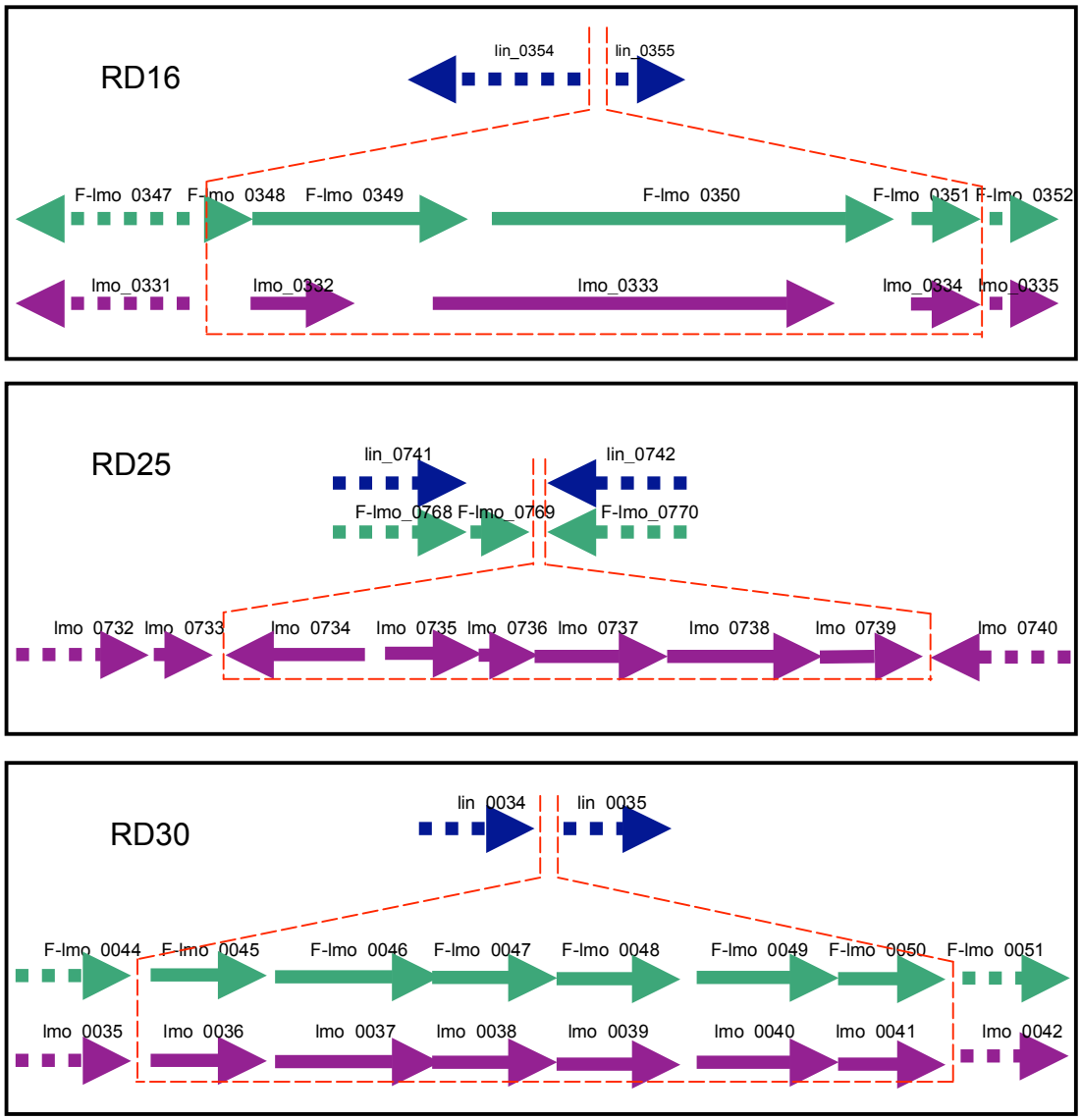
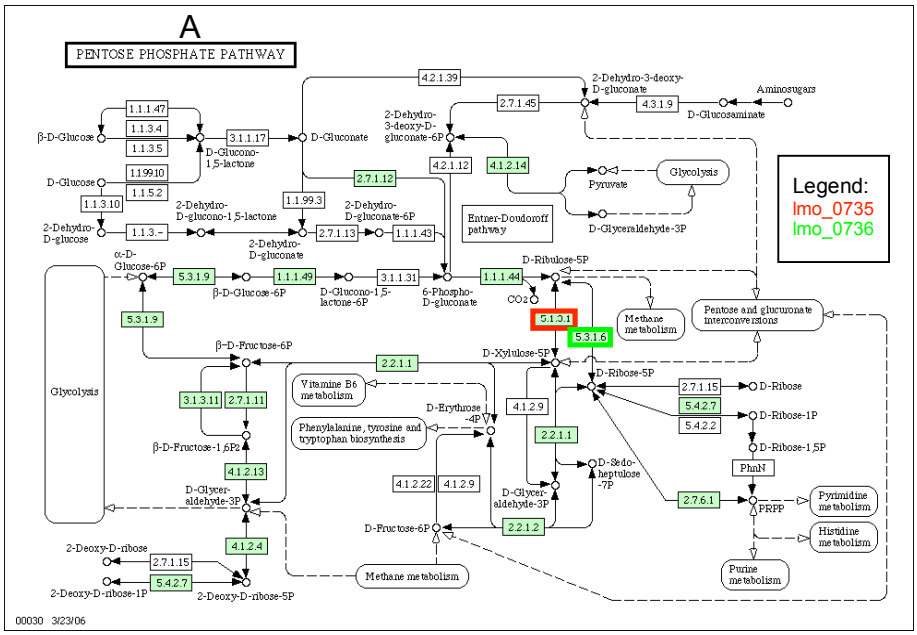
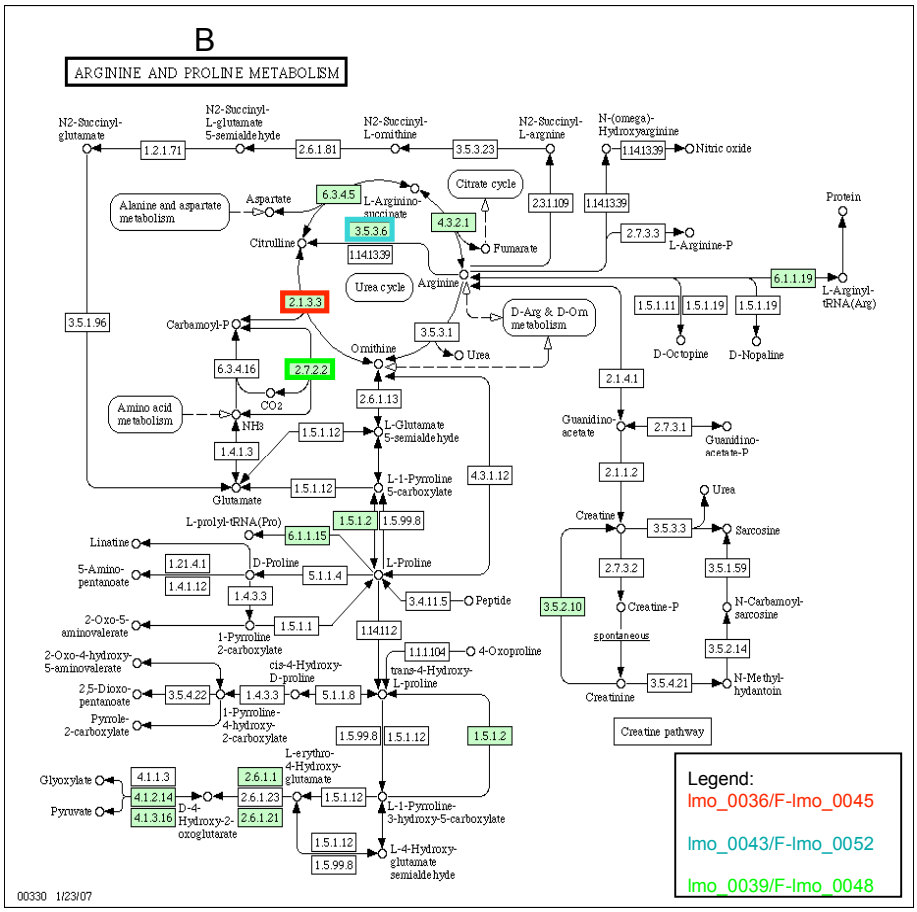

RD16
lin_0354
lin_0355
F-lmo_0347
F-lmo_0348
F-lmo_0349
F-lmo_0350
F-lmo_0351
F-lmo_0352
lmo_0331
lmo_0332
lmo_0333
lmo_0334
lmo_0335
RD25
lin_0741
lin_0742
F-lmo_0768
F-lmo_0769
F-lmo_0770
lmo_0732
lmo_0733
lmo_0734
lmo_0735
lmo_0736
lmo_0737
lmo_0738
lmo_0739
lmo_0740
RD30
lin_0034
lin_0035
F-lmo_0044
F-lmo_0045
F-lmo_0046
F-lmo_0047
F-lmo_0048
F-lmo_0049
F-lmo_0050
F-lmo_0051
lmo_0035
lmo_0036
lmo_0037
lmo_0038
lmo_0039
lmo_0040
lmo_0041
lmo_0042

Figure 4.3 Panels A and B picture biochemical pathways involving selected enzymes putatively encoded in genomic regions RD25 and RD30, respectively; pathway schematics were generated by KEGG (Kanehisa et al., 2006). Each box represents an enzyme and boxes shaded light green represent enzymes encoded in the *L. monocytogenes* EGD genome as determined by genome annotations. In panel A, enzyme 5.1.3.1 (ribulose-phosphate 3-epimerase, putatively encoded by *lmo0735* in RD 25) and enzyme 5.3.1.6 (ribose 5-phosphate isomerase, putatively encoded by *lmo0736* in RD 25) are indicated by color. In panel B, enzyme 3.5.3.6 (arginine deiminase, putatively encoded by *lmo0043* one gene downstream of RD30), enzyme 2.1.3.3 (ornithine carbamoyltransferase, putatively encoded by *lmo0036* in RD30) and enzyme 2.7.2.2 (carbamate kinase, putatively encoded by *lmo0039* in RD 30) are indicated by color.


