

## Abstract

YUE, LILI. Gene cassettes unique to epidemic-associated lineages of *Listeria monocytogenes* serotype 4b. (Under the direction of Dr. Sophia Kathariou)

*Listeria monocytogenes* is responsible for severe foodborne infections, with high mortality and morbidity. A clonal group of *L. monocytogenes* strains of serotype 4b designated Epidemic Clone I (ECI) has been responsible for numerous outbreaks of foodborne illness in the United States and elsewhere. A number of genes and gene cassettes unique to ECI strains have been identified through molecular approaches, and through the recently completed genome sequencing on ECI strain. However, the functions of most of these ECI specific sequences remain unclear. The purpose of this research has been to perform a comparative genomic analysis of the regions harboring four ECI specific gene cassettes, designated region 144, 133, 17B and 85, among five different *Listeria* strains for which genome sequencing projects have been undertaken. The strains included ECI strain F2365 (serotype 4b), ECII strain H7858 (serotype 4b), serotype 1/2a strains EGD-e and F6854, and *Listeria innocua* CLIP 11262. The genomic organization of the region harboring these cassettes in ECI was investigated in the different genomes, and transcriptional analysis by reverse-transcriptase polymerase chain reaction was pursued with three regions, 144, 133 and 17B. The comparative genomic organization data revealed that all four ECI specific regions have features typical of genomic islands (GEIs), being present in the genome of ECI strains but absent from the genome of other serotype 4b strains, suggesting horizontal insertion / deletion events and possible roles in pathogenicity and metabolism of the organism. Transcriptional data suggested that the six ECI specific open reading frames in region 133 were co-transcribed as a unit separate from adjacent genes which were highly conserved among different strains. To obtain information on the possible functions of the ECI specific genes and gene cassettes, a mutational approach was pursued. A deletion mutant of the ECI specific sequence in region 133 was constructed in two different ECI strains, and the mutants were

characterized bacteriologically and in terms of phenotypic microarrays. The deletion mutant F2365  $\Delta$ 133 was characterized in terms of basic bacteriological features including hemolytic activity, phage susceptibility, motility, cell morphology, growth at 37°C, 25°C and 4°C, bacitracin resistance, surface antigen detection, and with a panel of phenotypic microarrays. The parental strain F2365 was observed to grow better than the deletion mutant F2365  $\Delta$ 133 at 4°C, and the Phenotypic Microarrays identified certain differences in substrate utilization between the mutant and the wild type parental strain. The findings suggest that the ECI specific cassette in region 133 may contribute to bacterial growth at cold temperature and to the metabolism of certain carbon and nitrogen sources. Future studies employing animal and cell culture models will be needed to evaluate the possible impact of the ECI specific cassette in region 133 in virulence and pathogenesis of the bacteria.

**Gene cassettes unique to epidemic-associated lineages of  
*Listeria monocytogenes* serotype 4b**

**By**

**Lili Yue**

A thesis submitted to the Graduate Faculty of  
North Carolina State University  
in partial fulfillment of  
the requirements for the Degree of  
Master of Science

**FOOD SCIENCE**

Raleigh

2005

**APPROVED BY:**

---

**Dr. Eric Miller**

---

**Dr. Jenny Xiang**

---

Chair of Advisory Committee  
**Dr. Sophia Kathariou**

## **Biography**

Lili Yue was born in Beijing, China in 1977. She received her Bachelor degree in Food Engineering from Beijing Agricultural College in 2000. She came to the United State with her husband in 2002. She came to College of Agriculture and Life Sciences, North Carolina State University in 2003. She will receive her Master of Science in Food Science in fall 2005.

## **Acknowledgements**

I would like to acknowledge Dr. Sophia Kathariou for giving me the opportunity to pursue my Master's under his guidance. I appreciate the support and encouragement that Dr. Kathariou has given me through out all the years that I have known her. It has been a true honor to work with her through my graduate career. I would like to thank Dr. Eric Miller and Dr. Jenny Xiang for being on my committee and giving me their precious advice in my research. I would also like thank USDA for their financial support through my research.

I am very grateful to Dr. Driss Elhanafi for his help and discussion in my research, to all of other members in the lab and all my friends for their friendship and help.

Finally, I would like to thank my family for their support, and I would especially like to thank my husband for his love and support.

<b>Table of Contents</b>	<b>Page</b>
<b>List of Tables.....</b>	<b>viii</b>
<b>List of Figures.....</b>	<b>ix</b>
<b>CHAPTER I Literature Review .....</b>	<b>1</b>
1.1 Introduction.....	2
1.2 Detection methods.....	3
1.3 Pathogenic <i>Listeria monocytogenes</i> Vs. nonpathogenic <i>Listeria innocua</i> .....	4
1.4 Serotypes.....	6
1.5 Mobile genetic elements in the <i>L. monocytogenes</i> genome .....	9
1.5.1 Insertion elements (IS).....	9
1.5.2 Clustered regularly interspaced palindromic repeats (CRISPRs) .....	10
1.5.3 Prophages .....	10
1.5.4 Plasmids .....	10
1.5.5 Role of plasmids and transposons in antibiotic and heavy metal resistance in <i>Listeria</i> .....	11
1.6 Epidemic clones of <i>L. monocytogenes</i> .....	12
1.6.1 Epidemic clone I (ECI) .....	12
1.6.2 Epidemic clone II.....	14
1.6.3 Epidemic clone III.....	15
1.7 References.....	15
<b>CHAPTER II Genomic organization and transcriptional analysis of ECI specific cassettes 144, 133, 17B and 85 .....</b>	<b>25</b>
2.1 Abstract .....	26
2.2 Introduction.....	28
2.3 Materials and methods .....	29
2.3.1 Nucleotide sequence analysis for cassettes 144, 133, 17B and 85 .....	29

2.3.2 Isolation of genomic DNA of <i>Listeria</i> .....	29
2.3.3 Southern Blots.....	30
2.3.3.1 Cultures used in Southern Blots.....	30
2.3.3.2 Probe Labeling with DIG-dUTP (Roche) .....	30
2.3.3.3 Digestion of genomic DNA with restriction enzyme EcoRI and agarose gel electrophoresis .	31
2.3.3.4 Southern transfer.....	31
2.3.3.5 Hybridization.....	32
2.3.3.6 Detection .....	32
2.3.4 RNA Extraction.....	33
2.3.5 Reverse transcriptase-polymerase chain reaction (RT-PCR).....	34
2.3.5.1 Prepare RNA target and primer.....	34
2.3.5.2 Prepare reverse transcriptase mix.....	34
2.3.5.3 Reverse Transcription.....	34
2.3.5.4 PCR amplification.....	34
2.4 Results.....	35
2.4.1 Genomic organization of ECI specific regions and comparative genomic analysis among different <i>Listeria</i> genome. ....	37
2.4.2 G+C content.....	43
2.4.3 Southern blots .....	45
2.4.4 RT-PCR analysis .....	52
2.4.4.1 RT-PCR results for region 144 .....	52
2.4.4.2 RT-PCR results for region 133 .....	53
2.4.4.3 RT-PCR analysis of region 17B .....	55
2.5 Discussion.....	58
2.6 Reference: .....	63

## Chapter III Construction of deletion mutant lacking the ECI specific gene

<b>cassette in region 133 .....</b>	<b>67</b>
3.1 Abstract .....	68
3.2 Introduction.....	70
3.3 Materials and Methods.....	72
3.3.1 Bacterial strains and growth media .....	72
3.3.2 Mutant construction .....	72
3.3.2.1 Plasmid construction.....	74
3.3.2.1. (a) Enzyme restriction and ligation .....	74
3.3.2.1. (b) Electroporation .....	74
3.3.2.1. (c) Screen positive transformants.....	75
3.3.2.2 Conjugation of E.coli SM1076 harboring the recombinant plasmid and ECI strain F2381L- PHI-R .....	76
3.3.2.3 Isolation of integrants and recombinants .....	76
3.3.2.4 PCR screening and southern blots to identify deletion mutants.....	77
3.3.2.5 Sequencing confirmation of deletion mutants.....	77
3.3.2.6 Construction of ECI deletion mutant in strain F2365 .....	78
3.3.3 Mutant F2365 $\Delta$ 133 characterization identification.....	78
3.3.3.1 Hemolytic activity.....	78
3.3.3.2 Phage invasion .....	78
3.3.3.3 Motility.....	79
3.3.3.4 Growth at 37°C, 25°C and 4°C .....	79
3.3.3.5 Cell morphology .....	79
3.3.3.6 Bacitracin resistance.....	79
3.3.3.7 Phenotypic microarrays.....	80
3.3.3.8 Colony immunoblotting with monoclonal antibody MAb c74.22. ....	81

3.4 Result .....	82
3.4.1 Mutant construction result.....	82
3.4.2 Hemolytic activity.....	86
3.4.3 Phage invasion .....	86
3.4.4 Motility .....	86
3.4.5 Cell morphology .....	87
3.4.6 Growth at 37°C, 25°C and 4°C .....	87
3.4.7 Bacitracin resistance .....	88
3.4.8 Phenotypic Microarrays .....	89
3.4.9 Colony immunoblotting with monoclonal antibody MAb c74.22 .....	90
3.5 Discussion .....	95
3.6 References.....	96
<b>Appendix A PCR reaction system in Reverse Transcripton PCR.....</b>	<b>98</b>
<b>Appendix B Colony PCR reaction system.....</b>	<b>99</b>
<b>Appendix C Alignment in 133 region among <i>L. monocytogenes</i> strains F2365, EGD-e and <i>L.</i></b>	
<b><i>innouca</i>.....</b>	<b>100</b>
<b>Appendix D Restriction system.....</b>	<b>106</b>
<b>Appendix E Ligation system.....</b>	<b>106</b>
<b>Appendix F Working primers used in this study.....</b>	<b>107</b>

<b>List of Tables</b>	<b>Page</b>
<b>CHAPTER I Literature Review .....</b>	<b>1</b>
Table 1. Genomic summary of five <i>Listeria</i> strains. ....	<b>8</b>
<b>CHAPTER II Genomic organization and transcriptional analysis of ECI specific cassettes 144, 133, 17B and 85 .....</b>	<b>25</b>
Table 1. <i>Listeria</i> strains used in southern blotting with probes derived from different ORFs in the ECI specific regions 133 and 17B. ....	30
Table 2. Summary of southern blot results of <i>L. monocytogenes</i> and <i>L. innocua</i> with probes from three ORFs in region 133. ....	45
Table 3. Summary of southern blot results of <i>L. monocytogenes</i> and <i>L. innocua</i> with probes from ORFs in region 17B and flanking ORFs. ....	46
<b>Chapter III Construction of deletion mutant lacking the ECI specific gene cassette in region 133 .....</b>	<b>67</b>
Table 1. Primers for plasmid construction. Restriction enzyme sites were underlined. ....	73
Table 2. The growth of F2365 wild type and F2365 $\Delta$ 133 in series bacitacin concentration at 37°C and 4°C. ....	89
Table 3. PM1 carbon sources with significant value difference between F2365 wild type and mutant F2365 $\Delta$ 133. ....	91
Table 4. PM2 carbon sources with significant value difference between F2365 wild type and mutant F2365 $\Delta$ 133. ....	92
Table 5. PM3 nitrogen sources with significant value difference between F2365 wild type and mutant F2365 $\Delta$ 133. ....	93

## List of Figures

Page

### CHAPTER I Literature Review.....1

### CHAPTER II Genomic organization and transcriptional analysis of ECI

### specific cassettes 144, 133, 17B and 85.....25

Fig. 1 Location of the ECI specific cassettes 144, 133, 17B, and 85 in the genome of the ECI strain F2365. ....	36
Fig. 2 Genomic region harboring fragment 144 in F2365 (ECI) and other Listeria genomes. ....	38
Fig. 3.1 Genomic region harboring fragment 133 in F2365 (ECI) and other Listeria genome.....	39
Fig. 3.2 Secondary structure of palindromic repeat in region 133. $\Delta G = -20.4$ .....	39
Fig. 4 Genomic region harboring fragment 17B in F2365 (ECI) and other Listeria genomes.....	41
Fig. 5 Genomic region harboring fragment 85 in F2365 (ECI) and other Listeria genomes. ....	42
Fig. 6 ECI strain F2365 144 region G+C content map. Gray arrow: ECI specific ORF.....	43
Fig. 7 ECI strain F2365 133 region G+C content map. Gray arrow: ECI specific ORFs .....	43
Fig. 8 ECI strain F2365 17B region G+C content map. Gray arrow: ECI specific ORFs .....	44
Fig. 9 ECI strain F2365 85 region G+C content map. Gray arrow: ECI specific ORFs.....	44
Fig. 10.1 Southern blot of EcoRI-digested DNA from different strains, probed with ORF2630 (region 133). ....	46
Fig. 10.2 Southern blot of EcoRI-digested DNA from different strains, probed with ORF2628 (region 133). ....	47
Fig. 10.3 Southern blot of EcoRI-digested DNA from different strains, probed with ORF2627 (region 133). ....	47
Fig. 11.1 Southern blot of EcoRI-digested DNA from different strains, probed with ORF2800 (region 17B) .....	48
Fig. 11.2 Southern blot of EcoRI-digested DNA from different strains, probed with ORF2799 (region 17B). ....	48
Fig. 11.3 Southern blot of EcoRI-digested DNA from different strains, probed with intergenic sequence	

between ORFs 2798 and 2799 (region 17B).....	49
Fig. 11.4 Southern blot of EcoRI-digested DNA from different strains, probed with ORF 2798 (region 17B). .....	49
Fig. 11.5 Southern blot of EcoRI-digested DNA from different strains, probed with ORF 2797 (region 17B). .....	50
Fig. 11.6 Southern blot of EcoRI-digested DNA from different strains, probed with ORF 2796 (region 17B). .....	50
Fig. 11.7 Southern blot of EcoRI-digested DNA from different strains, probed with ORF 2795 (region 17B). .....	51
Fig. 11.8 Southern blot of EcoRI-digested DNA from different strains, probed with ORF 2794 (region 17B). .....	51
Fig. 12.1 Map of region 133 in strain F2365. ....	52
Fig. 12.2 RT-PCR results for ECI strain F2365. ....	52
Fig. 12.3 RT-PCR results for ECI strain F2365. ....	53
Fig. 13.1 Map of region 133 strain F2365 133 region map.....	54
Fig. 13.2 RT-PCR results for ECI strain F2365. Primer Pa was used to construct cDNA and primers Pa and Pb were used to detect transcription of ORF 2628, ORF 2629 and ORF 2630.....	54
Fig. 13.3 RT-PCR results for ECI strain F2365. Primer Pd was used to construct cDNA and primers Pc and Pd were used to detect transcription of ORF 2627, ORF 2628 and ORF 2629.....	54
Fig. 13.4 RT-PCR results for ECI strain F2365. Primer Pe was used to construct cDNA and primers Pe and Pf were used to detect transcription of ORF 2630 and ORF 2631; Primer Pg was used to construct cDNA, and primers Pg and Ph were used to detect transcription of ORF 2627 and ORF 2628; Primer Pi was used to construct cDNA, and primers Pi and Pj were used to detect transcription of ORF 2626 and ORF 2627. ....	55
Fig 14.1 Map of ECI specific region 17B in strain F2365.....	56
Fig. 14.2 RT-PCR results for ECI strain F2365. Primer P2 was used to construct cDNA, and primers P1 and P2 were used to detect transcription of ORF 2799 and ORF 2780; primer P3 was used to construct	

cDNA, and primer P1 and P3 were used to detect transcription of ORF 2799; primer P2 was used to construct cDNA, and primers P2 and P4 were used to detect transcription of ORF 2780..... 56

Fig. 14.3 RT-PCR results for ECI strain F2365. Primer P5 was used to construct cDNA, and primers P5 and P6 were used to detect transcription of ORF 2795 and ORF 2796; primer P5 was used to construct cDNA, and primers P5 and P7 were used to detect transcription of ORF 2795, ORF 2796 and ORF 2797; primer P5 was used to construct cDNA, and primers P5 and P8 were used to detect transcription ORF 2795, ORF 2796, ORF 2797 and ORF 2798..... 57

**Chapter III Construction of deletion mutant lacking the ECI specific gene cassette in region 133 .....67**

Fig. 1 Genomic region harboring fragment 133 in F2365 (ECI) and other Listeria genome..... 71

Fig. 2 Plasmid construction. Fragments upstream and downstream of the ECI specific cassette in region 133 were amplified by primers P1 and P5, P6 and P7 harboring restriction sites HindIII, BamHI and EcoRI, respectively. The two PCR fragments (1128 bp and 762 bp) for the upstream and downstream fragment, respectively) were ligated to the temperature sensitive shuttle vector pCON-1..... 73

Fig. 3 Homologous recombination-mediated 133 region deletion..... 74

Fig. 4 PCR screen of chloramphenicol sensitive colonies using primers Pm and Pn, located in the conserved ORFs ORF2625 and ORF2632, respectively, flanking the ECI specific region..... 83

Fig. 5 Mutant confirmation by southern blot, the probe was amplified by PCR with primers P1 and P7.83

Fig. 6 Sequence comparison of region 133 deletion in ECI strain F2381L-PHI-R..... 84

Fig. 7 PCR screen of chloramphenicol sensitive colonies to detect deletions in F2365..... 85

Fig. 8 F2365 Mutant confirmation by southern blot, done as described in the legend for Fig. 5. .... 85

Fig. 9 Sequence comparison of region 133 deletion in ECI strain F2365..... 86

Fig. 10 Swarming assay of F2365 and mutant F2365  $\Delta$ 133 on TSB+0.7% yeast extract agar (0.4% agar), following incubation at 37°C, 25°C and 4°C..... 87

Fig. 11 Growth of F2365 wild type and mutant F2365  $\Delta$ 133 on Trypticase Soy agar +0.7% yeast extract at 4°C..... 87

Fig. 12 Growth curve of F2365 wild type and mutant F2365  $\Delta$ 133 in Trypticase Soy agar +0.7% yeast extract at 4°C. Blue line: F2365 wild type; red line: mutant F2365  $\Delta$ 133 ..... 88

Fig. 13 PM1 carbon sources MicroPlate Value comparison of F2365 wild type Vs. mutant F2365  $\Delta$ 133.90

Fig. 14 PM2 carbon sources MicroPlate Value comparison of F2365 wild type Vs. mutant F2365  $\Delta$ 133.91

Fig. 15 PM3 nitrogen sources MicroPlate Value comparison of F2365 wild type Vs. mutant F2365  $\Delta$ 133..... 92

Fig. 16 PM4 Phosphorus & Sulfur sources MicroPlate Value comparison of F2365 wild type Vs. Mutant F2365 $\Delta$ 133.....93

**CHAPTER I**  
**Literature Review**

## 1.1 Introduction

The genus *Listeria* belongs to the *Clostridium* subbranch, together with *Staphylococcus*, *Streptococcus*, *Lactobacillus*, and *Bronchothrix*. This phylogenetic position of *Listeria* is consistent with its low G+C DNA content (36-42%) (3). The genus *Listeria* comprises six recognized species: *Listeria monocytogenes*, *Listeria innocua*, *Listeria ivanovii*, *Listeria seeligeri*, *Listeria welshimeri*, and *Listeria grayi* (39). Based on 16S and 23S rRNA sequence data, these six species can be divided into two lines of descent as follows: (i) *L. grayi* and (ii) the remaining five *Listeria spp.*. The latter subgroup has two distinct evolutionary branches, the first including *L. monocytogenes*, and *L. innocua* and the second encompassing *L. ivanovii*, *L. seeligeri*, and *L. welshimeri* (15, 47).

*Listeria* strains are divided into serotypes on the basis of somatic (O) and flagellar (H) antigens (3). *L. monocytogenes* comprises 13 serotypes: 1/2a, 1/2b, 1/2c, 3a, 3b, 3c, 4a, 4ab, 4b, 4c, 4d, 4e, and 7. Certain of the surface antigens among those employed for the serotyping scheme are shared among *L. monocytogenes*, *L. innocua*, *L. seeligeri*, and *L. welshimeri*. A kit for serotyping *Listeria* is now commercially available (Denka Seiken, Tokyo, Japan). Among 13 serotypes of *L. monocytogenes*, serotypes 1/2a, 1/2b and 4b account for 95% of human infections (3, 37). Serotype 1/2c strains are prevalent in foods and food processing environment, but rarely cause human infection. However, the majority of *L. monocytogenes* strains found in conjunction with large outbreaks of human food-borne listeriosis are serotype 4b (18, 25, 26, 30, 48, 49, 57). On the basis of multilocus enzyme electrophoresis (MLEE) typing, two genomic divisions have been identified in *L. monocytogenes*: division I includes serotypes 1/2a, 1/2c, 3a and 3c; and division II includes

serotypes 1/2b, 3b and 4b (5, 27, 41). *L. monocytogenes* can also be subtyped into three lineages by *EcoRI* ribotyping and other tools: lineage I, comprising serotypes 1/2b, 3b, 3c and 4b; lineage II, comprising serotypes 1/2a, 1/2c, and 3a; and lineage III, comprising serotypes 4a and 4c (8, 35).

*Listeriae* are saprophytic bacteria that are ubiquitous in nature (47). Among them, *L. monocytogenes* acts as an opportunistic pathogen for both humans and animals, mainly causing abortions in pregnant females, neonatal sepsis, and severe infections such as septicemia and meningoenzephalitis in susceptible hosts (18, 31). *L. monocytogenes* is a Gram positive, nonspore-forming, rod-shaped bacterium originally described in 1926 as being the causative agent in an outbreak of disease in laboratory rabbits and guinea pigs in England (50). *L. ivanovii* is principally responsible for abortions in sheep (31). Rarely, *L. ivanovii* and *L. seeligeri* can cause human listeriosis (16, 29). *L. innocua* has been implicated in rare cases of animal disease (45, 55).

## **1.2 Detection methods**

There are many methods for the isolation and identification of *Listeria*. Traditional methods are laborious and can take up to 1 to 2 weeks to complete, such as the CAMP (Christie, Atkins, Munch-Petersen) test, sugar fermentation, and nitrate reduction (39). Many methods involving monoclonal antibodies, DNA probes and PCR technology have been developed, for instance a molecular method based on restriction fragment length polymorphism (RFLP) of PCR-amplified fragments of the 23S rRNA gene was designed for rapid identification of *Listeria* species (39). Moreover, serotyping and pulsed-field gel

electrophoresis (PFGE) techniques have been frequently and successfully used in the characterization of *L. monocytogenes* (37). The Centers for Disease Control and Prevention (CDC) developed standardized PFGE methods, and in collaboration with the Association of Public Health Laboratories (APHL), created PulseNet. This allows scientists at public health laboratories throughout the country to rapidly compare the PFGE patterns of bacteria isolated from ill persons and determine whether they are similar. PulseNet plays a vital role in surveillance and identification of foodborne illness outbreaks that were previously difficult to detect.

### **1.3 Pathogenic *Listeria monocytogenes* Vs. nonpathogenic *Listeria innocua***

*L. monocytogenes* can survive in the extreme conditions encountered in the food chain, such as high salt concentrations and extremes of pH and temperature. These characteristics are shared by *L. innocua*, a nonpathogenic species found often in combination with *L. monocytogenes* in food and in the environment (20, 32). Although *L. innocua* is a nonpathogenic species, it is the species most closely related to *L. monocytogenes*, while the animal pathogen *L. ivanovii* and the nonpathogen *L. grayi* are more distantly related (23). However, *L. innocua* can overgrow *L. monocytogenes* in enrichment broths, which may lead to *L. monocytogenes* escaping detection (8, 21). The genome sequences of *L. monocytogenes* EGD-e (serotype 1/2a) and *L. innocua* include strain name CLIP 11262 that was sequenced have been compared. Both genomes encoded many putative surface and secreted proteins, transporters, and transcriptional regulators, consistent with the ability of both species to adapt to diverse environments (20). However, 30 of the 133 surface proteins identified in *L. monocytogenes* EGD-e were absent from *L. innocua* and 20 of these 30 *L. monocytogenes*

EGD-e specific surface proteins had the LPXTG Motif (anchoring proteins to the cell surface), which suggested that *L. monocytogenes* EGD-e may be more adaptable to the environments than *L. innocua*. The regulatory proteins are related to the ability of *Listeria* to adapt and respond to different environments and some global regulatory systems might also be associated with virulence. BglG regulatory proteins, one of the largest families of regulatory proteins, seems to be over-represented in *Listeria* and many are associated with the phosphoenolpyruvate-dependent phosphotransferase systems (PTS) involved in sugar transport and metabolism, which are abundant in *Listeria* (10). The comparison between the *L. monocytogenes* and *L. innocua* genomes reveals the conservation of all ABC transporters, but *L. monocytogenes* has many more PTS genes than *L. innocua*, suggesting that the *L. monocytogenes* EGD-e- specific PTS could be implicated in virulence (10, 20).

The presence of 270 *L. monocytogenes* EGD-e and 149 *L. innocua* strain-specific genes (clustered in 100 and 63 islets, respectively) suggests that virulence in *Listeria* results from multiple gene acquisition and deletion events (20). The G+C contents differ between species; the G+C contents of *Listeria* species are from 37% to 39%, while the G+C contents of other bacteria range from 25% to 75%. The G+C content of most of these islets are significantly lower or higher than the flanking regions, suggesting the acquisition by horizontal gene transfer. Other DNA fragments may also have been acquired by horizontal gene transfer from bacteria with a similar G+C content, or over time may have adapted to the *Listeria* genome. This seems to be one of the cases for the "virulence locus," whose G+C content is similar to that of the rest of the chromosome. Comparison of this region among *Listeria* species (11) and with the soil bacterium *Bacillus subtilis* indicates that this gene cluster was probably

acquired by a common ancestor of *Listeria* and that *L. innocua* subsequently lost most of it. Bacteriophages and plasmids may play a role in gene acquisition, and the sequenced *L. monocytogenes* and *L. innocua* strains contained one and five prophages, respectively. *L. innocua* contained a plasmid encoding heavy metal resistance. 2853 and 2973 protein-coding genes have been identified in *L. monocytogenes* and *L. innocua*, respectively. The encoded genes have a high similarity to those of *B. subtilis*, suggesting a common origin for these three species. Both *Listeria* genomes contained putative DNA uptake genes homologous to *B. subtilis* competence genes. As *Listeria* are not known to be naturally competent, the *Listeria* DNA uptake apparatus may have lost its original function. Alternatively, its regulation or the signals that induce competence may differ from those of *B. subtilis*. Gene transfer by transformation could explain most of the genomic differences between the two *Listeria* species as well as between *Listeria* and *B. subtilis* (20).

#### **1.4 Serotypes**

To better understand the molecular mechanisms of *L. monocytogenes* virulence in humans and survival of this bacterium in food and in the environment, several additional *L. monocytogenes* strains have been recently sequenced. To date, *L. innocua* (serotype 6a), *L. monocytogenes* strains EGD-e (serotype 1/2a), and F2365 (serotype 4b, genomic division II) have been fully sequenced. *L. monocytogenes* strains F6854 (serotype 1/2a, genomic division I), and H7858 (serotype 4b, genomic division II) have been partially sequenced (20, 23, 36).

Comparisons were made between serotype 4b and serotype 1/2a (23, 36). Before the genome of *L. monocytogenes* strain F2365 was fully sequenced, Herd and Kocks used *L.*

*monocytogenes* strain F4565 (serotype 4b, Isolated from the same outbreak as F2365) as the representative of serotype 4b to compare its genome with *L. monocytogenes* strain EGD-e. They found that 5 to 6% of the *L. monocytogenes* strain F4565 genome did not hybridize to the genome of *L. monocytogenes* strain EGD, and they deduced that about 150 to 190 kb of its DNA would be absent or substantially different from the genome of strain EGD-e. 39 different fragments, including 9 *L. monocytogenes* 4b specific fragments have been identified. Many of the isolated fragments had homology to bacterial surface components, and may offer the bacteria the ability to survive in variable environments. One of the 39 fragments, fragment number 85, and was homologous with a methyltransferase of the *Sau3A* restriction-modification system. This gene was confirmed and identified in the later paper (23, 36, 58).

*L. monocytogenes* strain F2365 genome was completed in 2004. It is a serotype 4b cheese isolate from the Jalisco (Mexican-style) cheese outbreak of 1985 in California (34, 36). *L. monocytogenes* strain F2365 is the first *L. monocytogenes* serotype 4b strain that has been fully sequenced. Its genome was compared to the genomes of *L. innocua*, *L. monocytogenes* strain EGD-e, *L. monocytogenes* strain F6854, which is a serotype 1/2a turkey frankfurter isolate from a sporadic case in 1988 in Oklahoma (1), and with *L. monocytogenes* strain H7858, which is a serotype 4b frankfurter isolate from the multistate outbreak of 1998-1999 in the USA (2). A summary of the key genomic findings of these five strains is presented in Table 1 (20, 36).

Table 1. Genomic summary of five *Listeria* strains. Modified from reference 19 and 36.

Strain	<i>L. monocytogenes</i> F2365	<i>L. monocytogenes</i> H7858	<i>L. monocytogenes</i> EGD-e	<i>L. monocytogenes</i> F6854	<i>L. innocua</i> Clip 11262
Serotype	4b	4b	1/2a	1/2a	6a
Chromosome Length	2 905 310	~2 893 921	2 944 528	~2 953 211	3 011 209
G+C content	38%	38%	39%	37.8%	37%
No. of ORFs	2847	3024	2853	2973	2973
Phage regions	2	2	1	3	5
Monocins	1	1	1	1	1
Plasmid	None	1	None	None	1
Plasmid Length	-	82 270	-	-	81 900
No. of ORFs on plasmid	-	94	-	-	79

The genes from all five *listeria* genomes were compared against each other by using BLAST. A total of 51, 97, 69 and 61 strain-specific genes were identified in strains F2365, F6854, H7858 and EGD-e, respectively. A cluster of F2365 specific genes were identified (LMOF2365\_0331±LMOF2365\_0323), including a putative type II restriction endonuclease with specificity for GATC sites, a DNA methylase specific for cytosines at GATC sites, and a DNA-binding protein. This result supported with the previous findings (23). Among the serotype 1/2a-specific genes were three clusters that encode pathways for the transport and metabolism of carbohydrates including ribose, and an unidentified pentose sugar. The serotype 1/2a-specific genes also included an operon that encodes the biosynthetic pathway for the antigenic rhamnose substituents that decorate the cell wall-associated teichoic acid polymer in serotype 1/2a strains (27), five glycosyl transferases and an adenine-specific DNA methyltransferase. The pleiotropic regulatory activator *prfA* (LMOF2365\_0211 in strain

F2365) and the four genes comprising the *agr* locus (*agrA-D*; LMOF2365\_0057-60 in strain F2365) were completely conserved across all four *L. monocytogenes* strains. Both of PrfA regulon and *agr* locus play very important roles in virulence of *Listeria* (4). The virulence factor of bacteria includes adhesion factors, colonization factors, invasion factors and toxins. The conservation of PrfA regulon and *agr* locus across these different strains suggests that they are under selective pressure to be maintained without mutations.

### **1.5 Mobile genetic elements in the *L. monocytogenes* genome**

Bacteria commonly carry a number of mobile elements (transposons, plasmid, prophages and integrons), the composition of which can vary widely among members of the same bacteria species. Horizontal transfer of these elements drives the evolution of pathogenicity (10).

#### **1.5.1 Insertion elements (IS)**

An intact insertion element, (IS) element (ISLmol), was present in the serotype 1/2a strains F6854 (two copies) and EGD-e (three copies). The chromosomes of serotype 4b strains (F2365 and H7858) lacked intact (IS) elements, but did contain four copies of transposase ORFA of the IS3 family, that were present in homologous locations in both strains. Three of the ORFAs were also present in the serotype 1/2a strains (EGD-e and F6854) in the same location; the additional copy of ORFA is considered a complete and a partial duplication in strains F2365 and H7858, respectively (36). It seems that all *L. monocytogenes* genomes contain the transposase ORFA, and this gene probably originated from the genome of an ancestral *Listeria* before the strains diverged. However, ISLmol appears to be a recent

acquisition and may still be mobile in the chromosome of the serotype 1/2a strains (36).