A
PENTOSE PHOSPHATE PATHWAY
Legend:
lmo_0735
lmo_0736
00030 3/23/06


B
ARGININE AND PROLINE METABOLISM
Legend:
lmo_0036/F-lmo_0045
lmo_0043/F-lmo_0052
lmo_0039/F-lmo_0048
00330 1/23/07

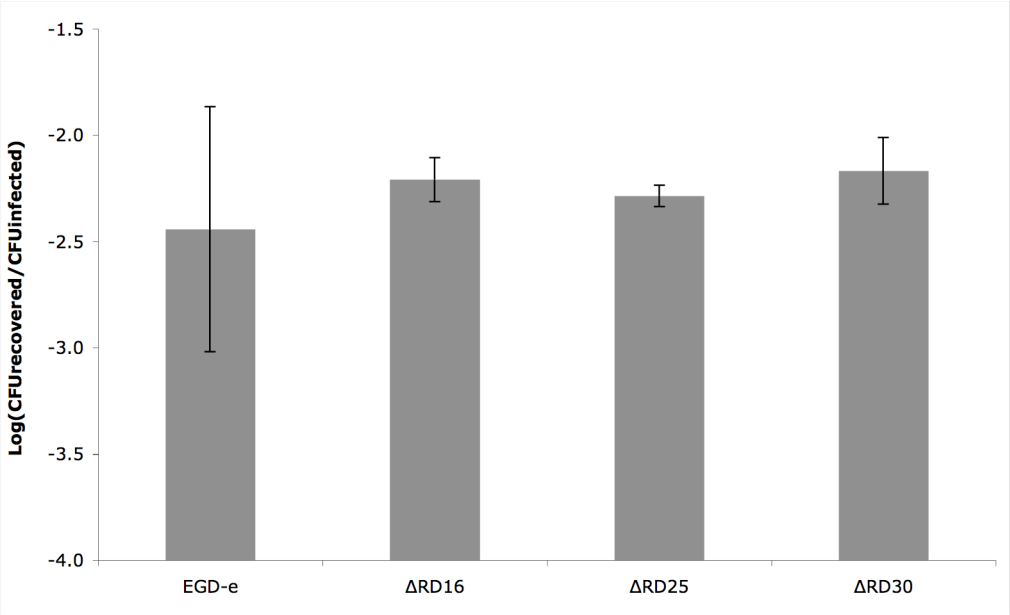
Overall, our data in combination with previous studies (Glaser et al., 2001; Doumith et al., 2004; Hain et al., 2006) showed that (i) RD16 is found in all/most lineage I and II isolates and absent from lineage III, (ii) RD 30 is found in all/most lineage I and II isolates and largely absent from lineage III, and (iii) RD25 is specific to lineage II and absent from lineage I, III, as well as *L. innocua and L. welshimeri*. As all three islands are unique to *L. monocytogenes* and show lineage specific distributions among *L. monocytogenes* isolates, we hypothesized that these RDs may contribute to the ability of *L. monocytogenes* to invade human intestinal epithelial cells an/or its ability to grow intracellularly in infected host cells and thus conducted appropriate tissue culture assays to characterized *L. monocytogenes* EGD-e null mutants in all the RDs as detailed below. Based on the ORFs identified in RD25 (i.e., ORFs related to ribose metabolism) and RD30 (i.e., ORFs related to arginine metabolism), we also hypothesized that (i) both RD25 and RD30 contribute to growth under nutrient limiting conditions, and (ii) RD25 and RD30 specifically contribute to growth and survival under ribose and arginine limiting conditions. The ΔRD25 and ΔRD30 mutants were thus also further characterized for their ability to survive and/or multiply under different nutrient limiting conditions.

***L. monocytogenes* ΔRD16, ΔRD25, and ΔRD30 have similar *in vitro* invasion efficiencies as well as intracellular growth abilities as their parent strain suggesting that neither RD16, RD25, or RD30 contribute to selected *in vitro* phenotypes related to virulence**. Initial characterization of *L. monocytogenes* ΔRD16, ΔRD25, and ΔRD30 strains was performed using two standard tissue culture assays that have previously been used successfully to characterize virulence-related characteristics of *L. monocytogenes* (Dussurget, 2004), including (i) an invasion assay in a human intestinal epithelial cell line (Caco2) cells, and (ii) intracellular growth assays in Caco2 cells and a stimulated mouse macrophage cell line (J774). Invasion

abilities of ΔRD16, ΔRD25, and ΔRD30 into Caco2 cells were not significantly different as compared to the parent strain EGD-e (Fig. 4.4; One way-ANOVA, $p>0.05$), indicating these genomic regions show limited contributions to invasion of human intestinal epithelial cells, a critical first step in the pathogenesis of foodborne listeriosis (Garner et al. 2006). While RD16, in particular, seemed a good candidate RD for a role in invasion (as it encodes a cell-wall associated LRR protein similar a internalins that have previously been shown to play a role in *L. monocytogenes* invasion of host cells), our results are consistent with data reported by Sabet et al. (2005), which also showed that deletion of one gene in RD16 (i.e., *inlI* or *lmo0333*) had no effect on invasion for a number of different cell lines, including Caco2 cells.

Characterization of *L. monocytogenes* ΔRD16, ΔRD25, and ΔRD30 strains in intracellular growth assays showed that all three mutants and the EGD-e parent strain did not differ significantly ($p>0.05$; One-way ANOVA for log [CFU] recovered at each time point) in their intracellular growth capabilities in either Caco2 cells (Fig. 4.5A) or in stimulated J774 cells (Fig. 4.5B). These findings suggest that none of these three genomic regions of interest are critical for intracellular growth *in vitro*. With regard to RD16, these results are consistent with findings by Sabet et al. (2005), which also showed that deletion of *inlI*, located in RD16 did not affect *L. monocytogenes* virulence after intragastric infection of guinea pigs, further supporting a limited role for genes in RD16 to virulence related phenotypes, at least in typically used tissue culture or animal models. While a number of studies have suggested that the host cell cytosol is not an environment that supports the growth of non-cytosol adapted bacteria (Bielecki et al., 1990, Goetz et al., 2001). Researchers determined that intracellularly replicating *Salmonella enterica* serovar Typhimurium required genes for uptake and synthesis of glutamine for survival (Klose and Mekalanos, 1997). Our data suggest

Figure 4.4 Caco2 cells were infected with ΔRD16, ΔRD25, and ΔRD30 strains as well as the *L. monocytogenes* parent strain (EGD-e) and a negative control strain (Δ*inlA*; Table 4.3) and intracellular *L. monocytogenes* were enumerated at 1.5 h post infection. Invasion efficiencies are reported as (log[CFU recovered/CFU infected]) Each bar depicts the mean invasion efficiency based on three independent replicates for each strain; error bars indicate the standard deviation. No statistical difference was observed between EGD-e and any of the mutant strains (ΔRD16, ΔRD 25, or ΔRD 30) using one-way ANOVA with Tukey HSD for multiple corrections ($p>0.05$).


-1.5
-2.0
-2.5
-3.0
-3.5
-4.0
Log(CFUrecovered/CFUinfected)
EGD-e
ΔRD16
ΔRD25
ΔRD30

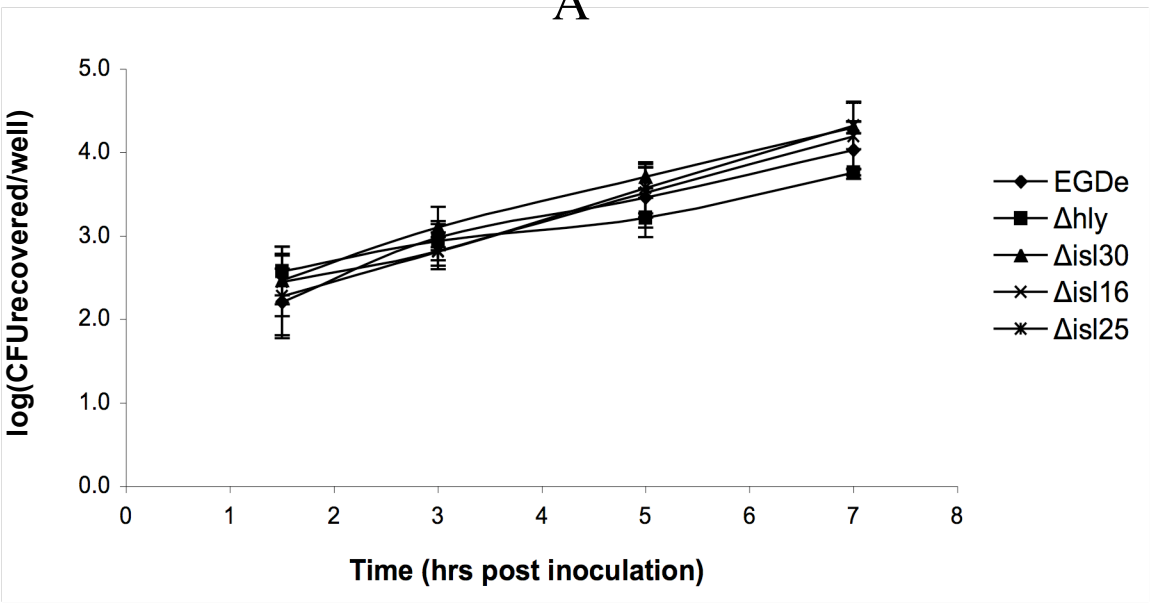
that the arginine and ribose genes encoded in RD 25 and 30 are not critical for intracellular growth of *L. monocytogenes*. Interestingly, a previous study (Klarsfeld et al., 1994) had shown, though, that genes involved in arginine uptake as upregulated during *L. monocytogenes* growth in J774 cells, supporting our initial hypothesis that arginine metabolism may contribute to intracellular survival and multiplication. Research with different *E. coli* strains has shown that ribose can be an important carbohydrate source during colonization of mouse intestines (Fabich, 2008). Joseph et al. (2006) found genes involved in ribose metabolism to be upregulated in *L. monocytogenes* growing in Caco2 cells, indicating a possible role for ribose metabolism in intracellularly replicating *L. monocytogenes*.