### **1.5.2 Clustered regularly interspaced palindromic repeats (CRISPRs)**

The clustered regularly interspaced palindromic repeats (CRISPRs) are short direct repeats, 25-50 nucleotides long, interspaced by unique sequences of similar size (7). Five *Listeria* genomes were compared by identifying polymorphisms using CRISPRs. *L. monocytogenes* serotype 1/2a strains and *L. innocua* had CRISPR repeats, but *L. monocytogenes* 4b did not. Thus, CRISPR elements may be involved in gene transfer events, and they can be considered a tool to differentiate among *Listeria* strains.

### **1.5.3 Prophages**

Nine putative prophages and five putative monocins / defective prophages were found in all five *Listeria* genomes, but F2365 does not contain an intact prophage in the genome. Six of the nine putative prophages have at least 11 ORFs that are homologous to ORFs of  $\Phi$ A118, a phage which is a temperate Siphovirus (morphotype B1), attacking predominantly serovars 1/2a (56). In *L. innocua*, a putative prophage ( $\Phi$ 11262.1) has inserted into tRNA-Lys-4, suggesting that this prophage may be a relatively recent acquisition in *L. innocua* by recombination with an existing prophage (36).

### **1.5.4 Plasmids**

No plasmids were found in *L. monocytogenes* serotype 1/2a strains (EGD-e and F6854) and *L. monocytogenes* strain F2365 (serotype 4b). However, *L. monocytogenes* H7858

(serotype 4b) was found to harbor a plasmid, pLM80, ca.80 kb, with high level of similarity to *L. innocua* pLI100 and *Bacillus anthracis* plasmid pXO2 (20, 44). Two distinct regions of similarity of pLM80 were identified. Region 1 is specific to *Listeria*; it encodes proteins responsible for the detoxification of arsenate and cadmium, and also contains six mobile genetic elements. Region 2 is similar to a region of pXO2, but it has diverged substantially from its counterparts in pXO2, suggesting that region may have been exchanged between *L. monocytogenes* and *B. anthracis* some time ago, or was transferred from a different source (36).

### **1.5.5 Role of plasmids and transposons in antibiotic and heavy metal resistance in *Listeria***

The plasmids in *Listeria* strains usually carry genes encode functions in carbohydrate fermentation, resistance to antibiotics, heavy metals or disinfectants, growth at low pH, NaCl tolerance or thermal inactivation by pasteurization (33). In most cases, antibiotic and heavy metal resistance attributes come from four types of mobile genetic elements: conjugative plasmids, mobilizable plasmids, transposons and integrons (12-14, 22, 28, 42, 43). Antibiotic resistance can be transferred by conjugation of enterococcal and streptococcal plasmids and transposons carrying antibiotic resistance genes from *Enterococcus-Streptococcus* to *Listeria* and between the different species of *Listeria*, such as plasmid pIP501, pAM $\beta$ 1, pRYC16 and Tn916(19, 40, 54). Multiple antibiotic resistance in *L. monocytogenes* can be acquired by self-transferable plasmids, such as pIP811(42), pUBX1 and pWDB100(22). Some *L. monocytogenes* strains have resistance to heavy metals which usually are toxic in bacteria. Transposon Tn5422 is plasmid-borne in *L. monocytogenes*, and contains two genes which

confer cadmium resistance(28).

## **1.6 Epidemic clones of *L. monocytogenes***

### **1.6.1 Epidemic clone I (ECI)**

*L. monocytogenes* has been implicated in numerous outbreaks and sporadic cases of human listeriosis. Since it can grow at low temperatures, it has the potential to contaminated highly processed, cold-stored, and ready-to-eat foods. At high-risk people for listeriosis are pregnant women and their fetuses, the very old and the very young people, and immunocompromised people, including HIV and cancer patients (18, 48). Serotypes 1/2a, 1/2b and 4b are involved in more than 95% of listeriosis cases. Furthermore, most publicized outbreaks have involved serotype 4b (9, 41).

Several closely related strains of serotype 4b, designated epidemic clone I (ECI), have been implicated in outbreaks in Nova Scotia, Canada (coleslaw, 1981), California (Mexican-style cheese, 1985), France (pork tongue in aspic, 1992), as well as several other outbreaks (17, 27). ECI strains were not identified in raw, refrigerated meat and poultry (46), and were rarely recovered from food (6). The foods are mostly contaminated by serotype 1/2 strains, especially 1/2a and 1/2c strains, not by serotype 4b strains, the reasons remained unknown (27). Several genetic and phenotypic markers specific to ECI strains have been identified. ECI strains share a unique restriction fragment length polymorphism in a genomic region essential for low-temperature (4°C) growth (27, 59). In addition, they also appear to methylate cytosines at GATC sites of their genome, thus make their genomic DNA resistant to digestion by the restriction enzyme *Sau3AI* (53, 60).

In 2001, Herd and Kocks showed that ECI strains have a number of unique genomic fragments and gene clusters commonly absent from other serotype 4b strains (23). Fragment 85 was specific to ECI strains and was putatively involved in cytosine methylation at GATC sites (23). Genome sequencing data recently revealed that this fragment was part of a gene cassette composed of three genes, 85R (5-methyl Cytosine restriction), 85M (methylase) and 85S (endonuclease *Sau3AI*). Thus, this cassette appears to have GATC-specific restriction-modification functions. The G+C contents of the genes 85M, 85R and 85S are 33, 28, and 31%, respectively, which is noticeably lower than the average of *L. monocytogenes* (38%). The low G+C content indicated that this cassette might have been horizontally introduced into the genome of ECI strains by lateral transfer from currently unidentified bacteria. In bacteria, restriction methylation systems are usually thought to associate with a defense against phage invasion, digesting the unmodified DNA injected into the cell by the phage (58); in some cases, DNA methylation is associated with regulation of genes involved in virulence and other functions (24, 51). In the recent publication, 51 ECI strain F2365 specific genes were identified, and the genes (LMOF2365\_001-LMOF2365\_0323) corresponding fragment 85R, 85M and 85S were also confirmed. The pleiotropic regulatory activator *prfA* (LMOF2365\_0211) was found completely conserved across four *L. monocytogenes* strains F2365 (serotype 4b, ECI), F6854 (serotype 1/2a), H7858 (serotype 4b, ECII) and EGD-e (serotype 1/2a). The PrfA regulon controls the major virulence genes (*hly*, *plcA*, *plcB*, *mpl*, *actA*, *inlA* and *inlB*) (36). And the four genes comprising the *agr* locus (*agrA-D*; LMOF2365\_0057-60) were found in all four *L. monocytogenes* strains and non-pathogenic strain *L. innocua*. *agr* locus was identified played important role in *Listeria* virulence (4)

### 1.6.2 Epidemic clone II

In 1998 - 1999, *L. monocytogenes* of serotype 4b was implicated in a multistate outbreak of listeriosis, which was caused by the consumption of contaminated hot dogs and was responsible for 101 human cases, including 21 deaths (2, 52). The ribotype and pulsed-field gel electrophoresis (PFGE) fingerprints of these strains had been rarely encountered before. Thus these bacteria appeared to represent a novel epidemic associated lineage, which is designated epidemic clone II (2, 17).

In recent publication, 69 of ECII strain H7858 specific genes were identified. 8605 high quality SNPs were found in the genome of H7858, suggesting a high mutation rate of genes involved in energy metabolism and transport, and having varying abilities to withstand adverse environments (36). One serotype 4b-specific region (LMOF2365\_0448 to LMOF2365\_0471) which was either divergent in or absent from the genome of ECII strains has been identified (17). Based on southern blotting result, several DNA fragments in this region are not present on ECII genome and only three ORFs are conserved between ECII and other serotype 4b strains. Except for conserved ORFs, the G+C contents of this region range from 19.44 to 35.78%, which is much lower than the average G+C content of *L. monocytogenes* (38%), suggesting that the corresponding genes in serotype 4b may have been acquired by horizontal transfer from another, unidentified source. This region is also adjacent to the internalin genes *inlA* and *inlB*, which are known to have important functions in cell invasion and virulence in *L. monocytogenes* (17). This suggested the possibility that

the serotype 4b-specific genes in this region may also be implicated in host-pathogen interactions and virulence. Comparative genomic analysis of the region suggests that it may have been introduced into serotype 4b by an insertion in the genome between LMOF2365\_0448 and LMOF2365\_0471 (*inlA*). Alternatively, the region could have been present in an ancestral *L. monocytogenes* lineage and maintained only in serotype 4b, becoming lost from the genomes of strains of other serotypes.

### 1.6.3 Epidemic clone III

Unlike ECI and ECII, the strains of ECIII are serotype 1/2a. Strain *L. monocytogenes* F6854, which caused the contamination of turkey franks in 1988 (2, 27, 52), belongs to this clone, and it has been partially sequenced. It is noticeable that ECIII strains appeared to have persisted in the processing plant for more than a decade (27). The strain identified in the 2000 multistate outbreak that was caused by contaminated turkey deli meat was the same PFGE genotype as the strain implicated in a human listeriosis case associated with the consumption of contaminated turkey in 1988 (27, 38).

## 1.7 References

1. 1989. Listeriosis associated with consumption of turkey franks. MMWR Morb Mortal Wkly Rep **38**:267-8.
2. 1998. Multistate outbreak of listeriosis--United States, 1998. MMWR Morb Mortal Wkly Rep **47**:1085-6.
3. **Allerberger, F.** 2003. Listeria: growth, phenotypic differentiation and molecular microbiology. FEMS Immunol Med Microbiol **35**:183-9.

4. **Autret, N., C. Raynaud, I. Dubail, P. Berche, and A. Charbit.** 2003. Identification of the agr locus of *Listeria monocytogenes*: role in bacterial virulence. *Infect Immun* **71**:4463-71.
5. **Bibb, W. F., B. G. Gellin, R. Weaver, B. Schwartz, B. D. Plikaytis, M. W. Reeves, R. W. Pinner, and C. V. Broome.** 1990. Analysis of clinical and food-borne isolates of *Listeria monocytogenes* in the United States by multilocus enzyme electrophoresis and application of the method to epidemiologic investigations. *Appl Environ Microbiol* **56**:2133-41.
6. **Boerlin, P., and J. C. Piffaretti.** 1991. Typing of human, animal, food, and environmental isolates of *Listeria monocytogenes* by multilocus enzyme electrophoresis. *Appl Environ Microbiol* **57**:1624-9.
7. **Bolotin, A., B. Quinquis, A. Sorokin, and S. D. Ehrlich.** 2005. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* **151**:2551-61.
8. **Bruhn, J. B., B. F. Vogel, and L. Gram.** 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**:961-7.
9. **Buchrieser, C., R. Brosch, B. Catimel, and J. Rocourt.** 1993. Pulsed-field gel electrophoresis applied for comparing *Listeria monocytogenes* strains involved in outbreaks. *Can J Microbiol* **39**:395-401.
10. **Buchrieser, C., C. Rusniok, F. Kunst, P. Cossart, and P. Glaser.** 2003. Comparison of the genome sequences of *Listeria monocytogenes* and *Listeria innocua*: clues for evolution and pathogenicity. *FEMS Immunol Med Microbiol* **35**:207-13.

11. **Chakraborty, T., T. Hain, and E. Domann.** 2000. Genome organization and the evolution of the virulence gene locus in *Listeria* species. *Int J Med Microbiol* **290**:167-74.
12. **Charpentier, E., and P. Courvalin.** 1999. Antibiotic resistance in *Listeria* spp. *Antimicrob Agents Chemother* **43**:2103-8.
13. **Charpentier, E., G. Gerbaud, and P. Courvalin.** 1999. Conjugative mobilization of the rolling-circle plasmid pIP823 from *Listeria monocytogenes* BM4293 among gram-positive and gram-negative bacteria. *J Bacteriol* **181**:3368-74.
14. **Chesneau, O., R. Lailier, A. Derbise, and N. El Solh.** 1999. Transposition of IS1181 in the genomes of *Staphylococcus* and *Listeria*. *FEMS Microbiol Lett* **177**:93-100.
15. **Collins, M. D., S. Wallbanks, D. J. Lane, J. Shah, R. Nietupski, J. Smida, M. Dorsch, and E. Stackebrandt.** 1991. Phylogenetic analysis of the genus *Listeria* based on reverse transcriptase sequencing of 16S rRNA. *Int J Syst Bacteriol* **41**:240-6.
16. **Cummins, A. J., A. K. Fielding, and J. McLauchlin.** 1994. *Listeria ivanovii* infection in a patient with AIDS. *J Infect* **28**:89-91.
17. **Evans, M. R., B. Swaminathan, L. M. Graves, E. Altermann, T. R. Klaenhammer, R. C. Fink, S. Kernodle, and S. Kathariou.** 2004. Genetic markers unique to *Listeria monocytogenes* serotype 4b differentiate epidemic clone II (hot dog outbreak strains) from other lineages. *Appl Environ Microbiol* **70**:2383-90.
18. **Farber, J. M., and P. I. Peterkin.** 1991. *Listeria monocytogenes*, a food-borne pathogen. *Microbiol Rev* **55**:476-511.
19. **Flamm, R. K., D. J. Hinrichs, and M. F. Thomashow.** 1984. Introduction of pAM beta 1 into *Listeria monocytogenes* by conjugation and homology between native L.

- monocytogenes plasmids. *Infect Immun* **44**:157-61.
20. **Glaser, P., L. Frangeul, C. Buchrieser, C. Rusniok, A. Amend, F. Baquero, P. Berche, H. Bloecker, P. Brandt, T. Chakraborty, A. Charbit, F. Chetouani, E. Couve, A. de Daruvar, P. Dehoux, E. Domann, G. Dominguez-Bernal, E. Duchaud, L. Durant, O. Dussurget, K. D. Entian, H. Fsihi, F. Garcia-del Portillo, P. Garrido, L. Gautier, W. Goebel, N. Gomez-Lopez, T. Hain, J. Hauf, D. Jackson, L. M. Jones, U. Kaerst, J. Kreft, M. Kuhn, F. Kunst, G. Kurapkat, E. Madueno, A. Maitournam, J. M. Vicente, E. Ng, H. Nedjari, G. Nordsiek, S. Novella, B. de Pablos, J. C. Perez-Diaz, R. Purcell, B. Remmel, M. Rose, T. Schlueter, N. Simoes, A. Tierrez, J. A. Vazquez-Boland, H. Voss, J. Wehland, and P. Cossart.** 2001. Comparative genomics of *Listeria* species. *Science* **294**:849-52.
  21. **Gnanou Besse, N., N. Audinet, A. Kerouanton, P. Colin, and M. Kalmokoff.** 2005. Evolution of *Listeria* populations in food samples undergoing enrichment culturing. *Int J Food Microbiol*.
  22. **Hadorn, K., H. Hachler, A. Schaffner, and F. H. Kayser.** 1993. Genetic characterization of plasmid-encoded multiple antibiotic resistance in a strain of *Listeria monocytogenes* causing endocarditis. *Eur J Clin Microbiol Infect Dis* **12**:928-37.
  23. **Herd, M., and C. Kocks.** 2001. Gene fragments distinguishing an epidemic-associated strain from a virulent prototype strain of *Listeria monocytogenes* belong to a distinct functional subset of genes and partially cross-hybridize with other *Listeria* species. *Infect Immun* **69**:3972-9.
  24. **Hernday, A., M. Krabbe, B. Braaten, and D. Low.** 2002. Self-perpetuating

- epigenetic pili switches in bacteria. Proc Natl Acad Sci U S A **99 Suppl 4**:16470-6.
25. **Ho, J. L., K. N. Shands, G. Friedland, P. Eckind, and D. W. Fraser.** 1986. An outbreak of type 4b *Listeria monocytogenes* infection involving patients from eight Boston hospitals. Arch Intern Med **146**:520-4.
  26. **Jacquet, C., B. Catimel, R. Brosch, C. Buchrieser, P. Dehaumont, V. Goulet, A. Lepoutre, P. Veit, and J. Rocourt.** 1995. Investigations related to the epidemic strain involved in the French listeriosis outbreak in 1992. Appl Environ Microbiol **61**:2242-6.
  27. **Kathariou, S.** 2002. *Listeria monocytogenes* virulence and pathogenicity, a food safety perspective. J Food Prot **65**:1811-29.
  28. **Lebrun, M., A. Audurier, and P. Cossart.** 1994. Plasmid-borne cadmium resistance genes in *Listeria monocytogenes* are present on Tn5422, a novel transposon closely related to Tn917. J Bacteriol **176**:3049-61.
  29. **Lessing, M. P., G. D. Curtis, and I. C. Bowler.** 1994. *Listeria ivanovii* infection. J Infect **29**:230-1.
  30. **Linnan, M. J., L. Mascola, X. D. Lou, V. Goulet, S. May, C. Salminen, D. W. Hird, M. L. Yonekura, P. Hayes, R. Weaver, and et al.** 1988. Epidemic listeriosis associated with Mexican-style cheese. N Engl J Med **319**:823-8.
  31. **Low, J. C., and W. Donachie.** 1997. A review of *Listeria monocytogenes* and listeriosis. Vet J **153**:9-29.
  32. **MacGowan, A. P., K. Bowker, J. McLauchlin, P. M. Bennett, and D. S. Reeves.** 1994. The occurrence and seasonal changes in the isolation of *Listeria* spp. in shop bought food stuffs, human faeces, sewage and soil from urban sources. Int J Food

- Microbiol **21**:325-34.
33. **Margolles, A., and C. G. de los Reyes-Gavilan.** 1998. Characterization of plasmids from *Listeria monocytogenes* and *Listeria innocua* strains isolated from short-ripened cheeses. *Int J Food Microbiol* **39**:231-6.
  34. **Mascola, L., L. Lieb, J. Chiu, S. L. Fannin, and M. J. Linnan.** 1988. Listeriosis: an uncommon opportunistic infection in patients with acquired immunodeficiency syndrome. A report of five cases and a review of the literature. *Am J Med* **84**:162-4.
  35. **Nadon, C. A., D. L. Woodward, C. Young, F. G. Rodgers, and M. Wiedmann.** 2001. Correlations between molecular subtyping and serotyping of *Listeria monocytogenes*. *J Clin Microbiol* **39**:2704-7.
  36. **Nelson, K. E., D. E. Fouts, E. F. Mongodin, J. Ravel, R. T. DeBoy, J. F. Kolonay, D. A. Rasko, S. V. Angiuoli, S. R. Gill, I. T. Paulsen, J. Peterson, O. White, W. C. Nelson, W. Nierman, M. J. Beanan, L. M. Brinkac, S. C. Daugherty, R. J. Dodson, A. S. Durkin, R. Madupu, D. H. Haft, J. Selengut, S. Van Aken, H. Khouri, N. Fedorova, H. Forberger, B. Tran, S. Kathariou, L. D. Wonderling, G. A. Uhlich, D. O. Bayles, J. B. Luchansky, and C. M. Fraser.** 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**:2386-95.
  37. **Okwumabua, O., M. O'Connor, E. Shull, K. Strelow, M. Hamacher, T. Kurzynski, and D. Warshauer.** 2005. Characterization of *Listeria monocytogenes* isolates from food animal clinical cases: PFGE pattern similarity to strains from human listeriosis cases. *FEMS Microbiol Lett*.

38. **Olsen, S. J., M. Patrick, S. B. Hunter, V. Reddy, L. Kornstein, W. R. MacKenzie, K. Lane, S. Bidol, G. A. Stoltman, D. M. Frye, I. Lee, S. Hurd, T. F. Jones, T. N. LaPorte, W. Dewitt, L. Graves, M. Wiedmann, D. J. Schoonmaker-Bopp, A. J. Huang, C. Vincent, A. Bugenhagen, J. Corby, E. R. Carloni, M. E. Holcomb, R. F. Woron, S. M. Zansky, G. Dowdle, F. Smith, S. Ahrabi-Fard, A. R. Ong, N. Tucker, N. A. Hynes, and P. Mead.** 2005. Multistate outbreak of *Listeria monocytogenes* infection linked to delicatessen turkey meat. *Clin Infect Dis* **40**:962-7.
39. **Paillard, D., V. Dubois, R. Duran, F. Nathier, C. Guittet, P. Caumette, and C. Quentin.** 2003. Rapid identification of *Listeria* species by using restriction fragment length polymorphism of PCR-amplified 23S rRNA gene fragments. *Appl Environ Microbiol* **69**:6386-92.
40. **Perez-Diaz, J. C., M. F. Vicente, and F. Baquero.** 1982. Plasmids in *Listeria*. *Plasmid* **8**:112-8.
41. **Piffaretti, J. C., H. Kressebuch, M. Aeschbacher, J. Bille, E. Bannerman, J. M. Musser, R. K. Selander, and J. Rocourt.** 1989. Genetic characterization of clones of the bacterium *Listeria monocytogenes* causing epidemic disease. *Proc Natl Acad Sci U S A* **86**:3818-22.
42. **Poyart-Salmeron, C., C. Carlier, P. Trieu-Cuot, A. L. Courtieu, and P. Courvalin.** 1990. Transferable plasmid-mediated antibiotic resistance in *Listeria monocytogenes*. *Lancet* **335**:1422-6.
43. **Poyart-Salmeron, C., P. Trieu-Cuot, C. Carlier, A. MacGowan, J. McLauchlin, and P. Courvalin.** 1992. Genetic basis of tetracycline resistance in clinical isolates of *Listeria monocytogenes*. *Antimicrob Agents Chemother* **36**:463-6.

44. **Read, T. D., S. L. Salzberg, M. Pop, M. Shumway, L. Umayam, L. Jiang, E. Holtzapple, J. D. Busch, K. L. Smith, J. M. Schupp, D. Solomon, P. Keim, and C. M. Fraser.** 2002. Comparative genome sequencing for discovery of novel polymorphisms in *Bacillus anthracis*. *Science* **296**:2028-33.
45. **Rocourt, J., H. Hof, A. Schrettenbrunner, R. Malinverni, and J. Bille.** 1986. [Acute purulent *Listeria seelingeri* meningitis in an immunocompetent adult]. *Schweiz Med Wochenschr* **116**:248-51.
46. **Ryser, E. T., S. M. Arimi, M. M. Bunduki, and C. W. Donnelly.** 1996. Recovery of different *Listeria* ribotypes from naturally contaminated, raw refrigerated meat and poultry products with two primary enrichment media. *Appl Environ Microbiol* **62**:1781-7.
47. **Sallen, B., A. Rajoharison, S. Desvarenne, F. Quinn, and C. Mabilat.** 1996. Comparative analysis of 16S and 23S rRNA sequences of *Listeria* species. *Int J Syst Bacteriol* **46**:669-74.
48. **Schuchat, A., B. Swaminathan, and C. V. Broome.** 1991. Epidemiology of human listeriosis. *Clin Microbiol Rev* **4**:169-83.
49. **Schwartz, B., D. Hexter, C. V. Broome, A. W. Hightower, R. B. Hirschhorn, J. D. Porter, P. S. Hayes, W. F. Bibb, B. Lorber, and D. G. Faris.** 1989. Investigation of an outbreak of listeriosis: new hypotheses for the etiology of epidemic *Listeria monocytogenes* infections. *J Infect Dis* **159**:680-5.
50. **Sewell, A. M., D. W. Warburton, A. Boville, E. F. Daley, and K. Mullen.** 2003. The development of an efficient and rapid enzyme linked fluorescent assay method for the detection of *Listeria* spp. from foods. *Int J Food Microbiol* **81**:123-9.

51. **Slominska, M., G. Konopa, J. Ostrowska, B. Kedzierska, G. Wegrzyn, and A. Wegrzyn.** 2003. SeqA-mediated stimulation of a promoter activity by facilitating functions of a transcription activator. *Mol Microbiol* **47**:1669-79.
52. **Stone, S. C., and J. Shoenberger.** 2001. Update: Multistate outbreak of listeriosis--United States, 2000. *Ann Emerg Med* **38**:339-41.
53. **Tran, H. L., and S. Kathariou.** 2002. Restriction fragment length polymorphisms detected with novel DNA probes differentiate among diverse lineages of serogroup 4 *Listeria monocytogenes* and identify four distinct lineages in serotype 4b. *Appl Environ Microbiol* **68**:59-64.
54. **Vicente, M. F., F. Baquero, and J. C. Perez-Diaz.** 1988. Conjugative acquisition and expression of antibiotic resistance determinants in *Listeria* spp. *J Antimicrob Chemother* **21**:309-18.
55. **Walker, J. K., J. H. Morgan, J. McLauchlin, K. A. Grant, and J. A. Shallcross.** 1994. *Listeria innocua* isolated from a case of ovine meningoencephalitis. *Vet Microbiol* **42**:245-53.
56. **Wendlinger, G., M. J. Loessner, and S. Scherer.** 1996. Bacteriophage receptors on *Listeria monocytogenes* cells are the N-acetylglucosamine and rhamnose substituents of teichoic acids or the peptidoglycan itself. *Microbiology* **142 ( Pt 4)**:985-92.
57. **Wiedmann, M., J. L. Bruce, C. Keating, A. E. Johnson, P. L. McDonough, and C. A. Batt.** 1997. Ribotypes and virulence gene polymorphisms suggest three distinct *Listeria monocytogenes* lineages with differences in pathogenic potential. *Infect Immun* **65**:2707-16.
58. **Yildirim, S., W. Lin, A. D. Hitchins, L. A. Jaykus, E. Altermann, T. R.**

- Klaenhammer, and S. Kathariou.** 2004. Epidemic clone I-specific genetic markers in strains of *Listeria monocytogenes* serotype 4b from foods. *Appl Environ Microbiol* **70**:4158-64.
59. **Zheng, W., and S. Kathariou.** 1995. Differentiation of epidemic-associated strains of *Listeria monocytogenes* by restriction fragment length polymorphism in a gene region essential for growth at low temperatures (4 degrees C). *Appl Environ Microbiol* **61**:4310-4.
60. **Zheng, W., and S. Kathariou.** 1997. Host-mediated modification of *Sau3AI* restriction in *Listeria monocytogenes*: prevalence in epidemic-associated strains. *Appl Environ Microbiol* **63**:3085-9.

## **CHAPTER II**

### **Genomic organization and transcriptional analysis of ECI specific cassettes 144, 133, 17B and 85**

## 2.1 Abstract

The genome of ECI strains has been known to harbor gene cassettes unique to this clonal group. Four such ECI specific gene cassettes, designated 144, 133, 17B and 85 were characterized further by taking advantage of the recently complete genome sequencing of the ECI strain F2365. The genomic organization of the regions harboring these ECI-specific cassettes was analyzed in five different *Listeria* strains, including ECI strain F2365 (serotype 4b), ECII strain H7858 (serotype 4b), serotype 1/2a strains EGD-e and F6854, and *Listeria innocua* CLIP 11262. The comparative genomic organization data revealed that all four ECI specific regions has features typical of genomic islands (GEIs), being present in the genome of ECI strains but absent from the genome of other serotype 4b strains, suggesting insertion / deletion events. The G+C contents of all four ECI specific cassettes were significantly lower (27-34%) than the average for *Listeria* (38%). Of the four cassettes, one (cassette 144) consisted of a single open reading frame (ORF), whereas cassettes 17B, 85, and 133 consisted of three, five, and six ORFs, respectively. Transcriptional analysis utilizing reverse transcriptase – polymerase chain reactions suggested that the single ECI - specific ORF in cassette 144 was co-transcribed with two upstream ORFs that were conserved among different strains, and that the six ORFs of cassette 133 were co-transcribed, and appeared to constitute a transcriptional unit separate from the flanking genes. A 42 bp palindromic sequence flanked the ECI specific ORFs on this cassette, possibly functioning as rho-independent transcriptional terminator based on secondary structure analysis, whereas a single copy of the sequence was found in a genomically equivalent location in the other strains. Two of the cassettes, 133 and 17B, were located immediately adjacent to genes with key functions in cell metabolism, *tuf* (translation elongation factor Tu) and *gidA* (glucose

inhibited division protein, involved in mortification of 23S rRNA), respectively. With the exception of cassette 85, which is involved in restriction and modification of DNA at GATC sites, the functions of the other ECI specific cassettes are unknown at this time. However, the genomic features of the cassettes suggest that they are genomic islands that may have been horizontally transferred from other organisms. Similarly to genomic islands in other bacteria, these cassettes may play important roles in bacterial pathogenicity and metabolism, in self-defense from phage infection, and in resistance to antibiotics, heavy metals or extreme environments.

## 2.2 Introduction

Most of the outbreaks of food borne listeriosis have involved a small number of genetically related strains, commonly of serotype 4b (3, 5, 7, 18). Epidemic clone I (ECI) strains have been implicated in numerous outbreaks in Europe and North America, including those in Nova Scotia, Canada (coleslaw 1981), California (Jalisco cheese, 1985), France (pork tongue in aspic, 1992), and others (4). ECI specific genetic and phenotypic markers have been identified, including a modification that rendered the genomic DNA of these strains resistant to digestion by the enzyme *Sau3AI*, and restriction fragment length polymorphisms (RFLPs) in several chromosomal markers (9, 20, 21). Recently, the genome sequencing of one ECI strain F2365, which was implicated in the 1985 Jalisco (Mexican style cheese) outbreak in California, has been completed and published (16).

The genome sequence of strain F2365 identified 51 genes that were present in the genome of this organism and absent from the other *Listeria* genomes that had been sequenced (8). Several of these putative ECI-specific genes were identified in a previous study that employed subtractive hybridization of one of the strains from the Jalisco cheese outbreak against the serotype 1/2a strain EGD-e (9). Subtractive hybridization is a method to identify DNA/RNA present in one sample but not in the others. In this study, we characterized the genomic regions harboring four of these fragments (fragments 144, 133, 17B, and 85) at the level of genomic organization and transcription utilizing the information from the complete genome sequence of F2365.

## 2.3 Materials and methods

### 2.3.1 Nucleotide sequence analysis for cassettes 144, 133, 17B and 85

For DNA and protein database searches and analysis, we used BLAST algorithms (<http://www.ncbi.nlm.nih.gov/BLAST/>) and Artemis software (<http://www.sanger.ac.uk/Software/Artemis>). For DNA and protein database comparisons among five different *Listeria* genomes, we used clustalW (<http://www.ebi.ac.uk/clustalw/>). The whole sequence data of *L. monocytogenes* strain EGD-e (serotype 1/2a), *L. monocytogenes* strain F2365 (serotype 4b), *L. innocua* strain CLIP 11262 and preliminary sequence data of *L. monocytogenes* strain H7858 (serotype 4b) and *L. monocytogenes* strain F2365 (serotype 1/2a) were obtained from (<http://www.ncbi.nlm.nih.gov/>). For RNA secondary structure analysis, we used RNA folding (mfold) (<http://www.bioinfo.rpi.edu/applications/mfold/old/rna/>). For protein prediction, we used protein families database alignments and HMMs (<http://www.sanger.ac.uk/Software/Pfam/>).

### 2.3.2 Isolation of genomic DNA of *Listeria*

A single colony was inoculated in 5 ml TSB+ 0.7%Yeast Extract, and was incubated at 37° C overnight. 3 ml of the cells were collected by centrifugation (11.9 rpm 3 minutes) in a 1.5 ml Eppendorf tube. The pellet was resuspended in 180 ul enzymatic lysis buffer (10 ml lysis buffer: 200 ul 1M tris-cl pH7, 40 ul 0.5 M ECTA, 120 ul Triton x 100, 200 mg lysozyme, 9.64 ml ddH<sub>2</sub>O) and incubated at 37° C for 3 hours. Qiagen DNeasy tissue kit (Valencia, CA) was used in the following steps (See Qiagen DNeasy tissue handbook from step 3 on page 19 to step 9 on page 29). Genomic DNA was resuspended in dH<sub>2</sub>O and stored at 4° C.

### 2.3.3 Southern Blots

#### 2.3.3.1 Cultures used in Southern Blots

Table 1. *Listeria* strains used in southern blotting with probes derived from different ORFs in the ECI specific regions 133 and 17B.

NO.	Strains	Serotypes	Origin	Sporadic (S)/ Epidemic (E)	Patient/food origin
1	F2381	4b (ECI)	CA, USA	E	Patient
2	G4026	4b (ECI)	ENGLAND	E	Patient
3	G3982	4b (ECI)	SWITZERLAND	E	Food
4	G3988	4b (ECI)	CANADA	E	Patient
5	G3992	4b (ECI)	FRANCE	E	Patient
6	G4030	4b (ECI)	FRANCE	E	Patient
7	264	4b (ECI)	HAWAII	S	Patient
8	266	4b	HAWAII	S	Patient
9	2001-126R	4b	NC, USA	S	Patient
10	4b1	4b	N/A	S	Patient
11	WS1	4b (ECIa)	NC, USA	E	Patient
12	2001-7R	4b	NC, USA	S	Patient
13	H7750	4b (ECII)	USA	E	Patient
14	H7738	4b	USA	E	Patient
15	F6854	1/2a	USA	S	Patient
16	J0161	1/2a (ECIII)	USA	E	Patient
17	G3986	3b	USA	S	Food
18	G3968	3b	USA	S	Patient
19	G3964	1/2b	CANADA	S	Patient
20	G4027	1/2b	USA	S	Patient
21	G4598	1/2b	Italy	E	Patient

#### 2.3.3.2 Probe Labeling with DIG-dUTP (Roche)

DNA fragments to be used as probes in Southern blots were labeled with random primers and digoxigenin, following the steps recommended by the vendor of the digoxigenin

labeling and detection kit (Roche). 10 ng- 1 ug DNA was added to autoclaved dH<sub>2</sub>O to a final volume of 15 ul. The DNA was denatured by heating in a boiling water bath at 95-100° C for 10 minutes, and immediately placed on ice. 2 ul 10X Hexanucleotide Mix, 2 ul 10X DIG DNA labeling mix and 1 ul Klenow (5u / ul) were added to the DNA and the mixture was incubated at 37° C overnight. The labeled DNA was stored at -20° C.

### **2.3.3.3 Digestion of genomic DNA with restriction enzyme EcoRI and agarose gel electrophoresis**

Genomic DNA was completely digested by the restriction enzyme EcoRI (New England, Biolabs, Beverly, MA) in a 37° C water bath overnight. The total volume of the restriction mixture was 20 ul, (1 ug genomic DNA, 1 ul enzyme EcoRI, 2ul 10X Buffer, supplied by the vendor). The whole volume (20 ul) of completely digested genomic DNA was separated in 1.0% agarose gel and 1 x TAE buffer (40 mM Tris acetate, 1 mM EDTA, pH 8.0).

### **2.3.3.4 Southern transfer**

After electrophoresis, the gel was placed in a glass or plastic container and soaked in 250 mM HCl for just 10 min at room temperature with gentle shaking on a shaker. The HCl was discarded, and the gel was rinsed briefly with dH<sub>2</sub>O, and soaked in denaturing solution (0.5 M NaOH, 1.5 M NaCl dissolved in dH<sub>2</sub>O) for 45 minutes, with gentle shaking, followed by gentle shaking in neutralization solution (0.5 M Tris-Cl, 1.5 M NaCl dissolved in dH<sub>2</sub>O) for 30 minutes. A nylon membrane (MSI, Gloucester, MA) was pre-wetted in H<sub>2</sub>O for 5 minutes, then transferred to 2 X SSC (dilution from 20 X SSC, 20 x SSC: 175. 3 g NaCl, 88.

2 g Na Citrate (Citric Acid Trisodium Salt), pH7.0, total volume 1L, autoclave) for several minutes. The DNA was transferred onto nylon membrane in 10X SSC by capillary action, overnight. Following transfer the nylon membrane was treated with a UV crosslinker (STRATAGENE, UV Strata linker 1800) to fix the DNA on the membrane.

### **2.3.3.5 Hybridization**

The membrane was prehybridized with the prehybridization solution (5X SSC, 0.1% of N-lauroylsarcosine, 0.02% of SDS, 2% of blocking reagent, and 50% of formamide) (Roche) for 3 hours and hybridized with the DNA probe overnight. Both prehybridization and hybridization were performed at 42° C in a hybridization tube under high stringency conditions.

### **2.3.3.6 Detection**

Before detection, the membrane was washed at room temperature with 2X SSC (containing 0.1% SDS) for 20 minutes, 1X SSC (containing 0.1% SDS) for 20 minutes, and 0.5X SSC (containing 0.1% SDS) for 20 minutes. The membrane was then blocked with blocking buffer (0.15 M NaCl, 0.1 M Tris-HCl, pH 7.5, 2% blocking reagent purchased from Roche) for 1 hr, and incubated with anti-DIG-alkaline phosphatase (1:5,000-1:10,000, v:v, in blocking buffer) for 30 minutes. The membrane was washed with Maleic buffer (0.1 M Maleic Acid, 0.15M NaCl, pH 7.5) twice. CSPD (Roche) (1:1000, v:v, in buffer 3: 0.1 M Tris-HCl, pH 9.5, 0.1 M NaCl, 0.05 M MgCl<sub>2</sub>) was used as the substrate for alkaline phosphatase. X-ray film (Fuji) was exposed to the chemiluminescent light resulting from hybridization of DNAs with the probe. The DNA on the nylon membrane can be re-probed

2-3 times. For re-probing, previous probe was removed (“stripped”) by incubating the membrane in alkaline probe-stripping solution (0.2 N NaOH, 0.1% SDS) at 42° C for 45 min twice, and then washing thoroughly in 6 x SSC. The membrane was then air dried, and sealed in a plastic bag and stored at 4° C until the next hybridization.

### **2.3.4 RNA Extraction**

A single colony was inoculated in 3 ml TSB+ 0.7%Yeast Extract, and incubated at 37° C overnight. All 3 ml of the cells were recollected by centrifugation (11.9 rpm, 3 minutes) in a 1.5 ml Eppendorf tube. And the pellet was suspended in 180 ul enzymatic lysis buffer and incubated at 37° C for 3 hours. 20 ul proteinase K (Qiagen) was added and incubated at room temperature for 45 minutes. 1 ml TRIzol (Invitrogen, Carlsbad, CA) (use 1 ml of TRIzol per  $1.0 \times 10^7$  cells to homogenize) was added and vortexed for 3 seconds. The mixture was incubated at room temperature for 2-3 minutes, and then was centrifuged (11400 rpm, 4° C, 15 minutes). The aqueous phase (the top clear phase) was transferred to a new 1.5 ml Eppendorf tube. The middle phase (white) was the protein phase, and the bottom phase (red) was the DNA phase. 0.5 ml Isopropyl alcohol was added and mixed by hand. The mixture was incubated at room temperature for 10 minutes, and was centrifuged (11400 rpm, 4° C, 10 minutes). The supernatant was discarded and the RNA pellet was washed once with 1 ml 75% ethanol, centrifuged (8900 rpm, 4° C, 5 minutes), and air dried briefly. The RNA was dissolved in 80 ul DEPC water (0.1% Diethylpyrocarbonate in dH<sub>2</sub>O, kept at room temperature overnight and then autoclaved), and incubated at 55-56° C for 10 min. RNA was stored at -70° C for later use.

### **2.3.5 Reverse transcriptase-polymerase chain reaction (RT-PCR)**

ImProm-II<sup>TM</sup> Reverse Transcription System (Promega, Madison, WI) was used to prepare complementary DNA (cDNA).

#### **2.3.5.1 Prepare RNA target and primer**

For each 20 ul reverse transcriptase reaction, RNA template up to 1 ug, primer (20 pmol/0.5 ug) and DEPC water (5 ul) were combined, incubated at 70° C for 5 minutes, and was quick-chilled at 4° C for 5 minutes and held on ice.

#### **2.3.5.2 Prepare reverse transcriptase mix**

For each 20 ul reverse transcriptase reaction, ImProm-II<sup>TM</sup> 5x reaction buffer (4ul), 25mM MgCl<sub>2</sub> (2.4 ul), dNTP mix (1ul), recombinant RNasin ribonuclease inhibitor (1ul) and DEPC water (5.6ul) were combined, the mixture was vortexed and 1 ul ImProm-II<sup>TM</sup> reverse transcriptase was added last.

#### **2.3.5.3 Reverse Transcription**

The template and reverse primer were add to the reaction mix. The final volume for each individual reaction was 20 ul. The reaction mix was annealed at 25° C for 5 minutes, the fist strand was extended at 37-42° C for 60 minutes, and the ImProm-II<sup>TM</sup> reverse transcriptase was heat-inactivated by incubating at 70° C for 15 minutes. cDNA was used immediately for PCR or stored at -80° C.

#### **2.3.5.4 PCR amplification**

To prepare PCR mix, the cDNA sample was subtracted by combining the

amplification reagents in a sterile 1.5 ml Eppendorf tube on ice. The components were combined in the order listed (Appendix A), and was mixed by vortexing and kept on ice.

## 2.4 Results

Five *Listeria* strains (four strains of *L. monocytogenes* and one of the non-pathogenic species *L. innocua*) with accessible genome sequencing data were included in the genomic characterization of the ECI-specific regions. The four *L. monocytogenes* strains were F2365 (serotype 4b ECI California Jalisco cheese isolate, 1985 outbreak, completely sequenced genome) (5, 11); H7858 (serotype 4b ECII 1998-1999 multistate outbreak frankfurter isolate, partially sequenced genome) (2, 11); EGD-e (serotype 1/2a, animal isolate, completely sequenced genome); and F6854 (serotype 1/2a Oklahoma turkey frankfurter isolate, 1988, partially sequenced genome) (1,11). The sequenced *L. innocua* strain was strain CLIP 11262 (serotype 6a, completely sequenced genome) (8).

BLAST analyses using the complete genome sequence for F2365 (NC\_002973) (9), revealed that the ECI-specific fragments 144, 133, 17B, and 85 that had been earlier identified (9) were internal to Open Reading Frames (ORFs) 0687, 2629, 2790, and 0327, respectively, in the genome BLAST analyses in addition showed that sequences with significant homology to these fragments could not be detected in any of the other *Listeria* genomes that had been sequenced, including the genome of H7858 which had the same serotype (serotype 4b) as F2365. Such analyses were in agreement with the earlier description of these fragments as ECI-specific (9). The location of the fragments in the genome of F2365 revealed that fragments 133, 17B and 85 were quite close to the origin of

replication, there are only 23 protein coding genes (55,622bp) from cassette 17B to the origin of replication, and all four fragments were within about  $\frac{1}{4}$  of the whole genome, the gene encoding one of the *Listeria monocytogenes* key virulence factor-listeriolysin O is also very closed to the origin of replication, listeriolysin O allows the bacterium to escape from a phagocytosis vacuole (Fig. 1). The location of these fragments in the genome appears to not be random, and their proximity to the origin of replication suggests that they may have high copy number and be involved in key cellular functions.

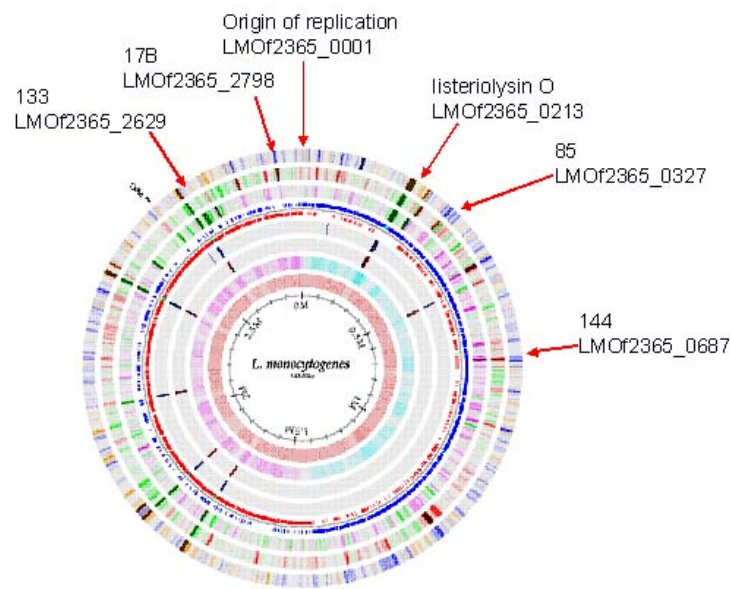


Fig. 1 Location of the ECI specific cassettes 144, 133, 17B, and 85 in the genome of the ECI strain F2365. The origin of replication and gene encoding key virulence factor Listeriolysin O are also indicated. LMOF2365\_xxxx designations are the ORF numbers of the corresponding locus, with the origin being ORF 0001. The ORF numbers in the map indicated the corresponding ORFs to the fragments 144, 133, 17B, and 85 identified in earlier study (9). Modified from ref. 9.

#### **2.4.1 Genomic organization of ECI specific regions and comparative genomic analysis among different *Listeria* genome.**

The genomic region harboring fragments 144, 133, 17B, and 85 was visualized by ARTEMIS and analyzed by BLAST in order to identify the boundaries of the F2365-specific region. This approach revealed that the regions harboring these fragments differed noticeably in terms of the amount and complexity of their ECI-specific content, as described below:

**Region harboring fragment 144.** The region harboring fragment 144 was the simplest in terms of organization, with a single ORF being specific to F2365, and being flanked by ORFs conserved among the different *L. monocytogenes* genomes as well as *L. innocua* (Fig. 2). The F2365-specific ORF (ORF 0687) was transcribed in the same direction as its upstream ORF, encoding a putative serine threonine protein phosphatase, whereas the downstream ORF was convergent to ORF 0687. Interestingly, in the non-pathogenic species *L. innocua* an unrelated ORF encoding an internalin like protein with the LPXTG motif characteristic of wall-associated proteins was identified in this region, and appeared to be unique to the genome of this non-pathogenic *Listeria*. In pathogenic *Listeria* internalins are associated with important virulence functions (10).

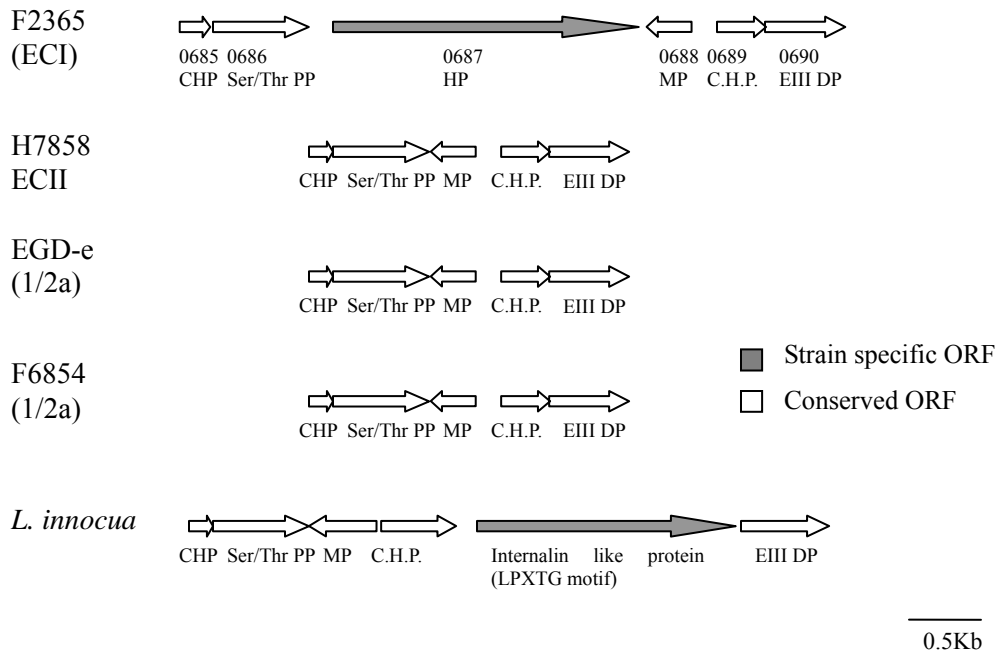


Fig. 2 Genomic region harboring fragment 144 in F2365 (ECI) and other *Listeria* genomes. CHP: Conserved hypothetical protein, HP: Hypothetical protein, Ser/Thr PP: Serine/threonine protein phosphatase, MP: Membrane protein, EIII DP: Endonuclease III domain protein. Arrows indicate direction of transcription. Gray arrows: strain specific ORFs, white arrows: conserved ORFs

**Region harboring fragment 133.** The region harboring fragment 133 was more complicated than region 144, with six ORFs being specific to F2365, and being flanked by ORFs conserved among the different *L. monocytogenes* genome as well as *L. innocua* (Fig. 3.1). The six F2365-specific ORFs were co-transcribed in the same direction. Interestingly, a 42 bp palindromic sequence flanked the ECI specific ORFs on this cassette, possibly functioning as rho- independent transcriptional terminator based on secondary structure analysis, whereas a single copy of the sequence was found in a genomically equivalent location in the other strains. The direct repeats have the potential to form a stem-loop structure ( $\Delta G = -20.4$ ) with a series of Ts at the 3' end, thus having the features of rho-independent terminator (Fig. 3.2).

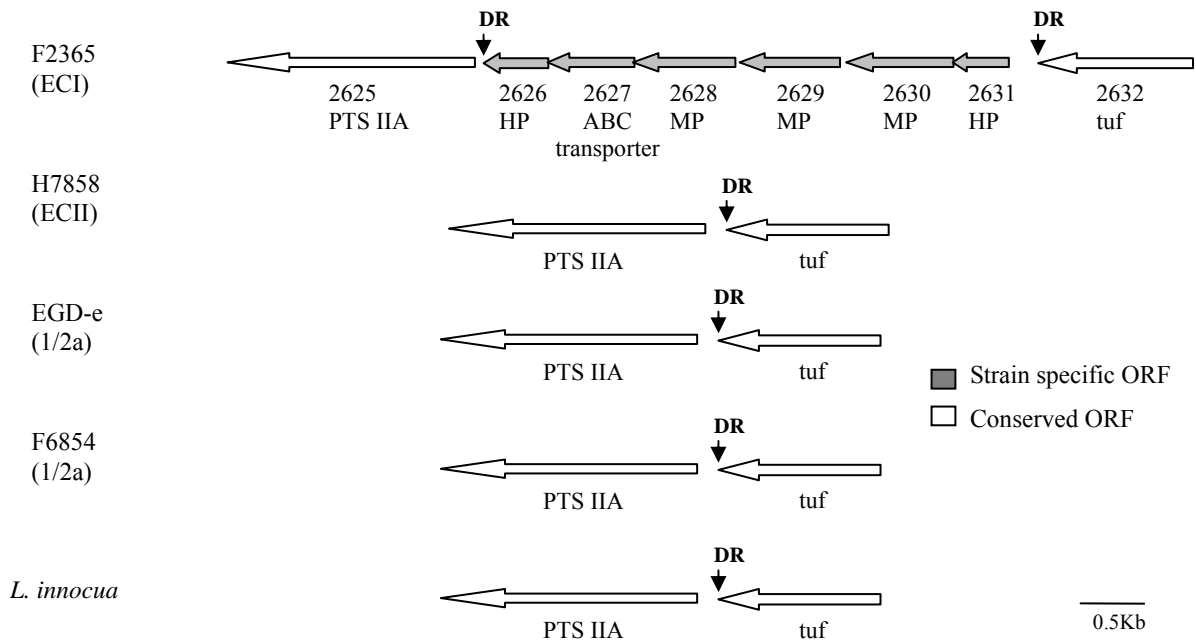


Fig. 3.1 Genomic region harboring fragment 133 in F2365 (ECI) and other *Listeria* genome. PTS IIA: putative PTS system, IIA component; HP: Hypothetical protein; MP: Membrane protein; Tuf: Translation elongation factor Tu, RF: Repetitive fragment (5' ATTCAAACCGACAAAGTCATTTGGCTTTGTTCGGTTTTTTTGT 3') Arrows indicate direction of transcription. Gray arrows: strain specific ORFs, white arrow: Conserved ORFs

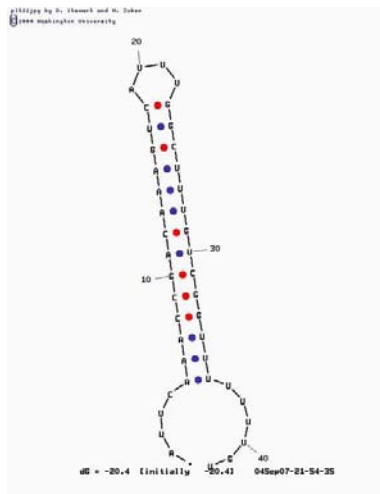


Fig. 3.2 Secondary structure of palindromic repeat in region 133.  $\Delta G = -20.4$   
Direct repeat sequence: 5' ATTCAAACCGACAAAGTCATTTGGCTTTGTTCGGTTTTTTTGT 3'

**Region harboring fragment 17B.** The region harboring fragment 17B has three ORFs being specific to F2365, and being flanked by ORFs conserved among the different *L. monocytogenes* genomes and as well as *L. innocua*. ORF2798 encodes a hypothetical protein with a conserved domain AbiF, abortive infection bacteriophage resistance protein. Abortive infection mechanism (Abi) affects prolate-headed-phage proliferation and provides resistance to bacteriophage by abortive infection. ORF2799 has a Helix-turn-helix XRE domain, suggesting that it encodes a prokaryotic DNA binding protein; it belongs to the xenobiotic response element family of transcriptional regulators. These two ORFs have opposite transcription orientations, and a large (1645 bp) intergenic sequence between them. ORF2800 has similarity with site-specific recombinase, of the phage integrase family; moreover, this region contains one or more premature stops and/or frameshifts. Interestingly, F2365 ORFs 2795, 2796 and 2797 were homologous ORFs with nucleotide similarity 68-74%, and had three or four homologues in other *L. monocytogenes* strains as well as *L. innocua*. *L. innocua* has two ORFs encoding hypothetical proteins in the same genomic location.

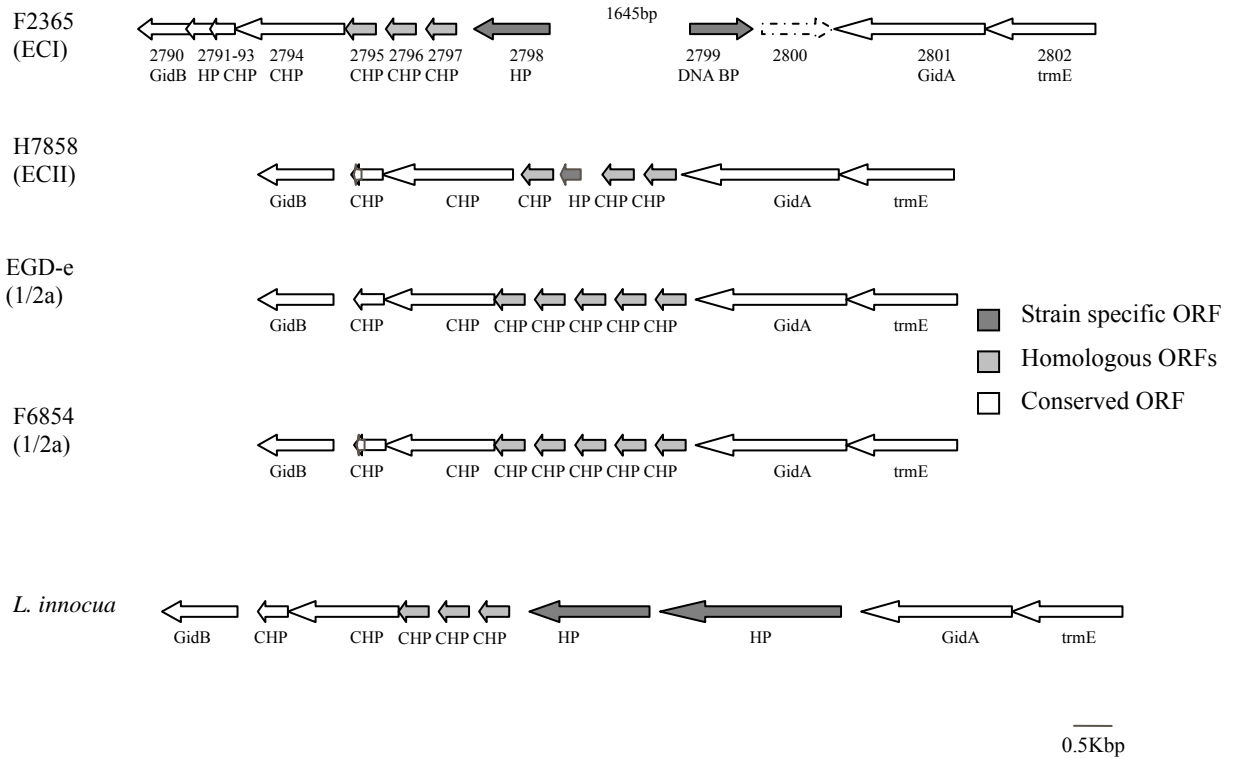


Fig. 4 Genomic region harboring fragment 17B in F2365 (ECI) and other *Listeria* genomes. HP: Hypothetical Protein; CHP: Conserved Hypothetical Protein; *Gid*: Glucose-inhibited division Protein; *trmE*: tRNA modification GTPase. Arrows indicate direction of transcription. Black arrows: strain specific ORFs, gray arrows: homologous ORFs, white arrows: conserved ORFs.

**Region harboring fragment 85.** The region harboring fragment 85 has five ORFs being specific to F2365, and being flanked by ORFs conserved among the different *L. monocytogenes* genomes and as well as *L. innocua*. There is one conserved hypothetical protein in the same place as cassette 85 in other four *Listeria* strains. *L. innocua* has one unrelated ORF encoding permease in this genome location. Comparative genomic analysis revealed five ECI specific ORFs in this region: ORF 0324 encodes a putative lipoprotein and contains premature stop codons; ORF 0325 encodes type II restriction enzyme *Sau3AI*, ORF 0326 encodes DNA-binding protein, ORF 0327 encodes C-5 cytosine-specific DNA

methylase and ORF 0328 encodes a hypothetical protein. There is an HNHc domain in ORF 0328. The domain HNHc (SMART ID: SM00507, SCOP nomenclature: HNH family) is associated with a range of DNA-binding proteins, involving in a variety of binding and cutting functions (14).

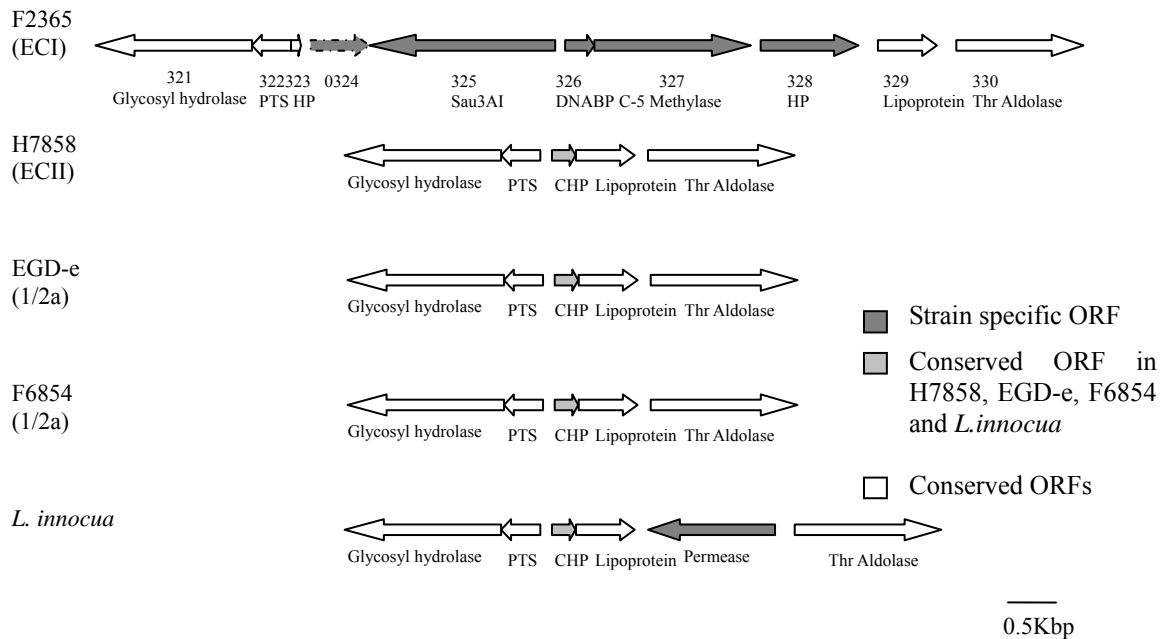


Fig. 5 Genomic region harboring fragment 85 in F2365 (ECI) and other *Listeria* genomes. HP: Hypothetical protein; PTS: PTS system, beta-glucoside-specific, IIA; ORF0324: Lipoprotein, putative, authentic frameshift with a premature stop codon in the ORF; CHP: Conserved hypothetical protein. Arrows indicate direction of transcription. Gray arrows: strain specific ORFs, white arrows: conserved ORFs.

### 2.4.2 G+C content

The G+C contents of each ECI specific region are significantly lower than *Listeria* average G+C content (38%) and the flanking ORFs. See Fig. 6-Fig. 9.