While our initial experiments found no evidence for roles in virulence associated phenotypes of the three targeted regions of differences, potential roles of these genes in virulence can obviously not be excluded. However, considering the wide host range of *L. monocytogenes* and the fact that no tissue culture or animal models can truly evaluate human virulence (e.g., as human species specific virulence genes may exist), comprehensive evaluation of these mutants will never be possible. Rather, future studies on the ecology, pathogens, and transmission of *L. monocytogenes* as well as novel genomic and post-genomic approaches (e.g., genome wide approaches to identify immunogenic proteins (Reid, 2002) may provide initial insights on the roles in virulence and host-pathogen interactions of the RDs identified here, thus providing clues on appropriate model systems for further studies using the null mutants characterized here as well as null mutants in other RDs.

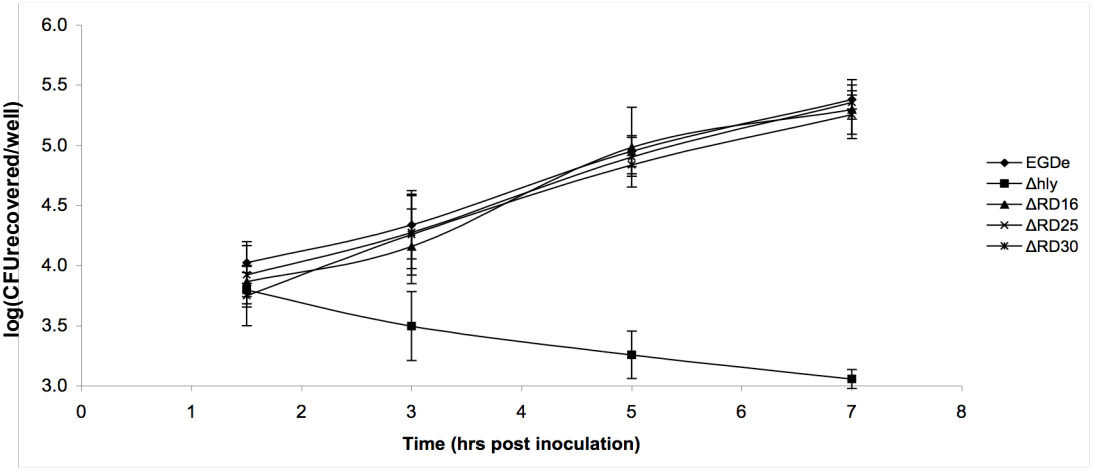
RD25 does not appear to contribute to *L. monocytogenes* growth in nutrient-deprived conditions, while RD30 may be involved in growth in nutrient-limited conditions at acidic pH. As both RD25 and RD30 encode ORFs involved in

Figure 4.5 Tissue culture cells were infected ΔRD16, ΔRD25, and ΔRD30 strains as well as the *L. monocytogenes* parent strain (EGD-e) and a negative control strain (Δ*hly*; Table 4.3). Intracellular *L. monocytogenes* were enumerated at 1.5, 3, 5 and 7 h post infection. Intracellular *L. monocytogenes* numbers for each time point are reported as log (CFU recovered)/well; numbers indicate mean intracellular bacterial numbers for three independent replicates; error bars indicate standard deviations. Using one-way ANOVA, no statistically significant differences ($p>0.05$) were observed between EGD-e and any of the three mutant strains at any of the four time points in Caco2 (panel A) or J774 (panel B) cells.

A


log(CFUrecovered/well)
5.0
4.0
3.0
2.0
1.0
0.0
0
1
2
3
4
5
6
7
8
Time (hrs post inoculation)
EGDe
Δhly
Δisl30
Δisl16
Δisl25

B


log(CFUrecovered/well)
6.0
5.5
5.0
4.5
4.0
3.5
3.0
0
1
2
3
4
5
6
7
8
Time (hrs post inoculation)
EGDe
Δhly
ΔRD16
ΔRD25
ΔRD30

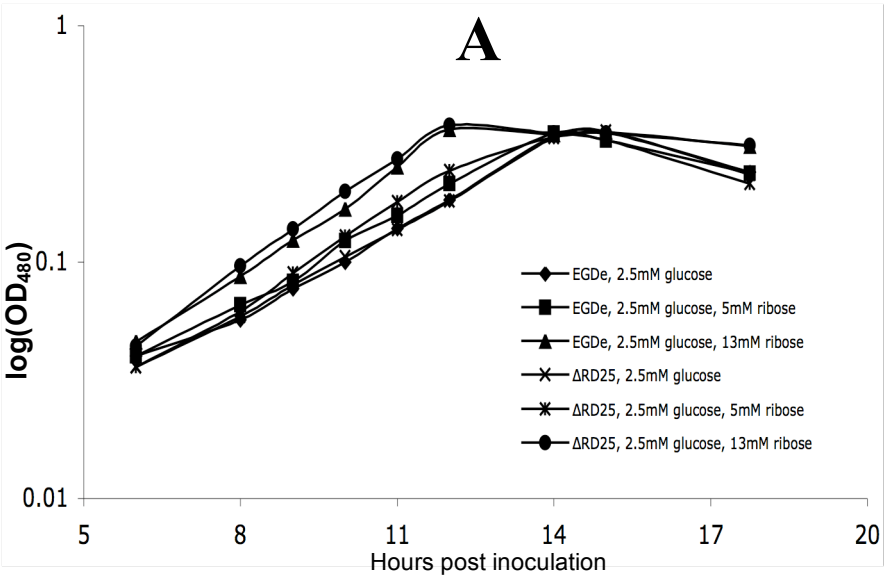
metabolic functions, we initially tested the ability of the DRD25 and DRD30 strains to grow in standard minimal media (MM), which was previously used to characterize *L. monocytogenes* (Phan-Thanh, 1997). In this minimal media, which contains arginine and glucose (as the sole carbohydrate), both mutant strains did not differ in their ability to grow at 37°C (with aeration), suggesting no generalize growth defects associated with the RD25 and RD30. Further experiments were thus conducted to specifically characterize the contributions of RD25 and RD 30 to growth and survival in arginine and ribose limiting conditions.

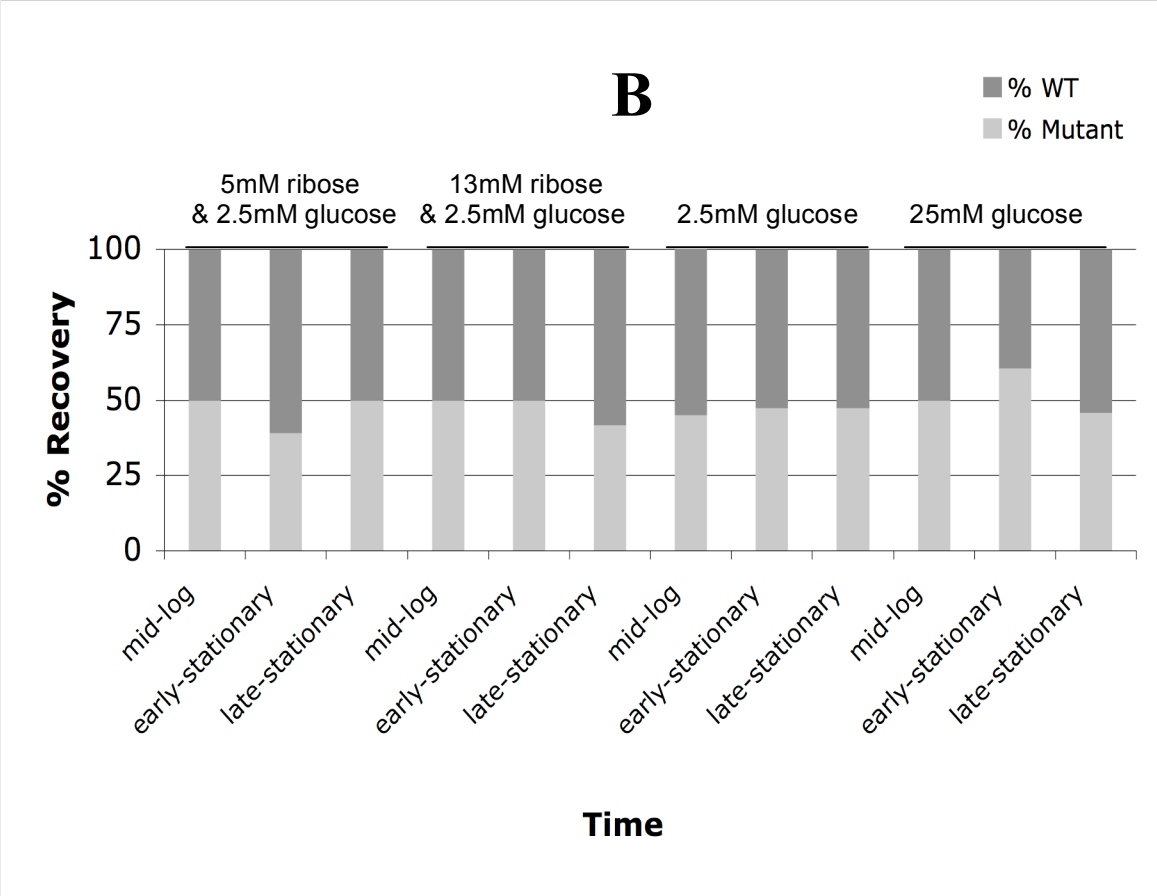
Initial experiments showed that our parent strain (*L. monocytogenes* EGD-e) could not grow in media containing ribose as the sole carbohydrate (data not shown), consistent with previous data for *L. monocytogenes* strain ScottA (Premaratne et al., 1991). Overall, both ΔRD25 and the parent strain EGD-e showed similar growth patterns in both MM with glucose and MM with glucose and ribose (Fig. 4.6A). The increased growth rate observed in MM with ribose likely reflects the fact that presence of ribose (in addition to glucose) reduces the need of bacteria for production of ribose sugars through the pentose phosphate pathway. This is consistent with observations that, in *L. monocytogenes* grown in MM that only contains glucose, glycolytic pathway genes are down-regulated while pentose phosphate pathways genes are upregulated (Joseph and Goebel, 2007). As this previously reported induction of pentose phosphate pathway genes in *L. monocytogenes* grown in minimal media with glucose as the sole carbohydrate supports a possible role for ribose metabolizing genes, such as those found in RD25 (i.e., genes that encode a putative ribulose-5-phosphate-3-epimerase and a putative ribose-5-phosphate isomerase; Table 4.4; Fig 4.3A), we conducted additional competition experiments to further define the importance for growth under ribose limiting conditions of genes located in RD25. In these experiments, growth of the ΔRD25 strain and its parent strain EGD-e, co-