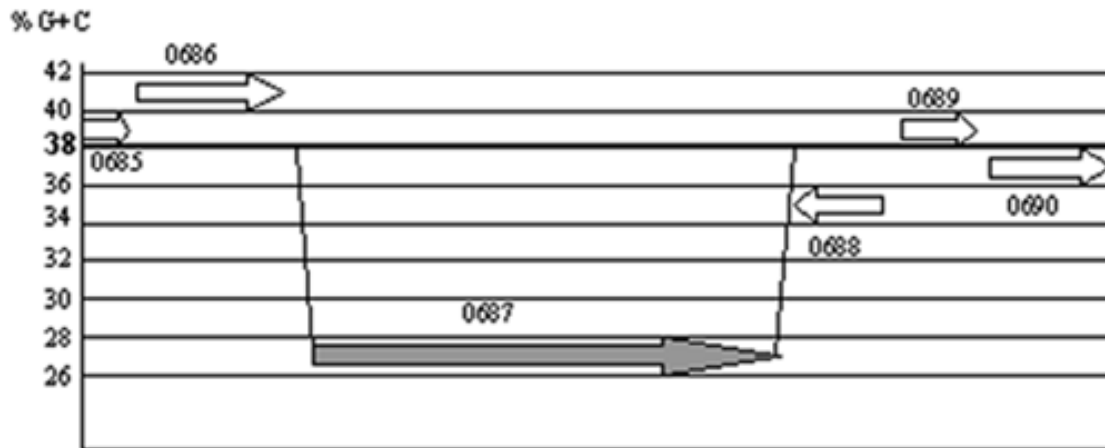


Fig. 6 ECI strain F2365 144 region G+C content map. Gray arrow: ECI specific ORF

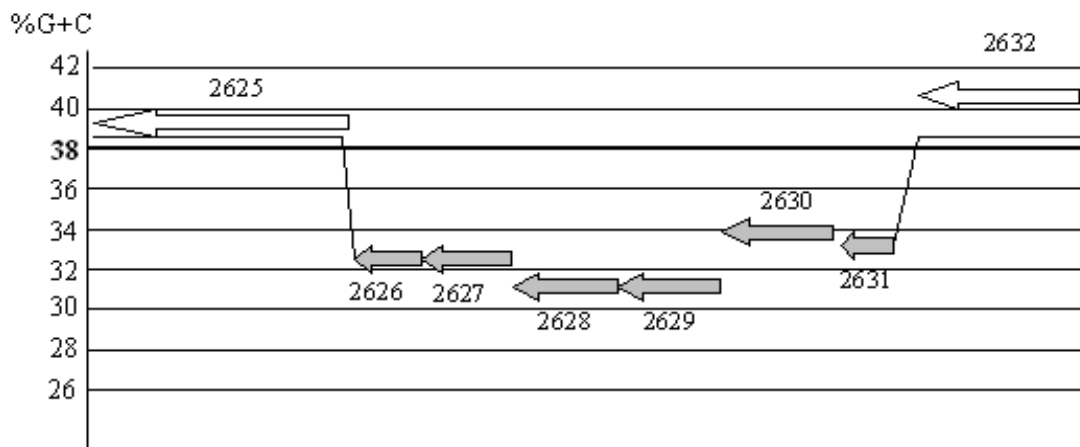


Fig. 7 ECI strain F2365 133 region G+C content map. Gray arrow: ECI specific ORFs

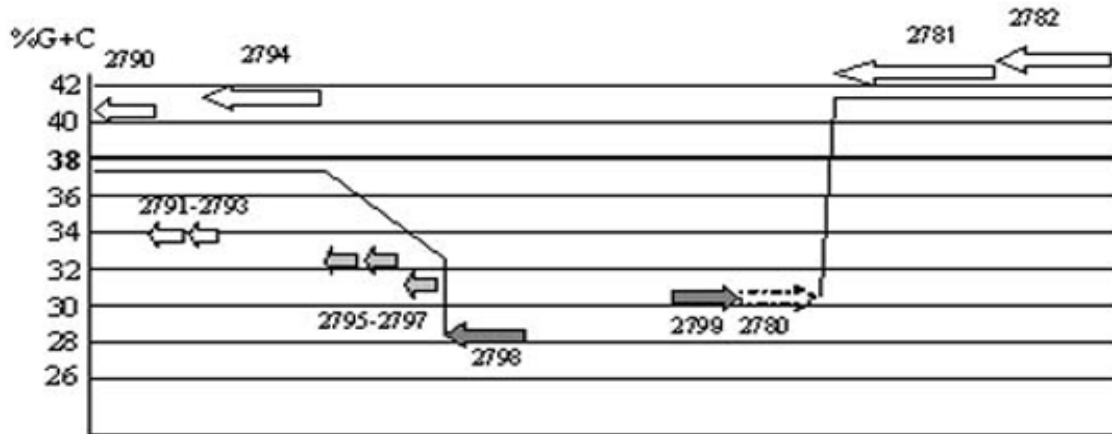


Fig. 8 ECI strain F2365 17B region G+C content map. Gray arrow: ECI specific ORFs

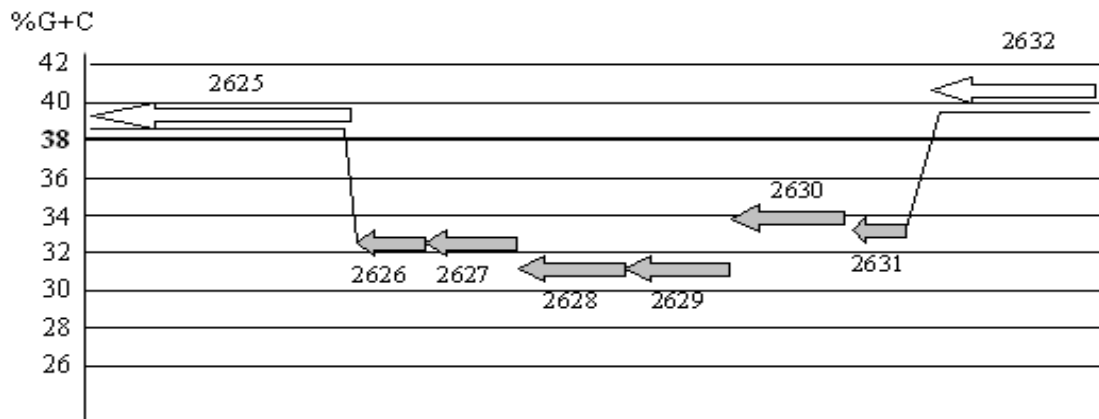


Fig. 9 ECI strain F2365 85 region G+C content map. Gray arrow: ECI specific ORFs

### 2.4.3 Southern blots

Southern blots were used to accurately determine the boundaries of the ECI specific regions 17B and 133 in multiple strains, representing different serotypes of *L. monocytogenes*. The *L. monocytogenes* strains used in southern blots are listed in table 2. Representative southern blots are shown in Figures 10 and 11, and the results are summarized in Table 3 and 4.

Table 2. Summary of southern blot results of *L. monocytogenes* and *L. innocua* with probes from three ORFs in region 133.

Probes	Strain category							
	ECI	ECIa	ECII	Sporadic 4b (266, 4b1	1/2a	<i>L. innocua</i>	1/2b	3b <sup>1</sup>
ORF2630	+	+	-	-	-	-	+	2/3+ 1/3-
ORF2628	+	+	-	-	-	-		
ORF2627	+	+	-	-	-	-	+	2/3+ 1/3-

<sup>1</sup> Of the three strains of serotype 3b that were tested, two (G3964 and G4598) reacted with probes derived from ORF2627 and ORF2630, whereas one strain (G4027) did not.

Table 3. Summary of southern blot results of *L. monocytogenes* and *L. innocua* with probes from ORFs in region 17B and flanking ORFs.

Probes	Strain category				
	ECI	ECIa	ECII	Sporadic 4b (266, 4b1	<i>L.innocua</i>
ORF2800	+	-	-	-	-
ORF2799	+	-	-	-	-
ORF2799-ORF2798	+	-	-	-	-
ORF2798	+	-	-	-	-
ORF2797	+	-	-	-	+
ORF2796	+	+	+	+	+
ORF2795	+	+	+	+	+
ORF2794	+	+	+	+	+

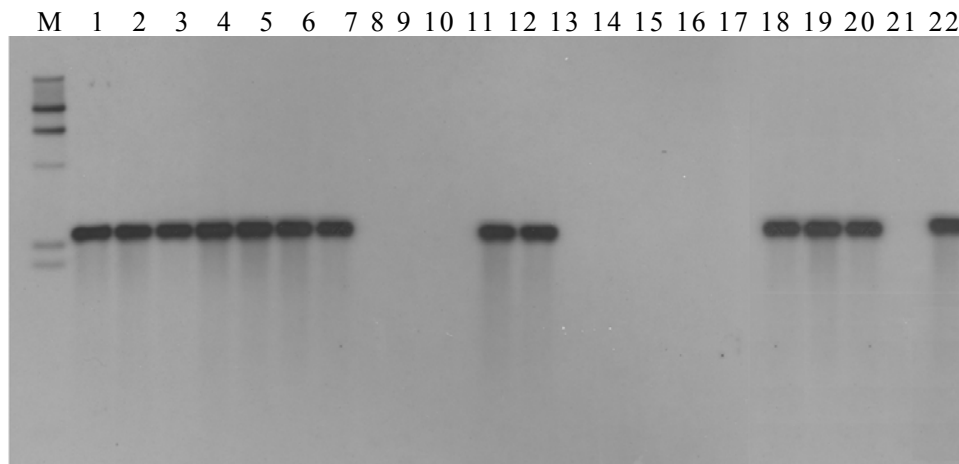


Fig. 10.1 Southern blot of EcoRI-digested DNA from different strains, probed with ORF2630 (region 133). Lanes are: 1: F2381 (ECI); 2: G4026 (ECI); 3: G3982 (ECI); 4: G3988 (ECI); 5: G3992 (ECI); 6: G4030 (ECI); 7: 264 (ECI); 8: 266 (serotype 4b); 9: 2001-126R (serotype 4b); 10: 4b1 (serotype 4b); 11: WS1 (serotype 4b); 12: 2001-7R (serotype 4b); 13: H7750 (ECII); 14: H7738 (ECII); 15: F6854 (serotype 1/2a); 16: J0161 (serotype 1/2a); 17: *L. innocua*; 18: G3986 (serotype 3b); 19: G3968 (serotype 3b); 20: G3964 (serotype 1/2b); 21: G4027 (serotype 1/2b); 22: G4598 (serotype 1/2b)

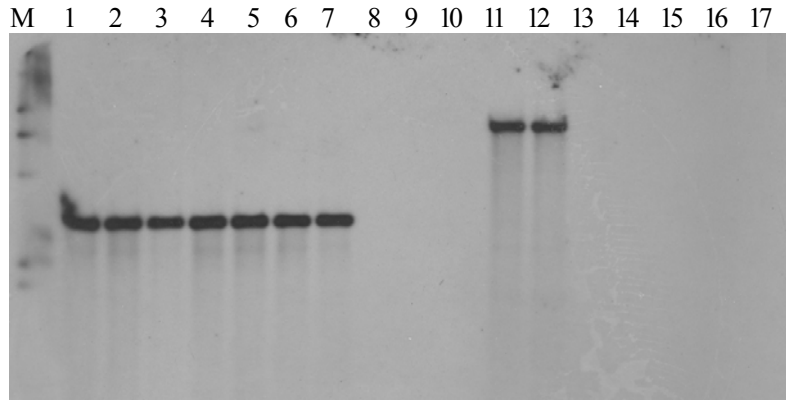


Fig. 10.2 Southern blot of EcoRI-digested DNA from different strains, probed with ORF2628 (region 133). Lanes are: 1: F2381 (ECI); 2: G4026 (ECI); 3: G3982 (ECI); 4: G3988 (ECI); 5: G3992 (ECI); 6: G4030 (ECI); 7: 264 (ECI); 8: 266 (serotype 4b); 9: 2001-126R (serotype 4b); 10: 4b1 (serotype 4b); 11: WS1 (serotype 4b); 12: 2001-7R (serotype 4b); 13: H7750 (ECII); 14: H7738 (ECII); 15: F6854 (serotype 1/2a); 16: J0161 (serotype 1/2a); 17: *L. innocua*

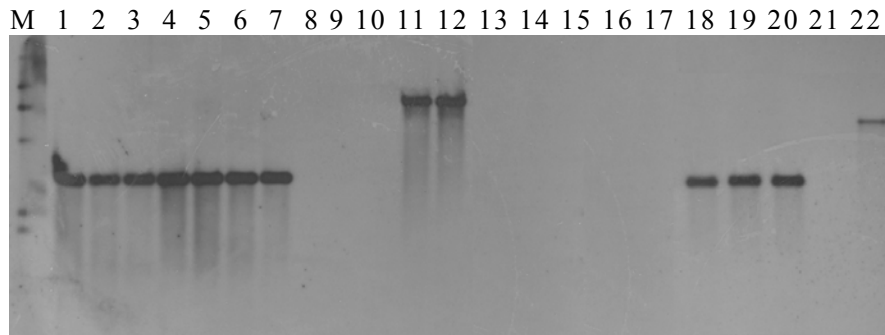


Fig. 10.3 Southern blot of EcoRI-digested DNA from different strains, probed with ORF2627 (region 133). Lanes are: 1: F2381 (ECI); 2: G4026 (ECI); 3: G3982 (ECI); 4: G3988 (ECI); 5: G3992 (ECI); 6: G4030 (ECI); 7: 264 (ECI); 8: 266 (serotype 4b); 9: 2001-126R (serotype 4b); 10: 4b1 (serotype 4b); 11: WS1 (serotype 4b); 12: 2001-7R (serotype 4b); 13: H7750 (ECII); 14: H7738 (ECII); 15: F6854 (serotype 1/2a); 16: J0161 (serotype 1/2a); 17: *L. innocua*; 18: G3986 (serotype 3b); 19: G3968 (serotype 3b); 20: G3964 (serotype 1/2b); 21: G4027 (serotype 1/2b); 22: G4598 (serotype 1/2b)

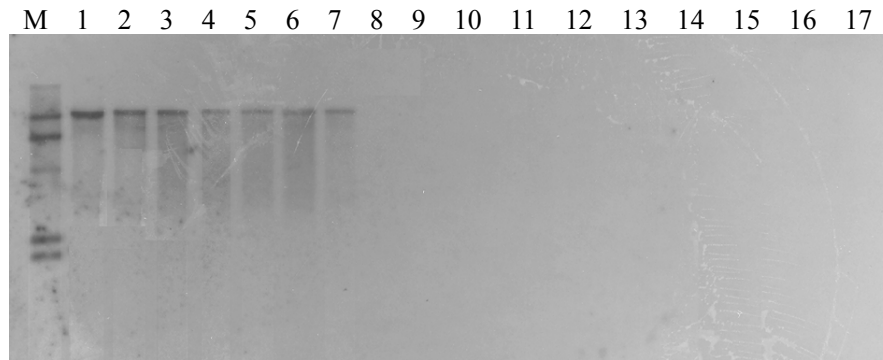


Fig. 11.1 Southern blot of EcoRI-digested DNA from different strains, probed with ORF2800 (region 17B). Lanes are: 1: F2381 (ECI); 2: G4026 (ECI); 3: G3982 (ECI); 4: G3988 (ECI); 5: G3992 (ECI); 6: G4030 (ECI); 7: 264 (ECI); 8: 266 (serotype 4b); 9: 2001-126R (serotype 4b); 10: 4b1 (serotype 4b); 11: WS1 (serotype 4b); 12: 2001-7R (serotype 4b); 13: H7750 (ECII); 14: H7738 (ECII); 15: F6854 (serotype 1/2a); 16: J0161 (serotype 1/2a); 17: *L. innocua*

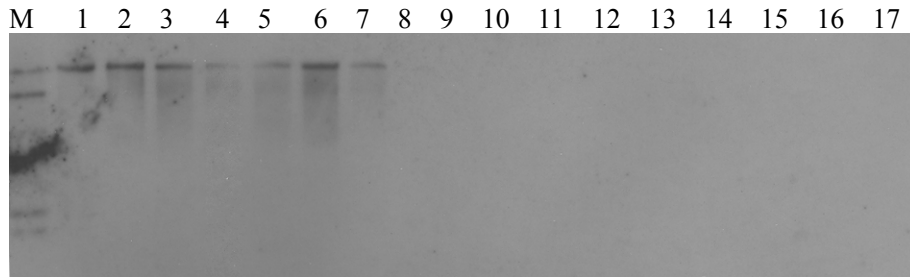


Fig. 11.2 Southern blot of EcoRI-digested DNA from different strains, probed with ORF2799 (region 17B). Lanes are: 1: F2381 (ECI); 2: G4026 (ECI); 3: G3982 (ECI); 4: G3988 (ECI); 5: G3992 (ECI); 6: G4030 (ECI); 7: 264 (ECI); 8: 266 (serotype 4b); 9: 2001-126R (serotype 4b); 10: 4b1 (serotype 4b); 11: WS1 (serotype 4b); 12: 2001-7R (serotype 4b); 13: H7750 (ECII); 14: H7738 (ECII); 15: F6854 (serotype 1/2a); 16: J0161 (serotype 1/2a); 17: *L. innocua*

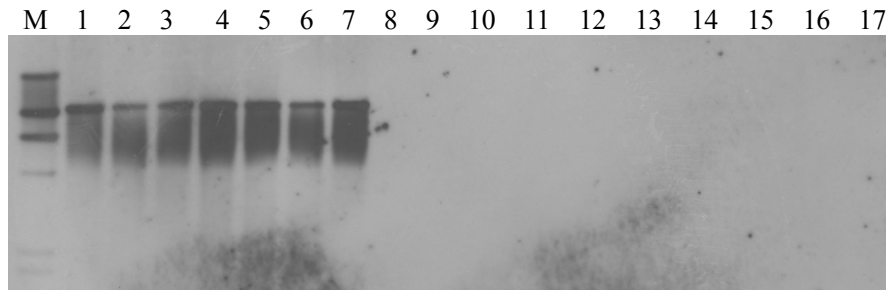


Fig. 11.3 Southern blot of EcoRI-digested DNA from different strains, probed with intergenic sequence between ORFs 2798 and 2799 (region 17B). Lanes are: 1: F2381 (ECI); 2: G4026 (ECI); 3: G3982 (ECI); 4: G3988 (ECI); 5: G3992 (ECI); 6: G4030 (ECI); 7: 264 (ECI); 8: 266 (serotype 4b); 9: 2001-126R (serotype 4b); 10: 4b1 (serotype 4b); 11: WS1 (serotype 4b); 12: 2001-7R (serotype 4b); 13: H7750 (ECII); 14: H7738 (ECII); 15: F6854 (serotype 1/2a); 16: J0161 (serotype 1/2a); 17: *L. innocua*

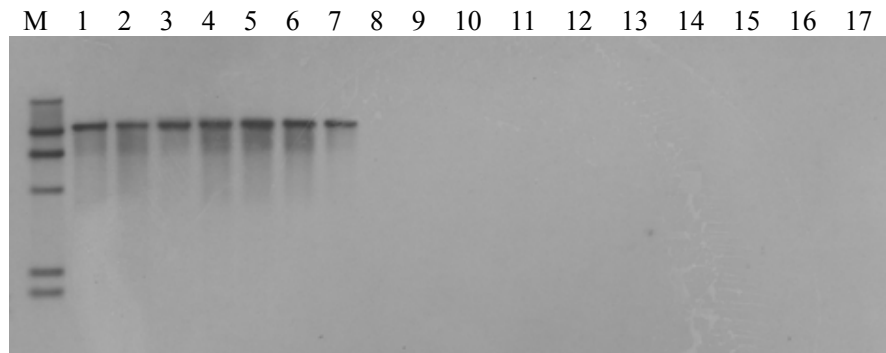


Fig. 11.4 Southern blot of EcoRI-digested DNA from different strains, probed with ORF 2798 (region 17B). Lanes are: 1: F2381 (ECI); 2: G4026 (ECI); 3: G3982 (ECI); 4: G3988 (ECI); 5: G3992 (ECI); 6: G4030 (ECI); 7: 264 (ECI); 8: 266 (serotype 4b); 9: 2001-126R (serotype 4b); 10: 4b1 (serotype 4b); 11: WS1 (serotype 4b); 12: 2001-7R (serotype 4b); 13: H7750 (ECII); 14: H7738 (ECII); 15: F6854 (serotype 1/2a); 16: J0161 (serotype 1/2a); 17: *L. innocua*

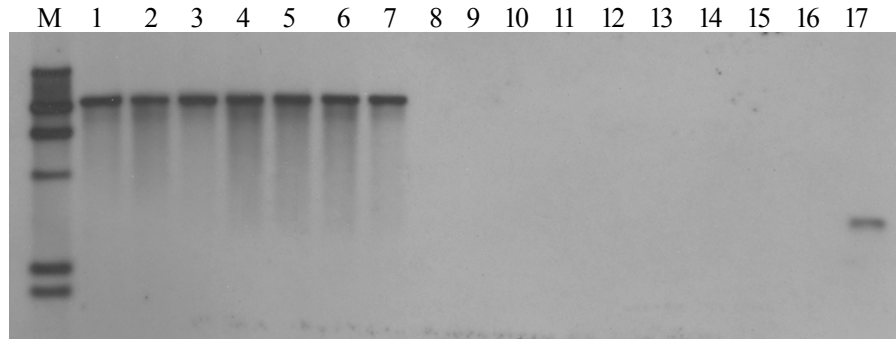


Fig. 11.5 Southern blot of EcoRI-digested DNA from different strains, probed with ORF 2797 (region 17B). Lanes are: 1: F2381 (ECI); 2: G4026 (ECI); 3: G3982 (ECI); 4: G3988 (ECI); 5: G3992 (ECI); 6: G4030 (ECI); 7: 264 (ECI); 8: 266 (serotype 4b); 9: 2001-126R (serotype 4b); 10: 4b1 (serotype 4b); 11: WS1 (serotype 4b); 12: 2001-7R (serotype 4b); 13: H7750 (ECII); 14: H7738 (ECII); 15: F6854 (serotype 1/2a); 16: J0161 (serotype 1/2a); 17: *L. innocua*

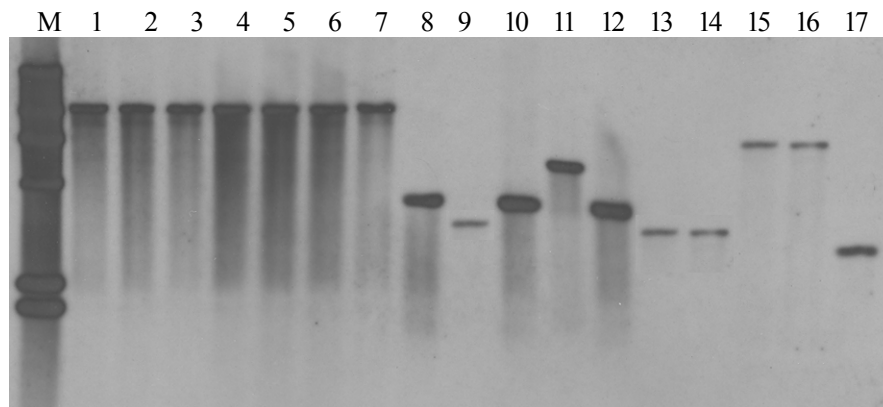


Fig. 11.6 Southern blot of EcoRI-digested DNA from different strains, probed with ORF 2796 (region 17B). Lanes are: 1: F2381 (ECI); 2: G4026 (ECI); 3: G3982 (ECI); 4: G3988 (ECI); 5: G3992 (ECI); 6: G4030 (ECI); 7: 264 (ECI); 8: 266 (serotype 4b); 9: 2001-126R (serotype 4b); 10: 4b1 (serotype 4b); 11: WS1 (serotype 4b); 12: 2001-7R (serotype 4b); 13: H7750 (ECII); 14: H7738 (ECII); 15: F6854 (serotype 1/2a); 16: J0161 (serotype 1/2a); 17: *L. innocua*

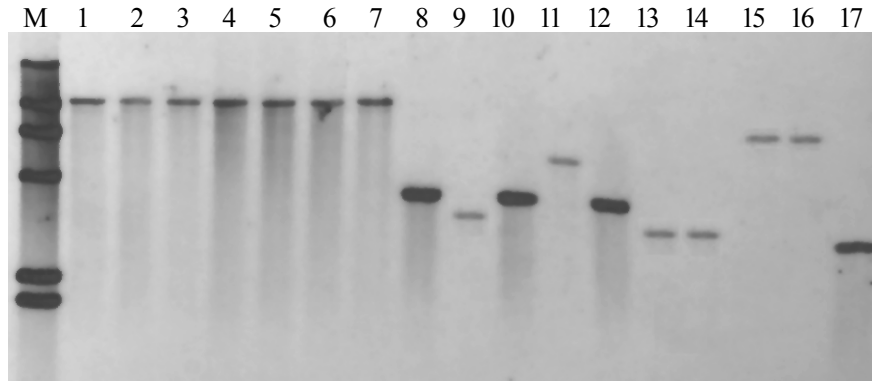


Fig. 11.7 Southern blot of EcoRI-digested DNA from different strains, probed with ORF 2795 (region 17B). Lanes are: 1: F2381 (ECI); 2: G4026 (ECI); 3: G3982 (ECI); 4: G3988 (ECI); 5: G3992 (ECI); 6: G4030 (ECI); 7: 264 (ECI); 8: 266 (serotype 4b); 9: 2001-126R (serotype 4b); 10: 4b1 (serotype 4b); 11: WS1 (serotype 4b); 12: 2001-7R (serotype 4b); 13: H7750 (ECII); 14: H7738 (ECII); 15: F6854 (serotype 1/2a); 16: J0161 (serotype 1/2a); 17: *L. innocua*

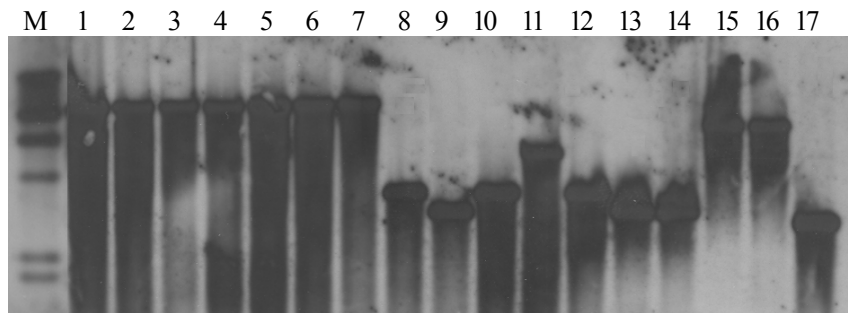


Fig. 11.8 Southern blot of EcoRI-digested DNA from different strains, probed with ORF 2794 (region 17B). Lanes are: 1: F2381 (ECI); 2: G4026 (ECI); 3: G3982 (ECI); 4: G3988 (ECI); 5: G3992 (ECI); 6: G4030 (ECI); 7: 264 (ECI); 8: 266 (serotype 4b); 9: 2001-126R (serotype 4b); 10: 4b1 (serotype 4b); 11: WS1 (serotype 4b); 12: 2001-7R (serotype 4b); 13: H7750 (ECII); 14: H7738 (ECII); 15: F6854 (serotype 1/2a); 16: J0161 (serotype 1/2a); 17: *L. innocua*

## 2.4.4 RT-PCR analysis

### 2.4.4.1 RT-PCR results for region 144

The results of RT-PCR suggested that conserved ORFs ORF0685 and ORF0686 were co-transcribed with ORF0687, which was ECI-specific ORF. See Fig. 12.1, Fig. 12.2 and Fig. 12.3.

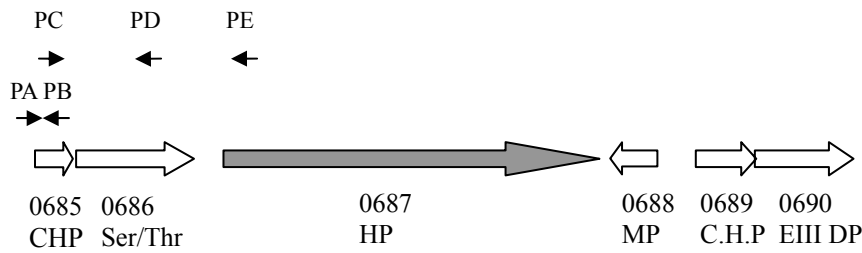


Fig. 12.1 Map of region 133 in strain F2365. Arrows indicate location of primers used in RT-PCR. Abbreviations for deduced polypeptides are as in Fig. 2.

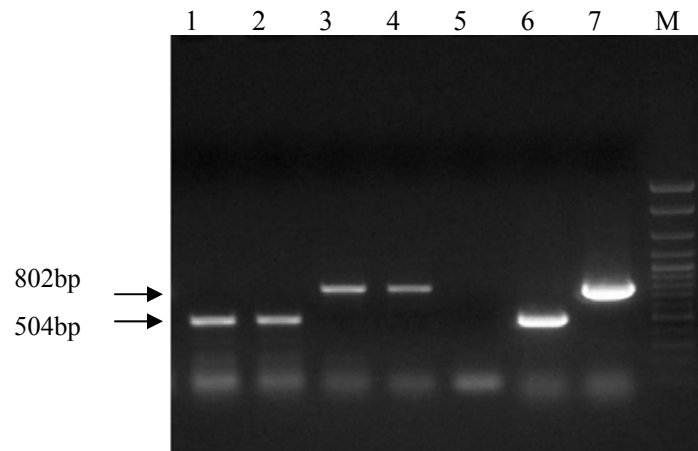


Fig. 12.2 RT-PCR results for ECI strain F2365. Primer PB was used to synthesize cDNA, and primers PA and PB were used to detect ORF0685 transcripts. Lanes: 1, PAPB (10ul cDNA); 2, PAPB (5ul cDNA); 3, PCPD (10ul cDNA); 4, PCPD (5ul cDNA); 5, RT-PCR negative control; 6, PAPB (DNA template); 7, PCPD (DNA template).

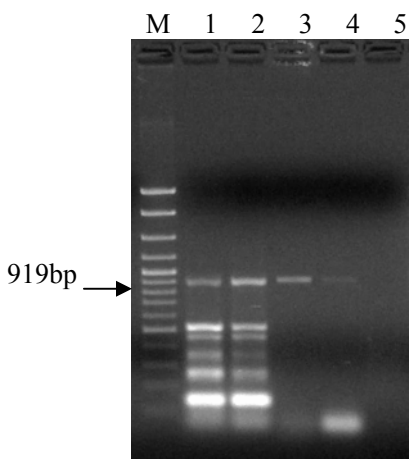


Fig. 12.3 RT-PCR results for ECI strain F2365. Primer PE was used to synthesize cDNA, and primers PC and PE were used to detect transcripts of ORF0685, ORF0686 and ORF0687. Lanes: 1, PCPE (10ul cDNA); 2, PCPE (5ul cDNA); 3, PCPE (10 ul cDNA, Self-priming); 4, PCPE (5ul cDNA, Self-priming); 5, negative control.

Self-priming: no primer was added to synthesize cDNA. The self-priming of the RNA could be explained as a result of secondary structure configurations that provide a suitable 3'- terminus to prime the reverse transcriptase.

#### 2.4.4.2 RT-PCR results for region 133

RT-PCR results suggested that the two conserved ORFs that flank the ECI specific ORFs were not co-transcribed with the ECI specific ORFs. Three of the ECI-specific ORFs that encode putative membrane proteins were co-transcribed. However, RT-PCR analysis of the ECI-specific ORFs in region 133 yielded inconsistent results, possibly suggesting low expression of the six ECI specific ORFs, or unstable messages. Some RT-PCR results were based on self-priming. The self-priming of the RNA perhaps could be explained as a result of secondary structure configurations that provide a suitable 3'- terminus to prime the reverse transcriptase. Suggested that all six ECI specific ORFs in this region were co-transcribed (Fig. 13.1-Fig. 13.4).

For the strains that have not been sequenced lacked these ORFs, regular PCR and RT-PCR was done to serotype 4b strains H7750, 4b1, 266 and serotype 1/2a strain F6854

with primers Pm and Pn. The results confirmed that these four strain did not harbor ECI-specific ORFs, and showed that the two conserved ORFs, which encode PTS system IIA component and translation elongation factor respectively, were co-transcribed (Data not shown).

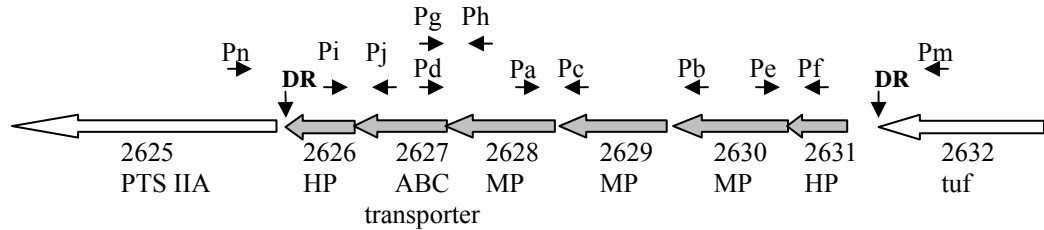


Fig. 13.1 Map of region 133 strain F2365 133 region map. Arrows indicate the primers used in RT-PCR. Abbreviations for deduced polypeptides are as in Fig. 3.

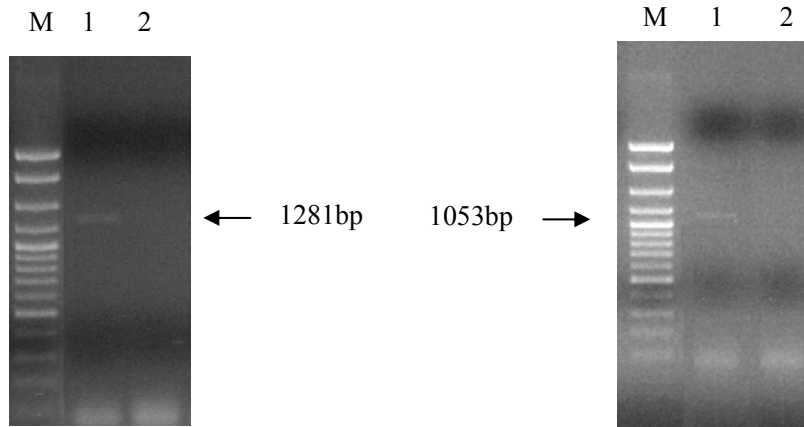


Fig. 13.2 RT-PCR results for ECI strain F2365. Primer Pa was used to construct cDNA and primers Pa and Pb were used to detect transcription of ORF 2628, ORF2629 and ORF2630. Lanes: 1, PaPb; 2, PaPb negative control

Fig. 13.3 RT-PCR results for ECI strain F2365. Primer Pd was used to construct cDNA and primers Pc and Pd were used to detect transcription of ORF2627, ORF2628 and ORF2629. 1: PcPd, 2: PcPd negative control

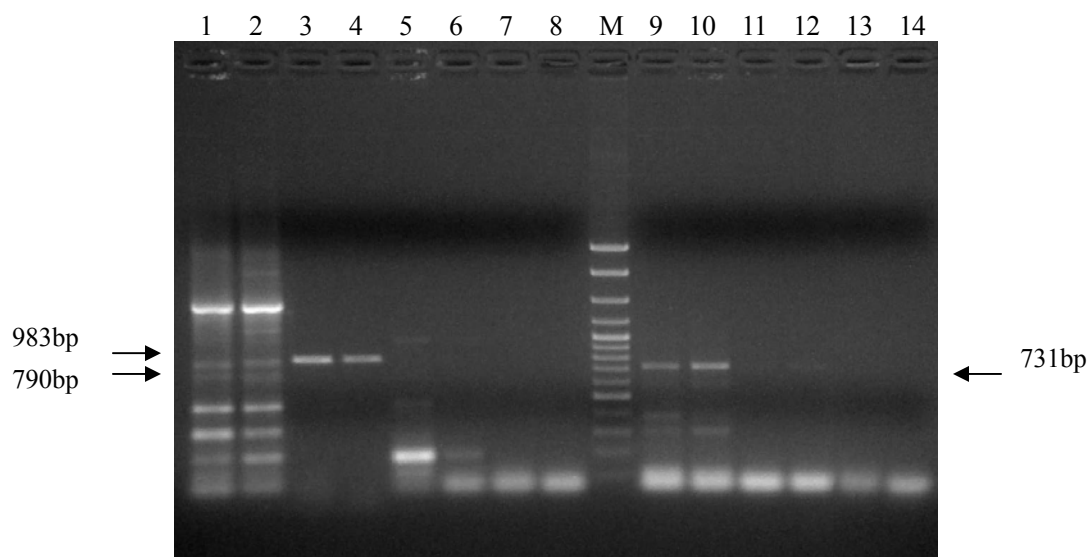


Fig. 13.4 RT-PCR results for ECI strain F2365. Primer Pe was used to construct cDNA and primers Pe and Pf were used to detect transcription of ORF2630 and ORF2631; Primer Pg was used to construct cDNA, and primers Pg and Ph were used to detect transcription of ORF2627 and ORF2628; Primer Pi was used to construct cDNA, and primers Pi and Pj were used to detect transcription of ORF2626 and ORF2627. 1: PePf (10ul cDNA), 2: PePf (5ul cDNA), 3: PePf (Self-priming, 10ul cDNA), 4: PePf (Self-priming, 5ul cDNA), 5: PgPh (10ul cDNA), 6: PgPh (5ul cDNA), 7: PgPh (Self-priming, 10ul cDNA), 8: PgPh (Self-priming 5ul cDNA), 9: PiPj (10ul cDNA), 10: PiPj (5ul cDNA), 11: PiPj (Self-priming, 10ul cDNA), 12: PiPj (Self-priming 5ul cDNA), 13: PePf negative control, 14: PgPh negative control.

#### 2.4.4.3 RT-PCR analysis of region 17B

The RT-PCR results suggested that two highly conserved ORFs, ORF 2802 and ORF 2801, which encode tRNA modification GTPase (TrmE) and glucose-inhibited division protein A (GidA), respectively, were co-transcribed (Fig. 14.5). The conserved ORFs from ORF 2797 to 2790, with ORF 2790 encoding glucose-inhibited division protein B (GidB), were also co-transcribed (Fig. 14.3, Fig. 14.4). In the ECI specific region, ORF 2800 and ORF 2799 were co-transcribed, whereas ORF 2798 was transcribed by itself (Fig. 14.2, Fig. 14.3). RT-PCR done for serotype 1/2a strain F6854 and serotype 4b strains H7858, 4b1 and 266 showed that the ORF encoding GidA protein was co-transcribed with the homologous

ORFs (Data not shown).

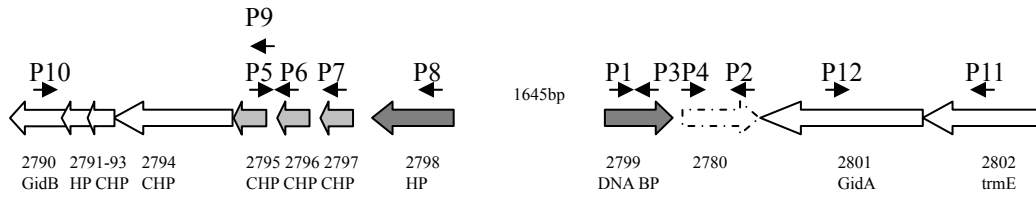


Fig 14.1 Map of ECI specific region 17B in strain F2365. Black arrows indicate location of primers used in RT-PCR. Abbreviations for deduced polypeptides are as in Fig. 4.

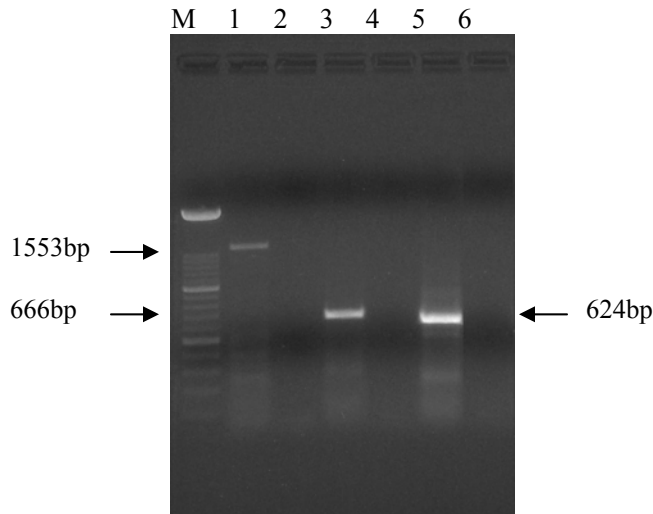


Fig. 14.2 RT-PCR results for ECI strain F2365. Primer P2 was used to construct cDNA, and primers P1 and P2 were used to detect transcription of ORF2799 and ORF2780; Primer P3 was used to construct cDNA, and primers P1 and P3 were used to detect transcription of ORF2799; Primer P2 was used to construct cDNA, and primers P2 and P4 were used to detect transcription of ORF2780. 1: P1P2, 2: P1P2 negative control, 3: P1P3, 4: P1P3 negative control, 5: P2P4, 6: P2P4 negative control.

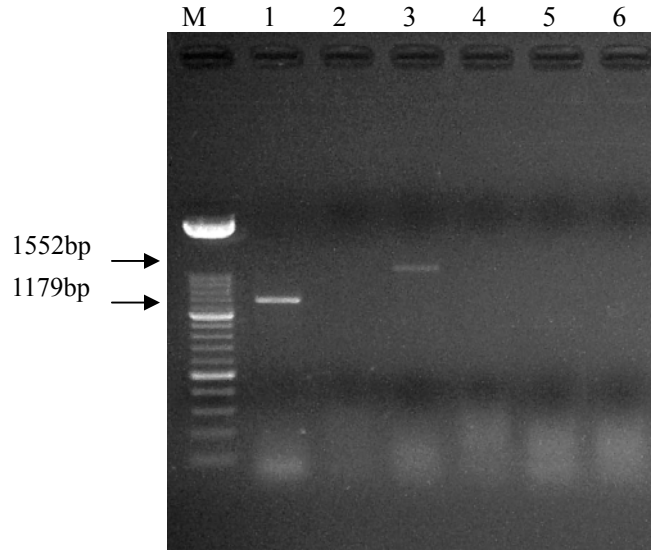


Fig. 14.3 RT-PCR results for ECI strain F2365. Primer P5 was used to construct cDNA, and primers P5 and P6 were used to detect transcription of ORF2795 and ORF2796; Primer P5 was used to construct cDNA, and primers P5 and P7 were used to detect transcription of ORF2795, ORF2796 and ORF2797; Primer P5 was used to construct cDNA, and primers P5 and P8 were used to detect transcription of ORF2795, ORF2796, ORF2797 and ORF2798. 1: P5P6, 2: P5P6 negative control, 3: P5P7, 4: P5P7 negative control, 5: P5P8, 6: negative control.

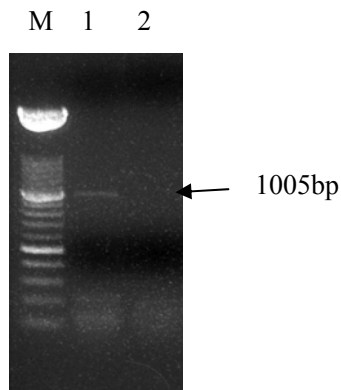


Fig. 14.4 RT-PCR results for ECI strain F2365. Primer P10 was used to construct cDNA and primers P9 and P10 were used to detect transcription of ORF2795 to ORF2790. 1: P9P10, 2: P9P10 negative control.

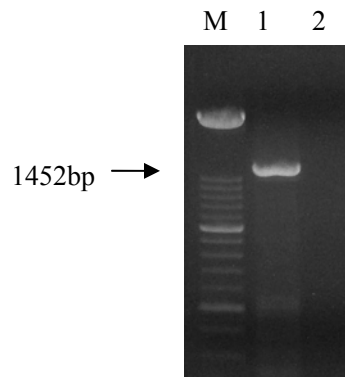


Fig. 14.5 RT-PCR result for ECI strain F2365. Primer P12 was used to construct cDNA and primers P11 and P12 were used to detect transcription of ORF2801 and ORF2802. 1: P11P12, 2: P11P12 negative control.

## 2.5 Discussion.

ECI specific fragments can also be referred to as ECI specific gene cassettes. The G+C content of these cassettes was noticeably lower than the average for the genome of *Listeria*, suggesting that the cassettes represent genomic islands (GEIs) that have been acquired by ECI strains by horizontal transfer. GEIs can have quite large size, often >10Kb , and are a prominent feature of bacterial genomes(6). Some of the GEIs may involve in bacterial pathogenicity and are called pathogenicity islands (PAIs), some may play a role in secretion (secretion islands), some confer resistance to antibiotics, (resistance islands), and some may be important in metabolism (metabolic islands).

Methods for identifying genomic islands can be divided into two groups: those for which bacterial isolation is required, and those for which it is not required. Examples of those for which bacterial isolation and cultivation is required include: (a) Genome comparison by subtractive hybridization; (b) Genome comparison by DNA–DNA hybridization, based on survey of known genomic regions; (c) tRNA screening; (d) Detection of horizontally acquired genetic information by identification of repeat structures, transposons and insertion sequence (IS), G+C content; (e) Island probing approach (3). Methods for which bacterial isolation and cultivation are not required include phenotypic screening of (meta-) genomic libraries for traits that are indicative of mobile and transferable DNA elements (3).

The ECI regions that were analyzed varied in complexity. Region 144 represented the simplest situation among the ECI specific regions; it has only one ECI specific Open Reading Frame (ORF), which encodes a hypothetical protein, whereas all other strains lacked

this ORF. Interestingly, the genome of the non-pathogenic species *L. innocua* harbored a unique gene encoding an internalin like protein in the vicinity of this region. Internalins are very important in virulence of pathogenic *Listeriae*, and this putative internalin gene in *L. innocua* may suggest that this genomic region harbored virulence-related genes in an ancestral listerial lineage that preceded the differentiation between *L. monocytogenes* and *L. innocua*.

The average G+C contents of *L. monocytogenes* strains F2365, H7858, EGD-e, F6854 and *L. innocua* are 38%, 38%, 39%, 37.8% and 37%. For 144 ECI specific region, the G+C content is about 26%, much lower than the average ECI G+C content. This unusual G+C content is a feature of pathogenicity islands, and suggests that this ECI specific region might have been horizontally transferred and may act as a virulence determinant in ECI strains.

There is a conserved domain-COG4594 in this ORF, the function of which is unknown. The protein BLAST result showed that the deduced polypeptide was similar (about 25% identity) with Rlo proteins of *Campylobacter jejuni*. In *C. jejuni* several *rlo* genes (*rloA* through *rloH*) are located between *hsdR* and *hsdS* (type I restriction-modification systems) (15).

RT-PCR results suggested that the ECI-specific ORF in region 144 was co-transcribed with two upstream ORFs, one of which encodes a serine/threonine protein phosphatase. Serine/threonine protein phosphatases catalyze the dephosphorylation of

phosphoserine and phosphothreonine residues and are involved in signal transduction.

The other three ECI specific cassettes that were analyzed were more complicated than region 144. Region 133 harbored six ECI specific ORFs, encoding two hypothetical proteins, three putative membrane proteins and one ABC transporter. The six ECI specific ORFs were flanked by palindromic direct repeats and by two conserved ORFs, encoding translation elongation factor Tu and putative PTS system IIA component, respectively. Comparative analysis of the sequenced genomes of other strains of *Listeria* showed no ORF between the translation elongation factor Tu and the putative PTS system IIA component, but all had one copy of the palindromic sequence between these two conserved ORFs.

The G+C content of the six ECI specific ORFs was about 32%, lower than the average for ECI strains (38%), and much lower than the two conserved flanking ORFs, whose G+C content was above 42%. This feature suggests the possibility that this region may be horizontally transferred from other organisms.

Protein BLAST failed to identify any proteins with significant similarity to the two hypothetical proteins. The ABC transporter was similar to *Bacillus cereus* G9241 the bacitracin transport ATP-binding protein BcrA (38% identity, 61% positives). BcrA is a part of the binding protein-dependent transport system for bacitracin that confers resistance to this antibiotic, possibly by mediating energy coupling to the transport system (12). *B. licheniformis*, a bacitracin producer, has an ABC transporter system which is hypothesized to pump out bacitracin for self-protection. The transporter is composed of two membrane

proteins, BcrB and BcrC, and two identical ATP-binding subunits of efflux transporter, BcrA (19). One of the membrane proteins-ORF2630 was similar to *Bacillus thuringiensis serovar israelensis* ATCC 35646 hypothetical membrane spanning protein (24% identities, 49% positives); the other two putative membrane proteins lacked any conserved domains. The sequence analysis results suggest that region 133 may be involved in antibiotic resistant mechanisms, possibly representing a resistance island.

The southern blot results suggested that region 133 is indeed present in all ECI strains that we screened, but could also be detected in a recently identified epidemic lineage of *L. monocytogenes* serotype 4b, implicated in an outbreak in Winston Salem, N. Carolina, in 2001. Other serotype 4b strains and strains of serotype 1/2a, as well as *L. innocua*, lacked this region. However, certain 1/2b and 3b strains also harbored the region. These findings were in agreement with the results of Herd and Kocks, who reported that fragment 133 in this region was unique to ECI and to strains of serotype 1/2b and 3b (ref). RT-PCR result suggested that the six ORFs in this region constituted a transcriptional unit separate from those in the flanking regions. This result is consistent with the hypothesis that the direct repeats might serve as terminators in this region.

17B region is the most complicated, as it has three ECI specific ORFs, in different transcriptional orientations. The presence of a putative phase integrase suggests that the ECI-specific genes in this region may have been horizontally transferred by a phage. Phage-mediated transfer is common in acquisition of pathogenicity islands (6). After the genes enter the new host cells, two genetic processes are important: stabilization of the new

elements, and optimized expression of the newly acquired genes. The high number of mutations often leads to stop codons in the ‘mobility’ genes of pathogenicity islands, such as integrases, origins of replication in plasmids or IS elements, thus, making the newly generated organisms stable and avoiding the loss of the new genes (13). In the ECI specific region, ORF2800 and ORF2799 were co-transcribed, and ORF2798 was transcribed by itself. It is of interest that ORF2800 was transcribed, considering that it harbors the premature stop codons, and likely not translated. ORF2801 and ORF2802 were co-transcribed, encoding glucose-inhibited division protein A (GidA) and tRNA modification GTPase (TrmE). The *gidA* genes are well conserved among prokaryotes and eukaryotes and are generally localized near the chromosomal replication origin. GidA has been identified to be involved in cell division (1, 2, 11); and involved in the biosynthesis of the hyper-modified nucleotide 5-methylaminomethyl-2-thiouridine which is found in the wobble position of bacterial tRNAs, and this modification is necessary to stabilize codon-anticodon interactions(4). RT-PCR done for serotype 1/2a strain F6854 and ECII strain H7858 showed that the ORF encoding GidA protein was co-transcribed with homologous ORFs. This suggests that all these conserved ORFs in region 17B were transcribed together originally, and that the ECI specific sequences were horizontally transferred from other organisms. *L. innocua* also acquired certain specific sequences in this region. The various G+C contents of ORFs in cassette 17B indicated that they might have been transferred from different organisms at different time.

Region 85 has been identified in previous investigations, and remains the only ECI specific cassette with a known function (9, 21). The function of this ECI specific region is to methylate cytosines at GATC sites, thus make ECI strains resistant to digestion by the

restriction enzyme Sau3AI. This is one of the key mechanisms of bacteria for self-protection against phage invasion (17). There is one conserved hypothetical protein in the same place as cassette 85 in other four *Listeria* strains, suggesting that this conserved hypothetical protein might have been replaced by cassette 85 in ECI strains.

In conclusion, the comparative genomic analysis revealed that the four ECI specific cassettes investigated here differ in complexity and gene content, but have certain common features: all four cassettes harbor ORFs with G+C content significantly lower than the average for *Listeria*; in at least two cases (regions 133 and 17B) the cassettes appear to have been inserted in genomic regions which are involved in key cellular functions (e.g. TuF); and all four cassettes are located in a well-defined portion of the genome, in relative proximity to the origin of replication. Further studies are needed to elucidate the mechanisms that have mediated acquisition and maintenance of these cassettes in ECI strains.

## **2.6 Reference:**

1. **Asai, T., D. B. Bates, E. Boye, and T. Kogoma.** 1998. Are minichromosomes valid model systems for DNA replication control? Lessons learned from *Escherichia coli*. *Mol Microbiol* **29**:671-5.
2. **Bates, D. B., E. Boye, T. Asai, and T. Kogoma.** 1997. The absence of effect of *gid* or *mioC* transcription on the initiation of chromosomal replication in *Escherichia coli*. *Proc Natl Acad Sci U S A* **94**:12497-502.
3. **Bibb, W. F., B. G. Gellin, R. Weaver, B. Schwartz, B. D. Plikaytis, M. W. Reeves, R. W. Pinner, and C. V. Broome.** 1990. Analysis of clinical and food-borne isolates of *Listeria monocytogenes* in the United States by multilocus enzyme electrophoresis

- and application of the method to epidemiologic investigations. *Appl Environ Microbiol* **56**:2133-41.
4. **Bregeon, D., V. Colot, M. Radman, and F. Taddei.** 2001. Translational misreading: a tRNA modification counteracts a +2 ribosomal frameshift. *Genes Dev* **15**:2295-306.
  5. **Buchrieser, C., R. Brosch, B. Catimel, and J. Rocourt.** 1993. Pulsed-field gel electrophoresis applied for comparing *Listeria monocytogenes* strains involved in outbreaks. *Can J Microbiol* **39**:395-401.
  6. **Dobrindt, U., B. Hochhut, U. Hentschel, and J. Hacker.** 2004. Genomic islands in pathogenic and environmental microorganisms. *Nat Rev Microbiol* **2**:414-24.
  7. **Evans, M. R., B. Swaminathan, L. M. Graves, E. Altermann, T. R. Klaenhammer, R. C. Fink, S. Kernodle, and S. Kathariou.** 2004. Genetic markers unique to *Listeria monocytogenes* serotype 4b differentiate epidemic clone II (hot dog outbreak strains) from other lineages. *Appl Environ Microbiol* **70**:2383-90.
  8. **Glaser, P., L. Frangeul, C. Buchrieser, C. Rusniok, A. Amend, F. Baquero, P. Berche, H. Bloecker, P. Brandt, T. Chakraborty, A. Charbit, F. Chetouani, E. Couve, A. de Daruvar, P. Dehoux, E. Domann, G. Dominguez-Bernal, E. Duchaud, L. Durant, O. Dussurget, K. D. Entian, H. Fsihi, F. Garcia-del Portillo, P. Garrido, L. Gautier, W. Goebel, N. Gomez-Lopez, T. Hain, J. Hauf, D. Jackson, L. M. Jones, U. Kaerst, J. Kreft, M. Kuhn, F. Kunst, G. Kurapkat, E. Madueno, A. Maitournam, J. M. Vicente, E. Ng, H. Nedjari, G. Nordsiek, S. Novella, B. de Pablos, J. C. Perez-Diaz, R. Purcell, B. Rimmel, M. Rose, T. Schlueter, N. Simoes, A. Tierrez, J. A. Vazquez-Boland, H. Voss, J. Wehland, and P. Cossart.** 2001. Comparative genomics of *Listeria* species. *Science* **294**:849-52.

9. **Herd, M., and C. Kocks.** 2001. Gene fragments distinguishing an epidemic-associated strain from a virulent prototype strain of *Listeria monocytogenes* belong to a distinct functional subset of genes and partially cross-hybridize with other *Listeria* species. *Infect Immun* **69**:3972-9.
10. **Kathariou, S.** 2002. *Listeria monocytogenes* virulence and pathogenicity, a food safety perspective. *J Food Prot* **65**:1811-29.
11. **Lobner-Olesen, A., and E. Boye.** 1992. Different effects of *mioC* transcription on initiation of chromosomal and minichromosomal replication in *Escherichia coli*. *Nucleic Acids Res* **20**:3029-36.
12. **Manson, J. M., S. Keis, J. M. Smith, and G. M. Cook.** 2004. Acquired bacitracin resistance in *Enterococcus faecalis* is mediated by an ABC transporter and a novel regulatory protein, BcrR. *Antimicrob Agents Chemother* **48**:3743-8.
13. **Martinez, J. L., and F. Baquero.** 2002. Interactions among strategies associated with bacterial infection: pathogenicity, epidemicity, and antibiotic resistance. *Clin Microbiol Rev* **15**:647-79.
14. **Mehta, P., K. Katta, and S. Krishnaswamy.** 2004. HNH family subclassification leads to identification of commonality in the His-Me endonuclease superfamily. *Protein Sci* **13**:295-300.
15. **Miller, W. G., B. M. Pearson, J. M. Wells, C. T. Parker, V. V. Kapitonov, and R. E. Mandrell.** 2005. Diversity within the *Campylobacter jejuni* type I restriction-modification loci. *Microbiology* **151**:337-51.
16. **Nelson, K. E., D. E. Fouts, E. F. Mongodin, J. Ravel, R. T. DeBoy, J. F. Kolonay, D. A. Rasko, S. V. Angiuoli, S. R. Gill, I. T. Paulsen, J. Peterson, O. White, W. C.**

- Nelson, W. Nierman, M. J. Beanan, L. M. Brinkac, S. C. Daugherty, R. J. Dodson, A. S. Durkin, R. Madupu, D. H. Haft, J. Selengut, S. Van Aken, H. Khouri, N. Fedorova, H. Forberger, B. Tran, S. Kathariou, L. D. Wonderling, G. A. Uhlich, D. O. Bayles, J. B. Luchansky, and C. M. Fraser.** 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**:2386-95.
17. **Pagie, L., and P. Hogeweg.** 2000. Individual- and population-based diversity in restriction-modification systems. *Bull Math Biol* **62**:759-74.
18. **Piffaretti, J. C., H. Kressebuch, M. Aeschbacher, J. Bille, E. Bannerman, J. M. Musser, R. K. Selander, and J. Rocourt.** 1989. Genetic characterization of clones of the bacterium *Listeria monocytogenes* causing epidemic disease. *Proc Natl Acad Sci U S A* **86**:3818-22.
19. **Podlesek, Z., A. Comino, B. Herzog-Velikonja, D. Zgur-Bertok, R. Komel, and M. Grabnar.** 1995. *Bacillus licheniformis* bacitracin-resistance ABC transporter: relationship to mammalian multidrug resistance. *Mol Microbiol* **16**:969-76.
20. **Zheng, W., and S. Kathariou.** 1995. Differentiation of epidemic-associated strains of *Listeria monocytogenes* by restriction fragment length polymorphism in a gene region essential for growth at low temperatures (4 degrees C). *Appl Environ Microbiol* **61**:4310-4.
21. **Zheng, W., and S. Kathariou.** 1997. Host-mediated modification of *Sau3AI* restriction in *Listeria monocytogenes*: prevalence in epidemic-associated strains. *Appl Environ Microbiol* **63**:3085-9.

## **Chapter III**

### **Construction of deletion mutant lacking the ECI specific gene cassette in region 133**

### 3.1 Abstract

A deletion was constructed in ECI specific region133 region in *Listeria monocytogenes* ECI strains F2381L-PHI-R. Bacteriological characterization of the mutant F2381L-PHI-R  $\Delta$ 133 revealed that both this strain and its wild type parent were negative with the serotype 4b-specific monoclonal antibody c74.22, and were both resistant to phage 20422-1. These findings were reproducible, and suggested that, unknown to us at the time that the mutant construction was initiated; the ECI parental strain had undergone a spontaneous mutation that was responsible for these phenotypes. The mutation had been apparently established while the strain had been serially passaged in the laboratory, because the stock culture (maintained at -70C) remained wild type in regard to these characteristics, being positive with c74.22 and susceptible to the phage. For this reason, we decided to construct the deletion in a different ECI strain, F2365, which is also the strain the genome sequence of which has been determined (3). Prior to mutant construction, F2365 was tested in terms of c74.22 and phage susceptibility, and was confirmed to be wild type in these regards. The deletion mutant, F2365  $\Delta$ 133, was characterized in terms of basic bacteriological features including hemolytic activity, phage susceptibility, motility, cell morphology, and growth at 37°C, 25°C and 4°C. In addition, the mutant was characterized in terms of bacitracin resistance and with a panel of Phenotypic Microarrays. No differences were detected between the deletion mutant and the wild type parental strain in hemolytic activity, phage susceptibility, motility, cell shape, bacitracin resistance and growth at 37°C and 25°C. However, at 4°C the parental strain F2365 was observed to grow faster than the deletion mutant F2365  $\Delta$ 133. The Phenotypic Microarrays identified certain differences in substrate utilization between the mutant and the wild type parental strain. These differences were not pronounced at the quantitative level,

but were consistently observed. The F2365 wild type respired better than the deletion mutant F2365  $\Delta$ 133 in 20 carbon sources, especially Uridine, and 11 nitrogen sources, especially N-Acetyl-D-Glucosamine, and the deletion mutant F2365  $\Delta$ 133 respired better than F2365 wild type in the nitrogen source Alloxan; no significant difference in the utilization of phosphorus and sulfur Sources. In conclusion, the gene expression in cassette 133 may involve in the bacteria growth at cold temperature and may work together on the membrane of ECI strains and involved in some carbon and nitrogen sources transportation or metabolism pathway.

### 3.2 Introduction

Subtractive hybridization studies and the genome sequencing of the ECI strain F2365 have allowed identification of several putative ECI specific regions in the genome (1). The genomic organization of four of these regions has been analyzed in detail (Chapter II). One of these regions, region 133, was of special interest, due to its location immediately adjacent to a gene with a key function in translation (Translation Elongation Factor Tu). The region 133 ECI specific cassette also harbored ORFs which encode putative membrane associated proteins, and an ABC transporter that may be involved in bacitracin transport, and resistance to this antimicrobial agent (Chapter II, and Fig. 1). The region 133 ECI cassette was also present in the ECIIa epidemic clonal group, which is related to ECI and appears to represent an emerging *L. monocytogenes* lineage, and in certain strains of serotypes 1/2b and 3b. As discussed in Chapter I, serotypes 1/2b, 3b, and 4b constitute one of the major genomic divisions (Genomic Division II also referred to as Lineage I). These results suggest that the gene cassette in this region may have been acquired by a *L. monocytogenes* strain ancestor to this division, and was subsequently lost by serotype 4b strains except for strains of specific epidemic associated clonal groups, i.e. ECI and ECIIa, or there was no cassette 133 in *L. monocytogenes* strain ancestor and this cassette was acquired by some serotype 1/2b and 3b strains, ECI strains and ECIIa strains. Interestingly, two copies of palindromic direct repeats (5' ATTCAAACCGACAAAGTCATTTGGCTTTGTTCGGTTTTT TTGT 3') were found flanking the ECI specific ORFs in ECI strains, while only one copy was found in strains which do not have the six ECI specific ORFs.

Currently the function of this gene cassette remains unknown. In order to investigate the possible function of the ORFs in this ECI specific region, we constructed a deletion in this region in two ECI strains, F2381L-PHI-R and F2365.

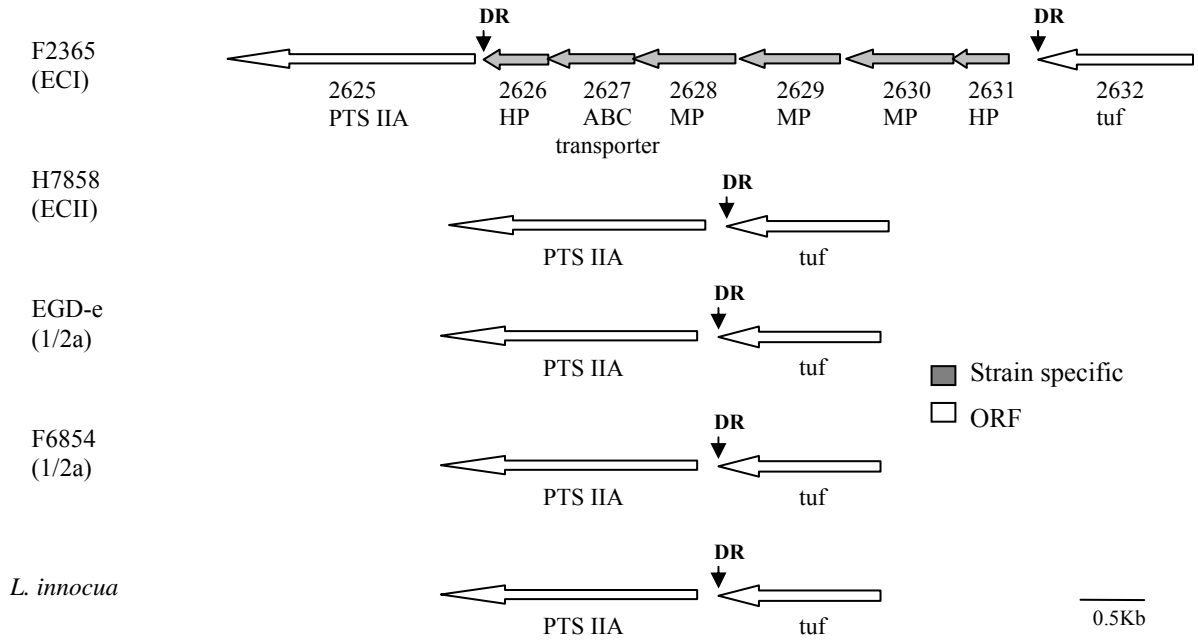


Fig. 1 Genomic region harboring fragment 133 in F2365 (ECI) and other *Listeria* genome. PTS IIA: putative PTS system, IIA component; HP: Hypothetical protein; MP: Membrane protein; Tuf: Translation elongation factor Tu, DR: Direct Repeat (5' ATTCAAACCGACAAAGTCATTTGGCTTTGTTCGGTTTTTTTGT 3') Arrows indicate direction of transcription. Gray arrows: strain specific ORFs, white arrow: Conserved ORFs

### **3.3 Materials and Methods**

#### **3.3.1 Bacterial strains and growth media.**

Strains of *Listeria* used in this work were described in Chapter II, along with the methods for growing and preserving the organisms.