inoculated into MM containing varying levels of ribose, was followed over time. Under all ribose concentrations tested and at all time points sampled (t=12, 20, and 23h, the ΔRD25 strain and EGD-e did not differ significantly in their numbers ($p>0.05$, chi-square test; Fig. 4.6B), suggesting that deletion of RD25 does not affect the ability of *L. monocytogenes* to grow in absence of ribose if glucose is available as a carbon source. Contributions of the genes in RD25 to growth under specific conditions (and carbohydrate availabilities) encountered in foods or other environments cannot be excluded though and it is tempting to speculate that these, and possible other lineage II specific metabolic genes, may contribute to an apparent enhanced ability of lineage II strains to grow, on average, to higher number in foods as compared lineage I strains (Chen, 2006) and to outcompete lineage I strains under certain conditions (Bruhn, 2005).

While RD30 encodes a number of genes involved in arginine metabolism, the ΔRD30 strain did not differ from its parent strain EGD-e in its ability to grow in minimal media with 0, 0.29, and 0.58 mM arginine (Fig. 4.7A). Even though EGD-e showed a higher growth rate in MM with 0.58 mM arginine (μ=0.279/h) as compared to MM without arginine (0.171/h) the final optical density reached by EGD-e was the same under all three arginine concentrations; by comparison arginine has been previously reported as an essential amino acid for *L. monocytogenes* ScottA (Premaratne et al., 1991). Since one ORF in RD30 appears to encode an ornithine carbamoyl-transferase (which converts citrulline into ornithine and carbamoyl-phosphate), we also evaluated the ability of the ΔRD30 strain to grow in MM with 0, 0.26, and 0.52 mM citrulline (a specialized amino acid derived from arginine, Fig. 4.3B). Under all citrulline concentrations, there was no apparent difference in growth between ΔRD30 strain and parent strain (Fig. 4.7A). In order to more sensitively determine whether deletion of RD30 affects the ability of *L. monocytogenes* to grow in

Figure 4.6 Panel A displays a representative growth curve of *L. monocytogenes* EGD-e and ΔRD25 in MM pH 7.1 with (i) 25 mM glucose and 0 mM ribose; (ii) 2.5mM glucose and 0 mM ribose; (iii) 2.5mM glucose and 5mM ribose, and (iv) 2.5mM glucose and 13mM ribose. This growth curve was only performed in duplicate as in neither trial were any differences in growth between EGD-e and the null mutant strain observed. A more sensitive competition experiment were subsequently performed in triplicate as detailed below.
Panel B shows data from a competitive growth experiment; this was performed by inoculating a 1:1 mixture of OD_{480}=0.4 cultures (grown in MM) of the ΔRD25 strain and the parent strain (EGD-e) as 1:100 dilution into MM containing (i) 25 mM glucose and 0 mM ribose; (ii) 2.5mM glucose and 0 mM ribose; (iii) 2.5mM glucose and 5mM ribose, and (iv) 2.5mM glucose and 13mM ribose. Bacteria were than plated on BHI after 12, 20, and 23 h of growth (representing mid-log, early stationary, and late stationary phase, respectively), followed by a PCR assay to differentiate mutant and parent strain colonies recovered. Results are shown as relative recovery of the ΔRD25 strain and the parent strain EGD-e; relative recovery was calculated based on characterization of 36 colonies (12 colonies from each of three independent replicates) at each time point. Recovery of the mutant and parent strains was not significantly different at any time point (Pearson's Chi-square, $p>0.05$).


A
1
0.1
0.01
log(OD480)
5
8
11
14
17
20
Hours post inoculation
EGDe, 2.5mM glucose
EGDe, 2.5mM glucose, 5mM ribose
EGDe, 2.5mM glucose, 13mM ribose
ΔRD25, 2.5mM glucose
ΔRD25, 2.5mM glucose, 5mM ribose
ΔRD25, 2.5mM glucose, 13mM ribose

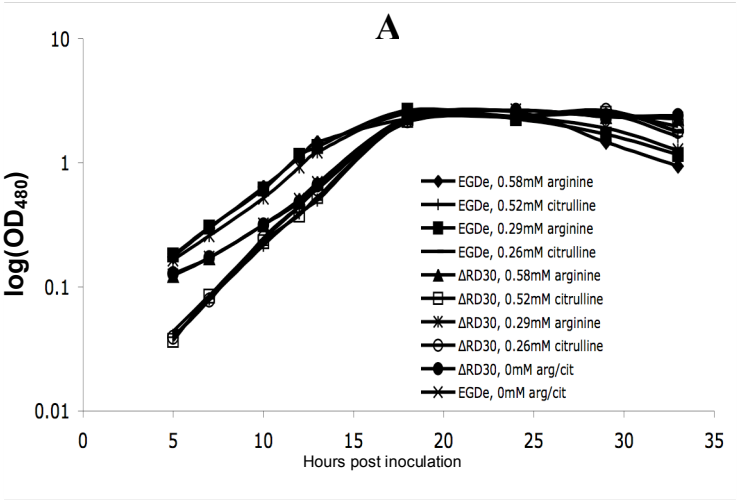

B
% WT
% Mutant
5mM ribose & 2.5mM glucose
13mM ribose & 2.5mM glucose
2.5mM glucose
25mM glucose
100
75
50
25
0
% Recovery
mid-log
early-stationary
late-stationary
Time

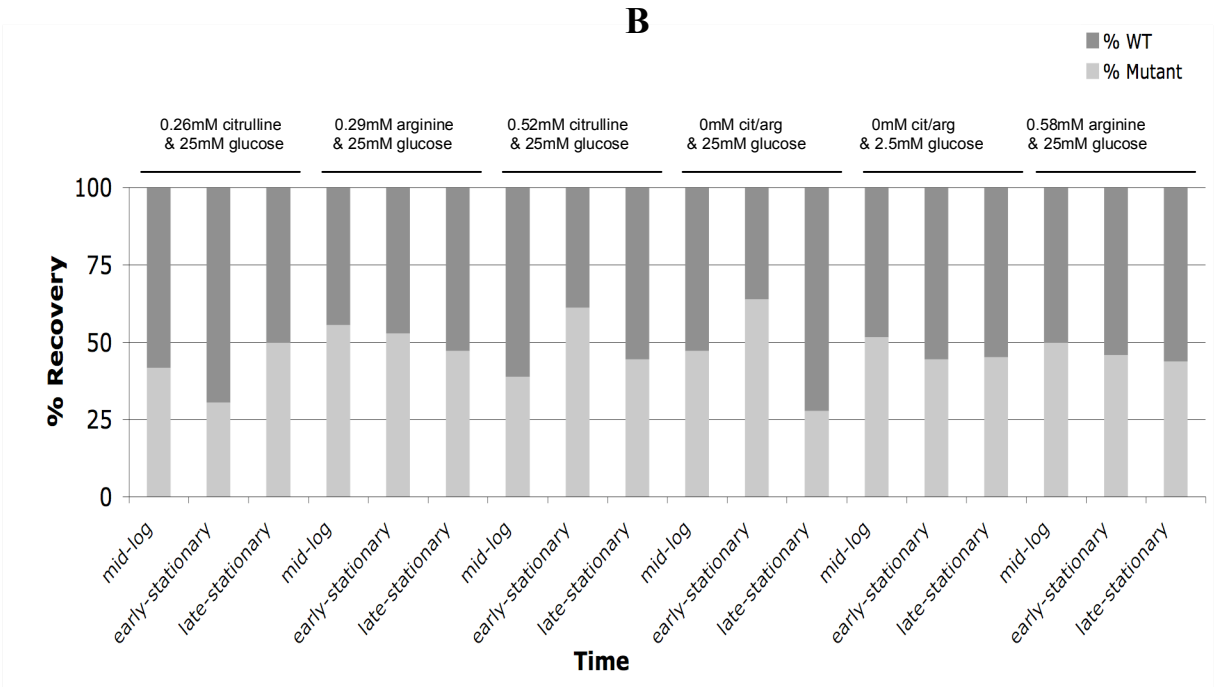
the presence of either arginine or citrulline, we also performed competitive growth experiments (between ΔRD30and EGD-e) in MM (pH 7.1) containing either (i) no arginine or citrulline, (ii) 0.29 or 0.58mM arginine, or (ii) 0.26 or 0.52 mM citrulline. Under all citrulline and arginine concentrations tested and at all time points sampled (t=12, 18, and 35h) the ΔRD30 strain and EGD-e did not differ significantly in their numbers ($p>0.05$, Chi-square test; Fig. 4.7B). Interestingly though, the *L. monocytogenes* ΔRD30 strain was recovered at significantly lower numbers than the parent strains EGD-e in the medium without arginine and citrulline at t=35 h (Fig 4.7B), suggesting some effect of the RD30 deletion on the ability of *L. monocytogenes* to grow and/or survive in the absence of arginine and citrulline. As the pH of the MM at t=35 had dropped to approx. pH 6.37, we hypothesized that the genes in RD30 may specifically contribute to the ability of *L. monocytogenes* to survive under acid stress in the absence of arginine.