#### **3.3.2 Mutant construction**

Fragments upstream and downstream of the ECI specific cassette in region 133 were amplified by primers P1 and P5, P6 and P7 harboring restriction sites HindIII, BamHI and EcoRI, respectively (Table 1). The two PCR fragments (1128 bp and 762 bp) for the upstream and downstream fragment, respectively) were ligated to the temperature sensitive shuttle vector pCON-1 (Fig. 2), and the recombinant plasmid was transformed to competent cells of *Escherichia coli* SM10 by electroporation. Transformants (designated *E. coli* SM1076) were selected on LB plates containing Ampicillin (0.5ug/ml) following 12 h of incubation at 37 °C. Conjugations were done between *E. coli* SM1076 and *L. monocytogenes* ECI strain F2381L-PHI-R, as well as between *E. coli* SM1076 and F2365. In the *L. monocytogenes* recipient, homologous recombination between the plasmid and the genome of *L. monocytogenes* resulted in integration of the plasmid, and the integrants were selected following incubation on BHI+ chloramphenicol (5ug/ml) plates, and then the selected integrants were incubated at 42°C, which is a temperature restrictive for replication of pCON-1, resulting in the loss of the plasmid, and in 133 region, about 50% of the chloramphenicol sensitive colonies in wild type region was replaced with the deletion version (Fig. 3).

Table 1. Primers for plasmid construction. Restriction enzyme sites were underlined.

Primer Name	Primer Sequence	Restriction Sites
P1	GCTGATGCACA <u>AAGCTT</u> ATGACC	HindIII
P5	CTTTGGATCCACAAAAAACCGACAAAGCC	BamHI
P6	TTGTGGATCCAATAAAAAATTA <u>ACTT</u> GAAAAATATC	BamHI
P7	CCTCGTCAGTATAACGA <u>ATTCCG</u>	EcoRI

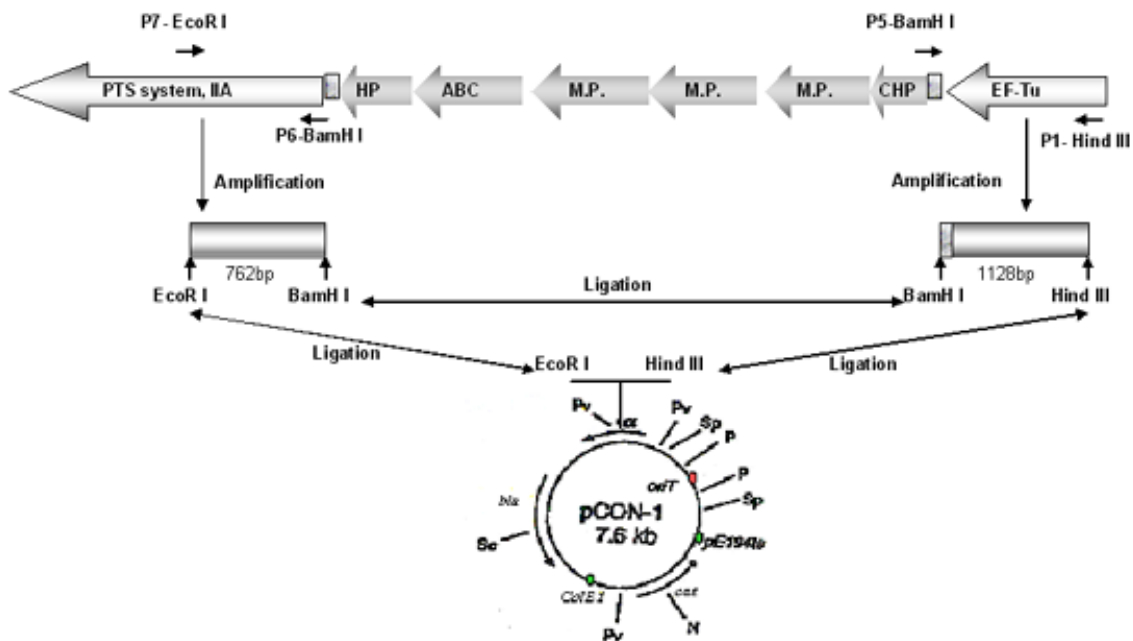


Fig. 2 Plasmid construction. Fragments upstream and downstream of the ECI specific cassette in region 133 were amplified by primers P1 and P5, P6 and P7 harboring restriction sites HindIII, BamHI and EcoRI, respectively. The two PCR fragments (1128 bp and 762 bp for the upstream and downstream fragment, respectively) were ligated to the temperature sensitive shuttle vector pCON-1. On the plasmid pCON-1, oriT is the origin of conjugation, *pE194Is* is the origin of replication in *L. monocytogenes*, *ColE1* is the origin of replication in *E. coli*, *cat* is the gene encodes chloramphenicol resistance protein, *bla* is the gene encodes ampicillin resistance protein.

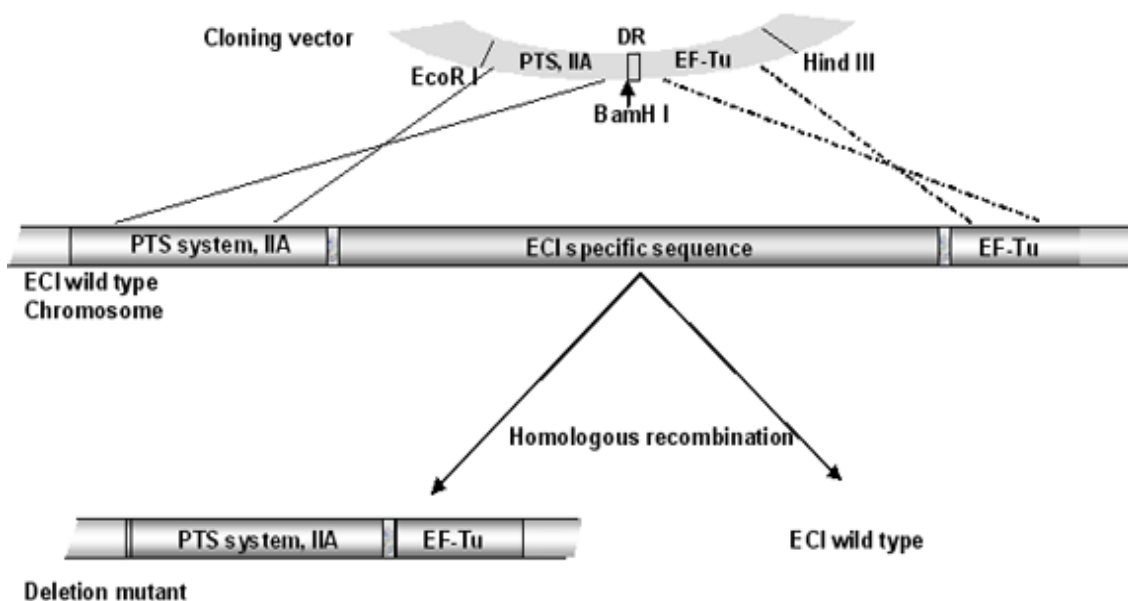


Fig. 3 Homologous recombination-mediated 133 region deletion. After the plasmid entering *L. monocytogenes* cell, homologous recombinations happened between the plasmid and the genome of *L. monocytogenes* under the temperature and antibiotics pressure.

### 3.3.2.1 Plasmid construction

#### 3.3.2.1. (a) Enzyme restriction and ligation

The upstream and downstream fragments were amplified with primers P1 and P5, P6 and P7 by Colony PCR (Appendix B, C). Plasmid pCon-1 and the PCR-amplified upstream and downstream fragments were digested by restriction enzymes EcoRI, HindIII and BamHI, respectively, at 37°C for 3 hours (Appendix D). The digestion products were confirmed by electrophoresis. The digested upstream and downstream fragments and the digested plasmid were ligated at room temperature overnight (Appendix E).

#### 3.3.2.1. (b) Electroporation

The plasmids were transformed into competent cell *E. coli* SM10 by electroporation. 1µl of the ligation reaction was gently mixed with 50µl competent cells using a pipette, in

duplicate for each transformation. Following electroporation (voltage max: 2.5; resistance: 200; capacitancy: 25; time const: 4-5), 1ml LB was added to resuspend the mixture in a 15-ml tube, and the cell suspension was incubated with shaking at 37°C for about 1hour. All transformants were plated on LB+Ampicillin (0.5ug/ml) agar and were incubated at 37°C overnight.

### **3.3.2.1. (c) Screen positive transformants**

For initial mass screenings, colonies of *E. coli* were picked with sterile toothpicks from the LB+Ampicillin (0.5ug/ml) agar plate, and resuspended in 75ul lysis solution in an Eppendorf Tube. (10 ml Lysis buffer consisted of 6ml H<sub>2</sub>O, 3ml 10% SDS, 1ml 1M Tris-HCL pH8 (autoclaved), 120ul 10M NaOH) in a total volume of 10 ml. Following incubation at room temperature for 5 min, an equal volume of phenol: chloroform: isoamyl alcohol (25:24:1) was added to the mixture. The mixture was then centrifuged at maximum speed in for 3 minutes. The upper aqueous phase was removed, and 15 ul of the supernatant was loaded on a 0.8% agarose gel. The colonies with plasmid size larger than the vector alone were chosen for restriction digestion screens. The plasmids were extracted (QIAprep miniprep handbook, plasmid DNA purification using the QIAprep spin miniprep kit and a microcentrifuge), and then digested by restriction enzymes HindIII and EcoRI at 37°C for 3 hours. The digested products were checked on 1% agarose gel, and two *E. coli* SM 1076 transformants (NO. 87 and NO. 88) were chosen for conjugation with ECI strains.

### **3.3.2.2 Conjugation of *E.coli* SM1076 harboring the recombinant plasmid and ECI strain F2381L-PHI-R**

*E.coli* SM1076 NO.87 and No.88 were inoculated separately in 5ml BHI +Ampicillin (100ug/ml) and incubated at 37°C overnight. *L. monocytogenes* ECI strain F2381L-PHI-R was inoculated in 5ml BHI and incubated at 37°C overnight. 1.5 ml of F2381L-PHI-R, *E.coli* SM1076 NO.87 and No.88 were centrifuged at 14000 rpm for 1min, the supernatant was removed, and the pellet was washed with 1ml BHI twice. 1ml BHI was added to resuspend F2381L-PHI-R, and then F2381L-PHI-R was mixed with *E.coli* SM1076 NO.87 and No.88 pellets by pipetting, separately. The mixture was centrifuged at 1400 rpm for 1min, the pellet was washed with 1ml BHI, and the mixture was resuspended in 1ml BHI. One nitrocellulose membranes (MSI, Gloucester, MA) (3cmx3cm) were placed on a BHI+Penicillin G (0.5ug/ml) plate, 50µl of the mixture was added onto each membrane, and the plate was incubated at 30°C overnight. The nitrocellulose membrane from each plate was placed into a 15-ml tube with 4ml BHI, the bacteria were washed into the BHI medium by vortexing, and 50 µl of the mixture was streaked or spread-plated onto BHI+ Nalidixic acid (30ug/ml) + chloramphenicol (5ug/ml) plates. The plates were air-dried dry in the biosafety laminar flow cabinet for 30min, and then incubated at 30°C overnight.

### **3.3.2.3 Isolation of integrants and recombinants**

Isolated colonies were streaked on BHI+ Nalidixic acid (30ug/ml) + chloramphenicol (5ug/ml) agar, and incubated at 30°C overnight. The colonies were confirmed by PCR using primers P1 and P7 and checked on a 1% agarose gel. Two transconjugants (NO.1 and NO.22) were inoculated in 5ml BHI+ chloramphenicol (5ug/ml), respectively, incubated at 30 °C overnight, diluted (1:100) dilution in BHI, plated (50µl) on BHI+ chloramphenicol (5ug/ml)

agar, and incubated at 42°C for 48 hours. Three colonies were streaked on BHI+ chloramphenicol (5ug/ml) agar, and incubated at 42°C for 24 to 48 hours. Two colonies from these plates were inoculated in BHI, and incubated at 30 °C overnight for 3 to 4 serial transfers. 300µl of the culture from the fourth transfer was inoculated in 5ml BHI, incubated at 42°C overnight, diluted ( $10^{-7}$ ), plated (50µl) on BHI agar, and incubated at 37°C overnight. 90 isolated colonies from this plate were tested for chloramphenicol resistance by replica plating on BHI agar and BHI + chloramphenicol (5ug/ml) agar, and incubated at 30°C 24-48 hours.

#### **3.3.2.4 PCR screening and southern blots to identify deletion mutants**

Chloramphenicol sensitive colonies from the previous screen were examined by PCR using primers Pm and Pn which were in ORF2625 and ORF2632, respectively, flanking region 133 (Appendix F). Cultures from colonies NO.9, NO.17 and NO.19 were confirmed by Southern blots.

#### **3.3.2.5 Sequencing confirmation of deletion mutants**

Culture NO.19 was chosen for sequencing. The deletion region was amplified with primers P1 and P7 by colony PCR. The PCR product was purified by QIAquick PCR purification kit (See QIAquick Spin Handbook, QIAquick PCR purification kit protocol using a microcentrifuge) and diluted in water, then quantified on a 1% agarose gel. PCR DNA (in water, approximate 50 ng, 10ng / 100 base pair) and primer Pm / Pn (in water, 10pmoles) were mixed and put in one 2-ml microcentrifuge tube, total volume was 20 ul. Primers Pm and Pn which were in ORF2625 and ORF2632, respectively, flanking region 133. The DNA sample was sequenced by UNC-CH Genome Analysis Facility.

### **3.3.2.6 Construction of ECI deletion mutant in strain F2365**

The procedures employed for construction of the deletion in F2365 were the same as described above for Chloramphenicol sensitive colonies were confirmed by PCR using primers P1 and P7. Colonies NO.12, NO.13 and NO.16 were confirmed by Southern blot and colony NO.12 was chosen for sequencing. The sequencing results confirmed that the ECI specific region was deleted with the first direct repeat remaining, and only four base pairs were changed at the BamHI site.

### **3.3.3 Mutant F2365 $\Delta$ 133 characterization identification**

#### **3.3.3.1 Hemolytic activity**

ECI strain F2365 wild type and F2365  $\Delta$ 133 cells were streaked on sheep blood agar plates (Remel), and were incubated at 37°C for 1day, 25°C for 2 days, and 4°C for 30 days (each incubation done in duplicate).

#### **3.3.3.2 Phage invasion**

Phage 20422-1 is a *Listeria* genus specific phage isolated in 2004 by Jae-Won Kim in our laboratory from a North Carolina turkey processing plant. This phage infects all serotypes of *L. monocytogenes*. F2365 wild type and F2365  $\Delta$ 133 cells were incubated in LB broth at 37°C and 20°C overnight. *L. monocytogenes* cells (F2365 wild type/ F2365  $\Delta$ 133), phage 20422-1 and MgSO<sub>4</sub> were mixed in soft LB agar (0.75% agar). The mixture was poured onto a hard LB agar plate (1.5% agar), incubated in 37°C, and examined for plaques.

### **3.3.3.3 Motility**

ECI strain F2365 wild type and F2365  $\Delta$ 133 cells were grown in TSB+0.7% yeast extract at 37°C overnight, spotted (5 ul, done in triplicate for each strain) on a soft agar plate (TSB+0.7% yeast extract, with 0.4% agar), and incubated at 37°C, 25°C and 4°C.

### **3.3.3.4 Growth at 37°C, 25°C and 4°C**

F2365 wild type and F2365  $\Delta$ 133 cells were inoculated in TSB+0.7% yeast extract, and incubated at 37°C overnight. Dilutions ( $10^{-8}$ ) were made in TSB+0.7% yeast extract and OD<sub>600</sub> value was determined with the spectrophotometer (BIO-RAD, Smart Spec 300) to get an almost equal OD<sub>600</sub> value for each culture. 10 ul of each culture was plated onto TSB+0.7% yeast extract agar, and was incubated at 37°C for 1.5 days, 25°C for 3 days, and 4°C for 35 days (done in duplicate).

### **3.3.3.5 Cell morphology**

F2365 wild type and F2365  $\Delta$ 133 cells were grown in TSB+0.7% yeast extract at 37°C for 10 hours (observed every two hours), 25°C for 48 hours (observed every 5 hours), and 4°C for 30 days (observed every two days). Bacteria were observed with phase contrast microscopy (Leica).

### **3.3.3.6 Bacitracin resistance**

F2365 wild type and F2365  $\Delta$ 133 cells were grown in TSB+0.7% Yeast Extract to early log phase (OD<sub>530</sub> = 0.3). A 0.5 ml culture was added to a test tube that contained 0.5ml of TSB with serially diluted concentrations of bacitracin (indicate range of concentrations and

source of bacitracin). Growth was monitored by measuring OD530 after 5h at 37°C, after 5 days at 4°C.

In the presence of ZnSO<sub>4</sub>, the sensitivity to bacitracin will be increased, ZnSO<sub>4</sub> can mediate the binding between bacitracin and isoprenyl pyrophosphate (IPP) (6), the antibiotic action of bacitracin is to bind to IPP to prevent IPP from dephosphorylation into isoprenyl phosphate (IP) by specific phosphatase(s), thus reducing the amount of IP that is available for carrying sugar-peptide units (7). In this study, we also did bacitracin test in the presence of ZnSO<sub>4</sub> (40 ug/ml).

### **3.3.3.7 Phenotypic microarrays**

Phenotypic MicroArrays (Biolog, Hayward, CA) can be used to evaluate carbon (PM1-2), nitrogen (PM3), and phosphorus/sulfur (PM4) metabolism of several bacterial species (<http://www.biolog.com/accesstoPMTech.html>). *L. monocytogenes* strain F2365 wild type, mutant strain F2365  $\Delta$ 133 were grown at 37°C overnight on BUG+B Agar (Biolog, Hayward, CA). The following solutions were prepared to add to IF-0 inoculating fluid (Biolog, Hayward, CA): 100x stock solution of solution D (menadione sodium bisulfite [Sigma Cat# M5750] 5.524 mg + sodium salicylate [Sigma Cat# S3007] 320.2mg), solution E (sodium pyruvate [Sigma Cat# P2256] 2.2 g + sodium thioglycollate [Sigma Cat# T0632] 570.5 mg + uridine [Sigma Cat# U3750] 24.42 mg). The chemicals were dissolved in 10 ml of purified ddH<sub>2</sub>O, filter sterilized, and stored in a refrigerator in 50-ml tubes wrapped in aluminum foil. Solution D was added to IF-0 for PM1 and PM2 plates, whereas solution E was added to IF-0 for PM3 and PM4 plates. 150  $\mu$ l of the 100x stock solution was added to 15 ml of IF-0 inoculating fluid immediately before preparing the cell suspensions. A dry sterile cotton

swab was used to remove a small amount of cells from the agar surface and to completely resuspend the cells in inoculating fluid (IF-0, solution D/E). Turbidity was adjusted to  $OD_{590}=0.2$  and each cell suspension was poured into a sterile filling reservoir and was inoculated into the PMs by pipetting 100  $\mu$ l per well. The PMs were incubated at 37° C for 24 hr (PM1 and PM2) or for 36-48 hr (PM3 and PM4), and the results were read and scored. The formation of purple color indicated utilization of the C, N, P, or S source in a particular well. The A-1 well was used as a reference well for a “negative” reaction. Any well with more color than the reference well was considered “positive”. The PM plates were scanned and software ImageJ was used to prepare the scoring tables.

#### **3.3.3.8 Colony immunoblotting with monoclonal antibody MAb c74.22.**

ECI strain F2365 and mutant F2365  $\Delta$ 133 were grown in BHI at 37° C overnight. A piece of nitrocellulose membrane was placed on the BHI agar plate, 5  $\mu$ l of the culture was spotted in triplicate on the membrane, and the plates were incubated at room temperature for 36-48 hr. The colonies were washed off the membrane with Towbin transfer buffer (3.03g Tris, 14.4g glycine, 200 ml methanol, dH<sub>2</sub>O to 1000 ml, adjusted to pH8.3, with 10 ml of 10%SDS added), with shaking, and membrane was then washed twice in TTBS buffer for 5-10 min (TTBS, Tris-buffered saline with Tween-20: 2.42 g Tris base, 8 g NaCL, 3.8ml 1M Hydrochloric acid, dH<sub>2</sub>O to 1000ml, adjusted to pH7.6, with 0.1% Tween-20 added). The membrane was incubated in Blotto blocking solution (5% non-fat milk powder in TTBS) with shaking for 1 hour. The membrane was then incubated with primary antibody (Blotto blocking solution + primary antibody, 1:1000), ascites antibody MAb c74.22 (isotype IgG1) (2) with shaking for one hour. The membrane was washed in TTBS, using two quick rinses

followed by three 5 min washes. The membrane was then incubated with secondary antibody (1:500 goat-anti mouse immunoglobulin, polyvalent, horseradish peroxidase-conjugated; Fisher) for one hour. The membrane was washed in TTBS using 2 quick rinses followed by three 5 min washes, and incubated in freshly prepared substrate solution (15 ml of 3 mg/ml of 4-Chloro-1-Naphthol in 95% ethanol, 25 ml of 0.01M Tris-HCL, pH7.5, and 300 ul of 3% of H<sub>2</sub>O<sub>2</sub>) for 5 to 30 minutes, with a cover to avoid light until color was fully developed. After color was developed, the membrane was rinsed and soaked in dH<sub>2</sub>O for 5 minutes, dried in air, and kept in the dark.

### **3.4 Result**

#### **3.4.1 Mutant construction result**

The mutant F2381L-PHI-R  $\Delta$ 133 was confirmed by PCR screening with primers Pm and Pn, located in the conserved ORFs ORF2625 and ORF2632, respectively, flanking the ECI specific region (Fig. 4); and confirmed by southern blots, the genomic DNA was digested by restriction enzyme PvuII and styI, separately, the probe was amplified by PCR with primers P1 and P7 (Fig. 5). The sequencing result was compared with the wild type showed in Fig. 6.

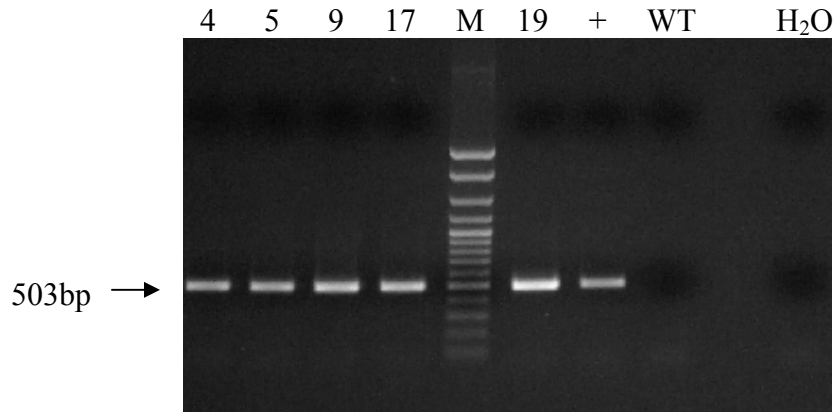


Fig. 4 PCR screen of chloramphenicol sensitive colonies using primers Pm and Pn, located in the conserved ORFs ORF2625 and ORF2632, respectively, flanking the ECI specific region. Lanes 1-5 and 6 were chloramphenicol sensitive colonies, lane 7 was positive control, lane 8 was F2365 wild type, and lane 9 was negative control for the PCR.

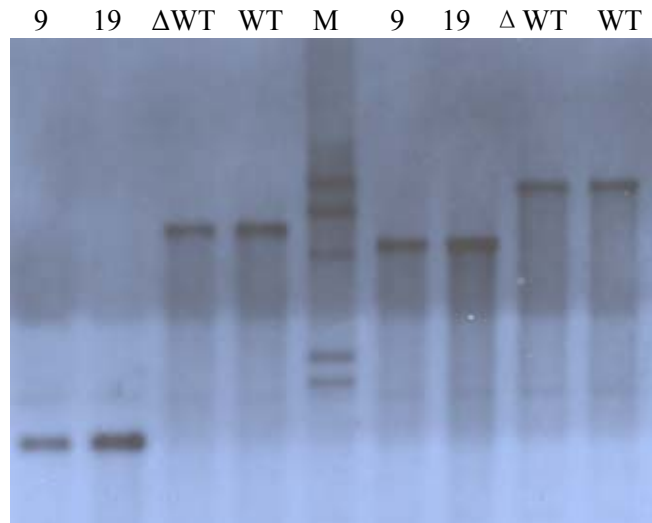


Fig. 5 Mutant confirmation by southern blot, the probe was amplified by PCR with primers P1 and P7. Lanes 9 and 19: region 133 deletion mutants; ΔWT: chloramphenicol sensitive wild type obtained at the same time as the deletion mutants, after homologous recombination; WT: F2381L-PHI-R wild type, Lanes 1-4 were digested by restriction enzyme PvuII, and lanes 6-9 were digested by restriction enzyme StyI.

```

Wildtype      CAGTTGAACTAATTGCACCAATCGCTATCGAAGACGGTACTAAATTCTCTATCCGTGAAG 60
Deletion      CAGTTGAACTAATTGCACCAATCGCTATCGAAGACGGTACTAAATTCTCTATCCGTGAAG 60
*****

Wildtype      GCGGACGTACAGTAGGCGCTGGCGTTGTTTCTAACATCAGCAAATAATATCTGATACGAT 120
Deletion      GCGGACGTACAGTAGGCGCTGGCGTTGTTTCTAACATCAGCAAATAATATCTGATACGAT 120
*****

Wildtype      TCAAACCGACAAAGTCATTTGGCTTTTGTTCGGTTTTTTTTGT
Deletion      TCAAACCGACAAAGTCATTTGGCTTTTGTTCGGTTTTTTTTGT
*****

Deleted sequence

GCCTGAAATAAAAAATTAACT 180
(BamHI site) GGATCCAATAAAAAATTAACT 180
* * *****

Wildtype      TGAAAAATATCACATTTTCGCATAACCTTAAAGTAGACATATCTTTTTACCTTCTAGACC 240
Deletion      TGAAAAATATCACATTTTCGCATAACCTTAAAGTAGACATATCTTTTTACCTTCTAGACC 240
*****

Wildtype      GAAATCAAGGAGGTAGGCCAAGTGGTACAATTTGATGCTCGAAATATGGCGTTGCTCGAA 300
Deletion      GAAATCAAGGAGGTAGGCCAAGTGGTACAATTTGATGCTCGAAATATGGCGTTGCTCGAA 300
*****

Wildtype      TCACTCGTTGTGGCGAATGTATATCTTGCACCCGAGAAATTACAAGAAGA 350
Deletion      TCACTCGTTGTGGCGAATGTATATCTTGCACCCGAGAAATTACAAGAAGA 350
*****

```

Fig. 6 Sequence comparison of region 133 deletion in ECI strain F2381L-PHI-R. ECI specific region was deleted with the first direct repeat remaining (in italics), and four base pairs were changed at the BamHI site (indicated in bold).

The mutant F2365  $\Delta$ 133 was confirmed by PCR screening with primers P1 and P7 (Fig. 7); and confirmed by southern blots, the genomic DNA was digested by restriction enzyme PvuII and styI, separately, the probe was amplified by PCR with primers P1 and P7 (Fig. 8). The sequencing result was compared with the wild type showed in Fig. 9.

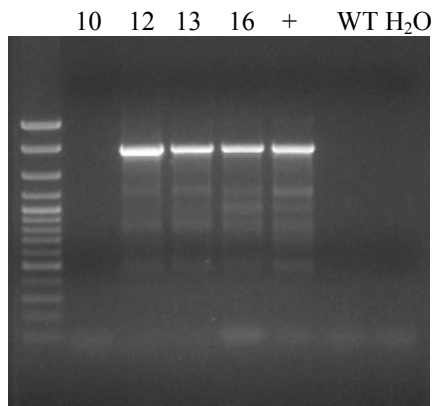


Fig. 7 PCR screen of chloramphenicol sensitive colonies to detect deletions in F2365. Primers P1 and P7 were in ORF2625 and ORF2632, respectively, flanking region 133. Lanes 1-4 were chloramphenicol sensitive colonies, lane 5 was positive control, lane 6 was F2365 wild type, and lane 7 was negative control for the PCR.

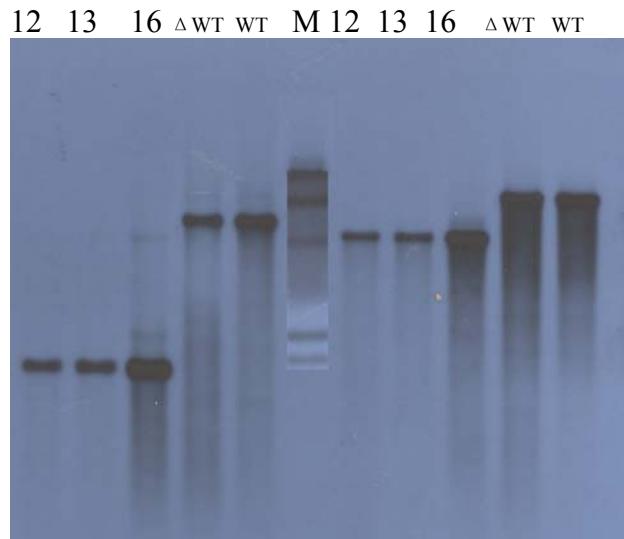


Fig. 8 F2365 Mutant confirmation by southern blot, done as described in the legend for Fig. 5.  $\Delta$ WT: wild type obtained after homologous recombination, as in Fig. 5; WT: F2365 wild type; lanes 1-5 were digested by restriction enzyme PvuII; lanes 6-10 were digested by restriction enzyme StyI.

```

Wildtype      CAGTTGAACTAATTGCACCAATCGCTATCGAAGACGGTACTAAATTCTCTATCCGTGAAG 60
Deletion     CAGTTGAACTAATTGCACCAATCGCTATCGAAGACGGTACTAAATTCTCTATCCGTGAAG 60
*****

Wildtype      GCGGACGTACAGTAGGCGCTGGCGTTGTTTCTAACATCAGCAAATAATATCTGATACGAT 120
Deletion     GCGGACGTACAGTAGGCGCTGGCGTTGTTTCTAACATCAGCAAATAATATCTGATACGAT 120
*****

Wildtype      TCAAACCGACAAAGTCATTTGGCTTTGTTCGGTTTTTTTGT
Deletion     TCAAACTGACAAAGTCATTTGGCTTTGTTCGGTTTTTTTGT
*****

```

**Deleted sequence**

```

GCCTGAAATAAAAAATTAACT 180
(BamHI site) GGATCCAATAAAAAATTAACT 180
* * *****

Wildtype      TGAAAAATATCACATTTTCGCATAACCTTAAAGTAGACATATCTTTTTACCTTCTAGACC 240
Deletion     TGAAAAATATCACATTTTCGCATAACCTTAAAGTAGACATATCTTTTTACCTTCTAGACC 240
*****

Wildtype      GAAATCAAGGAGGTAGGCCAAGTGGTACAATTTGATGCTCGAAATATGGCGTTGCTCGAA 300
Deletion     GAAATCAAGGAGGTAGGCCAAGTGGTACAATTTGATGCTCGAAATATGGCGTTGCTCGAA 300
*****

Wildtype      TCACTCGTTGTGGCGAATGTATATCTTGCACCCGAGAAATTACAAGAAGA 350
Deletion     TCACTCGTTGTGGCGAATGTATATCTTGCACCCGAGAAATTACAAGAAGA 350
*****

```

Fig. 9 Sequence comparison of region 133 deletion in ECI strain F2365. ECI specific region was deleted with the first direct repeat remaining (in italics), and four base pairs were changed at the BamHI site (in bold).

### 3.4.2 Hemolytic activity

F2365 and mutant F2365 Δ133 had identical hemolytic activity at 37°C and 25°C. No activity was detected at 4°C (data not shown).

### 3.4.3 Phage invasion

Infections with phage 20422-1 revealed that F2365 wild type and mutant F2365 Δ133 were equally susceptible, with similar numbers of plaques formed both at 37°C and 20°C (data not shown).

### 3.4.4 Motility

When spotted on soft agar and grown at 37°C, 25°C and 4°C, F2365 wild type and mutant F2365 Δ 133 were indistinguishable in terms of their swarming patterns (Fig. 10).