To further test the hypothesis that genes in RD30 contribute to *L. monocytogenes* growth and/or survival under acidic conditions (in the absence of arginine), we performed separate growth as well as competition experiments with the ΔRD30 strain and EGD-e in pH 5.5 MM with 0, 0.29, and 0.58 mM arginine. When ΔRD30 and EGD-e were grown separately in pH5.5 MM with the different arginine concentrations, no apparent differences in growth were observed (Fig. 4.8). In the competition experiments in pH 5.5 MM with 0.58 mM arginine, the ΔRD30 strain and EGD-e did not differ significantly in their numbers ($p>0.05$, chi-square test; Fig. 4.9A). However, in the competition experiments in pH5.5 mM without arginine, significantly lower number of the ΔRD30 strains (as compared to EGD-e) were recovered at both the early- and late-stationary phase time points (Pearson's Chi-

Figure 4.7 Panel A displays a representative growth curve of the ΔRD30 strains and the *L. monocytogenes* parent strain EGD-e in MM pH 7.1 containing (i) 0.58 mM arginine;(ii) 0.29 mM arginine; (iii) 0.52 mM citrulline; (iv) 0.26 mM citrulline; and (v) 0 mM arginine and 0 mM citrulline. Note that several of the growth curve lines overlap making the graph appear to have fewer than ten lines. Only two independent growth curves were performed as both consistently showed no effect of the null mutation; more sensitive competition experiments were subsequently performed in triplicate as detailed below.

Panel B shows data from a competitive growth experiment; this was performed by inoculating a 1:1 mixture of OD_{480}=0.4 cultures (grown in MM) of the ΔRD30 strain and the parent strain (EGD-e) as 1:100 dilution into MM containing (i) 0.58 mM arginine;(ii) 0.29 mM arginine; (iii) 0.52 mM citrulline; (iv) 0.26 mM citrulline; and (v) 0 mM arginine and 0 mM citrulline. Bacteria were than plated on BHI after 12, 18, and 35 h of growth (representing mid-log, early stationary, and late stationary phase, respectively), followed by a PCR assay to differentiate mutant and parent strain colonies recovered. Results are shown as relative recovery of the ΔRD30 strain and the parent strain EGD-e; relative recovery was calculated based on characterization of 36 colonies (12 colonies from each of three independent replicates) at each time point. Recovery of the mutant and parent strains was not significantly different at any time point (Pearson's Chi-square, $p>0.05$), except for significantly lower recovery of the ΔRD30 strain in media without arginine or citrulline at the late-stationary phase time point (bar marked by an asterisk; Pearson's Chi-square, $p<0.05$).


A
10
1
0.1
0.01
log(OD480)
0
5
10
15
20
25
30
35
Hours post inoculation
EGDe, 0.58mM arginine
EGDe, 0.52mM citrulline
EGDe, 0.29mM arginine
EGDe, 0.26mM citrulline
ΔRD30, 0.58mM arginine
ΔRD30, 0.52mM citrulline
ΔRD30, 0.29mM arginine
ΔRD30, 0.26mM citrulline
ΔRD30, 0mM arg/cit
EGDe, 0mM arg/cit


B
% WT
% Mutant
0.26mM citrulline & 25mM glucose
0.29mM arginine & 25mM glucose
0.52mM citrulline & 25mM glucose
0mM cit/arg & 25mM glucose
0mM cit/arg & 2.5mM glucose
0.58mM arginine & 25mM glucose
% Recovery
100
75
50
25
0
mid-log
early-stationary
late-stationary
Time

square, $p<0.05$; Fig. 4.9C). In the competition experiments in pH5.5 mM with 0.29 mM arginine, significantly lower number of the ΔRD30 strains (as compared to EGD-e) were recovered at the early-stationary phase time point (Pearson's Chi-square, $p<0.05$; Fig. 4.9B). These data suggest that RD30 contributes to *L. monocytogenes* growth and survival at low pH in environments with no or limited arginine. This hypothesis is consistent with previous reports that genes involved in the glutamate decarboxylase [GAD] system are the major contributors to *L. monocytogenes* survival at low pH (Cotter et al., 2001A and B) and suggestions that genes connected to arginine catabolism may also contribute to stabilzing the cytoplasmic pH in *L. monocytogenes* (Gahan and Hill, 2005). Specifically, arginine catabolism can generate of two units of ammonia (which can sequester hydrogen ions) via the arginine deiminase (ADI) system, through (i) catabolism of arginine to citrulline and NH_3 and (ii) subsequent degradation of citrulline to ornithine and carbamoyl phosphate, followed by ATP-generating carbamoyl phosphate breakdown to NH_3, and CO_2 (both enzymes involved in step two are putatively encoded in RD30 [*lmo0036* and *lmo0039*]; Fig. 4.3B) (Arena et al., 2002). Given that *L. monocytogenes* EGD-e is able to synthesize its own arginine (as demonstrated by its growth in media without arginine), we hypothesize that during combined acid-stress and arginine-limited conditions the RD30 null mutant is less efficient at metabolizing the available endogenously produced arginine (to generate ammonia) and, thus, less capable of stabilizing cytoplasmic pH. Some previous research has linked the presence of arginine metabolism genes with acid resistance and/or virulence of other pathogens. For example, genome comparisons of two *Staphylococcus aureus* strains, one methicillin sensitive (MSSA) and one methicillin resistant (MRSA), found an arginine

Figure 4.8 This figure displays a representative growth curve of the ΔRD30 strain and the *L. monocytogenes* parent strain EGD-e in MM pH 5.5 containing (i) 0.58 mM arginine;(ii) 0.29 mM arginine; (iii) and 0 mM arginine. Note that several of the growth curve lines overlap making the graph appear to have fewer than six lines. Only two independent growth curves were performed as both consistently showed no effect of the null mutation; more sensitive competition experiments were subsequently performed in triplicate (Fig. 4.9).

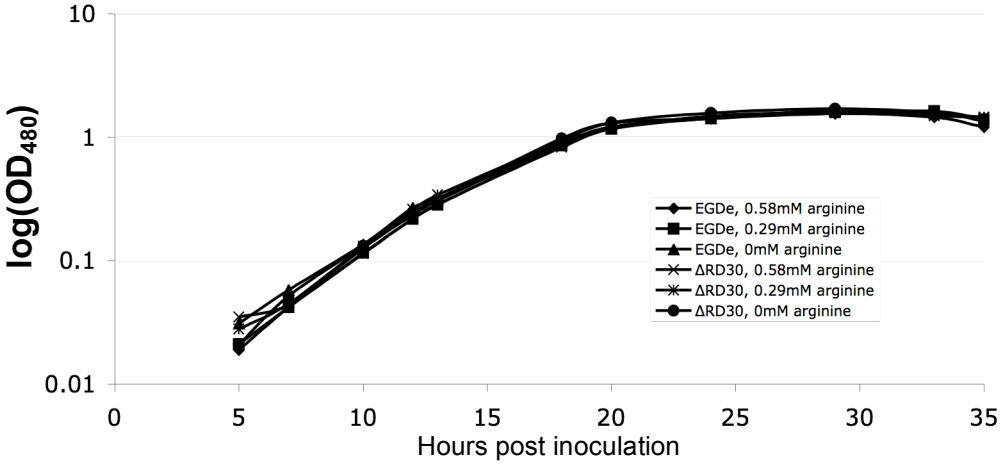
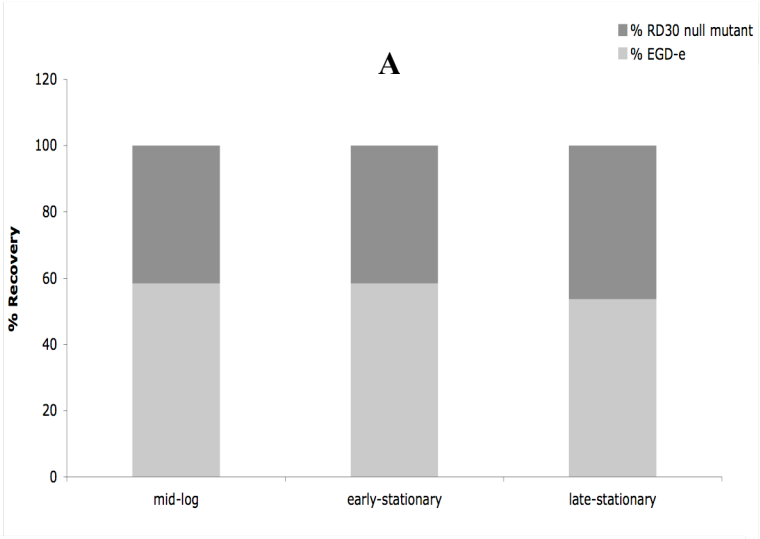
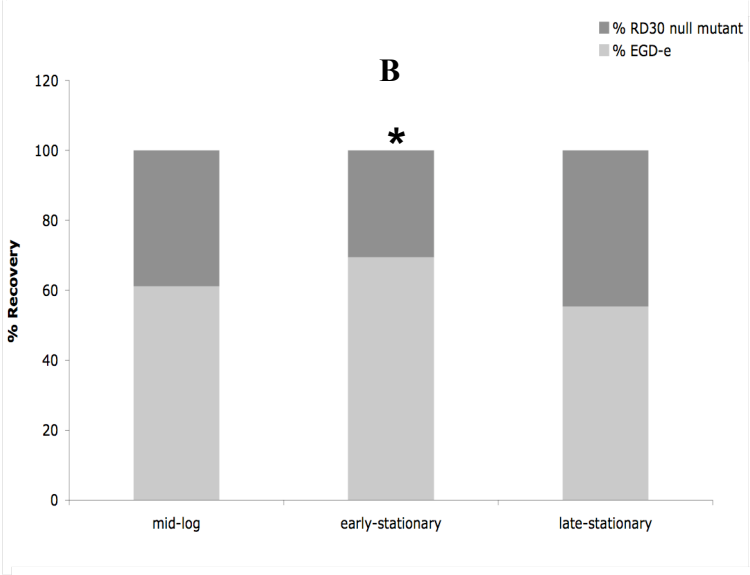
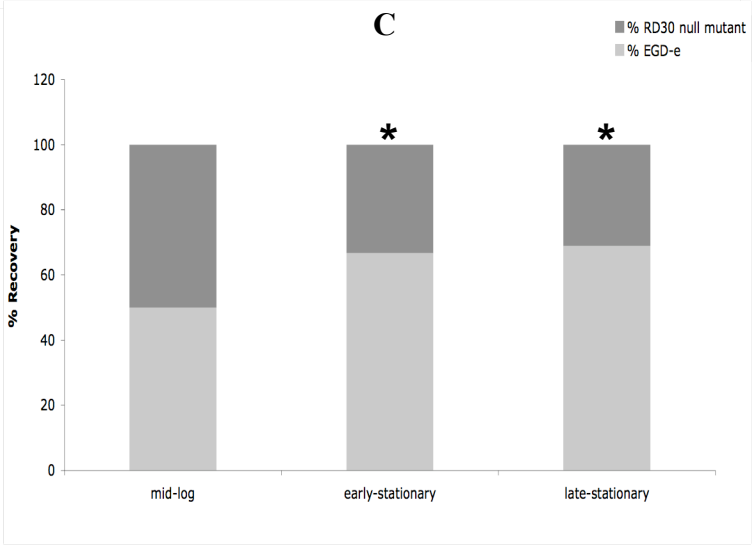