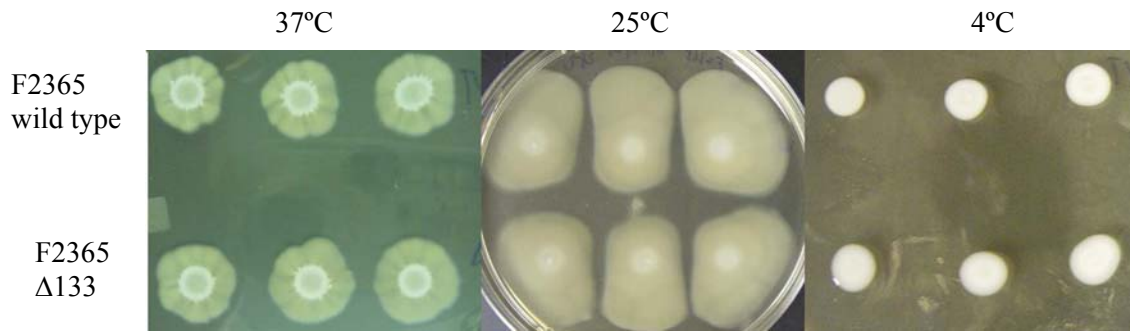


Fig. 10 Swarming assay of F2365 and mutant F2365  $\Delta$ 133 on TSB+0.7% yeast extract agar (0.4% agar), following incubation at 37°C, 25°C and 4°C.

### 3.4.5 Cell morphology

Phase contrast microscopy revealed no morphologic difference between F2365 wild type and mutant F2365  $\Delta$ 133 cells grown at 37°C, 25°C and 4°C (data not shown).

### 3.4.6 Growth at 37°C, 25°C and 4°C

F2365 wild type and mutant F2365  $\Delta$ 133 formed similar size colonies at 37°C and 25°C. However, F2365 wild type grew better than mutant F2365  $\Delta$ 133 at 4°C (Fig.11, 12).

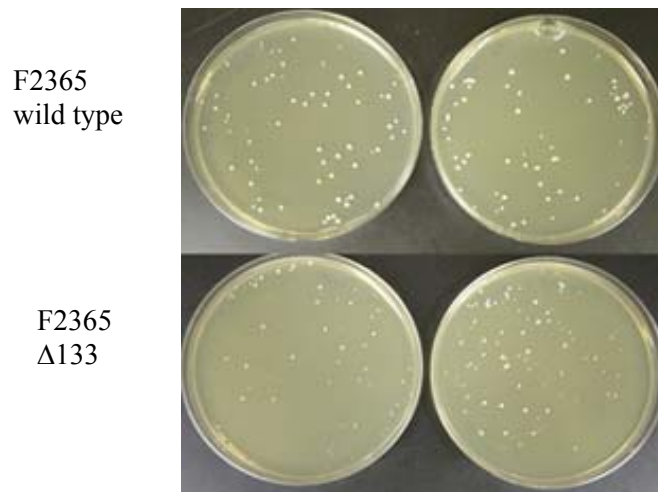


Fig. 11 Growth of F2365 wild type and mutant F2365  $\Delta$ 133 on Trypticase Soy agar +0.7% yeast extract at 4°C.

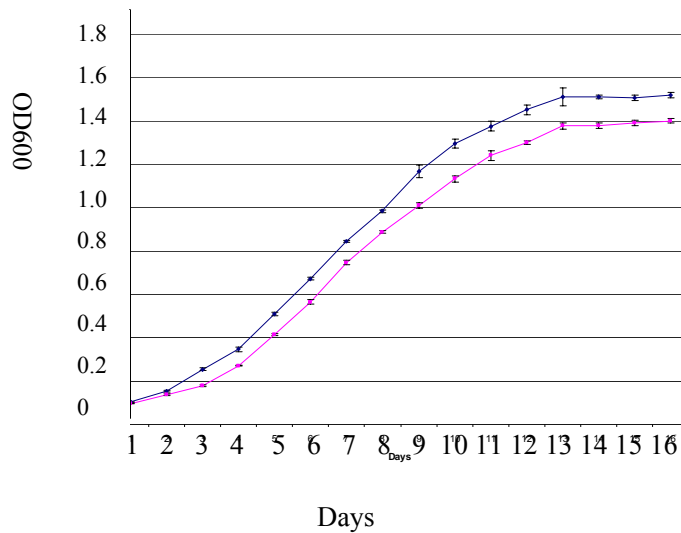


Fig. 12 Growth curve of F2365 wild type and mutant F2365 Δ133 in Trypticase Soy agar +0.7% yeast extract at 4°C. Blue line: F2365 wild type; red line: mutant F2365 Δ133

### 3.4.7 Bacitracin resistance

No significant difference could be noticed between F2365 wild type and mutant F2365 Δ133 in terms of resistance to bacitracin. The minimum inhibitory concentrations (MICs) of bacitracin at 37°C was 25ug/ml, MIC at 4°C was 5ug/ ml, both for the mutant and for the wild type (Table. 2). There is no significant difference between F2365 wild type and mutant F2365 Δ133 in the presence of ZnSO<sub>4</sub> (40 μg/ml).

Table 2. The growth of F2365 wild type and F2365  $\Delta$ 133 in series bacitracin concentration at 37°C and 4°C.

Bacitracin (Unit/ml)	F2365 Wild type OD530 at 37°C	F2365 $\Delta$ 133 OD530 at 37°C	F2365 Wild type OD530 at 4°C	F2365 $\Delta$ 133 OD530 at 4°C
0	1.1185	1.1075	1.117	1.108
0.005	1.1365	1.143	1.0995	1.146
0.01	1.1435	1.159	1.137	1.1225
0.05	1.126	1.1205	1.101	1.0675
0.1	1.1345	1.1385	1.132	1.073
0.5	1.1525	1.0905	1.092	1.0585
1	1.099	1.119	1.017	1.008
2.5	1.067	1.0705	0.852	0.771
5	1.032	1.0365	0.442	0.4385
7.5	0.99	0.969	0.3535	0.3645
10	0.921	0.8345	0.293	0.3395
25	0.4145	0.3685	0.2435	0.266
50	0.271	0.252	0.268	0.2755
75	0.2575	0.2235	0.2355	0.2675
100	0.237	0.202	0.23	0.263
125	0.213	0.1855	0.2145	0.2485
150	0.2015	0.182	0.3625	0.251
200	0.1955	0.1725	0.2025	0.2305

### 3.4.8 Phenotypic Microarrays

ECI strain F2365 wild type utilized the carbon sources better than mutant F2365  $\Delta$ 133 in N-Acetyl-D-Glucosamine, D-Trehalose, D-Mannose, Glycerol, D-Fructose,  $\alpha$ -D-Glucose, Maltose,  $\beta$ -Methyl-D-Glucoside, Maltotriose, Adenosine, D-Cellobiose,  $\beta$ -Cyclodextrin,  $\gamma$ -Cyclodextrin, Dextrin, Gentiobiose,  $\alpha$ -Methyl-D-Galactoside, Xylitol, especially in Uridine medium. In Arbutin and Dihydroxy Acetone medium, the mutant utilized the carbon sources slightly better than the wild type (Fig. 13, Table 3). For nitrogen sources, F2365 wild type utilized nitrogen sources better than the mutant F2365  $\Delta$ 133 in L-Cysteine, L-Threonine, L-Threonine, Thymidine, Uridine, Inosine, Xanthosine, Uric Acid medium, especially in D-Glucosamine, N-Acetyl-D-Glucosamine medium. However, the mutant utilized Alloxan medium slightly better than the wild type (Fig. 14, Table 4). For phosphorus and sulfur sources, there was no significant difference between F2365 wild type and the mutant F2365

$\Delta 133$ . (Figure 15, 16, Table 5) Uridin is a nucleoside, and nucleosides could serve as carbon source under certain induction in *E. coli* and *B. subtilis*. In *B. subtilis*, the *nupC* locus encodes a pyrimidine-nucleoside transporter that mediates the transport of uridine, thymidine and deoxyuridine (5). N-acetyl-D-glucosamine (NAG) is a good energy and nitrogen source for bacteria, is a main constituent of bacterial cell wall peptidoglycan. NAG is transported via a NAG-specific phosphotransferase system (PTS<sup>NAG</sup>). In the cytosol of *Bacillus subtilis*, ~90% of the NAG taken up was converted to cell wall precursors (4).

### 3.4.9 Colony immunoblotting with monoclonal antibody MAb c74.22

Both of ECI strain F2365 wild type and mutant F2365  $\Delta 133$  were positive with monoclonal antibody MAb c74.22.

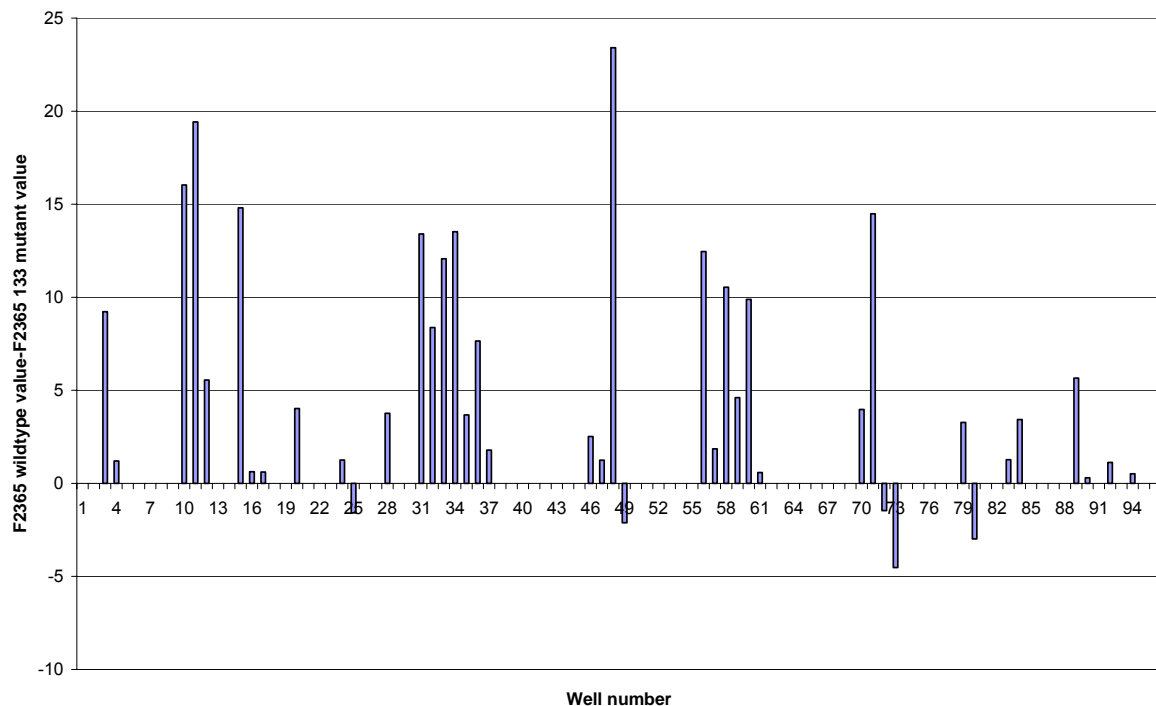


Fig. 13 PM1 carbon sources MicroPlate Value comparison of F2365 wild type Vs. mutant F2365  $\Delta 133$ . X axis indicated well number in 96-well plate, Y axis indicated the value of F2365 wild type value minus mutant F2365  $\Delta 133$  value. Well NO.1 was negative control.

Table 3. PM1 carbon sources with significant value difference between F2365 wild type and mutant F2365  $\Delta$ 133

Well Number	Difference	Carbon Sources
3	9.22	N-Acetyl-D-Glucosamine
10	16.033	D-Trehalose
11	19.42	D-Mannose
15	14.808	Glycerol
31	13.401	D-Fructose
33	12.073	$\alpha$ -D-Glucose
34	13.53	Maltose
48	23.409	Uridine
56	12.459	$\beta$ -Methyl-D-Glucoside
58	10.538	Maltotriose
60	9.887	Adenosine
71	14.489	D-Cellobiose

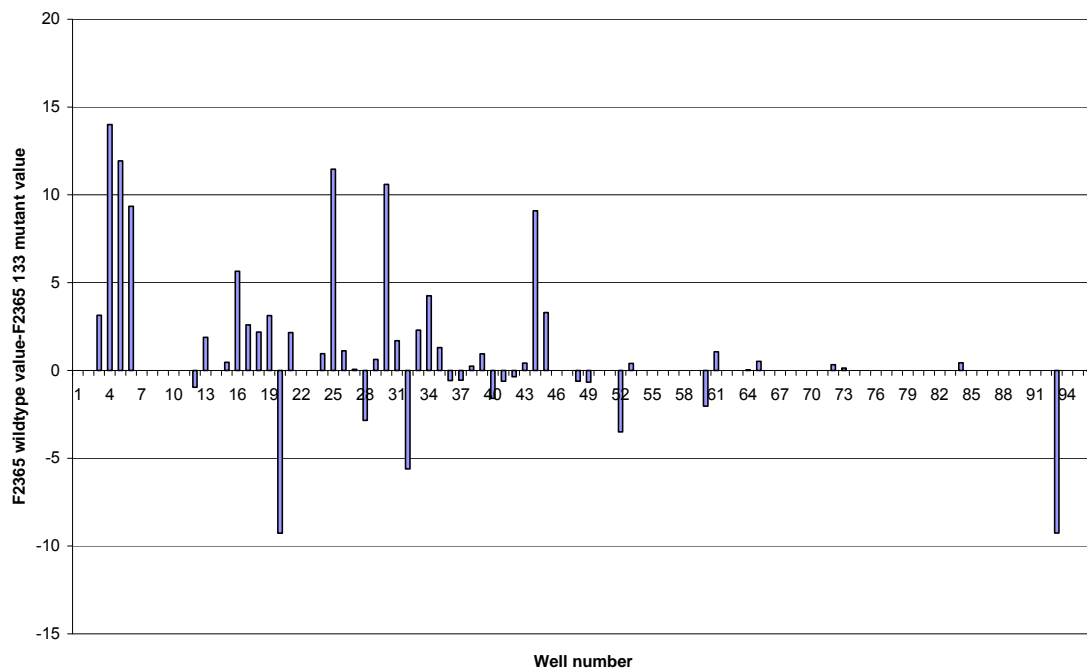


Fig. 14 PM2 carbon sources MicroPlate Value comparison of F2365 wild type Vs. mutant F2365  $\Delta$ 133. X axis indicated well number in 96-well plate, Y axis indicated the value of F2365 wild type value minus mutant F2365  $\Delta$ 133 value. Well NO.1 was negative control.

Table 4. PM2 carbon sources with significant value difference between F2365 wild type and mutant F2365  $\Delta$ 133

Well number	Difference	Carbon resource
4	14.006	$\beta$ -Cyclodextrin
5	11.933	$\gamma$ -Cyclodextrin
6	9.348	Dextrin
20	-9.267	Arbutin
25	11.455	Gentiobiose
30	10.598	$\alpha$ -Methyl-D-Galactoside
44	9.092	Xylitol
93	-9.253	Dihydroxy Acetone

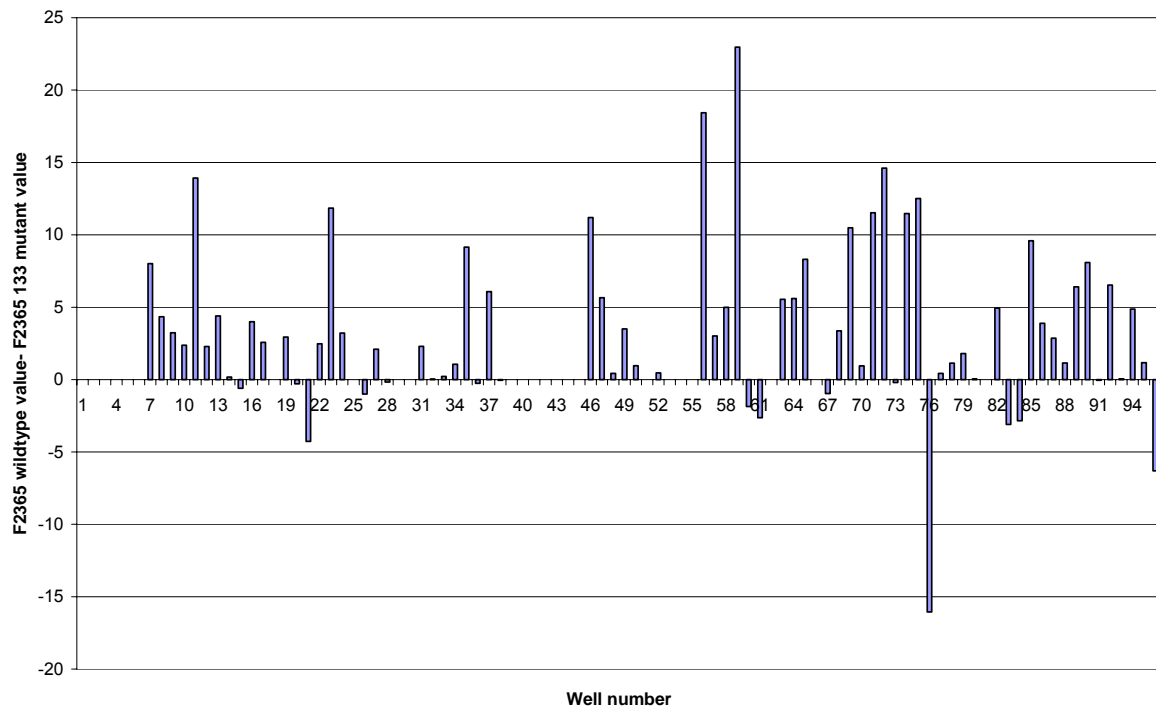


Fig. 15 PM3 nitrogen sources MicroPlate Value comparison of F2365 wild type Vs. mutant F2365  $\Delta$ 133. X axis indicated well number in 96-well plate, Y axis indicated the value of F2365 wild type value minus mutant F2365  $\Delta$ 133 value. Well NO.1 was negative control.

Table 5. PM3 nitrogen sources with significant value difference between F2365 wild type and mutant F2365  $\Delta$ 133.

Well number	Difference	Nitrogen resource
11	13.927	L-Cysteine
23	11.849	L-Threonine
46	11.196	Ethylenediamine
56	18.438	D-Glucosamine
59	22.967	N-Acetyl-D-Glucosamine
69	10.489	Thymidine
71	11.53	Uridine
72	14.61	Inosine
74	11.472	Xanthosine
75	12.509	Uric Acid
76	-16.053	Alloxan

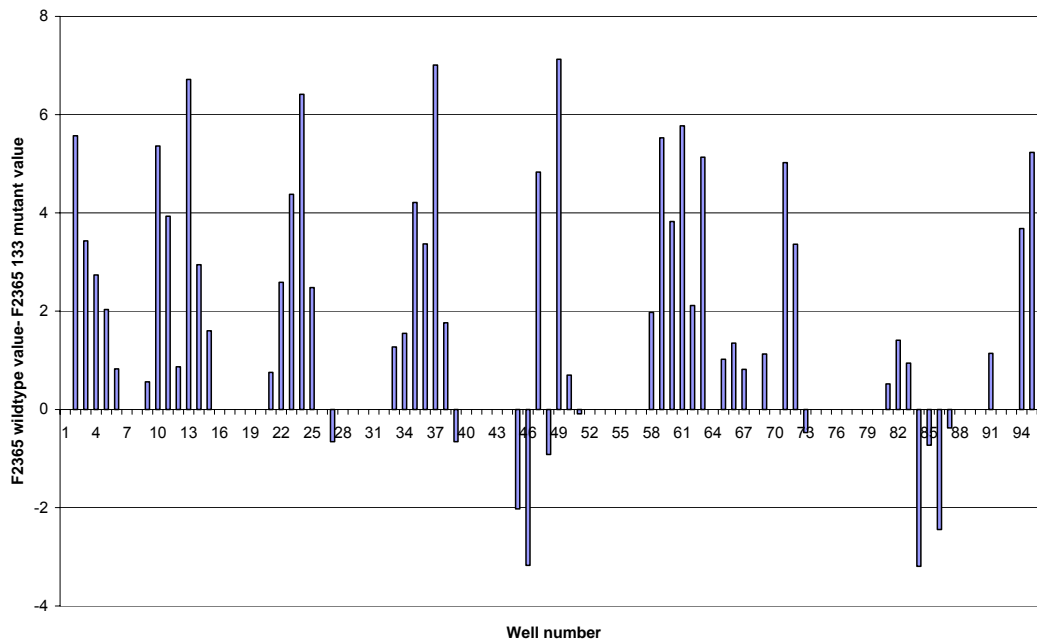


Fig. 16 PM4 Phosphorus & Sulfur sources MicroPlate Value comparison of F2365 wild type Vs. mutant F2365  $\Delta$ 133. X axis indicated well number in 96-well plate, Y axis indicated the value of F2365 wild type value minus mutant F2365  $\Delta$ 133 value. Well NO.1 was negative contro

### 3.5 Discussion

In this study, a deletion of the ECI specific region 133 was constructed in ECI strain F2365 by homologous recombination. Phenotypic characterization of the deletion mutant, F2365  $\Delta$ 133, revealed that it was indistinguishable from its wild type parental strain in terms of basic bacteriologic characteristics (hemolytic activity, cell morphology, motility, phage susceptibility, serotype 4b-specific surface antigen expression). At 37°C and 25°C mutant and wild type grew similarly. However, the growth of the mutant F2365  $\Delta$ 133 was impaired at 4°C, which indicated that one or more of the genes in region 133 may contribute to the organism's ability to grow at low temperature. Even though growth at cold temperatures is a typical characteristic of *L. monocytogenes* and other *Listeria* spp., it is possible that some of the determinants contributing to cold growth may differ among different strains. It is noteworthy that of the six ECI specific ORFs in region 133, three encoded putative membrane proteins. It is possible that the absence of membrane proteins normally encoded by the genes in this cassette may adversely affect membrane structure and function, especially at low temperature. Even though the deletion did not abolish the ability of the bacteria to grow in the cold, the observed reduction in cold growth suggests that this gene cassette may contribute to the food safety threat that ECI bacteria pose. Cold tolerance, which allows *L. monocytogenes* to grow in high risk (highly processed, cold stored, ready-to-eat) foods at refrigerator temperature, is a key factor for the ability of *L. monocytogenes* to cause food-borne outbreaks.

The mutant F2365  $\Delta$ 133 was surveyed by Phenotypic Microarray technology, and the metabolism of carbon, nitrogen, Phosphorus and Sulfur Sources were tested. The utilization

of 20 carbon sources, especially Uridine, and 11 nitrogen sources, especially N-Acetyl-D-Glucosamine and Alloxan, were altered significantly by the deletion of the ECI specific genes in region 133, whereas no significant changes were noticed in the utilization of phosphorus and sulfur sources. The products of the ECI specific genes in region 133, including the putative ABC transporter and the three putative membrane proteins may work together to mediate transport of certain carbon and nitrogen sources, thus affecting the metabolism of the organism. It is conceivable that the structural changes responsible for the impaired growth at low temperature are also mediating the altered metabolic reactions detected by the phenotypic microarrays.

At this time, the possible impact of the mutation in virulence or pathogenicity of the organism is not known. The antibiotics sensitivity and biofilm formation of F2365 and deletion mutant are going to be tested; and F2365 and the deletion mutant are going to be investigated using an oral murine infection model developed by Drs. Monica Boruki and Cristina Cunha, at Washington State University. Other animal and cell culture models will also need to be employed to further characterize the possible impact of this gene cassette in interactions of the bacterium with potential host cells, and in overall virulence.

### 3.6 References

1. Herd, M., and C. Kocks. 2001. Gene fragments distinguishing an epidemic-associated strain from a virulent prototype strain of *Listeria monocytogenes* belong to a distinct functional subset of genes and partially cross-hybridize with other *Listeria* species. *Infect Immun* 69:3972-9.
2. Kathariou, S., C. Mizumoto, R. D. Allen, A. K. Fok, and A. A. Benedict. 1994. Monoclonal antibodies with a high degree of specificity for *Listeria monocytogenes* serotype 4b. *Appl Environ Microbiol* 60:3548-52.
3. Nelson, K. E., D. E. Fouts, E. F. Mongodin, J. Ravel, R. T. DeBoy, J. F. Kolonay, D. A. Rasko, S. V. Angiuoli, S. R. Gill, I. T. Paulsen, J. Peterson, O. White, W. C. Nelson, W. Nierman, M. J. Beanan, L. M. Brinkac, S. C. Daugherty, R. J. Dodson, A. S. Durkin, R. Madupu, D. H. Haft, J. Selengut, S. Van Aken, H. Khouri, N. Fedorova, H. Forberger, B. Tran, S. Kathariou, L. D. Wonderling, G. A. Uhlich, D. O. Bayles, J. B. Luchansky, and C. M. Fraser. 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* 32:2386-95.
4. Riemann, L., and F. Azam. 2002. Widespread N-acetyl-D-glucosamine uptake among pelagic marine bacteria and its ecological implications. *Appl Environ Microbiol* 68:5554-62.
5. Schuch, R., A. Garibian, H. H. Saxild, P. J. Piggot, and P. Nygaard. 1999. Nucleosides as a carbon source in *Bacillus subtilis*: characterization of the *drm-pupG* operon. *Microbiology* 145 ( Pt 10):2957-66.

6. **Scogin, D. A., H. I. Mosberg, D. R. Storm, and R. B. Gennis. 1980. Binding of nickel and zinc ions to bacitracin A. *Biochemistry* 19:3348-52.**
7. **Tsuda, H., Y. Yamashita, Y. Shibata, Y. Nakano, and T. Koga. 2002. Genes involved in bacitracin resistance in *Streptococcus mutans*. *Antimicrob Agents Chemother* 46:3756-64.**

## Appendix A PCR reaction system in Reverse Transcripton PCR

Component	Final concentration	Volume per 100 ul reaction
DEPC water		X ul
10x thermophilic polymerase reaction buffer (without MgCl <sub>2</sub> )	1x	8.0 ul
MgCl <sub>2</sub> , 25mM	2mM	5.6 ul
dNTP mix, 10mM	0.2mM	1 ul
Upstream primer	1 uM	X ul
Downstream primer	1 uM	X ul
Promega <i>Taq</i> DNA polymerase	5.0u	1.0 ul
PCR mix volume per reaction		80 ul
Volume of cDNA added		20 ul
Total PCR volume		100 ul

## Appendix B Colony PCR reaction system

10X Ex Taq <sup>TM</sup> Buffer	2ul
dNTP Mixture	1.6ul
Template	Add single colony to reaction mix by picking the colony with either a sterile loop or pipette tip and twirling this in the liquid in the tube
Primer 1 (10pmol)	0.5ul
Primer 2 (10pmol)	0.5ul
TaKaRa Ex Taq <sup>TM</sup>	0.125ul
H <sub>2</sub> O	Up to 20ul

## Appendix C Alignment in 133 region among *L. monocytogenes* strains

### F2365, EGD-e and *L. innouca*

Multiple sequence alignment among *L. monocytogenes* strains F2365, EGD-e and *L. innouca*, Start and stop codons were high lighted, primers P1, P5, P6 and P7 were high lighted, restriction sites HindII, BamHI and EcoRI were underlined, repetitive sequences were italic marked.

```
(ORF2632 start codon)
F2365 2632-2625  ATGGCAAAGAAAAATTGACCGCTCTAAACCCCATGTTAACATTGGTACTATTGGACAC 60
EGD-e 2653-2652  ATGGCAAAGAAAAATTGACCGCTCTAAACCCCATGTTAACATTGGTACTATTGGACAC 60
L.innocua 2802-2801 ATGGCAAAGAAAAATTGACCGCTCTAAACCCCATGTTAACATTGGTACTATTGGACAC 60
*****

F2365 2632-2625  GTTGACCATGGTAAACAACCTTTAACTGCTGCAATTACAACCTGACTTGCTAAAAAAGGC 120
EGD-e 2653-2652  GTTGACCATGGTAAACAACCTTTAACTGCTGCAATTACAACCTGACTTGCTAAAAAAGGC 120
L.innocua 2802-2801 GTTGACCATGGTAAACTACTTTAACTGCTGCAATTACAACCTGACTTGCTAAAAAAGGC 120
*****

(P1 including HindIII site)
F2365 2632-2625  TATGCTGATGCACAAGCTTATGACCAAATTGATGGTCTCCGGAAGAAAGAGAACGTGGT 180
EGD-e 2653-2652  TATGCTGATGCACAAGCTTATGACCAAATTGATGGTCTCCGGAAGAAAGAGAACGTGGG 180
L.innocua 2802-2801 TTTGCTGATGCACAAGCTTATGATCAAATTGATGGTCTCCGGAAGAAAGAGAACGTGGT 180
* *****

F2365 2632-2625  ATCACAATCTCTACTGCTCACGTTGAGTACCAAACCTGACAGCCGTCACATATGCACACGTT 240
EGD-e 2653-2652  ATCACAATCTCTACTGCTCACGTTGAGTACCAAACCTGACAGCCGTCACATATGCACACGTT 240
L.innocua 2802-2801 ATCACAATCTCTACTGCTCACGTTGAGTACCAAACCTGACAACCGTCACATATGCACACGTT 240
*****

F2365 2632-2625  GACTGCCCAGGACATGCCGATTACGTTAAAAACATGATCACTGGTCTGCACAAATGGAC 300
EGD-e 2653-2652  GACTGCCCAGGACATGCCGATTACGTTAAAAACATGATCACTGGTCTGCACAAATGGAC 300
L.innocua 2802-2801 GACTGCCCAGGACATGCCGATTACGTTAAAAACATGATCACTGGTCTGCACAAATGGAC 300
*****

F2365 2632-2625  GGAGCTATCTTAGTAGTATCTGCTGCTGATGGCCCAATGCCACAAACTCGTGAACATATC 360
EGD-e 2653-2652  GGAGCTATCTTAGTAGTATCTGCTGCTGATGGCCCAATGCCACAAACTCGTGAACATATC 360
L.innocua 2802-2801 GGAGCTATCTTAGTAGTATCTGCTGCTGATGGCCCAATGCCACAAACTCGTGAACATATC 360
*****

F2365 2632-2625  TTACTTTCACGTCAAGTTGGTGTTCATACATCGTTGTATTTCATGAACAAATGTGACATG 420
EGD-e 2653-2652  TTACTTTCACGTCAAGTTGGTGTTCATACATCGTTGTATTTCATGAACAAATGTGACATG 420
L.innocua 2802-2801 TTACTTTCACGTCAAGTTGGTGTTCATACATCGTTGTATTTCATGAACAAATGTGACATG 420
*****

F2365 2632-2625  GTTGACGATGAAGAATTACTAGAATTAGTTGAAATGGAAATTCGTGATCTATTAAC TGAA 480
EGD-e 2653-2652  GTTGACGATGAAGAATTACTAGAATTAGTTGAAATGGAAATTCGTGATCTATTAAC TGAA 480
L.innocua 2802-2801 GTTGACGATGAAGAATTACTAGAATTAGTTGAAATGGAAATTCGTGATCTATTAAC TGAA 480
*****

F2365 2632-2625  TATGAATTCCTGGCGATGACATTCTGTAAATCAAAGGTTTCAGCTCTTAAAGCATTCAA 540
EGD-e 2653-2652  TATGAATTCCTGGCGATGACATTCTGTAAATCAAAGGTTTCAGCTCTTAAAGCATTCAA 540
L.innocua 2802-2801 TATGAATTCCTGGCGATGACATTCTGTAAATCAAAGGTTTCAGCTCTTAAAGCATTCAA 540
*****

F2365 2632-2625  GGTGAAGCTGACTGGGAAGCTAAAATTGACGAGTTAATGGAAGCTGTAGATTCTTACATT 600
EGD-e 2653-2652  GGTGAAGCTGACTGGGAAGCTAAAATTGACGAGTTAATGGAAGCTGTAGATTCTTACATT 600
L.innocua 2802-2801 GGTGAAGCTGACTGGGAAGCTAAAATTGACGAGTTAATGGAAGCTGTAGATTCTTACATT 600
*****

F2365 2632-2625  CCAACTCCAGAACGTGATACTGACAAACCATTTCATGATGCCAGTTGAGGATGTATTCTCA 660
EGD-e 2653-2652  CCAACTCCAGAACGTGATACTGACAAACCATTTCATGATGCCAGTTGAGGATGTATTCTCA 660
L.innocua 2802-2801 CCAACTCCAGAACGTGATACTGACAAACCATTTCATGATGCCAGTTGAGGATGTATTCTCA 660
*****
```