10
1
0.1
0.01
log(OD480)
0
5
10
15
20
25
30
35
Hours post inoculation
EGDe, 0.58mM arginine
EGDe, 0.29mM arginine
EGDe, 0mM arginine
ΔRD30, 0.58mM arginine
ΔRD30, 0.29mM arginine
ΔRD30, 0mM arginine

Figure 4.9 This competitive growth experiment was performed by inoculating a 1:1 mixture of OD_{480}=0.4 cultures (grown in MM) of the ΔRD30 strain and the parent strain (EGD-e) as 1:100 dilution into MM containing 0 (panel A), 0.29 (panel B), and 0.58 mM (panel C) arginine. Bacteria were than plated on BHI after 15, 29, and 35 h of growth (representing mid-log, early stationary, and late stationary phase, respectively), followed by a PCR assay to differentiate mutant and parent strain colonies recovered. PCR was performed using multiplex allele specific PCR with three primers (two primers external to the deleted genomic region and one located just internal to the 5' end of the region; see Table A4.2). This PCR yields a product of approx. 670 bp (with the two external primers) for the ΔRD30 strain and a product of approx. 400 bp (with the 5' external primer and the 5' internal primer) for the parent strain. Results are shown as relative recovery of the ΔRD30 strain and the parent strain EGD-e; relative recovery was calculated based on characterization of 36 colonies per media formulation per time point (12 colonies from each of three independent replicates) at each time point. Recovery of the mutant and parent strains was significantly different (bars marked with asterisks) at early and late stationary phase time points for MM with 0mM arginine (Pearson's Chi-square, $p<0.05$).


A
% RD30 null mutant
% EGD-e
% Recovery
120
100
80
60
40
20
0
mid-log
early-stationary
late-stationary


B
% RD30 null mutant
% EGD-e
% Recovery
120
100
80
60
40
20
0
*
mid-log
early-stationary
late-stationary


C
% RD30 null mutant
% EGD-e
% Recovery
120
100
80
60
40
20
0
*
*
mid-log
early-stationary
late-stationary

deiminase operon present in the MRSA genome but not in the MSSA genome (Highlander, 2007). In *Streptococcus suis*, researchers have noted that genes in the arginine deiminase system are considered factors that may contribute to virulence as these genes are present in outbreak associated strains (Jing, 2008) and have been linked to survival of *S. suis* in acidic environments (Gruening, 2006). Therefore, the presence of RD30 in *L. monocytogenes* may contribute to enhanced acid resistance under environmental stress conditions (e.g., in foods) or acid stress conditions encountered during infection of the host (e.g., gastric acid).

Conclusions

While a number of *L. monocytogenes* species specific genomic regions have previously been shown to include genes that contribute to virulence (e.g., *inlA* and *inlB*, the virulence genes in the main *Listeria* pathogenicity island [Vazquez-Boland, 2001]) and resistance to environmental stresses (e.g., *sigC*), a comprehensive understanding of the distribution and phenotypic contributions of species and lineage specific genomic regions still remains to be achieved. We thus combined *in silico* genome analyses with construction of large deletion mutants that completely removed each of three whole regions of difference (ranging from 6.5 to 5.8 kb). Surprisingly, expect for RD30, none of these mutants showed any apparent phenotypic effects in the assays used, suggesting that many of the main virulence associated phenotypes (e.g., invasion of human intestinal epithelial cells) may largely dependent on already identified virulence genes, with other species- and lineage-specific genes playing supportive or more subtle roles in *L. monocytogenes* virulence and transmission, including possible lineage specific virulence and transmission traits. Interestingly though, we did show that the genes in RD 30 contribute to acid resistance of *L. monocytogenes* in arginine restricted media, suggesting possible roles for these genes

in transmission (e.g., survival in low pH environments and foods) or virulence (e.g, by facilitating survival in at low pH host environments, such as the gastric environments or the intravacuolar environment.

APPENDIX

Table A4.1 Diversity set of *L. monocytogenes* and *L. innocua* isolates used to confirm specific distributions of selected RDs

FSL Isolate ID[a]	Lineage	Serotype	Ribotype	Description
L. monocytogenes isolates				
E1-119	I	1/2b	DUP 1042B	animal clinical
E1-123	II	1/2a	DUP 1039C	animal clinical
E1-124	I	4b	DUP 1038B	animal clinical
E1-125	I	4b	DUP 1042B	animal clinical
F2-086	III	4a	DUP 10142	human sporadic
F2-270	III	4a	DUP 18007A	human sporadic
F2-318	III	-	DUP 1059A	sheep
F2-373	II	1/2a	DUP 1039C	RTE meat product
F2-493	I	1/2a	DUP 1024A	ground beef
F2-501	III	4b	116-110-S-2	human sporadic
F2-590	II	1/2a	DUP 1057A	raw fish product
F2-601	I	4b	DUP 1042B	human clinical
F2-602	I	3b	DUP 1042A	human clinical
F2-634	II	1/2a	DUP 1053A	human sporadic
F2-637	I	4b	DUP 1044A	human sporadic
F2-639	II	1/2a	DUP 1039B	human sporadic
F2-640	II	1/2a	DUP 1039C	RTE meat product
F2-658	I	4b	DUP 1038B	human clinical
F2-661	I	4b	DUP 1042B	human sporadic
F2-663	II	1/2a	DUP 1054C	human sporadic
F2-666	I	1/2b	193-65-S-4	raw fish product
F2-667	I	1/2b	DUP 1042B	RTE meat product
F2-672	I	4b	DUP 1044A	human sporadic
F2-695	III	4a	DUP 1061A	human sporadic
F2-723	II	1/2a	DUP 1039C	raw meat product
F2-897	I	1/2b	DUP 1025A	RTE fish product
F2-898	I	1/2b	DUP 1052A	RTE meat product
J1-168	III	4a	116-110-S-2	human sporadic
J2-074	III	4c	DUP 10146	animal clinical
M1-001	III	4b	DUP 10142	animal clinical
N4-288	I	1/2b	DUP 1052A	animal clinical
N4-289	I	4b	DUP 1038B	animal clinical
N4-290	II	1/2a	DUP 1039C	animal clinical
N4-291	I	1/2b	DUP 16635B	animal clinical
N4-292	II	1/2a	DUP 1030B	animal clinical
N4-293	II	1/2a	DUP 1045A	animal clinical
R2-128	III	-	DUP 1061A	RTE bagged salad
R2-142	III	-	DUP 18036	RTE bagged salad
S4-019	II	-	DUP 1045A	pristine environment, soil
S4-295	II	-	DUP 1045B	pristine environment, soil
S4-304	II	-	DUP 1046A	vegetation, leaves
S4-497	II	-	DUP 1039C	soil
S4-766	II	-	DUP 1039E	vegetation, leaves
S4-821	II	-	DUP 1039C	pristine water
S4-880	II	-	DUP 1039C	vegetation, grass
S4-887	II	-	DUP 1039E	pristine water
S4-941	I	-	DUP 1038B	vegetation, leaves
S6-072	II	-	DUP 1045B	pristine water
L. innocua isolates				
S4-051	-	-	-	lake water
S6-022	-	-	-	urban sidewalk
S4-551	-	-	-	park mulch
S4-235	-	-	-	urban sidewalk
S6-035	-	-	-	urban puddle water

Table A4.2 Primers used for PCR screens and null mutant construction.