F2365 2632-2625 ATCACTGGTCGTGGAACAGTTGCAACTGGACGTGTTGAACGTGGACAAGTTAAAGTTGGT 720  
EGD-e 2653-2652 ATCACTGGTCGTGGAACAGTTGCAACTGGACGTGTTGAACGTGGACAAGTTAAAGTTGGT 720  
L.innocua 2802-2801 ATCACTGGTCGTGGAACAGTTGCAACTGGACGTGTTGAACGTGGACAAGTTAAAGTTGGT 720  
\*\*\*\*\*

F2365 2632-2625 GACGAAGTAGAAGTTATCGGTATGGAAGAAGAAAGCAAAAAGTAGTAGTAAGTGGAGTA 780  
EGD-e 2653-2652 GACGAAGTAGAAGTTATCGGTATGGAAGAAGAAAGCAAAAAGTAGTAGTAAGTGGAGTA 780  
L.innocua 2802-2801 GACGAAGTAGAAGTTATCGGTATGGAAGAAGAAAGCAAAAAGTAGTAGTAAGTGGAGTA 780  
\*\*\*\*\*

F2365 2632-2625 GAAATGTTCCGTAAATTACTAGACTACGCTGAAGCTGGCGACAACATTGGCGCACTTCTA 840  
EGD-e 2653-2652 GAAATGTTCCGTAAATTACTAGACTACGCTGAAGCTGGCGACAACATTGGCGCACTTCTA 840  
L.innocua 2802-2801 GAAATGTTCCGTAAATTACTAGACTACGCTGAAGCTGGCGACAACATTGGCGCACTTCTA 840  
\*\*\*\*\*

F2365 2632-2625 CGTGGTGTGCTCGTGAAGATATCCAACGTGGTCAAGTATTAGCTAAACCAGGTTTCGATT 900  
EGD-e 2653-2652 CGTGGTGTGCTCGTGAAGATATCCAACGTGGTCAAGTATTAGCTAAACCAGGTTTCGATT 900  
L.innocua 2802-2801 CGTGGTGTGCTCGTGAAGATATCCAACGTGGTCAAGTATTAGCTAAACCAGGTTTCGATT 900  
\*\*\*\*\*

F2365 2632-2625 ACTCCACACACTAACTTCAAAGCTGAAACTTATGTTTAACTAAAGAAGAAGGTGGACGT 960  
EGD-e 2653-2652 ACTCCACACACTAACTTCAAAGCTGAAACTTATGTTTAACTAAAGAAGAAGGTGGACGT 960  
L.innocua 2802-2801 ACTCCACACACTAACTTCAAAGCTGAAACTTATGTTTAACTAAAGAAGAAGGTGGACGT 960  
\*\*\*\*\*

F2365 2632-2625 CACACTCCATTCTTCAACAACCTACCGCCCAATTCTATTTCCGTACTACTGACGTAAC 1020  
EGD-e 2653-2652 CACACTCCATTCTTCAACAACCTACCGCCCAATTCTATTTCCGTACTACTGACGTAAC 1020  
L.innocua 2802-2801 CACACTCCATTCTTCAACAACCTACCGCCCAATTCTATTTCCGTACTACTGACGTAAC 1020  
\*\*\*\*\*

F2365 2632-2625 GGTATTGTTACACTTCCAGAAGTACTGAAATGGTAATGCCTGGTGATAACATTGAGCTT 1080  
EGD-e 2653-2652 GGTATTGTTACACTTCCAGAAGTACTGAAATGGTAATGCCTGGTGATAACATTGAGCTT 1080  
L.innocua 2802-2801 GGTATTGTTACACTTCCAGAAGTACTGAAATGGTAATGCCTGGTGATAACATTGAGCTT 1080  
\*\*\*\*\*

F2365 2632-2625 GCAGTTGAACTAATTGCACCAATCGCTATCGAAGACGGTACTAAATTCTCTATCCGTGAA 1140  
EGD-e 2653-2652 GCAGTTGAACTAATTGCACCAATCGCTATCGAAGACGGTACTAAATTCTCTATCCGTGAA 1140  
L.innocua 2802-2801 GCAGTTGAACTAATTGCACCAATCGCTATCGAAGACGGTACTAAATTCTCTATCCGTGAA 1140  
\*\*\*\*\*

(ORF2632 stop codon)

F2365 2632-2625 GGCGGACGTACAGTAGGCGCTGGCGTTGTTTCTAACATCAGCAAAATAATCTGATACGA 1200  
EGD-e 2653-2652 GGCGGACGTACAGTAGGCGCTGGCGTTGTTTCTAACATCAGCAAAATAATCTGATACGA 1200  
L.innocua 2802-2801 GGCGGACGTACAGTAGGCGCTGGCGTTGTTTCTAACATCAGCAAAATAATCTGATACGA 1200  
\*\*\*\*\*

(P5 including BamHI site)GGATCC

F2365 2632-2625 TTCAAACCGACAAAGTCATTTGGCTTTGTCGGTTTTTTTGTCTAAAAAAGCTGAATAAT 1260  
EGD-e 2653-2652 TTCAAACCGACAAAGTCATTTGGCTTTGTCGGTTTTTTTGT----- 1241  
L.innocua 2802-2801 TTCAAACCGACAAAGTCATTTGGCTTTGTCGGTTTTTTTGT----- 1241  
\*\*\*\*\*

F2365 2632-2625 CTAGTAATAAAGAATTAGTTCATATTTCTAAAAAATTAGATATATACAATTTTTGATAAT 1320  
EGD-e 2653-2652 -----  
L.innocua 2802-2801 -----

(ORF2631 start codon)

F2365 2632-2625 TATGCTATATTTGGAATGCACATAAATAATTACAGAGGAGAGAGAATGAAATGAAAAAG 1380  
EGD-e 2653-2652 -----  
L.innocua 2802-2801 -----

F2365 2632-2625 AAATAGTTGTTTTAGGGTTAGCGGGTGCAATATTTTTATCAGGAATGGGAATCTTAAAC 1440  
EGD-e 2653-2652 -----  
L.innocua 2802-2801 -----

F2365 2632-2625 GTAAGCGCAGCAAATTATTCGGATACAAAGTTTAGTTTACTACTAGGTAAACTTGGTGCA 1500  
EGD-e 2653-2652 ----- 1248  
L.innocua 2802-2801 ----- 1248

F2365 2632-2625 AATGATTACACTGGTTCTAGACAAAAACAGAACACTACATCATCTTATGTGAAATTAAT 1560  
EGD-e 2653-2652 -----  
L.innocua 2802-2801 -----

F2365 2632-2625 TCAATTGGAAAAGGAACAATGGATACATGGCTTTTAAAAATCTAATGGAGCTAGCGTTCGA 1620  
EGD-e 2653-2652 -----  
L.innocua 2802-2801 -----

F2365 2632-2625 AGTAAATATGTGACTGTGAGGCAAGGTGAAAGCAAAAAAATTGCTAACTACGCTTATGAA 1680  
EGD-e 2653-2652 -----  
L.innocua 2802-2801 -----

F2365 2632-2625 GATTATGGCAAATGTAATGTTAAATTAGCTGCTGAACTAGTAAACTCAATTTGTTAGA 1740  
EGD-e 2653-2652 -----  
L.innocua 2802-2801 -----

F2365 2632-2625 GTTACTGCTACAGGCCTTTGGAGTCCAGATAGTATTTAA(ORF2631 Stop codon)  
EGD-e 2653-2652 -----  
L.innocua 2802-2801 -----

Omit sequence of F2365 from ORF2630 to ORF2627

F2365 2632-2625 (ORF2626 start codon)ATGAAAAAATACTTGGCGGA 1800  
EGD-e 2653-2652 -----  
L.innocua 2802-2801 -----

F2365 2632-2625 ATCTTAATTATTTTTGTAATAGGCCTGTTTGCTTGGCGGTATATGATGTAAAAGCTAAT 1860  
EGD-e 2653-2652 -----  
L.innocua 2802-2801 -----

F2365 2632-2625 TCTTTTTCATATGAAAATAAGGCACATGCTGAACAGGAAAAATTTCAACTTGAAGCTCA 1920  
EGD-e 2653-2652 -----  
L.innocua 2802-2801 -----

F2365 2632-2625 ACAATCAGTGCTGGAAAAGCATTGTTGTAAGTGATGCAGACTTAAATAAATATGTGACA 1980  
EGD-e 2653-2652 -----  
L.innocua 2802-2801 -----

F2365 2632-2625 AAAGATTATTTTAAGCAGGAAAATAAGACTCTTTTGCTCGTTCAATTAGAATCTACAGAA 2040  
EGD-e 2653-2652 -----  
L.innocua 2802-2801 -----

F2365 2632-2625 AAAGATATAAGAATTTCTGATTTTCAGTTAGGGTATAAAGAATTTGTAACGTTGTCAGAT 2100  
EGD-e 2653-2652 -----  
L.innocua 2802-2801 -----

F2365 2632-2625 ACTTCGGCAGCAAGTTATGAATTTGAAGATGGGGTGTATAAAATGGTGTAGGTTTTAAT 2160  
EGD-e 2653-2652 -----

L.innocua 2802-2801 -----

F2365 2632-2625 ATACCTAAAGAATTACTAGCAACTAACAAGACATTCACATTAGTTACTCCGAGTAAATAT 2220  
EGD-e 2653-2652 -----  
L.innocua 2802-2801 -----

(ORF2626 stop codon)

F2365 2632-2625 TGGAAGAACGGAGCAAGGGATGTTGTGGAATCAGTTTATAAAAAAGTAGCTTTATTCAAA 2280  
EGD-e 2653-2652 -----  
L.innocua 2802-2801 -----

GGATCC(P6 including BamHI site)

F2365 2632-2625 CCGACAAAGTCATTTGGCTTTGTTCGGTTTTTTTGCCTGAAATAAAAAATTAACCTTGAAA 2340  
EGD-e 2653-2652 -----ATACAAAAATAAAAAATATACTTGAAA 1267  
L.innocua 2802-2801 -----GCCTGAAATAAAAAATTAACCTTGAAA 1266  
\*\*\*\*\* \*\*

F2365 2632-2625 AATATCACATTTTCGCATAACCTTAAAGTAGACATATCTTTTTACCTTCTAGACCGAAAT 2400  
EGD-e 2653-2652 AATATCACATTTTCGCATAACCTTAAAGTAGACATATCTTTTTACCTTCTAGACCGAAAT 1327  
L.innocua 2802-2801 AATATCACATTTTCGCATAACCTTAAAGTAGACATATCTTTTTACCTTCTAGACCGAAAT 1326  
\*\*\*\*\*

(ORF2625 start codon)

F2365 2632-2625 CAAGGAGGTAGGCCAAGTGGTACAATTTGATGCTCGAAATATGGCGTTGCTCGAATCACT 2460  
EGD-e 2653-2652 CAAGGAGGTAGGCCAAGTGGTACAATTTGATGCTCGAAATATGACATTACTCGAATCGCT 1387  
L.innocua 2802-2801 CAAGGAGGTAGGCCAAGTGGTACAATTTGATGCTAGAAATATGATGCTGCTCGAATCGCT 1386  
\*\*\*\*\* \* \*\*\*\*\* \*\*

F2365 2632-2625 CGTTGTGGCGAATGTATATCTTGCACCCGAGAAATTACAAGAAGAACTAGGAATTTCCAA 2520  
EGD-e 2653-2652 CGTCGTGGCGAATGTATATCTTGCACCCGAGAAATTACAAGAAGAACTAGGAATTTCCAA 1447  
L.innocua 2802-2801 CGTCGTGGCGAATGTATATCTTGCACCCGAGAAATTACAAGAAGAACTAGGCATTTCCAA 1446  
\*\*\* \*\*\*\*\*

F2365 2632-2625 ACGAACACTTCAATATGATGTGAAAAAATAATAAAGAATTAGATAATATAGGACTCGA 2580  
EGD-e 2653-2652 ACGAACACTTCAATATGATGTAGAAAAAATAATAAAGAATTAGATAATATAGGACTTGA 1507  
L.innocua 2802-2801 ACGAACACTACAATATGATGTTGAAAAAATAATAAAGAATTAGATGATATAGGGCTTGA 1506  
\*\*\*\*\* \*\*\*\*\* \*\* \*\*

F2365 2632-2625 TGGTATTCAATCTGTTCTGTTGACAAAGGGTATTATTTATTAGAAGATGAAAAATCGACAAT 2640  
EGD-e 2653-2652 TGGTATTCAATCCGTTCTGTTGACAAAGGGTATTATTTATTAGAAGAGAAAAAACGACAAT 1567  
L.innocua 2802-2801 TGGTATTCAATCCGTTCTGTTGACAAAGGGTATTATTTATTAGAAGATGAAAAATCGACAAT 1566  
\*\*\*\*\* \*\*\*\*\*

F2365 2632-2625 TAAAGAAATCCTTGAAATAGGGAAGCGAGCCACAAAGTCTTTTCAGCAAGTGAGCGTCTG 2700  
EGD-e 2653-2652 TAAAGAGATCCTTGAAATAGGGAAGCGAGCCACAAAGTCTTTTCAGCTAGTGAACGCCG 1627  
L.innocua 2802-2801 TAAAGAAATCCTTGAAATAGGGAAGCGAGCCACAAAGTCTTTTCAGCAAGTGAGCGTCTG 1626  
\*\*\*\*\* \*\* \*\*\*\*\* \*\* \*\*

F2365 2632-2625 CATCCGTATTTTATTTTCTGCTCGTAACAGATGCGCGAGTAATCATTGATACGATTAA 2760  
EGD-e 2653-2652 CATCCGTATTTTATTTTCTGCTCGTAACAGATGCGCGAGTAATATCGATACGATTAA 1687  
L.innocua 2802-2801 TATCCGTATTTTATTTTCTGCTCGTAACAGATGCGCGTGTAAATATTGATACAATTAA 1686  
\*\*\*\*\* \*\*\*\*\* \*\* \*\* \*\*

F2365 2632-2625 TGAATGCAATGAAGTCAGTCGTAATACAAGTTTGCAGGACATTAACAATTAATAATTAGC 2820  
EGD-e 2653-2652 TGAATGCAATGAAGTCAGTCGTAATACCAGTTTACAGGATATTAACAATTAATAATTAGC 1747  
L.innocua 2802-2801 TGCGTGCAATGAAGTTAGTCGTAACACTAGTTTACAGGATATTAACAAGTTAAATAATTAGC 1746  
\*\* \*\*\*\*\* \*\* \*\*\*\*\*

F2365 2632-2625 GCTAAAACAGTTTAATTTAGAACTTTCGCTATGACCGTAAAAATGGCAATATGGTTCTCGG 2880  
EGD-e 2653-2652 GCTAAAACAGTTTAATTTAGAGCTCGCTTATGATCGTAAAAATGGGAATATGGTTCTGGG 1807  
L.innocua 2802-2801 GCTCAAACAGTTTAACCTTAGAGCTCGCTTATGACCGGAAAAATGGGAATATGGTTCTTGG 1806  
\*\*\* \*\*\*\*\* \*\* \*\* \*\*\*\*\* \*\* \*\*\*\*\*

F2365 2632-2625 TGATGAGCGCAGTATTCGCCAATTTTTTATTCACTATTGTATGAATAACGAAGAAATCGC 2940  
EGD-e 2653-2652 TGATGAGCGAGTATACGCCAATTTTTTATTCACTATTGTATGAATAACGAAGAAATCGC 1867  
L.innocua 2802-2801 TGATGAACGTAGTATTCGTCAGTTTTTCACTATTGTATGAATAACGAAGAAATAGC 1866  
\*\*\*\*\* \*\* \*\*\*\*\* \*\* \*\*\*\*\*

F2365 2632-2625 AACCGCAGACCAGTTGCTCGATTTAATGAAAATTAATCCGATGATAAAAAATCAGGAGTT 3000  
EGD-e 2653-2652 AACCGCGGACCAGTTGCTCGATTTAATGAAAATTAATCCGATGATAAAAAATCAGGAGTT 1927  
L.innocua 2802-2801 CACGGCAGACCAGTTACTCGATTTAATGAAAATAAATCCAATGATTAATAAATAAAGAACT 1926  
\* \* \* \* \*

F2365 2632-2625 ATTCCCAAATTAGATACGATTTTCGAAATTTTAGCAGTAACGAGAAAAAATCGGAAT 3060  
EGD-e 2653-2652 ATTCCCAAATTAGATATGATTTTCGAAATTTTAGCAGTCACAGAGAAAAAATCGGTAT 1987  
L.innocua 2802-2801 TTTCCACATTTAGATACAATTTTGAATATTAGCAGTGACAGAGAAAAAATCGGTAT 1986  
\* \* \* \* \*

(P7 including EcoRI site)

F2365 2632-2625 TCGTTATACTGACGAGGTCCTTGAGCGCATTGGTATTATGATTTTCTTCTTTAAAGAACG 3120  
EGD-e 2653-2652 TCGTTACACAGATGAAGTTATTGAGCGTATTGGTATTATGATTTTCTTCTTTAAGGAACG 2047  
L.innocua 2802-2801 TCGTTATACTGACGAGGTCATTGAACGAATCGGAATTATGATTTTCTTCTTTAAGGAACG 2046  
\* \* \* \* \*

F2365 2632-2625 AATGAAGCGTAATTGCTATTTAAATGAACAAGAAGAGCATGAAGTAGAATCTTTTGAGAT 3180  
EGD-e 2653-2652 AATGAAACGCAATTGTTATTTAAATGAGAAAAGAAGACATGAAGTAGAATCATTGGTAT 2107  
L.innocua 2802-2801 AATGAAACGCGATTGTTATTTAAATGAGCACGAAGAACATGAAGTAGAATCTTTAATAT 2106  
\* \* \* \* \*

F2365 2632-2625 TGCGCAGGAAATTTATCAGCAATTGCAGCAAAGTGAGAATTTTAAATCAATCATGCGGA 3240  
EGD-e 2653-2652 TGCGCAGGAAATTTATCAGCAATTGCAGCAAAGCGAGAATTTTAAATTAATCATGCGGA 2167  
L.innocua 2802-2801 TGCTCAGGAAATTTACGAGCAATTACAGCAAAAAGAGCATTTTAAGATCAATCATGCGGA 2166  
\* \* \* \* \*

F2365 2632-2625 AATCACTTATTTAGGTAACCTTTTGCTTGGGGCGAGTCGTTTGAATGATGATGCGGCTGC 3300  
EGD-e 2653-2652 AATCACTTATTTAGGTAACCTTTTGCTTGGTGCAGTCGTTTGAATGATGATGCGGCTGC 2227  
L.innocua 2802-2801 AATTACTTATTTAGGCAACCTTTTACTTGGCGCTAGTCGTTTGAATGATGATGCGGCTGC 2226  
\* \* \* \* \*

F2365 2632-2625 TGAAGGAAAATTAGATCTCATCGTTGAAAAAATCATCGCCGAATTTGAACGCTTGCTTG 3360  
EGD-e 2653-2652 TGAAGGAAAATTAGATCTAATCGTTGAAAAAATCATGCTGAATTTGAACGCTTGCTTG 2287  
L.innocua 2802-2801 TGAAGGAAAGCTGGATATAATTGTTGAGAAAATTATTGCCGAATTTGAGCGCCTTGCTTG 2286  
\* \* \* \* \*

F2365 2632-2625 CGTGAATTTTGAAGACCATCGCAATTTGAAAAAAGATTTATTGCTTCATTTACAACCAGC 3420  
EGD-e 2653-2652 TGTGAATTTTGAAGACCATCGCAATTTGAAAAAAGATTTATTACTCCATTTACAACCAGC 2347  
L.innocua 2802-2801 TGTGAATTTTGAAGACCATCGTAGTTTAAAAAAGATTTATTGCTCCATTTACAACCAGC 2346  
\* \* \* \* \*

F2365 2632-2625 CTACTATCGGCTTAAATTCCAAATGAGTGGATTAACCCGCTAAGGACAGACATTAACA 3480  
EGD-e 2653-2652 ATACTACCGTCTTAAATTCCAAATGAGTGGATTAACCCGCTGCGTACAGACATTAACA 2407  
L.innocua 2802-2801 ATATTATCGGCTTAAATTCCAAATGAAATGGATTAATCCGCTGCGAACAGATATTAAGCA 2406  
\* \* \* \* \*

F2365 2632-2625 AAGTTACAGTGATGTGTACGAGATTACGAAAAAATCACTCGAACCGCTAGAAGATTTACT 3540  
EGD-e 2653-2652 AAGTTACAGTGATGTGTACGAGATTACGAAAAAATCACTCGAACCGCTAGAAGATTTGCT 2467  
L.innocua 2802-2801 AAGTTATAGTGACGTGTATGAGATTACAAAAAATCATTAGAACCCTAGAAGATTTACT 2466  
\* \* \* \* \*

F2365 2632-2625 TGGTGAACAATTCAGAGGATGAAATAGCGTACGTAACAATCTTGTTCGGTGGCTATCT 3600  
EGD-e 2653-2652 TGGTGAAGCAATACCAGAGGATGAAATAGCGTACGTAACAATCTTGTTCGGCGGCTATCT 2527  
L.innocua 2802-2801 CGGTGAAAAAATCCCCGAAGATGAGATAGCGTACGTAACGATTTTGTTCGGCGGTTATCT 2526  
\* \* \* \* \*

F2365 2632-2625 TTCACGCAAAAAATAACTTTAGTTGAACGAAAAAATCTTAAATCGTTTGTCTCAAAGG 3660  
EGD-e 2653-2652 TTCACGCAAAAAATAACTTTAGTTGAACGAAAAAATCTTAAATCGTTTGTCTCAAAGG 2587  
L.innocua 2802-2801 TTCTCGCAAAAAATAACTTTAGTTGAACGAAAAAATCTTAAATCGTTTGTCTCAAAGG 2586  
\* \* \* \* \*

F2365 2632-2625 TGTCGGAACGTCGCGGATGATTGAACGGCAATTATCGCAATTACTTGGTGAGCGAGTAGA 3720  
EGD-e 2653-2652 CGTCGGAACGTCGCGGATGATTGAACGGCAATTATCGCAATTACTTGGTGAAACGAGTAGA 2647  
L.innocua 2802-2801 GGTAGGAACGTCGCGGATGATTGAGCGGCAACTTGCACAATTACTCGGTGAACGTTGTA 2646  
\* \* \* \* \*

F2365 2632-2625 GATATTAGAACCAATTTCCATCCGTGAATTTGAAAAAGGGTTATACGCACCAGACTTTAT 3780  
EGD-e 2653-2652 GATATTAGAACCAATTTCCATCCGTGAGTTTGAAAAAGGGCTGTATGCACCAGATTTTAT 2707  
L.innocua 2802-2801 AATATTAGAACCAATTTCCATCCGTGAATTTGAAAAAGGGTTATACGCACCAGACTTTAT 2706  
\*\*\*\*\*

F2365 2632-2625 TGTGTCGACTTTGCCAATTATGGAACCAAAAGCACCAGTTTTTATCGTTAGTCCAATTTT 3840  
EGD-e 2653-2652 CGTATCTACTTTGCCAATTATGGAACCAAAAGCACCAGTTTTTATTGTTAGTCCGATTTT 2767  
L.innocua 2802-2801 TGTGTCGACTTTGCCAATTATGGAACCGAAAGCACCAGTTTTTATCGTTAGTCCGATTTT 2766  
\*\* \* \*

F2365 2632-2625 GACAGAAGCGCAGAAACAACAACCTGATGAAGGCGATTGCCCCGCACATTTTGCAAAAGGA 3900  
EGD-e 2653-2652 AACAGAAGCGCAGAAAGCAACAATTGATGAAGGTAATTGCTCCGCATATTTTACAAAAAGA 2827  
L.innocua 2802-2801 GACTGAGGCGCAGAAACAACAACCTGATGAAGGCGATTGCCCCGCACGTTTTTGCAAAAGGA 2826  
\*\* \* \*

F2365 2632-2625 CTCGGATGCGCGCATGTTGTCTTCTGTGCTTGATGTAGTAGATCAATATGCCAAAGTGGGA 3960  
EGD-e 2653-2652 TTCTGATGCGCGCATGTTATCGTCCGTGCTCGATGTAGTAGATCAATATGCCAAAGTGGGA 2887  
L.innocua 2802-2801 TTCGGATGCGCGCATGTTGTCTTCTGTGCTTGATGTAGTGGATCAATATGCCAAAGTGGGA 2886  
\*\* \* \*

F2365 2632-2625 AGACCGCGAGAAAATTGGCGGCAAACTGAAATCGGTATTGTTCCAAGTGAATCAGACAG 4020  
EGD-e 2653-2652 AGACCGCGAAAAATTAGCGGCCAAGCTAAAGTCCGATATTATTTCAAGTGAATCAGACAG 2947  
L.innocua 2802-2801 AGACCGCGAAAAATTGGCGGCCAAGCTGAAGTCCGATATTGTTCCAAGTGAATCAGACAG 2946  
\*\*\*\*\*

F2365 2632-2625 CCAACTAGAAAAATCTCCAACACTCGAAGAACTTTTACCGAAAGAACGAATTACTTTTAA 4080  
EGD-e 2653-2652 CCAGCTAGAGAAATCGCCATCTCTTGATGAACTATTACCGCAAGAACGAATTATTTTAA 3007  
L.innocua 2802-2801 CCAACTAGAAAAATCGCCAACACTGGAAGAACTGTTGCCAAAGAACGAATTATTTTAA 3006  
\*\* \* \*

F2365 2632-2625 AGAAAGTGTGGCGGACTGGCGTGAAGCTATCCTAGTTGCCCTCGAAATCGTTCGCAACAAGA 4140  
EGD-e 2653-2652 AGAAAGTGTGACGGACTGGCAAGAAGCTATCCGAGTTGCTTCAAAACCACTACAGCAAGA 3067  
L.innocua 2802-2801 AGATAGTGTGGCGGATTGGCATGAAGCTATCCGAGTTGCGTCAAAACCACTACAACAAGA 3066  
\*\* \* \*

F2365 2632-2625 AGGCTACATTTTCGAGAAATTATCAGCATGCGATGATTGAAAATATTGAGAAACTCGGACC 4200  
EGD-e 2653-2652 AGGTTATATCTCGAAGAAGCTATCAGCATGCAATGATTGAAAATATTGAGAAACTCGGACC 3127  
L.innocua 2802-2801 AGGCTATATATCAGGAAACTATCAACAAGCTATGATTGAAAACATTGAAAAGCTAGGACC 3126  
\*\* \* \*

F2365 2632-2625 ATACATCGTTATTGCGCCAGGGATTGCACCTGCCACATGCGTCTGTGGATGACGGGGCATA 4260  
EGD-e 2653-2652 ATATATCGTTATTGACCAGGGATTGCACCTCCCATGCTTCCGTTAGATGACGGGGCATA 3187  
L.innocua 2802-2801 ATATATCGTTATTGACCAGGGATAGCACTGCCACACGCGTCCGTTGGATGACGGGGCATA 3186  
\*\* \* \*

F2365 2632-2625 CCGAGTCGGAATGAGCTTACTGCGTTTAAATCAGCCAGTATCATTTTCAAGTAAAGCGAA 4320  
EGD-e 2653-2652 CCGTGTGGAATGAGCTACTACGGTTAGACCAGCCGTTTCATTTTCAAGTAAAGCGAA 3247  
L.innocua 2802-2801 TCGAGTCGGAATGAGCTTACTGCGTTTAAATCAGCCAGTATCATTTTCAAGTAAAGCGAA 3246  
\*\* \* \*

F2365 2632-2625 AGATCAAGTGAAGTTAATTATTGTGCTCGCTTCCATTGACTCCTACACATATTAATGC 4380  
EGD-e 2653-2652 GGATCAGGTTAAATTAATTATTGTGCTCGCTTCCATTGACTCCTACACATATTAATGC 3307  
L.innocua 2802-2801 GGATCAAGTGAAGTTAATTATTGTGCTCGCTTCCATTGACTCCTATACACATATTAATGC 3306  
\*\*\*\*\*

F2365 2632-2625 GCTGAGCCAACCTACTAATTTAATCATGAAACATCACTTGCTAGAGCAGATTGAACAAGC 4440  
EGD-e 2653-2652 ACTTAGCCAACCTACTAATTTAATCATGAAACATCACTTGCTAGAGCAGATTGAACAAGC 3367  
L.innocua 2802-2801 GCTAAGCCAGCTTACTAATTTAATTTATGAAACATCACTTACTTGAGCAGATTGAACAAGC 3366  
\*\* \* \*

(ORF2625 stop codon)

F2365 2632-2625 GACATCAGCAGCGGAAATTGCCGCAATGTTAACGATACAATAA 4483  
EGD-e 2653-2652 GACATCAGCCGCGGAAATTGCCGCAATGTTAACCATAAAATAA 3410  
L.innocua 2802-2801 CGAATCAGCAGCAGAAATTGCCGCAATGTTAACATAAAAATAA 3409  
\*\*\*\*\*

### Appendix D Restriction system

Plasmid	10ul	Upstream	10μl	Downstream	10μl
pCon-1		fragment		fragment	
Buffer2	6μl	10xBufferE	6μl	10xBufferMC	6μl
EcoRI	1.5μl	BamHI	1.5μl	BamHI	1.5μl
HindIII	1.5μl	HindIII	1.5μl	EcoRI	1.5μl
H2O	41μl	H2O	41μl	H2O	41μl
Total	60μl	Total	60μl	Total	60μl

### Appendix E Ligation system

10x ligase buffer	2μl	Negative control:	
pCON-1	3ul		
Upstream fragment	6ul	10x ligase buffer	2μl
Downstream fragment	6ul	pCON-1	3ul
H2O	2μl	H2O	14μl
T4 Ligase	1μl	T4 Ligase	1μl
Total	20μl	Total	20μl

## Appendix F Working primers used in this study

Primer Name	Oligo sequence: 5'-3'
PA	GAAAACTGGGATAAAGAGCAAGAA
PB	ATGTAAATTTTGTACCGGGGTATG
PC	TGAATTGCGACAAACATATCAAG
PD	TGTAAATTTTGTACCGGGGTATG
PE	GTGATGACTTTTTCTTCCACGAC
Pa	ATATGTAGTTTTATTGAGTCCACTAGCTATTTA
Pb	TACTAGCCCAATAATTACAAAATAATAAATCAA
Pc	TGCAGCTAATTTCCAATCTGTTTA
Pd	AGCCATTATGACCTATGAACGAAT
Pe	CACTAGGTAAACTTGGTGCAAATG
Pf	TGCTACGAAATAATCAAGTGGTGT
Pg	AAGTTTGCCAGTTTTTCTATTTGG
Ph	TCCATAATAGCTTGAGCAATACCA
Pi	GTGGGATTTATTCGTTTCATAGGTC
Pj	CACTTACAACAAATGCTTTTCCAG
Pm	GTACTACTGACGTAAGTGGTATTGTTACACTT
Pn	GAATACCATCGAGTCCTATATTATCTAATTCT
P1	CTCCAGGAGCTGACAAAAGC
P2	GCTTGCTCCCGATTCAAATA
P3	CCTCAACAAGTTCGGGGTTA
P4	GGTTGCTAGCAGACTTCGAGA
P5	ACTATCCTTATGAAACGGCGTAAG
P6	GGAATTACAGTGATGACCATTTTG
P7	ATGATTGTTGCGGGATTAGG
P8	AACGCTTTTCACTCTTTTACACC
P9	TAGCCAAACAAATAGGGAGCAT
P10	TACATAAAACGCGGCAGAGATA
P11	ACTTGAAGCTTTAAATGGCGTTAC
P12	GTTCTTCCAGGTGTTCTGAAAGTT