Target Region	Primer Name	Sequence (5'-3')[a,b]
RD16	SRM86 SOE-A[c]	*cgggatcc* tac ttt agc taa ttc ctcc
	SRM87 SOE-B	taa cta aat atg aaa ctg tcc
	SRM88 SOE-C	ggacagtttcatatttagtta aat tgt ttc aca gaa tgc tc
	SRM89 SOE-D	*gctctaga*ttt agc ata agc aat acc ca
	SRM94 exF[d]	ctg tta tgg ttt caa agg tt
	SRM95 exR	agt att gca gat gtt ttc gc
	SRM94 *L. innocua*-F[e]	ttattgtaaaaattttggacagt
	SRM95 *L. innocua*-R	ttcgttgacttttatttctctc
	JMB13 LII-F[f]	ttgcgtactcatttaagacg
	JMB16 LII-R	aaacaattcaatgaaaaagagc
	JMB17 LI-F	ttttacgcgaattttagggc
	JMB20 LI-R	gaaaaacagcttaaaatcagc
RD25	SRM90 SOE-A	*cgggatcc* att tcc gaa aaa ctg gag at
	SRM91 SOE-B	aaa agc atc aaa cat agt acg a
	SRM92 SOE-C	tcgtactatgtttgatgctttt aaa aaa gta att gaa act aat gga
	SRM93 SOE-D	*gctctaga* tgg aat gat tgc caa aac g c
	SRM96 exF	agt caa ata ttc gat ccc ag
	SRM97 exR	gat tga ttt ttt gtc cca tc
	SRM96 *L. innocua*-F	acttatagctttggcggc
	SRM97 *L. innocua*-R	taccaaactttatcaatcagtc
	JMB21 LII-F	tttggggatgagtcaggc
	JMB24 LII-R	aacgcagccaaacatatcc
RD30	SRM80 SOE-A	*g ggatcc* atg aag agt aaa aca aaa cca g
	SRM81 SOE-B	aga aaa atc ttt agt gtc gat
	SRM82 SOE-C	atc gac act aaa gat ttt tct att atg gag cta atc tta tac cag
	SRM83 SOE-D	*caa ctgcag* ggt aaa aat gga aat act acc
	SRM84 exF	ttc ctt tcc aaa ttt tag cgc
	SRM100 int1	tta ata gat tct ttc att ttg cg
	SRM85 exR	caa tat gta aaa tcg aac cat c
	SRM98 *L. innocua*-F	ttagctcatcacattgcaga
	SRM99 *L. innocua*-R	aagaacaaaatgatgtaagtcc
	JMB05 LII-F	gggtgattcggtctgttc
	JMB08 LII-R	ggtaataaaatccaccaggc
	JMB09 LI-F	gggtgattccgtctgttc
	JMB12 LI-R	taataaaatccaccaggcac

[a]Italicized regions at the beginning of a primer used in mutant making represent an enzyme cut site

[b]Underlined portions of primers used in mutant making indicate the complementary overlap of the SOE-C and SOE-B primers

[c]SOE-A/D indicate primers used in the null mutant making process.

[d]exF/R and int1 primers designate those used to differentiate the parent and mutant strains in competitive growth curve assays

[e]L. innocua-F/R primers were used to screen L. innocua isolates for presence of the genomic regions of interest

[f]LII-F/R primers were used to screen for the presence of the genomic regions of interest in a set of unsequenced *L. monocytogenes* lineage II isolates (LI-F/R primers screen lineage I isolates).

Table A4.3 Protein BLAST searches of the predicted protein sequences corresponding to each ORF of selected RDs.

RD	No. of ORF	Relevant Locus Tags [a]	Organism with Similar Genetic Loci	Relevant Locus Tags in Corresponding Organism	% aa identity per ORF
5	5	*lmo_0447-0448*	*Enterococcus faecium* DO	*Efae_1316-1317*	68, 60
7	1	*F-lmo_1974*	*Enterococcus faecalis* V583	*EF_2505*	37
11	1	*lmo_0842*	*Enterococcus faecalis* V583	*EF_2505*	42
12	1	*lmo_2576*	*Bacillus halodurans* C-125	*BH_2014*	40
22	5	*lmo_2733,2734,2735*	*Vibrio vulnificus* YJ016/CMCP6	*VV_1395-1393*	38,35,48
25	6	*lmo_0732-0734,0739*	*Lactococcus lactis subsp. cremoris* SK11	*LACR_0625-0628*	63,66,79,67
26	7	*F-lmo_0267-0269*	*Enterococcus faecalis* V583	*EF_1342-1345*	42,64,67
27	3	*lmo_0752,0754*	*Staphylococcus haemolyticus* JCSC1435	*SH_0474-0475*	57,55
28	7	*lmo_1030-1034,1036*	*Lactobacillus acidophilus* NCFM	*LBA_1488-1490, 1478,1480*	36,54,54,49 ,51
30	6	*lmo_0036-0041*	*Pediococcus pentosaceus* ATCC25745	*PEPE_1418-1423*	69,67,75,56 ,63,58
48	2	*lmo_2788-2789*	*Clostridium longisporum*	*abgF-abgG*	41,41
49	7	*lmo_1081-1083*	*Bacillus clausii* KSM-K16	*ABC_3689-3691*	78,66,68

[a] Locus tags refer to *L. monocytogenes* EGD-e or, for RDs not found in EGD-e, to *L. monocytogenes* F2365

REFERENCES

Amano, F. & Akamatsu, Y. (1991). A lipopolysaccharide (lps)-resistant mutant isolated from a macrophagelike cell line, j774.1, exhibits an altered activated-macrophage phenotype in response to LPS. Infect. Immun., 59, 2166-2174.

Arena, M. E. & Manca de Nadra, M. C. (2002). Comparative survey in Lactobacillus plantarum of the growth and metabolism of arginine and citrulline in different media. J. Agric. Food Chem., 50, 6497-6500.

Bielecki, J., Youngman, P., Connelly, P. & Portnoy, D. A. (1990). Bacillus subtilis expressing a haemolysin gene from Listeria monocytogenes can grow in mammalian cells. Nature, 345, 175-176.

Broad Institute of Harvard and MIT. Listeria monocytogenes sequencing project. (http://www.broad.mit.edu).

Bruhn, J. B., Vogel, B. F. & Gram, L. (2005). Bias in the Listeria monocytogenes enrichment procedure: Lineage 2 strains outcompete lineage 1 strains in University of Vermont selective enrichments. Appl. Environ. Microbiol., 71, 961-967.

Camilli, A., Portnoy, A. & Youngman, P. (1990). Insertional mutagenesis of Listeria monocytogenes with a novel tn917 derivative that allows direct cloning of DNA flanking transposon insertions. J. Bacteriol., 172, 3738-3744.

Chen, Y., Ross, W. H., Gray, M. J., Wiedmann, M., Whiting, R. C. & Scott, V. N. (2006). Attributing risk to Listeria monocytogenes subgroups: Dose response in relation to genetic lineages. Journal of Food Protection, 69,

Conte, M. P., Petrone, G., Di Biase, A. M., Ammendolia, M. G., Superti, F. & Seganti, L. (2000). Acid tolerance in Listeria monocytogenes influences invasiveness of enterocyte-like cells and macrophage-like cells. Microbial Pathogenesis, 29, 137-144.

Cotter, P. D., O'Reilly, K. & Hill, C. (2001). Role of the glutamate decarboxylase acid resistance system in the survival of Listeria monocytogenes lo28 in low ph foods. J Food Prot., 64, 1362-1368.

Cotter, P. D., Ryan, S., Gahan, C. G. & Hill, C. (2005). Presence of gadD1 glutamate decarboxylase in selected Listeria monocytogenes strains is associated with an ability to grow at low ph. Appl Environ Microbiol., 71, 2832-2839.

Darling, A. C. E., Mau, B., Blattner, F. R. & Perna, N. T. (2004). Mauve: Multiple alignment of conserved genomic sequence with rearrangements. Genome Res, 14, 1394–1403.

Doumith, M., Cazalet, C., Simoes, N., Frangeul, L., Jacquet, C., Kunst, F., Martin, P., Cossart, P., Glaser, P. & Buchrieser, C. (2004). New aspects regarding evolution and virulence of Listeria monocytogenes revealed by comparative genomic and DNA arrays. Infect Immun., 72, 1072-1083.

Dussurget, O., Cabanes, D., Dehoux, P., Lecuit, M., Buchrieser, C., Glaser, P. & Cossart, P. (2002). Listeria monocytogenes bile salt hydrolase is a prfa-regulated virulence factor involved in the intestinal and hepatic phases of listeriosis. Mol Microbiol, 45, 1095-1106.

Gahan, C. G. M. & Hill, C. (2005). Gastrointestinal phase of Listeria monocytogenes infection. J Appl Microbiol, 98, 1345-1353.

Garner, M., Njaa, B., Wiedmann, M. & Boor, K. (2006). Sigma b contributes to Listeria monocytogenes gastrointestinal infection but not to systemic spread in the guinea pig infection model. Infect Immun., 74, 876-886.

Glaser, P., Frangeul, L., Buchrieser, C., Rusniok, C., Amend, A., Baquero, F., Berche, P., Bloecker, H., Brandt, P., Chakraborty, T., Charbit, A., Chetouani, F., Couve, E., de Daruvar, A., Dehoux, P., Domann, E., Dominguez-Bernal, G., Duchaud, E., Durant, L., Dussurget, O., Entian, K., Fsihi, H., Garcia-del Portillo, F., Garrido, P., Gautier, L., Goebel, W., Gomez-Lopez, N., Hain, T., Hauf, J., Jackson, D., Jones, L., Kaerst, U., Kreft, J., Kuhn, M., Kunst, F., Kurapkat, G., Madueno, E., Maitournam, A., Vicente, J., Ng, E., Nedjari, H., Nordsiek, G., Novella, S., de Pablos, B., Perez-Diaz, J., Purcell, R., Remmel, B., Rose, M., Schlueter, T., Simoes, N., Tierrez, A., Vazquez-Boland, J., Voss, H., Wehland, J. & Cossart, P. (2001). Comparative genomics of Listeria species. Science, 294, 849-852.

Goetz, M., Bubert, A., Wang, G., Chico-Calero, I., Vazquez-Boland, J.-A., Beck, M., Slaghuis, J., Szalay, A. A. & Goebel, W. (2001). Microinjection and growth of bacteria in the cytosol of mammalian host cells. PNAS, 98, 12221-12226.

Gray, M., Zadoks, R., Fortes, E., Dogan, B., Cai, S., Chen, Y., Scott, V., Gombas, D., Boor, K. & Wiedmann, M. (2004). Listeria monocytogenes isolates from foods and humans form distinct but overlapping populations. Appl Environ Microbiol., 70, 5833-5841.

Gruening, P., Fulde, M., Valentin-Weigand, P. & Goethe, R. (2006). Structure, regulation, and putative function of the arginine deiminase system of Streptococcus suis. J Bacteriol, 188, 361-369.

Hain, T., Steinweg, C., Kuenne, C. T., Billion, A., Ghai, R., Chatterjee, S. S., Domann, E., Karst, U., Goesmann, A., Bekel, T., Bartels, D., Kaiser, O., Meyer, F., Puhler, A., Weisshaar, B., Wehland, J., Liang, C., Dandekar, T., Lampidis, R., Kreft, J., Goebel, W. & Chakraborty, T. (2006). Whole-genome sequence of Listeria welshimeri reveals common steps in genome reduction with Listeria innocua as compared to Listeria monocytogenes. J. Bacteriol., 188, 7405-7415.

Highlander, S., Hulten, K., Qin, X., Jiang, H., Yerrapragada, S., Mason, E., Shang, Y., Williams, T., Fortunov, R., Liu, Y., Igboeli, O., Petrosino, J., Tirumalai, M., Uzman, A., Fox, G., Cardenas, A., Muzny, D., Hemphill, L., Ding, Y., Dugan, S., Blyth, P., Buhay, C., Dinh, H., Hawes, A., Holder, M., Kovar, C., Lee, S., Liu, W., Nazareth, L., Wang, Q., Zhou, J., Kaplan, S. & Weinstock, G. (2007). Subtle genetic changes enhance virulence of methicillin resistant and sensitive Staphylococcus aureus. BMC Microbiology, 7, 99.

Horton, R., Cai, Z., Ho, S. & Pease, L. (1990). Gene splicing by overlap extension: Tailor-made genes using the polymerase chain reaction. Biotechniques, 8, 528-535.

Jeffers, G. T., Bruce, J. L., McDonough, P. L., Scarlett, J., Boor, K. J. & Wiedmann, M. (2001). Comparative genetic characterization of Listeria monocytogenes isolates from human and animal listeriosis cases. Microbiology, 147, 1095–1104.

Jing, H. B., Yuan, J., Wang, J., Yuan, Y., Zhu, L., Liu, X. K., Zheng, Y. L., Wei, K. H., Zhang, X. M., Geng, H. R., Duan, Q., Feng, S. Z., Yang, R. F., Cao, W. C., Wang, H. L. & Jiang, Y. Q. (2008). Proteome analysis of Streptococcus suis serotype 2. Proteomics, 8, 333-349.

Joseph, B. & Goebel, W. (2007). Life of Listeria monocytogenes in the host cells' cytosol. Microbes Infect 9, 1188-1195.

Joseph, B., Przybilla, K., Stuhler, C., Schauer, K., Slaghuis, J., Fuchs, T. M. &

Goebel, W. (2006). Identification of Listeria monocytogenes genes contributing to intracellular replication by expression profiling and mutant screening. J Bacteriol, 188, 556-568.

Kanehisa, M., Goto, S., Hattori, M., Aoki-Kinoshita, K. F., Itoh, M., Kawashima, S., Katayama, T., Araki, M. & Hirakawa, M. (2006). From genomics to chemical genomics: New developments in kegg. Nucl. Acids Res., 34, D354-357.

Kim, H., Boor, K. & Marquis, H. (2004). Listeria monocytogenes sigmaB contributes to invasion of human intestinal epithelial cells. Infect Immun., 72, 7374-7378.

Klarsfeld, A., Goossens, P. L. & Cossart, P. (1994). Five Listeria monocytogenes genes preferentially expressed in infected mammalian cells: plcA, purH, purD, pyrE and an arginine ABC transporter gene, arpJ. Mol Microbiol, 13, 585-597.

Klose, K. E. & Mekalanos, J. J. (1997). Simultaneous prevention of glutamine synthesis and high-affinity transport attenuates Salmonella Typhimurium virulence. Infect. Immun., 65, 587-596.

Nelson, K., Fouts, D., Mongodin, E., Ravel, J., DeBoy, R., Kolonay, J., Rasko, D., Angiuoli, S., Gill, S., Paulsen, I., Peterson, J., White, O., Nelson, W., Nierman, W., Beanan, M., Brinkac, L., Daugherty, S., Dodson, R., Durkin, A., Madupu, R., Haft, D., Selengut, J., Van Aken, S., Khouri, H., Fedorova, N., Forberger, H., Tran, B., Kathariou, S., Wonderling, L., Uhlich, G., Bayles, D., Luchansky, J. & Fraser, C. (2004). Whole genome comparisons of serotype 4b and 1/2a strains of the food-borne pathogen Listeria monocytogenes reveal new insights into the core genome components of this species. Nuc Acids Res, 32, 2386-2395.

Pell, A. N. (1997). Manure and microbes: Public and animal health problem? J Dairy Sci, 80, 2673-2681.

Phan-Thanh, L. & Gormon, T. (1997). A chemically defined minimal medium for the optimal culture of Listeria. Int J Food Microbiol., 35, 91-95.

Piffaretti, J.-C., Kressebuch, H., Aeschbacher, M., Bille, J., Bannerman, E., Musser, J. M., Selander, R. K. & Rocourt, J. (1989). Genetic characterization of clones of the bacterium Listeria monocytogenes causing epidemic disease. PNAS, 86, 3818-3822.

Premaratne, R. J., Lin, W. J. & Johnson, E. A. (1991). Development of an improved

chemically defined minimal medium for Listeria monocytogenes. Appl Environ Microbiol, 57, 3046-3048.

Rasmussen, O. F., Skouboe, P., Dons, L., Rossen, L. & Olsen, J. E. (1995). Listeria monocytogenes exists in at least three evolutionary lines: Evidence from flagellin, invasive associated protein and listeriolysin o genes. Microbiology, 141, 2053-2061.

Reid, S. D., Green, N. M., Sylva, G. L., Voyich, J. M., Stenseth, E. T., DeLeo, F. R., Palzkill, T., Low, D. E., Hill, H. R. & Musser, J. M. (2002). Postgenomic analysis of four novel antigens of group a streptococcus: Growth phase-dependent gene transcription and human serologic response. J. Bacteriol., 184, 6316-6324.

Sabet, C., Lecuit, M., Cabanes, D., Cossart, P. & Bierne, H. (2005). LPXTG protein InlJ, a newly identified internalin involved in Listeria monocytogenes virulence. Infect Immun., 73, 6912–6922.

Sauders, B. (2005). Molecular epidemiology, diversity, and distribution, and ecology of Listeria. Ph.D. thesis, Cornell University, Ithaca, NY.

Smith, K. & Youngman, P. (1992). Use of a new integrational vector to investigate compartment-specific expression of the bacillus subtilis spoiim gene. Biochemie, 74, 705-711.

Stothard, P. & Wishart, D. S. (2005). Circular genome visualization and exploration using cgview. Bioinformatics, 21, 537-539.

Swaminathan, B. & Gerner-Smidt, P. (2007). The epidemiology of human listeriosis. Microbes Infect, 9, 1236-1243.

Vazquez-Boland, J. A., Dominguez-Bernal, G., Gonzalez-Zorn, B., Kreft, J. & Goebel, W. (2001). Pathogenicity islands and virulence evolution in Listeria. Microbes Infect, 3, 571-584.

Wiedmann, M., Bruce, J. L., Keating, C., Johnson, A. E., McDonough, P. L. & Batt, C. A. (1997). Ribotypes and virulence gene polymorphisms suggest three distinct listeria monocytogenes lineages with differences in pathogenic potential. Infect Immun., 65, 2707-2716.

Zhang, C., Nietfeldt, J., Zhang, M. & Benson, A. K. (2005). Functional consequences of genome evolution in listeria monocytogenes: The lmo0423 and lmo0422 genes encode {sigma}c and lstr, a lineage II-specific heat shock system. J. Bacteriol, 187, 7243-7253.

Zhang, C., Zhang, M., Ju, J., Nietfeldt, J., Wise, J., Terry, P. M., Olson, M., Kachman, S. D., Wiedmann, M., Samadpour, M. & Benson, A. K. (2003). Genome diversification in phylogenetic lineages I and II of Listeria monocytogenes: Identification of segments unique to lineage ii populations. J Bacteriol., 185, 5573-5584.

CHAPTER 5

SUMMARY AND CONCLUSIONS

Listeria monocytogenes can cause serious foodborne disease in certain human populations. *L. monocytogenes* is widely distributed in nature and because of its ability to grow and persist under a variety of conditions, it is difficult to prevent *its* contamination of and growth in foods. To determine *L. monocytogenes* mechanisms for persistence in nature (e.g. soil or silage) and food environments (e.g. RTE meats and produce) as well as clarify its mechanisms of virulence, more research investigating the inherant properties of *L. monocytogenes* and the differences among its lineages is necessary.

In the second chapter of this dissertation, *L. monocytogenes* ability to attach to and persist on the plant model *Arabidopsis thaliana* was investigated. This research determined that *L. monocytogenes* strains were well adapted to attach to and even proliferate on *A. thaliana* leaves, regardless of lineage. These findings were consistent with previous research of *L. monocytogenes* interactions with specific food-plants. This research sets the stage for the development of a model system for studying *L. monocytogenes* persistence on plant-derived foods, using a well-characterized plant model with a sequenced genome.

Chapters three and four encompass research devoted to exploring the function of unique genes and regions of the *L. monocytogenes* genome. As it has been commonly noted that *L. monocytogenes* strains are not equally virulent despite the fact that they all contain the main virulence gene cluster, the research in chapers three and four

sought to better define *L. monocytogenes* virulome. None of the genes or regions investigated appeared to contribute directly to *L. monocytogenes* virulence, as evaluated using tissue culture models of infection and intracellular growth. Therefore, this research indicates that it is likely that the differing virulence potential among the *L. monocytogenes* strains does not result from the presence or absence of unique virulence factors. Rather, it is hypothesized here that the differing virulence potential is a consequence of sequence divergence, possibly contributing to differences among the lineages in stress-response and virulence gene regulation perhaps affecting the potency of the known virulence factors. The comparative genomic analyses and functional characterizations described in chapters three and four may, however, be useful for application to future studies of *L. monocytognes* virulence that focus on detecting different patterns of gene regulation or presence and distribution of specific virulence factor alleles.

Future research efforts should concentrate on the study of *L. monocytogenes* strains associated with human clinical cases using comparative and functional genomic analyses to further the understanding of *L. monocytogenes* virulence potential, facilitating refinement of risk assessments and improving food safety.